STRUCTURAL AND MECHANISTIC STUDIES OF

NUCLEIC ACID DEMETHYLASE FTO

TOH DEWEI JOEL

NATIONAL UNIVERSITY OF SINGAPORE

2017

STRUCTURAL AND MECHANISTIC STUDIES OF

NUCLEIC ACID DEMETHYLASE FTO

TOH DEWEI JOEL

B.Sc (Hons.), Nanyang Technological University

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
NUS GRADUATE SCHOOL FOR INTEGRATIVE
SCIENCES AND ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2017

Supervisors:
Professor Wanjin Hong, Main Supervisor
Associate Professor Yong-Gui Gao, Co-Supervisor
Dr Esther Woon, Co-Supervisor

Examiners:
Professor Haiwei Song
Professor Jayaraman Sivaraman
Associate Professor Yuguang Mu, Nanyang Technological University

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

_____

Toh Dewei Joel

Acknowledgements

This journey is not possible without all the help out I received along the way. First, I like to give thanks to my supervisor Professor Wanjin Hong. I like to thank him for allowing me to be part of the IMCB family and the help and scientific advices he has given me over the past four years. Next, I like to show my sincerest gratitude to my co-supervisor Associate Professor Yong-Gui Gao. I am deeply appreciative as he provided me the platform for experimentations, the room to explore my scientific curiosity and at the same time, close guidance to keep me on track over the past four years. I would not have achieved what I have achieved during my PhD training if not for his patience and continuous support. I also like to thank Dr Esther Woon for allowing me to participate in her research project, which complemented my work, making this thesis possible.

I would also like to thank my TAC members Professor Haiwei Song and Prof Vinay Tergaonkar for taking out their valuable time to give me scientific advice and monitoring of my research progress.

I also want to thank Prof Gao's lab members for making my four years enjoyable. Particularly, I like to thank Dr Yun Chen for providing me with the invaluable mentorship when I first picked up protein purification and protein crystallisation. I would also like to thank Dr Veerendra Kumar for his guidance in X-ray crystallography structure solvation.

I would like to thank A*STAR Graduate Academy for funding my scholarship. I like to acknowledge NGS allowing me to partake in their graduate school program. I also like to thank IMCB and Nanyang Institute of Structure Biology (NISB), NTU for allowing me to use their state-of-the-art facility for my research work. I like to thank

Lastly, I would like to thank my family for the endless support and patience afforded to me over the past four years.

TABLE OF CONTENTS

Summary

FTO gene was the first gene that provided the strongest direct link to obesity uncovered by genome-wide association studies. Functional FTO is crucial for normal growth and development. Later on, FTO was identified as an m6A demethylase. M6A modification was known to be a prevalent mark across all kingdoms of organisms. In mammals, m6A occurred within a consensus sequence and was initially believed to be a permanent post-transcription modification in RNA. The ability of FTO to reverse m6A methylation prompted revisitation on the role of m6A in RNA transcripts. The status of m6A in RNA transcripts are further dynamically regulated by methyltransferases and m6A-binding proteins. This affects RNA splicing, translation and degradation, representing epigenetic regulation at the RNA level. FTO and ALKBH5 belonging to the AlkB 2OG oxygenases family are the only m6A demethylases known so far. They are implicated in unrelated diseases such as obesity for FTO and spermatogenesis for ALKBH5, revealing very different biological roles. It is not clear how FTO is selectively expressed over ALKBH5 and how FTO regulate the RNA methylation status *in vivo*. Thus, the mechanistic and structural studies of FTO will aid in the elucidation of FTO roles and functions.

We developed and reported the first potent, selective and cell-active inhibitor for FTO through a combined approach of X-ray crystallography, differential scanning fluorimetry (DSF) and biochemical assays. Human AlkB oxygenases have conserved 2OG-binding site but not substrate-binding site. This is supported by DSF which revealed m3T substrate only preferentially binds to FTO. Encouraged by this, we envisioned that selectivity and potency could be achieved by exploiting the substrate-binding site. The inhibitor we have developed has at least 30-fold selectivity over other human AlkB oxygenases. Particularly, the inhibitor demonstrated a 100-fold

preference for FTO (IC$_{50}$ = 0.81 µM, T$_m$ shift = 11.2°C) over ALKBH5 (IC$_{50}$ = 108.1 µM; T_m = 2.7°C). The co-crystal structure of FTO complexed to inhibitor solved at atomic resolution revealed the basis of selectivity. The position of the pyridyl nitrogen of the inhibitor and Glu234$_{FTO}$ is crucial for the selectivity. This is important as we can potentially, selectively perturb FTO in cells. We further validated that the cell-activity of the inhibitor. Potentially, this inhibitor can serve as a functional probe for mechanistic studies of FTO in cells.

Next, we asked what determined the substrate specificity of FTO. Through *in vitro* demethylation studies, we concluded that FTO does not need m6A-containing consensus sequence for activity. Various secondary structural conformations are adopted by mRNA. Although m6A do not prevent duplex formation, the methylation was found to destabilise Watson-Crick base pairing by 0.5-1.7 kcal/mol depending on sequential context. We purposefully designed various m6A-containing RNA oligonucleotides that are highly similar in sequences that were able to form duplexes or hairpins *in vitro*. We showed that FTO and ALKBH5 are able to selectively demethylate m6A-containing hairpins. Demethylation caused a conformational change of the corresponding RNA oligonucleotides, adopting a duplex conformation. With this validation, it is fair to ask whether FTO performs its biological role by affecting secondary structures of m6A-containing mRNA.

List of tables

List of figures

List of abbreviations

2OG	2-oxoglutarate
2,4-PDCA	Pyridinedicarboxylic acid
5mC	5-methylcytosine
BMI	Body mass index
CD	Circular dichroism
DNA	Deoxyribonucleic acid
DSBH	Double-stranded β-helix
DSF	Differential scanning fluorimetry
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
fm6A	N^6-formyladenosine
FTO	Fat mass obesity-associated protein
GST	Glutathione S-transferase
GWAS	Genome-wide association study
HATs	Histone acetyltransferases
HDACs	Histone deacetylases
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	Hypoxia-inducible transcription factor
HMTs	Histone methyltransferases
HPLC	High performance liquid chromatography
hm6A	N^6-hydroxymethyladenosine
JmjC	Jumonji domain-containing
lncRNA	Long noncoding RNA

LLG	Log-likelihood gain
m1A	N^1-methyladenine
m1G	N^1-methylguanine
m3C	N^3-methylcytosine
m3T	N^3-methylthymidine
m6A	N^6-methyladenosine
METTL	Methyltransferase-like
miRNA	microRNA
MMS	Methylmethane sulphonate
MR	Molecular replacement
MST	Microscale thermophoresis
NOG	N-oxalylyglycine
NRL	Nucleotide recognition loop
PEG	Polyethylene glycol
PHD	Prolyl hydroxylase domain-containing protein
PDB	Protein data bank
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RRM	RNA-recognition motif
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNP	Single nucleotide polymorphism
TBE	Tris/borate/EDTA
Tris	Tris(hydroxymethyl)aminomethane
TET	Ten-eleven translocation

TFZ	Translation function Z
UTR	Untranslated region
WTAP	Wilm's tumor 1-associating protein
XDS	X-ray Detector Software
T_m	Melting temperature
Å	Angstrom (10^{-10} m)
β-ME	2-mercaptoethanol

CHAPTER 1

LITERATURE REVIEW

1.1 Nonheme-Fe(II)/2-oxoglutarate (2OG)-dependent dioxygenases

1.1.1 Overview

Oxygenases that utilise ferrous iron are prevalent in nature and are utilised by organisms carrying out aerobic activities [1, 2]. Oxygenases are enzymes that use atmospheric diatomic oxygen and incorporate one or two oxygen atoms into the substrates. Monooxygenases, which incorporate one oxygen atom into their substrates, are arguably the first major class of oxygenases to be characterised and studied. An example is cytochrome P450, widely found in mammals, plants, insects, yeast and bacteria, oxidises a variety of functional groups through ferrous diatomic oxygen complex in the heme prosthetic group.

Nonheme dioxygenases represent another class of oxygenases that perform oxidation by incorporating both atoms of diatomic oxygen into their substrates [3]. Nonheme-Fe(II)/2-oxoglutarate (2OG)-dependent dioxygenases are a superfamily of oxidising enzymes utilised by nature and are found in organisms including mammals, plants and bacteria (Fig.1) [4-10]. Just like cytochrome P450, 2OG oxygenases are able to perform a variety of oxidation reactions on a wide range of substrates. The first 2OG oxygenase, collagen proline hydroxylase, was discovered in 1967 by Hutton *et al*. despite its prevalence in nature [11]. Prior to the report by Hutton *et al*., it was thought that hydroxyproline of collagen was incorporated by hydroxyproline amino acid [12]. With the advancement of genome sequencing technology, more 2OG oxygenases were identified in diverse organisms. In addition, the biological functions of some human 2OG oxygenases have been intensely studied.
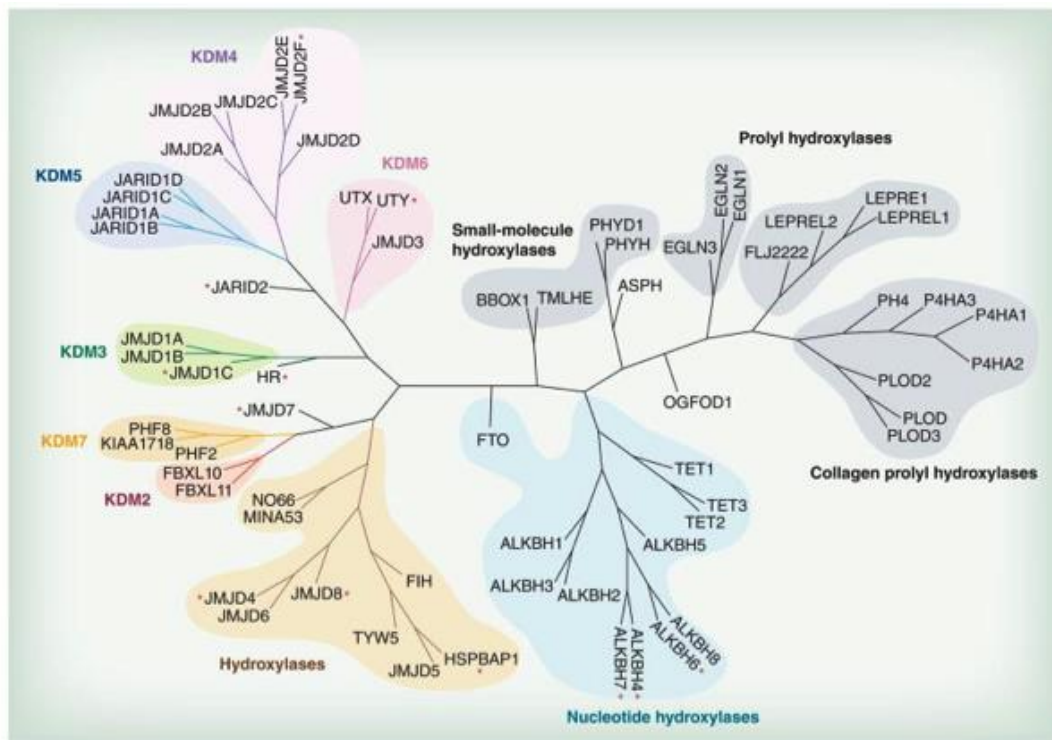
Figure 1. Phylogenetic tree of human 2OG oxygenases. Despite low sequence identity, human ALKBH1-8, FTO and TET1-3 share highly similar structural elements and conserved residues. (Picture adapted from 'The roles of Jumonji-type oxygenases in human disease' [10].)

The α,β-heterodimeric hypoxia-inducible transcription factor (HIF) binds to various target genes upon exposure to low ambient oxygen environment, inducing responses to counter the anaerobic environment [13-16]. Prolyl hydroxylase domain-containing proteins (PHDs) are 2OG oxygenases and play a key role in hypoxic response. PHD hydrolyses the proline residues 402 and 564 of HIFα and therefore marks HIF for degradation via the ubiquitin-proteosome pathway. PHD acts as direct oxygen sensor, whereby in an ambient oxygen environment, HIF is not needed and thus degraded.

Another family of human 2OG oxygenases (about 30 subfamily members) are the human jumoji domain-containing (JmjC)-type 2OG oxygenases [13, 14, 17, 18]. Most JmjC oxygenases are known to play a role in epigenetic regulation by oxidising

methylated lysine and arginine residues of histones. Histones modification via methylation is able to affect the chromatin structure and thus affect downstream gene expression [18, 19]. As such, JmjC oxygenases are of interest in cancer therapeutics.

The ten eleven translocation (TET) family of proteins is composed of 3 known subfamily members of human 2OG oxygenases [14, 20, 21]. Tet has a smaller scope of substrate specificity and is only known to oxidise 5-methylcytosine and its derivative of DNA and RNA [22]. The Tet family proteins are believed to play an epigenetic role in embryonic stem cell pluripotency and cancer.

The last major family of human 2OG oxygenases is the AlkB family of proteins [14, 23-25]. This family consists of 9 subfamily members (Fig. 1). Some of the subfamily members are well characterised and they are known to play diverse biological roles, for instance, cancer, fertility and obesity. AlkB family of proteins can oxidise a range of methylated DNA and RNA nucleotides. Fat mass and obesity-associated protein (FTO), an obesity related protein, is one member of the AlkB subfamily [26]. The structural and mechanistic study of substrate recognition by FTO is the major subject of my PhD study depicted in this thesis.

1.1.2 The double-stranded β-helix (DSBH) domain

The double-stranded β-helix (DSBH) fold is a characteristic structural feature of 2OG oxygenases (Fig. 2) [27-29]. DSBH fold is synonymous with JmjC, jelly-roll and cupin fold, these alternative names were derived from different communities of scientists. The DSBH core possesses an asymmetrical barrel-like structure, with one end of the barrel bigger than the other. The structure is made up of eight β-strands forming a pair of four-stranded antiparallel β-sheets. The antiparallel β-sheets are sometimes numbered in Roman numerals, from I to VIII and further classified into major and minor β-sheets. The major β-sheet generally has longer amino acid chains consisting of β-strands I, VIII, III and VI, whereas the minor β-sheet has shorter amino acid chains consisting of β-strands II, VII, IV and V.

It is common to observe the presence of α-helixes preceding the DSBH fold of 2OG-oxygenases. The α-helixes are generally situated on the exterior of the DSBH core and are believed to contribute to the stability of DSBH. In addition, β-strands at the N-terminal of DSBH fold could participate in substrate selection and binding. Within the DSBH fold, loops and short helical inserts connect between β-strands. Particularly, the βIV/V insert that is located between β-strand IV and V displays 2OG-oxygenases subfamily dependent characteristics. Note that βIV/V inserts vary in length, and play various roles among 2OG oxygenases such as acting as dimerisation interface and substrate interaction (Fig. 2B) [30-32]. For instance, the βIV/V insert of mammalian TET2 DSBH contains 379 amino acids and is highly disordered [33]. It is situated at the exterior of the DSBH fold surface and does not affect the demethylation activity of TET2. The precise function of the βIV/V insert of mammalian TET2 remains unknown. On the other hand, the βIV/V insert of PHD2, a subfamily member of PHD family, consists less than 10 amino acids in length.
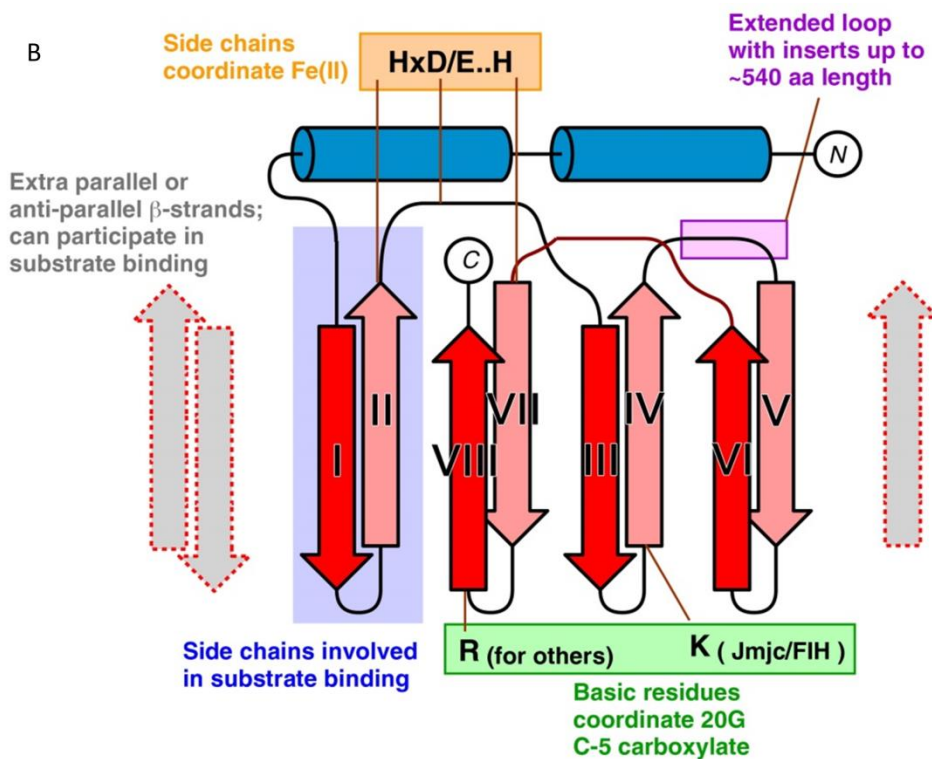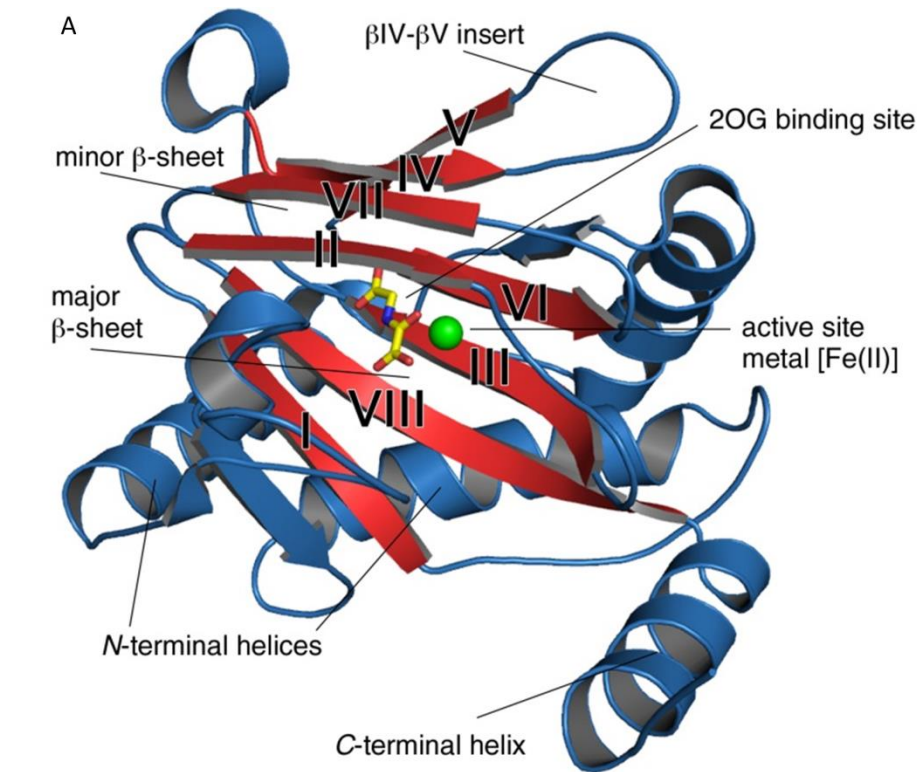
Figure 2. The double-stranded β-helix (DSBH) domain. A. Crystal structure depicting the typical DSBH jelly-roll motif made up of eight β-strands. **B.** 2D topology diagram for 2OG oxygenases. (Picture is adapted from 'Structural studies on human 2-oxoglutarate dependent oxygenases' [28].)

The crystal structures of 2OG oxygenases solved so far ubiquitously showed that the catalytic sites are located at the bigger opening end of the DSBH fold, where the substrates approach the 2OG oxygenases [4, 27-29]. Co-factor and co-substrate, ferrous iron and 2-oxoglutarate respectively, are located in the catalytic site [34]. The ferrous iron is coordinated by the iron binding amino acid motif, a HxD/E...H triad (x being any amino acid residue), which is generally conserved across the 2OG oxygenases family (Fig. 3) [35, 36]. The histidine is located at the end of β-strand II and an aspartate or glutamate residue is resided one amino acid away in the loop that connects the β-strands II and III. The second ferrous iron binding histidine is located at the start of the β-strand VII. The characteristic fold of the DSBH is such that the triad of amino acid residues are brought in close proximity, ligand to the ferrous iron metal center, forming one side of a triangular face of the octahedron.

Co-substrate 2OG coordinates to the ferrous iron in a bidentate manner through the 1-carboxylate and 2-oxo groups. The 1-carboxylate is observed to be *trans* to either one of the histidine residues while the 2-oxo group is *trans* to the aspartate or glutamate residue in the available crystal structures. However, for some 2OG oxygenases that perform halogenation, the aspartate or glutamate residue proved not to be crucial for catalysis [37, 38]. The 5-carboxylate group of 2OG that points towards the smaller opening end of the DSBH fold interacts via electrostatic and hydrogen bonding with basic residues, such as a lysine (for JmjC family) located at the start of β-strand IV, or an arginine (for PHD and AlkB family) located at the start of β-strand VIII [39, 40]. The 5-carboxylate group is further stabilised by an alcoholic residue from the surrounding such as serine/threonine or a tyrosine, further enriching the hydrogen bonding network.

βI   L1 loop region   βII

```
FTO      .VTLLNFMDPQKMPYLKEEPYFGMGKMAVSWHHDEN..LVD   238
AlkB     DACLINRYAP.GA..............KLSLHQDKDE.PDL   139
ALKBH2   .FVLINRYKDGCD.............HIGEHRDDERELAP    180
ALKBH3   NSLLCNLYRNEKD.............SVDWHSDDEPSLGR    200
ALKBH5   .SAVINDYQP.GG.............CIVSHVDPIH..IF    211
```

βIII   βIV

```
FTO      RSAVAVYSYSCEGPEEESEDDSHLEGRDPDIWHVGFKIS..   277
AlkB     RAPIVSVSLG.L...................PAIFQFGG...   158
ALKBH2   GSPIASVSFG.A...................CRDFVFRHK..   200
ALKBH3   CPIIASLSFG.A...................TRTFEMRKK..   220
ALKBH5   ERPIVSVSFF.S...................DSALCFGCKFQ   233
```

βV   βVI

```
FTO      ....WD.I.........E.....TPGLAIPLHQGDCYFMLDD   300
AlkB     ...LK..R.............NDPLKRLLLEHGDVVVWGGE   181
ALKBH2   ....DSRG........KSPSRRVAVVRLPLAHGSLLMMNHP   229
ALKBH3   ....P..PPEENGDYTY.....VERVKIPLDHGTLLIMEGA   250
ALKBH5   FKPIR..V.............SEPVLSLPVRRGSVTVLSGY   259
```

βVII   βVIII

```
FTO      LNATHQHCVLAG........SQPRFSSTHR.........   322
AlkB     SRL.FYHGIQPLKAGFHPLTIDCRYNLTFRQAGKKE...   216
ALKBH2   TNTHWYHSLPVRKKV....LAPRVNLTFRKILLT....   259
ALKBH3   TQADWQHRVPKEYHS....REPRVNLTFRTVYP.....   279
ALKBH5   AADEITHCIRPQ......DIKERRAVIILRKTRLDAPRL   292
```

Figure 3. Structure-based sequence alignment of AlkB oxygenases. Alignment depicting the DSBH domains of various crystallised AlkB oxygenases. The eight β-strands (βI-VIII) that made up the DSBH domains are relatively conserved, making up of similar amino acids except for βIV. Only FTO has a L1 loop (highlighted green). The amino acids between βIII and βIV are disordered for FTO (underlined in red) and absent in other homologs. All AlkB oxygenases shown here have short βIV/V inserts (5-12 amino acids in length). Columns marked with asterisk (*) depict conserved residues that bind to 2OG. Catalytic triad HxD…H are involved in chelation to Fe(II) while arginine is involved in hydrogen bonding to the 5-carboxylate group of 2OG The alignment was performed by T-coffee and the diagram was prepared by ESPript 3.0 [35, 36].

1.1.3 Catalytic cycle

Presence of diatomic oxygen is essential for the catalysis of 2OG dioxygenases [27-29, 41]. At the start of the catalytic cycle, the ferrous iron is at the rest state, where it is ligand to the HxD/E...H triad motif with three water molecules *trans* to each of the amino acid, forming a octahedral coordination geometry (Fig. 4). Co-substrate 2OG then comes in and displaces two water molecules (step 1). Co-crystallised structures of 2OG oxygenases with 2OG or *N*-oxalylglycine (NOG) are often observed in this conformation whereby the metal centre is hexa-coordinated by the triad motif, the bidentate coordination of 2OG or NOG and water molecule [31-33, 39, 42]. The water molecule is displaced by diatomic oxygen upon the onset of the catalytic reaction (step2).

The 2OG oxygenase superfamily is able to oxidise a wide range of substrates [8, 9, 11, 22, 23, 43, 44]. The proposed consensus catalytic mechanism that is discussed below is for *N*-methylated substrates that have direct relevance to this thesis [45-47]. Binding of diatomic oxygen leads to the formation of superoxo radical anion (O_2^-) coordinating to the oxidised Fe(III). The anionic diatomic oxygen then performs a nucleophilic attack at the α-carbon of 2-oxo group of 2OG forming a bridged peroxo intermediate (step 3). After which, decarboxylation occurs and two oxygen atoms from the diatomic oxygen are each incorporated into the succinate by-product and Fe(IV)-oxo intermediate (step 5). Upon formation of Fe(IV)-oxo intermediate, the oxygen is brought close to the substrate for hydroxylation (step 6). The hydroxylation of substrate occurs through a radical formation mechanism whereby Fe(III) is reduced to Fe(II) with eventual incorporation of a hydroxyl group onto the substrate (step 7). Spontaneous deformylation occurs giving the demethylated substrate (step 8).

Succinate leaves the ferrous iron allowing the incorporation of co-substrates for another round of catalytic cycle.
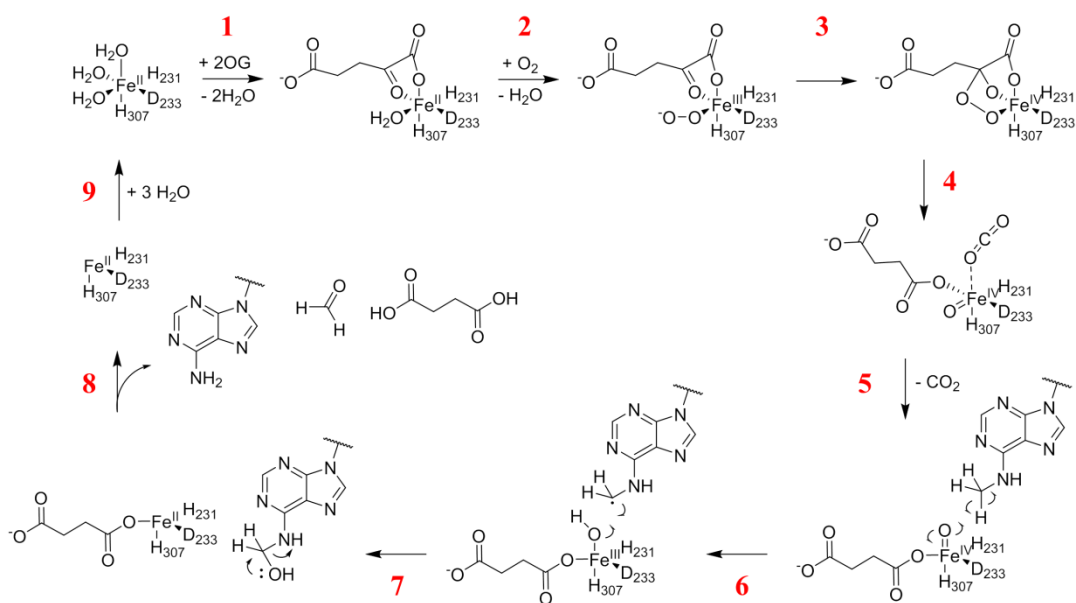
Figure 4. Catalytic cycle for *N*-demethylation. (Modified from 'Nucleic acid oxidation in DNA damage repair and epigenetics' [41].)

1.2 Types of substrates for AlkB family of 2OG oxygenases

1.2.1 Alkylated nucleobases due to oxidative damage

Genomic DNA in cells is exposed to various environmental alkylating factors that non-enzymatically introduce methylations or exocyclic adducts, which are cytotoxic [48-51]. Alkylating agents can be classified into S_N2 and S_N1 types. Commonly known S_N2 agents include methylmethane sulphonate (MMS) and methyl halides, which come from combusted biomass, cigarette smoke and industrial by-products. On the other hand, S_N1 agents include nitroso-compounds like N-methyl-N-nitrosourea, which come from reaction of nature occurring nitrite with decaying food and waste matters. Methylation by methyl radicals could also occur due to environmental factors.

Across the four classical bases in DNA, there are twelve possible sites for methylation, which include exocyclic oxygen and nitrogen and the endocyclic nitrogen (Fig. 5) [50]. In addition, the oxygen atom of negatively charged phosphate backbone of DNA is also susceptible to methylation. When the DNA bases are base paired, the oxygen and nitrogen from the nucleobases are protected from alkylation damage agents. Thus, it is believed that alkylation damages occur during transcription and translation, when ssDNA is present due to unwinding of the DNA helix.

Hydrogen bonding between complementary nucleobases in Watson-Crick base pairings can be disrupted if the hydrogen bond acceptors, which consist of a lone pair of electrons, are unavailable due to alkylation such as methylation. The inability of DNA to form proper base pairing will inevitably result in the normal functions of the cells leading to cytotoxicity [49]. Alkylated nucleobases happen to be the substrates of AlkB family of 2OG-oxygenases [52, 53]. Cells utilise the AlkB family of 2OG-

oxygenases to reverse the effect of methylation damage on their DNA and RNA. A review of the AlkB repair enzymes will be discussed later in this thesis.

C:G base pair   T:A base pair
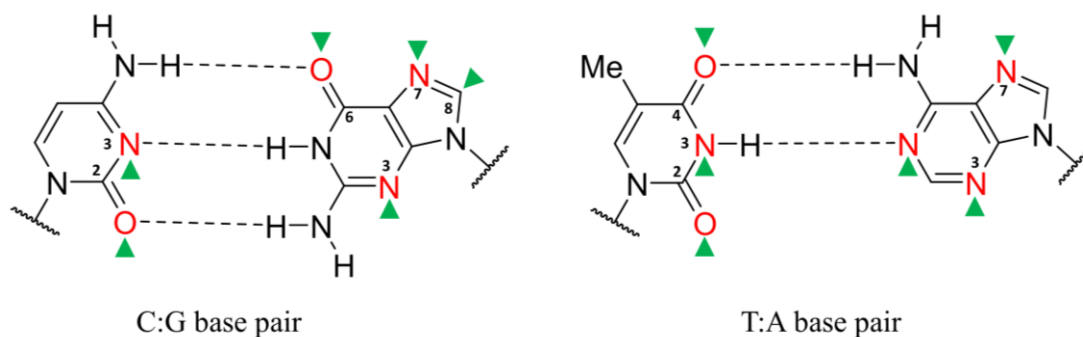
Figure 5. Twelve possible sites for alkylation damage. Atoms marked with green arrows reflect sites of that are susceptible to methylation damage. Alkylations at certain heteroatoms prevent duplex formation thus toxic to cells. (Modified from 'Repairing DNA-methylation damage' [50].)

1.2.2 Methylation due to purposed addition

Purposed chemical modifications that directly affect transcription or translation of DNA and RNA, respectively, or indirectly via histones modifications, provide an additional layer of control for downstream biological functions [54, 55]. The term 'epigenetics' is coined for the regulation of genetic expression through dynamic chemical modifications of the nucleobases that made up the genetic codes.
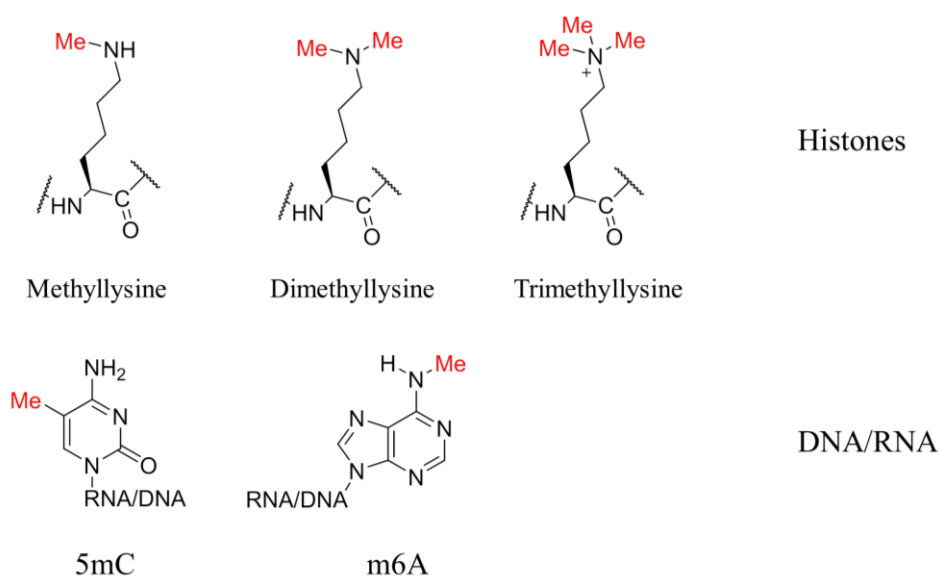
Figure 6. Methylation marks. Purposeful methylations involved in epigenetic regulation can be found in histones, DNA and RNA. (Modified from 'The control of histone lysine methylation in epigenetic regulation' [58].)

1.2.3 Methylated histones

A well-studied example is the covalent modifications by a variety of chemical functional groups on amino acids located at the N-terminal tails of the histones (Fig. 6) [18, 56-58]. Well-documented chemical modifications on histones tail include methylation, acetylation and phosphorylation. Common amino acid residues that are subjected to chemical modifications are lysine, arginine, threonine and serine. Those chemical modifications are dynamic as they can be covalently added and removed enzymatically. Acetyl functional groups are added and removed by histone acetyltransferases (HATs) and histone deacetylases (HDACs) respectively. On the other hand, methylation are added and removed by histone methyltransferases (HMTs) and (JmjC)-type 2OG oxygenases respectively. The presence of chemical modifications affects the packaging of DNA in the chromatin, asserting an effect on downstream gene expression and regulation by recruitment of various regulatory enzymes.

1.2.4 Methylated nucleobase of DNA

DNA methylation, catalysed by DNA methyltransferases [59], is known to play roles in genomic imprinting, X chromosome inactivation, embryogenesis and cancer in mammalian cells. Interestingly, DNA methylation was first known to only exist as 5-methylcytosine (5mC) in animals, plants and fungi [60]. 5mC occurs in the CpG islands, which are regions in the genome with repeating cluster of cytosine and guanine (Fig. 6) [55, 61-64]. Methylated CpG consists of about 1% of the entire DNA bases in the genome and they are mostly found methylated in a symmetrical manner together with their complementary counterpart. Approximately 60% of the human gene is believed to be regulated by promoters associated with CpG islands. 5-hydroxymethylcytosine (5hmC), the oxidised form of 5mC, was also reported to be a stable epigenetic mark of cytosine [65-67]. It was of great interest as 5mC was thought to be the only epigenetic modification. 5hmC was found to link to the maintenance of mammalian embryonic stem cells (ESCs) state. Further reports confirmed the role of 5hmC in playing a part in the differentiation potential of pluripotent ESCs through epigenetic reprogramming.

TET1 was the first member of the TET family to be identified as the 2OG oxygenase that oxidises the substrate 5mC to its product 5hmC, this finding has implicated the involvement of TET1 in epigenetic reprogramming of ESCs [65, 68]. TET2 and TET3 were subsequently reported to be able to perform a similar oxidation reaction [69]. 5hmC is further oxidised by the TET family of proteins to 5-formylcytosine (5fmC) and finally to 5-carboxylcytosine (5caC) [70-72]. 5fmC and 5caC are stable modification products that are also involved in ESCs state maintenance.

Mammalian *N*6-methyladenosine (m6A) in DNA was recently reported to be present in ESCs and is involved in epigenetic regulation by preventing the expression of genes [73]. M6A in DNA is a substrate of ALKBH1, a member of AlkB family of 2OG oxygenase.

1.2.5 Methylated nucleobase (m6A) of RNA

*N*6-methyladenosine (m6A) is the most prevalent form of methylated nucleobase in mammalian RNA, with an average of three m6A modifications per mRNA (Fig. 6) [74-76]. It is believed that m6A was first identified in mRNA of isolated mammalian cell lines in 1974 [77, 78]. Desrosier *et al*. reported that m6A was exclusively found in mRNA and not in any other RNA types such as ribosomal RNA. Subsequently, m6A was also identified in RNA of viruses, plants and yeasts [79-82]. M6A was initially believed to be a permanent post transcription modification by a methyltransferase and remain unaltered for the rest of the mRNA short half-life [83, 84]. In 2014, it was established that the methylation in the *N*6 position of adenosine in human cells is regulated by human methyltransferase-like 14 (METTL14) and METTL3 [85]. The METTL3-METTL14 heterodimer complex is regulated by Wilm's tumor 1-associating protein (WTAP) that localises the heterodimer complex to the nuclear speckles, where the target mRNA is located [86]. Earlier this year, structural and biochemical studies revealed that METTL3 contained the active domain for methylation while METTL14 structurally stabilised METTL3 for efficient catalysis, although both proteins were predicted to contain the methyltransferase domain [87-89].

Demethylation of m6A: Research on the function of mammalian m6A progressed slowly, and it was reignited when the *N*6-methylation mark was found to be removed by a member of the AlkB 2OG oxygenase family, FTO, in 2011 (Fig. 7) [26, 43]. The activity of FTO towards m6A is confined to a single stranded oligonucleotide, with almost negligible activity towards a duplex oligonucleotide. Prior to this discovery, FTO was thought to be a 3-methylthymine (3-meT) (for ssDNA) and 3-methyluricil (3-meU) (for ssRNA) demethylase [90, 91]. This is of huge significance as it shows that the *N*6-methylation can be dynamically regulated, potentially playing a role similar to that of 5mC of the CpG island. Furthermore, in 2013, ALKBH5, a member of the AlkB 2OG oxygenase family was reported to be able to demethylate m6A, sharing a same substrate as FTO [44, 92]. Like FTO, ALKBH5 is only active towards single stranded oligonucleotides. However, FTO and ALKBH5 perform oxidative demethylation through different mechanisms [93]. Oxidative demethylation of FTO goes through two oxidative derivatives of m6A. They are namely *N*6-hydroxymethyladenosine (hm6A) and *N*6-formyladenosine (f6A), both are stable species with half-life of about 3 hours. They resemble the oxidised derivatives of 5mC, namely 5hmC and 5fmC respectively. Despite resembling the cytosine counterparts, the exact biological significance of hm6A and f6A is unknown. In contrast, the presence of hm6A and f6A was not observed for the oxidative demethylation of ALKBH5. This might be due to the difference between structural alignment of the residues in FTO and ALKBH5 active domain, thus suggesting the different ways of how FTO and ALKBH5 influence their RNA substrates for downstream biological functions.
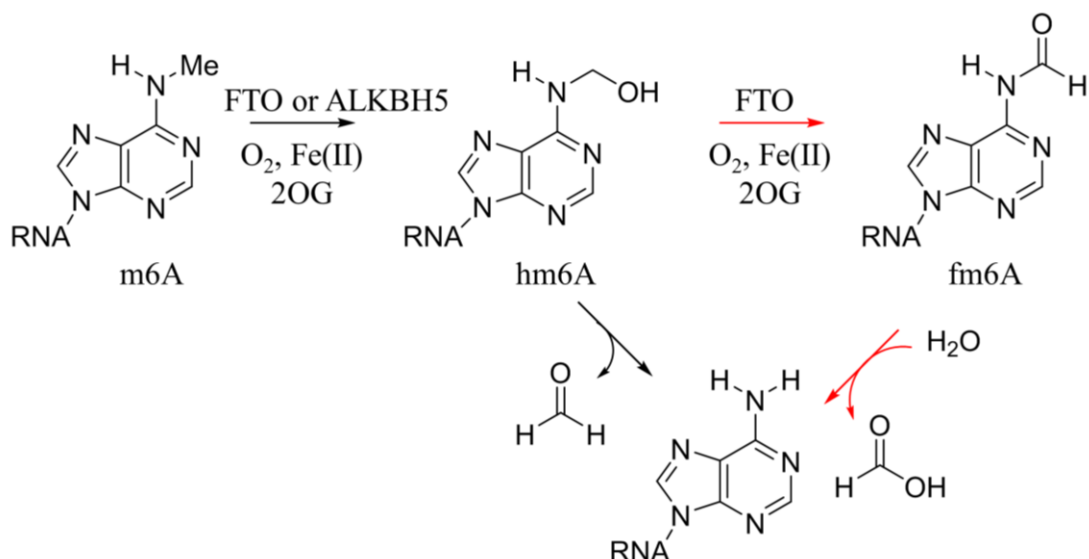
Figure 7. Oxidative demethylation of m6A. Black arrows represent the reaction pathway taken by both FTO and ALKBH5. Red arrows represent the reaction pathway that can be taken only by FTO. (Adapted and modified from 'FTO-mediated formation of *N*6-hydroxymethyladenosine and *N*6-formyladenosine in mammalian RNA' [93].)

Consensus sequence: M6A methylation marks RNA in a non-random manner and is almost exclusively found within a set of consensus sequence motif, DR(m6A)CH (where D = A, G or U; R = G or A and H = A, C or U) (Fig. 8) [74, 94-96]. M6A is unevenly distributed along a RNA transcript with increasing prevalence near the stop codon and 3' untranslated region (3' UTR) (Fig. 8A). The elucidation of the motif was largely benefitted by the advancement of high throughput sequencing, bioinformatics analysis software packages and specificity of m6A antibodies (Fig. 8B and C) [97, 98]. The general strategy undertook was fragmentation of mRNA, followed by immunoprecipitation of m6A-containing RNA and finally, sequencing. Nevertheless, credit also must be given to pioneers whom reported m6A mostly occurring in G(m6A)C and A(m6A)C motifs, shedding light on non-random distribution of m6A [99, 100]. Not all consensus motifs will be methylated at the same time in RNA transcripts. This represents the dynamic nature of m6A modifications.
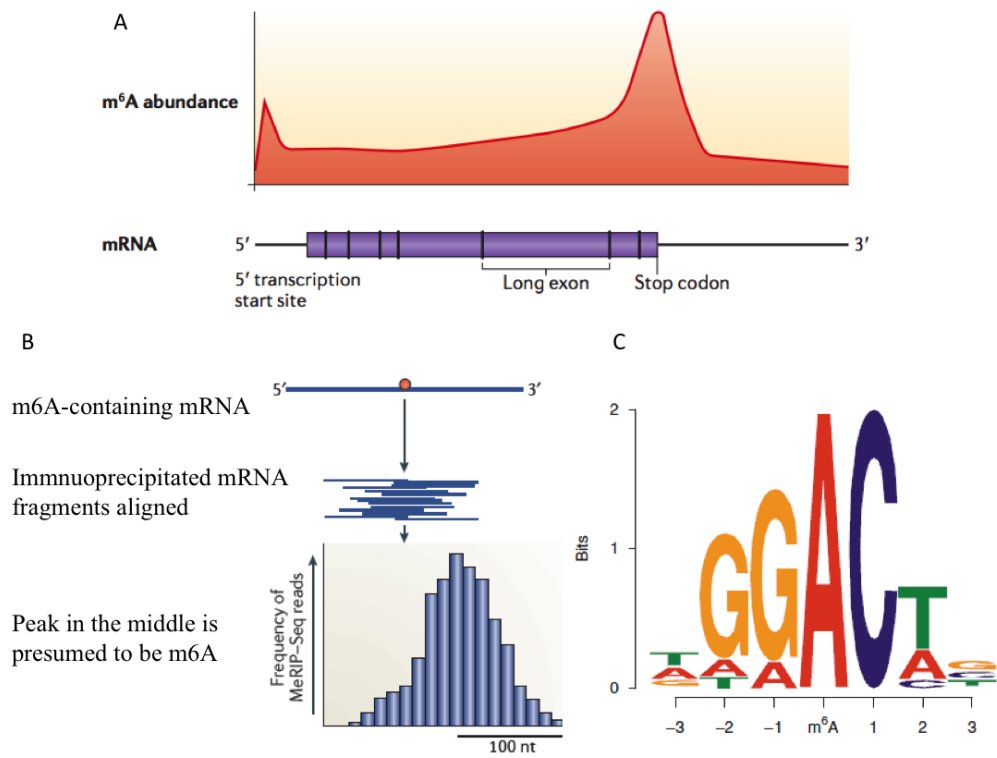
Figure 8. M6A profiling in RNA transcript. A. Prevalence of m6A in the stop codon region in a RNA transcript. **B.** Key steps in m6A-immunoprecipitation sequencing. **C.** Sequence motifs containing m6A peaks. Size of nucleotide represents frequency of reads. (Adapted and modified from 'Gene expression regulation mediated through reversible m6A RNA methylation' [96], 'The dynamic epitranscriptome: N6-methyladenosine and gene expression control' [75] and 'Transcriptome-wide mapping of *N*6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing.' [94])

M6A regulating proteins: M6A is found in various types of RNAs such as mRNA, microRNA (miRNA) and long noncoding RNA (lncRNA) [101-105]. M6A in RNA is found to affect mRNA splicing, stability, translation and localisation [106-109]. This represents an additional layer of epigenetic regulation via covalent modification of RNAs. For instance, m6A in mRNA can directly affect and determine cell fate of ESCs and pre-adipocytes [110-116]. The position of m6A found in mature mRNA affects downstream biological activities. M6A in 5' UTR is involved in translation and localisation upon binding by eukaryotic initiation factor 3 (eIF3) complex or YTHDF2 respectively [112, 117]. M6A in the 3' UTR regulates translation and mRNA stability by the binding of YTH family of proteins [106, 117]. Human YTH family of proteins (YTHDF 1-3 and YTHDC 1 and 2) has high affinity (nM range) towards the methylated consensus motif [106, 108, 118-120]. Different members of the YTH family can bind to m6A in different parts of mRNA affecting different downstream biological consequences. As such, methylation in the *N*6 position of adenosine (by METTL3-METTL14 complex) resulted in m6A, which can be demethylated (by FTO or ALKBH5) and recognised (by eIF3 or YTH domains). Thus, these proteins are termed as writers, erasers and readers respectively (Fig. 9) [75, 121-128].
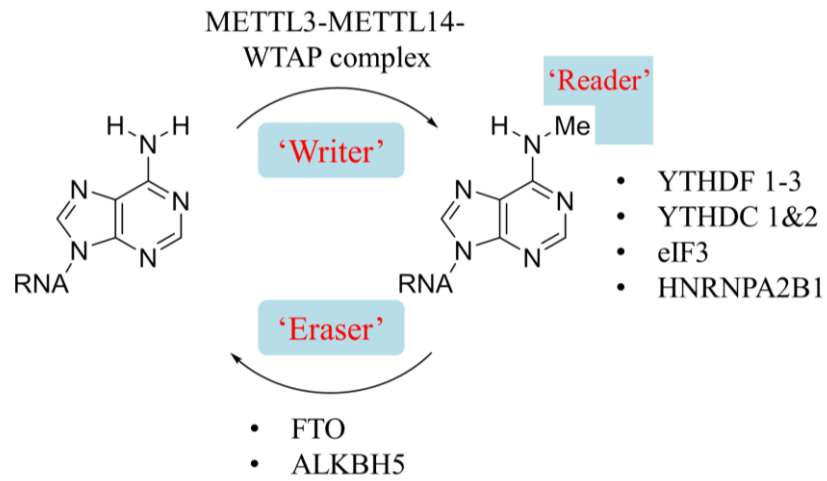
Figure 9. Writer, eraser and reader proteins of m6A in human. M6A is dynamically regulated in RNA transcripts. (Modified from 'Update: Mechanisms underlying methyladenosine modification of eukaryotic mRNA.' [128])

Duplex involving m6A: RNA oligonucleotides are able to form duplexes via Watson-Crick base pairing [129]. The presence of a methyl group on the *N*6 of the adenine nucleobase does not prevent m6A from the formation of Watson-Crick base pairing with its uracil counterpart [130]. However, the presence of a single methylation can destabilise the RNA duplex by approximately 1.4 kcalmol^{-1} [131]. The destabilisation occurs as the methyl group is forced to rotate, leading to adopting a high-energy *anti* conformation. It is also reported that the amount of destabilisation is dependent on sequential context [132]. In living system, dynamic modification of m6A can affect RNA duplex structure (close hairpin to an open hairpin) (Fig. 10). An example will be the RNA binding protein, heterogeneous nuclear ribonucleoprotein C (HNRNPC), which asserts a downstream biological effect upon binding [104, 105, 133]. This is known as the m6A-switch mechanism.
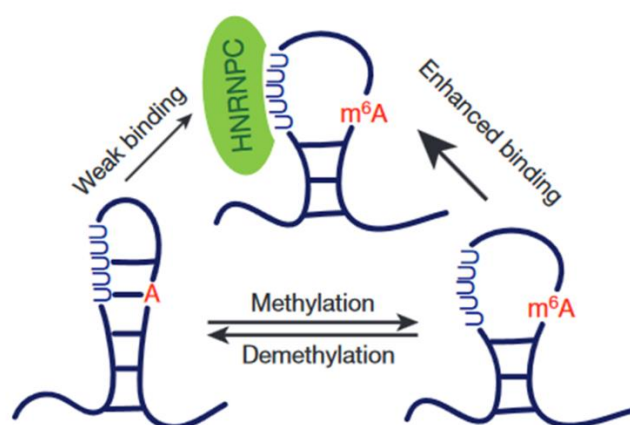
Figure 10. M6A-switch mechanism. The presence of m6A enhanced binding of HNRNPC. Demethylation of m6A caused a structural change to the hairpin and disfavoured association of HNRNPC. (Adapted from '*N*6-methyladenosine-dependent RNA structural switches regulate RNA-protein interaction.' [104])

1.3 Human fat mass and obesity-associated protein (FTO)

1.3.1 FTO and obesity

Arguably, *FTO* is the most well-studied and characterised member in the human AlkB family at the genetic and protein level [134-142]. *FTO* is the first genome-wide association studies (GWAS) identified gene that has direct implication in obesity. *FTO* gene is located on chromosome 16 and is more than 400 kilobase long. Previous attempts to link obesity to genetic basis proved to be futile before the GWAS. Three independent studies reported that various single nucleotide polymorphisms (SNPs) have a positive correlation to body mass index (BMI), a common indicator for obesity [143-145]. Individuals homozygous for each risk allele SNP contributes to about 3 kg and approximately 1.67-fold risk of being obese to an individual who does not possess the risk allele. Following the reports in 2007, many groups around the world reported that FTO gene SNPs indeed have significant association with obesity via BMI [146-155]. Prior to the discovery of *FTO* as an obesity gene, studies of *FTO* were done and reported on mouse [156-159]. Heterozygous *FTO* deletion in mouse caused fused toes phenotype and homozygous *FTO* knockout cause malformation and mid-gestation death. There was no report on FTO effect on obesity and thinness. However the mouse models were not accurate models for the study of FTO as the mice lacked several important genes. *FTO* is reported to be only found in vertebrates and marine algae [160, 161].

Physiological effects of FTO: As mentioned earlier in this thesis, *FTO* codes for a 2OG oxygenase belonging to the AlkB family of proteins [26, 162]. However, the exact biology of *FTO* and action pathway that leads to obesity was not well known. Reports of *FTO* started to accumulate after the three landmark GWAS in 2007 in bid to elucidate the role of *FTO* in obesity [143-145]. Evidence pointed to mammalian FTO being expressed in various regions that has relations to energy expenditure regulation such as skeletal muscles and liver adipose tissue [163]. In addition, *FTO* mRNA is found in cerebellum, hippocampus and hypothalamus [164, 165]. Although *FTO* has direct link to obesity, it is noteworthy that *FTO* variants were also reported to be linked to acute coronary syndrome, elevated blood pressure and hyperandrogenemia in woman with polycystic ovary syndrome [166-168].

Energy regulation by FTO: The first direct evidence how *FTO* affects obesity by energy haemostasis was observed when a *FTO* knockout mouse model displayed a lean body mass with significant reduction in fat mass via increasing energy expenditure [169]. In addition, *FTO* knockout mouse was observed to have relatively high postnatal mortality and if the mouse survived, exhibited growth retardation. This provided evidence that *FTO* might play a role more than just energy homeostatic that leads to obesity. The following year, 2010, it was reported that mice models with an additional knock-in of one and two *FTO* gene copies exhibited obesity due to increased fat mass from higher food intake [170]. The authors suggested that *FTO* induced obesity might be a due to *FTO* influenced hormonal leptin release through adipose tissues. With increasing evidence that FTO influences obesity by the propensity of food intake, some focus was turned to how and which part of the brain is affected by *FTO* expression level [160, 165, 171]. *FTO* expression level changes upon increase and decrease food deprivation. It is in the hypothalamus of the brain

where the expression of FTO seems to affect food intake preferences [172-175]. In the event of fasting, there is a 60% decrease of *FTO* mRNA level in the arcuate situated in the hypothalamus. Level of FTO in hypothalamus that affects food intake is also consistent with various reports by other groups.

FTO demethylase and obesity: Some investigations were performed to probe the link between catalytic activity of FTO and obesity [176, 177]. One such investigation was performed on a mouse model, which had a point mutation (isoleucine to phenylalanine) at position 367 of *FTO* gene. The point mutation occurred in a conserved region of around 20 amino acids across vertebrates not in the catalytic domain. The mouse model exhibited reduced FTO expression with corresponding reduction in catalytic activity of FTO. The mouse models did not suffer growth retardation unlike their *FTO* knockout counterparts. However, the mouse models exhibited maturity onset reduction in fat mass and body weight. This indirectly reflects the influence of the catalytic activity of FTO to obesity. Reviews provided thus far are from animal models. There are little documented cases of how catalytic impairment of FTO affects human beings. A loss-of-function mutation at position 316 of human FTO where arginine was mutated to glutamine was reported in some members of a Palestinian Arab consanguineous family [177]. Arginine 316 is important for the binding of co-substrate 2OG. It is an autosomal recessive disease where affected members displayed post-natal growth retardation, dysmorphism of head and face, psychomotor malfunction, cardiac and genital functional defects. Incidentally, there is a recent report of a patient who had homozygous mutation of histidine 271 to proline that catalytically impaired FTO [178]. The patient suffered very much the same physiology. No obesity was reported in those affected members but it is evident that a functioning FTO plays a greater role than just obesity.

SNPs in *FTO*: Much research focus of *FTO* in affecting obesity has been placed in the relative expression level of the protein in animal models. Although evidences pointed to FTO expression level is indeed able to affect obesity, questions still remain on how the non-coding first intronic variants of the *FTO* gene are able to control the expression level of FTO and thus, affect obesity. In 2014, it was reported that the intronic variant of *FTO* acts as a long range enhancer for a neighbouring gene named *Irx3* [141, 179]. The promoter of *IRX3* interacts with the first intronic variant of FTO and regulates IRX3 expression. IRX3-deficient mice have reduced body weight up to 30% with less fat mass and higher metabolism. This report offered direct evidence at molecular level of how the first intronic SNPs of *FTO* gene is linked to obesity. FTO expression level was reproducibly showed to affect obesity, however the authors did not address how *IRX3* is able to affect expression of FTO.

FTO and adipogenesis: In recent years, much attention was also turned to how FTO regulate the development of adipose tissues which causes obesity [113, 180, 181]. In 2014, it was reported that FTO expression and the level of m6A in mouse 3T3-L1 pre-adipocytes have inverse relationship during adipogenesis. Catalytically active FTO is able to promote adipogenesis through the regulation of m6A level. On the other hand, FTO knockdown is able to prevent differentiation of pre-adipocytes. FTO influences adipogenesis by affecting the expression of two isoforms of an adipogenesis regulating factor, Runt-related transcription factor 1 (RUNX1T1) through the control of the m6A level, which is positioned near the edge of an exon close to a splice site. In the absence of FTO, methylation near the splice site results in the recruitment of a protein named SRSF2, which induces exon inclusion and thus the expression of RUNX1T1 long isoform (RUNX1T1-L). RUNX1T1-L prevents differentiation of 3T3-L1 pre-adipocytes. However, SRSF2 is not recruited in the absence of m6A near

the splice site due to FTO demethylation. This causes exon skipping that results in the expression of RUNX1T1 short isoform (RUNX1T1-S). RUNX1T1-S promotes differentiation of 3T3-L1 pre-adipocytes. Subsequently, there are more reports supporting the role of FTO in regulating adipogenesis in adipocytes [182, 183]. However, there was no report on how the first intronic SNPs of *FTO* play in adipogenesis.

There is also no information of the direct link of how the first intronic SNPs of *FTO* could affect FTO expression level. So far, obesity caused by the first intronic SNPs of *FTO* and the expression level of functional FTO are proved to be independent of each other. It remains to be elucidated whether any links on the molecular biology level could be drawn from the effect of intronic SNPS on FTO expression.

1.3.2 Structural basis of substrate recognition by FTO

The first crystal structure of FTO was reported in 2010 (Fig.11) [91]. FTO was recombinantly expressed in *Escherichia coli* (*E.coli*) and was crystallised with an N-terminal nickel affinity tag containing six histidine residues followed by a thrombin cleavage site (amino acid residues: LVPRGS). The first 31 amino acid residues from the N-terminal of FTO were truncated for crystallisation. This was not surprising as about twenty residues from the N-terminal were predicted to be disordered through prediction using various servers [184, 185]. Despite the truncation, crystal structure of FTO allowed us to study the DSBH domain, characteristic of AlkB 2OG oxygenase family and the C-terminal helical bundle domain in great details.

Structural features of FTO: Like many 2OG oxygenases, β-strands and α-helixes precedes the DSBH domain in FTO. Two long α-helixes, α-helix 3 and α-helix 4 lie at the exterior of the DSBH domain, the side consisting of major β-sheets, presumably to provide structural integrity to the DSBH domain (Fig. 11). Two NRLs are located at the bigger opening end of the DSBH core domain, forming part of the entrance into the catalytic domain of FTO. NRL1 is longer, making up of about fifteen amino acid residues (Fig. 3). The tip of NRL1 comprises a short α-helix with five residues. Of the five residues, two of them are basic residues arginine and lysine, believed to interact with oligonucleotides. NRL2 is made up of two β-strands joined by a short loop containing two residues. L1 loop (17 amino acid residues), named by the authors, is located between β-strand I and II of the DSBH domain (Fig. 3 and 11). Notably, L1 loop is unique and highly conserved among FTO proteins across different species, and it is absent in other AlkB proteins. Although it is highly unstructured, seemingly, the loop is rigidified by hydrophobic interaction between residues in L1 loop (Tyr220 and Phe221) and second β-strand (Trp230) of the DSBH domain (Fig. 12). In addition, Glu218 of L1 interacts with Trp230, conferring additional stability. This might be the reason why the electron density of L1 loop could be observed as it was predicted to have high conformational flexibility. Superimposition of AlkB:dsDNA (PDB ID 4NID) and ALKBH2:dsDNA (PDB ID 3BUC) complexes with FTO revealed a steric clash of the non-methylated strand with L1 loop (Fig. 13) Taken into consideration the rigidity of L1 loop, L1 loop may play a part in the substrate selection, preferring single-stranded oligonucleotides over duplexes. βIV/V insert of FTO is about seven residues long and do not appear to participate in any substrate selectivity event. Notably, L1 loop of FTO occupies an almost similar position as βIV/V insert of

hALKBH5. Thus, βIV/V insert of hALKBH5 is believed to be responsible for single-stranded oligonucleotides selectivity [186].

FTO has a characteristic C-terminal domain helical bundle. The C-terminal domain is made up of four main helixes and represents a new type of fold. Notably, C terminal domain is absence in the members of AlkB 2OG oxygenase family discussed in this thesis. The exact function of C-terminal domain helical bundle is unknown. However, the presence of the C-terminal domain is essential in maintaining the activity of the DSBH domain, presumably, through stabilisation. Co-expressed DSBH and the C-terminal domain were able to form stable complexes *in vitro*, and subsequently, able to recover some of the lost enzymatic activity.

Like other AlkB 2OG oxygenase family members, FTO has a highly conserved iron binding amino acid motif, a HxD/E...H triad. The first crystal structure of FTO (PDB ID: 3LFM) provided structural insights into substrate binding mechanism of FTO. The carboxylate group of NOG occupy deep into the 2OG binding domain, forming hydrogen bonding interaction with Arg316 (Fig. 14). The co-crystallised substrate m3T is tucked in between NRL2 and the 2OG binding domain. Tyr108 of NRL2 and His231 of the catalytic triad sandwiches the nucleobase ring of m3T through a slightly distorted π-π stacking interaction. The methyl group of m3T is situated just beside the co-factor iron(II), facilitating the hydroxylation reaction via the Fe(IV)-oxo intermediate. Furthermore, m3T is stabilised by the hydrogen bonding between the amide nitrogen of Glu234 and O4-m3T. M3T and NRL1 are unable to make meaningful interaction as they are not in close proximity. However, as NRL1 is placed nearer to the surface of FTO, it is believed to be able to interact with oligonucleotide substrate.

It is noteworthy that FTO has never been crystallised with ssRNA although it has been widely acclaimed that FTO is able to demethylate its natural substrate m6A found in RNA. Arguably, a cross-linking strategy must be used to observe the binding mechanism of FTO with oligonucleotides [43]. This will shed light on the exact role of L1 loop, NRL1, NRL2 and C-terminal domain helical bundle play in substrate recognition and selection.
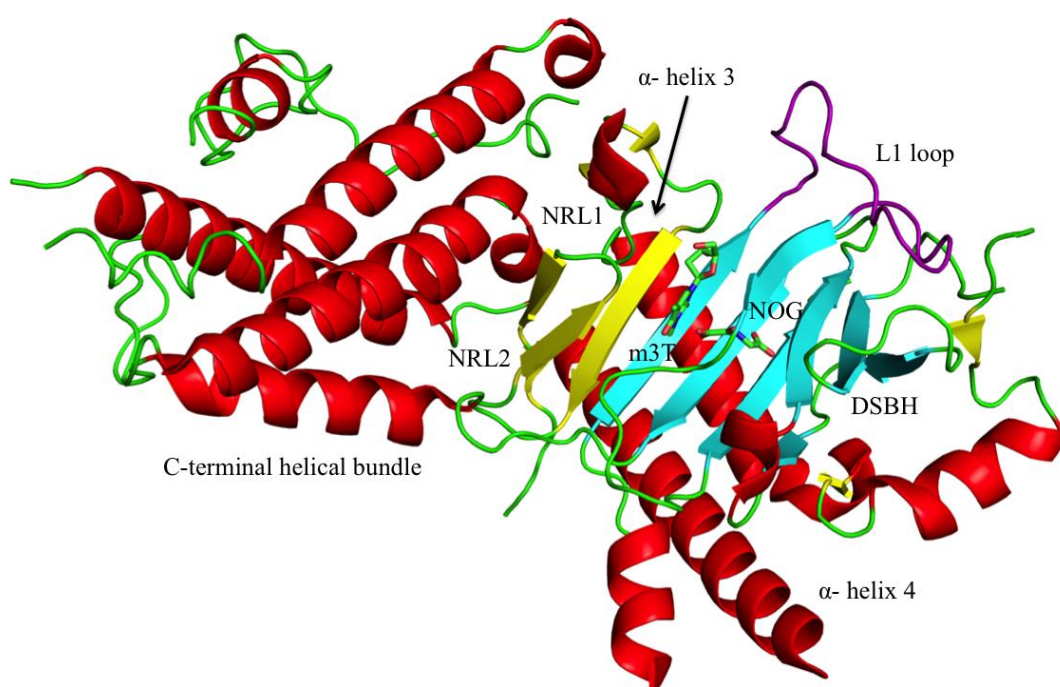
Figure 11. Overall structure of FTO (PDB ID 3LFM). Characteristic structural features of FTO include L1 loop (purple) and the C-terminal helical bundle (red helices on the left of FTO). NRL 1 & 2 is indicated in the diagram as a pair of yellow hairpins situated between the helical bundle and the DSBH domain (cyan). Two long α-helixes, α-helix 3 and α-helix 4 lie at the exterior of the DSBH domain. M3T is situated near the opening of the catalytic domain, lying next to NRL2. NOG, analog of natural co-substrate 2OG is embedded in the DSBH domain.
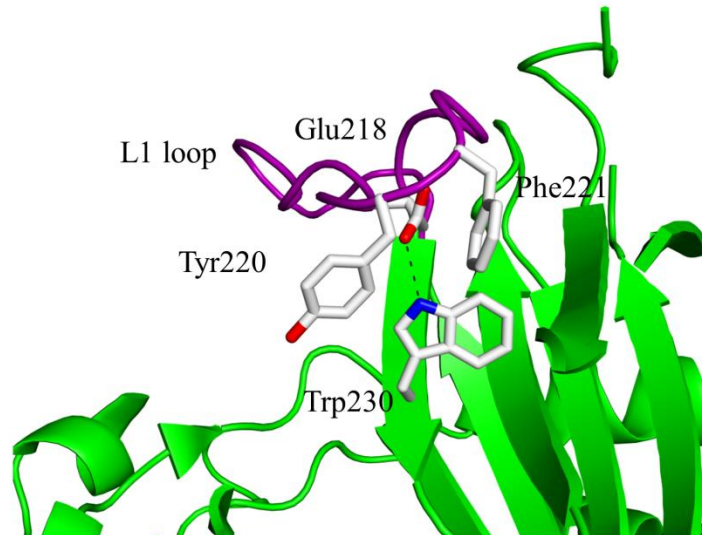
Figure 12. Hydrophobic interaction between L1 loop and DSBH of FTO. The loop is rigidified by hydrophobic interaction between residues (Tyr220 and Phe221) in L1 loop (purple) and βII (Trp230) of the DSBH domain. In addition, Glu218 of L1 interacts with Trp230, conferring additional stability. Figure adapted from PDB ID 3LFM.
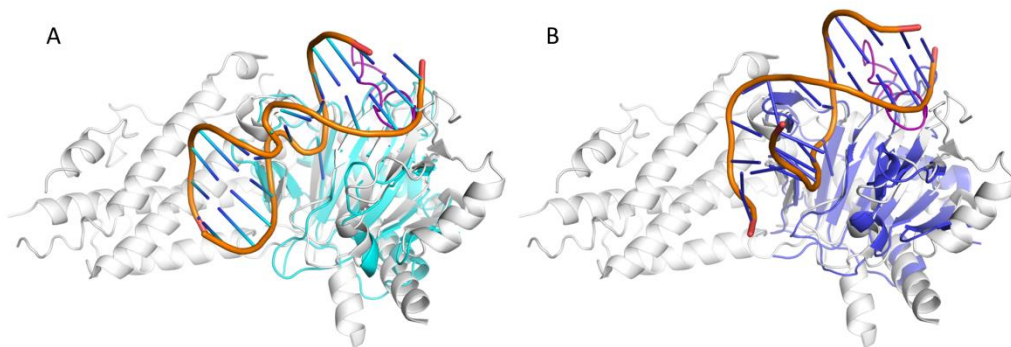
Figure 13. Steric clash of FTO L1 loop (purple) with dsDNA. A. Alignment of FTO (white) (PDB ID 3LFM) with AlkB:dsDNA (cyan) (PDB ID 4NID). **B.** Alignment of FTO (white) (PDB ID 3LFM) with ALKBH2:dsDNA (blue) (PDB ID 3BUC).
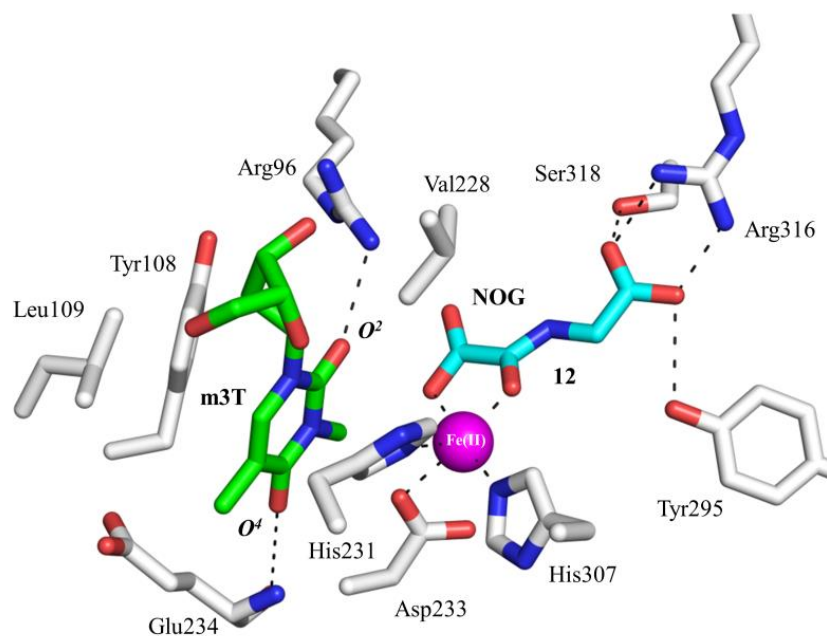
Figure 14. Catalytic domain of FTO with m3T and NOG. The carboxylate group of NOG occupy deep into the 2OG-binding domain, forming hydrogen bonding interaction with Arg316. Nucleobase of m3T is sandwiched between Tyr108 and His231. Figure adapted from PDB ID 3LFM.

1.4 Other AlkB nucleic acid demethylases

1.4.1 *Escherichia coli* (*E.coli*) AlkB

AlkB protein from *E.coli* is arguable to be the first identified member of the AlkB 2OG oxygenase family [187]. AlkB protects bacteria from oxidative alkylating damage [188, 189]. AlkB was identified in 1983 when Sekiguchi and co-workers discovered that *E.coli* with AlkB mutation sensitised *E.coli* to methylation agent methyl methanesulfonate (MMS). The subsequent reports revealed that AlkB expression is able to suppress alkylation damage of DNA in bacteria. In addition, AlkB introduced into mammalian cells reverse methylation damage to DNA [190, 191]. However, its molecular mechanism was not known due to lacking biochemical data although there were some reports of AlkB being iron dependent and its substrates were DNA nucleobases [192]. In 2001, the mystery of AlkB function was revealed through sequence profiling analysis, namely a 2OG dependent oxygenase protein [8].

Given the discovery of AlkB belonging to the 2OG dependent oxygenase superfamily, repair mechanisms of AlkB then started to emerged [193-196]. It was found that AlkB directly repair the cytotoxic damage on both single and double stranded DNAs by demethylating 1-methyladenosine (m1A) and 3-methylcytosine (m3C). Moreover, AlkB was found to have a range of substrates preference as it was later reported to be able to reverse methylation damage of nucleobases, 1-methylguanosine (m1G) and m3T in DNA and RNA. Arguably, AlkB represent the most promiscuous member of the AlkB family of 2OG oxygenases known so far [197-201].

Crystal structures of AlkB revealed information regarding how NRL and βIV/V inserts accommodate a range of oligonucleotide substrates [39, 40, 202, 203]. AlkB was also solved with DNA duplex crosslinked via disulphide bond revealing

mechanism of DNA damage repair. AlkB from *E.coli* serves as a 'model' protein for the study of human AlkB family of 2OG oxygenases. AlkB facilitated our understanding of substrate recognition mechanism by AlkB family of 2OG oxygenases. In addition, AlkB was used as a proof-of-concept for the development of potent and selective inhibitors that discriminate other family 2OG oxygenases, paving way for functional probes and drug design for human AlkB family of 2OG oxygenases [204-207].

1.4.2 Human ALKBH1 (ALKBH1)

ALKBH1 was identified in 1996, the first human AlkB family of oxygenases discovered, sharing high similarity (52%) with AlkB from *E. coli* [208]. However, the function of ALKBH1 *in vivo* is not as well characterised as its counterparts FTO and ALKBH5. In addition, the structure of ALKBH1 is not available, and it will be interesting to know how much the DSBH domain of ALKBH1 resembles that of AlkB from *E.coli*.

ALKBH1 is believed to be a DNA demethylase as it is able to partially rescue the phenotype of *E.coli* with loss-of-function mutation of AlkB when exposed to alkylating agent MMS. However, the natural substrate of ALKBH1 is still under debate. The substrates of ALKBH1 are m3C, methylation marks on histone 2A and most recently, m6A in DNA [73, 209-212]. If the aforementioned methylated substrates were proved to be natural substrates, ALKBH1 could be the most promiscuous human AlkB 2OG oxygenase family member. It is noteworthy that m3C is the only oxidative damage caused lesion and the other two substrates are considered to be epigenetic marks. Several reports have pointed out that ALKBH1 could be more than an alkylated damage repair enzyme [209, 210, 212-214]. A few studies have

reported that ALKBH1 is a mitochondrial protein and play a crucial role in stem cells differentiation via epigenetic regulation. Those reports showed that ALKBH1 regulate transcription epigenetically via histones and DNA methylation. However, we cannot completely rule out a possible role of ALKBH1 in epigenetic regulation at RNA level as it does not discriminate ssDNA or ssRNA substrates [73].

1.4.3 Human ALKBH2 (ALKBH2)

ALKBH2 was first discovered in 2002 as a DNA lesion repairing enzyme [215]. It was identified through analysing the sequence identity and secondary structure prediction, further validated by enzymatic assays. ALKBH2 is reported by various groups to demethylate m1A, m3T and m3C with a slight preference to act on dsDNA over ssDNA [50, 197, 198, 216-220]. The first crystal structure of ALKBH2 was solved with dsDNA cross-linked to the active domain and it represents the only one in complex with nucleic acids for the entire human AlkB family [40]. In the crystal structure of ALKBH2:dsDNA complex, notably 3 basic residues RKK (Arg241 to Lys243) interacted with the negatively charged dsDNA and Phe102 finger residue maintains the π- π stacking interaction of the dsDNA duplex when the modified base of dsDNA is flipped out into the catalytic domain. The ALKBH2:dsDNA complex crystal structure gives a glimpse of the mechanism of how ALKBH2 binds and repairs dsDNA lesions.

Like many other human AlkB oxygenase family members, biological function of ALKBH2 is not well characterised. ALKBH2 primarily act as DNA repair enzyme, which maintains cellular environment caused by alkylation damage thus preventing cell death [221]. Despite the perceived usefulness of ALKBH2, the presence of its active form is persistent in bladder cancers and glioblastoma [222-225]. Drugs used

during cancer therapy are often alkylating agents, damaging the genome of cancer cells leading to apoptosis. ALKBH2 often contributes to drug resistance in chemotherapy, which is probably caused by cancer cell taking advantage of the repair property of ALKBH2 that can reverse damaged DNA.

1.4.4 Human ALKBH3 (ALKBH3)

The first report of the characterisation of ALKBH3 as a DNA repair enzyme was made at about the same time as ALKBH2 with the similar types of experimental validations[215]. Similar to ALKBH2, ALKBH3 has very much diminished activity toward short oligonucleotides (trimer)[197]. Both ALKBH3 and ALKBH2 are able to demethylate similar nucleotide substrates. In contrast to ALKBH2, ALKBH3 is unable to demethylate dsDNA. The structural basis of why ALKBH3 does not act on dsDNA was elucidated through comparing crystal structure of ALKBH2:dsDNA complex and the crystal structure of ALKBH3 [42, 226, 227]. Glu123 of ALKBH3 corresponds to Phe102 of ALKBH2 in the crystal structures. In the crystal structure of ALKBH2:dsDNA complex, Phe102 finger residue plays an important role by facilitating the maintenance of the structural integrity of dsDNA by π- π stacking interaction when ALKBH2 performs its enzymatic activity. However, Glu123 of ALKBH3 is negatively charged in nature, presumably, repelling the phosphate backbone of dsDNA, destabilising the formation of ALKBH3:dsDNA complex. In addition, when Glu123 of ALKBH3 was mutated to Phe123, ALKBH3 gained dsDNA demethylation property. This elucidated the substrate selectivity mechanism of ALKBH2 and ALKBH3.

Compared to its ALKBH2 counterpart, biological roles and functions of ALKBH3 are better characterised. Particularly, ALKBH3 is also known as prostate cancer antigen-1

(pca-1) as it was first observed high expression level in prostate carcinoma samples [228-232]. ALKBH3 transfected cell lines displayed resistance against MMS-induced cell death, most probably through the dealkylation properties of ALKBH3. Other than prostate cancer, ALKBH3 is also implicated in non-small-cell-lung cancer, pancreatic cancer, papillary thyroid cancer and human rectal carcinoma [229, 233, 234]. The mechanism of how ALKBH3 maintains genome integrity via resistance to alkylation and thus continuous proliferation of various types of tumours and healthy cells was elucidated in 2011 [235]. In a pull down experiment, ALKBH3 was associated with activating signal cointegrator complex 3(ASCC3). ASCC3 was then determined to be a 3'-5' helicase that unwinds DNA, allowing the methylated base in a ssDNA to be repaired by ALKBH3. Due to overexpression and the hijacked of ALKBH3 for prostate cancers' own advantage, ALKBH3 serves as a viable therapeutic target. As such, efforts have been invested to develop a drug that targets ALKBH3. There is only one crystal structure of ALKBH3 complexed with NOG thus far and more crystal structures solved with small molecular inhibitors will facilitate the rational drug design target ALKBH3.

1.4.5 Human ALKBH5 (ALKBH5)

The first attempted characterisation of ALKBH5 as a human AlkB 2OG oxygenase was in 2011 [92]. Despite testing a range of known methylated substrates for human AlkB family of 2OG oxygenases, the authors failed to identify a suitable substrate for recombinant ALKBH5. Finally, characterisation of ALKBH5 as an AlkB family of 2OG oxygenase was based on ^{14}C isotopic labelling of carbon in a 2OG turnover assay, namely ^{14}C isotopic labelled CO_2 was produced in the presence of co-factors, Fe(II) and co-substrates, oxygen and ascorbate. Less CO_2 was produced when ALKBH5 was inhibited by generic inhibitor for 2OG oxygenases, pyridine-2,4-

dicarboxylate (2,4-PDCA), reflecting amenability to inhibition. In addition, the authors also identified ALKBH5 as a HIF-1α responsive gene where *ALKBH5* gene has two HIF binding sites at its promoter. Lastly, the authors discovered that ALKBH5 localise mainly in the nucleus of cells, similar to FTO, and postulated that ALKBH5 is a DNA repair enzyme.

In 2013, it was reported that ALKBH5 shares the same substrate as FTO and was inactive against dsDNA [44]. ALKBH5 is able to demethylate m6A in mRNA and is linked to fertility. They further confirmed that ALKBH5 is localised in the nuclear speckle. However, ALKBH5 does not play a DNA repair function as previously proposed. The discovery of ALKBH5 represents the second enzyme after FTO, which is able to remove the dynamic modification, m6A in mRNA[236]. Despite the identification of ALKBH5 as epigenetic regulatory enzyme, the exact molecular mechanism of how ALKBH5 performs its biology roles remains elusive.

The first crystal structure of ALKBH5 was first reported in 2014 [186, 237-240], revealing a crystallographic dimer. The dimer-dimer interaction was aided by the π- π stacking interaction of IOX3 (compound name given by the authors), covalently linked to the surface of the active domain. ALKBH5 displayed the characteristic DSBH fold of human AlkB family of 2OG oxygenases. The manganese divalent ion (in place of ferrous iron) is ligand by the motif, a HxD/E...H triad (x being any amino acid residue) which are generally conserved across the 2OG oxygenases family. Nevertheless, the first crystal structure of ALKBH5 gave some insights towards its substrate preference. Similarly to FTO, two NRLs precede the DSBH fold. NRL1 of ALKBH5 (3 residues) is much shorter than that of FTO (~15 residues) while NRL2 of ALKBH5 longer and disordered (~14 residues) than that of FTO (~10 residues). Both longer NRLs of ALKBH5 and FTO contain some positively charged and aromatic

residues such as arginine, lysine, and tyrosine, which interact with the oligonucleotide. An additional interesting observation was that the βIV/V loop of ALKBH5 is stabilised by disulphide bond formation between Cys230 and Cys267 of βVII. This rigidified loop might serve to block out dsDNA as observed when ALKBH5 is superimposed with AlkB and ALKBH2 complexed with dsDNA. This could confer stricter substrate specificity towards single stranded oligonucleotide. The first reported crystal structure of ALKBH5 was followed by co-crystal structures of ALKBH5 with various generic inhibitors such as NOG, 2,4-PDCA and small molecules from crystallisation conditions.

Just recently, ALKBH5 is found to have a direct role in breast cancer stem cells (BCSCs) proliferation [241]. Exposure of BCSCs to hypoxic environment stimulates expression of HIF-1α and HIF-2α which binds to the promoter upstream of *ALKBH5* gene. ALKBH5 demethylate m6A found in NANOG mRNA, a pluripotency factor, which then promote the BCSCs proliferation. Knockdown of ALKBH5 will prevent BCSCs proliferation. Thus, this finding demonstrated the potential for ALKBH5 to be a therapeutic target.

For a better understanding of the biological function of ALKBH5, some questions need to be addressed. For instance, what is the difference between substrate preference of ALKBH5 and FTO given that they play distinct biological roles? In addition, the same preference of both ALKBH5 and FTO to single stranded m6A-containing oligonucleotide substrates *in vitro* and *in vivo* adds to another layer of complications.

1.5 Amenability to small molecule inhibition

As aforementioned, human 2OG oxygenases are implicated in many diseases. Thus, it is of significant interest to investigate whether human 2OG oxygenases are viable targets for therapeutic intervention. As such, over the years, there are many reports, patents and reviews of the design of small molecule inhibitors and drugs that target the enzymatic activities of human 2OG oxygenases such as (JmjC)-type 2OG oxygenases and PHDs [242-244]. Crystal structures of human 2OG oxygenases greatly facilitate the rational drug design. Thus, it has opened up the possibility to develop specific and potent small molecules that target human 2OG oxygenases. Some of the potential drugs have an IC_{50} at nM range. In addition, some of the drugs have entered clinical trials and the clinical studies effort is still ongoing [242, 243].

We are starting to gain a foothold regarding the understanding of the human AlkB family of 2OG oxygenases. There are documentations of some of the human AlkB family members implicated in various diseases, most probably through covalent modifications of nucleobases in DNA and RNA. These warrant the investigation on the potential of inhibiting the activity of human AlkB family of 2OG oxygenases by small molecules for probes and drugs development. The *E.coli* AlkB can be enzymatically inhibited by NOG, with an IC_{50} of ~700 μM [204]. Arguably, this is the first indication of the ability of AlkB family of 2OG oxygenases to be enzymatically inhibited. NOG is able to weakly inhibit all the human members of 2OG oxygenases family and is considered a 'generic' inhibitor. The promiscuity of NOG is not surprising as the binding mechanism primarily depends on the binding to the catalytic triad, very much well conserved across the members of 2OG oxygenases family forming similar interactions as 2OG [245]. NOG is a competitive inhibitor, occupying the same position as 2OG in the catalytic domain, chelating to the co-factor iron (II).

In 2012, AlkB was crystallised with various inhibitors that is more complex than NOG [205]. These inhibitors have extra aromatic functional groups extended from the alkyl backbone of NOG, embedding into a hydrophobic sub-pocket next to a His187 (one of the catalytic triad residues of AlkB) and Trp178 of AlkB. This shows that high potency and selectivity could potentially be obtained for human AlkB family of 2OG oxygenases' inhibitors, by taking advantage of the residues in adjacent sub-domains, which may not be as well conserved as the catalytic triad. Just recently, another class of inhibitor for AlkB was reported. Rhein, a natural product, occupies the substrate binding site of AlkB [207]. This further illustrates that surrounding residues, other than the 2OG binding pocket, could be exploited for inhibitors and probes design.

FTO is directly linked to obesity. Hence, it has gathered much interest in developing probes and drugs in hope to shed light on the biology and potential therapeutic purposes. Reports of potent and selective inhibitors have steadily increased since 2012 [245-253]. Different strategies such as *in silico* screening and rational design were employed to identify inhibitors for FTO such as rhein and meclofenamic acid [246, 249]. Various approaches such as biophysical assays, biochemical assays, cell-based assays and X-ray crystallography were used for hit validation and characterisation. One challenge was to discover potent and selective inhibitors that can specifically inhibit FTO through different mode of binding. Gratifyingly, FTO demonstrated the ability to be inhibited by different classes of inhibitors through various binding mechanisms. We are possibly the first to contribute, via a novel tethering methodology to discover a class of inhibitor [245]. The inhibitors act through a unique binding mechanism, which will be further discussed in this thesis.

As mentioned earlier in this thesis, ALKBH3 has been a subject for therapeutic intervention. The research group designed and attempted to optimise the

bioavailability of the ALKBH3 inhibitor, HUHS015 [232]. However, it will be interesting to understand the binding mechanism of HUHS015 through structural approaches. There are reports that claimed the presence of inhibitors for ALKBH5. However, both are weak and non-selective. In addition, it will be valuable to develop inhibitors for ALKBH1, ALKBH2 and ALKBH5. Nevertheless, FTO is the first human AlkB family of 2OG oxygenases to be potently and selectively inhibited in my thesis. This successful demonstration paves way for the probe and drug design of other members of human AlkB family of 2OG oxygenases, mentioned above, despite their high structural homology.

1.6 Aims and significance of the project

Human AlkB family of 2OG oxygenases are demethylases mainly involved in nucleic acid repair and epigenetic modification. They have attracted much interest due to their implication in various human diseases making them potential therapeutic targets. FTO is a member of the AlkB family, which demethylates m6A [43]. M6A is a recently identified epigenetic mark on RNA that plays a role in transcript RNA splicing, stability, translation and localisation [126]. *FTO* is the first GWAS identified obesity gene [143]. Particularly, functional FTO is important for normal growth and development in humans [177]. However, the biology of m6A demethylation by FTO is unclear. Furthermore, human AlkB homolog ALKBH5 is also able to perform m6A demethylation, making it challenging to elucidate the roles and functions of FTO [44].

Presence of m6A in mRNA weakens Watson-Crick base pairings but does not prevent duplex formation [130, 131]. Depending on sequential context, the m6A mark might influence the different secondary structures adopted by mRNA, therefore affecting the availability of binding factors to intearact with mRNA to trigger downstream biological events [127]. FTO and ALKBH5 are the only known m6A demethylases in human and they seemingly demethylate m6A selectively. However, it remains unclear about the enzymes' substrate recognition mechanism towards m6A-induced conformation in mRNA.

This thesis aims to address the following issues:

1) To repeat and establish crystallisation condition for FTO that can achieve routine diffraction to atomic resolution for structure-based small molecules inhibitory studies. This was achieved by optimisation of the purification protocol and systematic screening of precipitants and additives of the crystal hit condition.

2) To develop a potent and selective inhibitor for FTO and understand the mechanism of inhibition, setting a platform for probe development and therapeutic purposes. We managed to solve the structures of co-crystal complexes of FTO with various inhibitors. Next, we were able to analyse and understand the mechanism for potent and selective inhibition of FTO.

3) To investigate how FTO and hALKBH5 recognise different conformations of mRNA mediated by m6A. We discovered that FTO and hALKBH5 were able to recognise and discriminate m6A-containing oligonucleotides of almost similar sequences but different conformations *in vitro*. In addition, FTO and hALKBH5 were able to affect a conformational change of the oligonucleotides upon demethylation.

FTO gene has a direct link to obesity and FTO is an m6A demethylase with epigenetic implications. Structural understanding of the potency and selectivity of an FTO inhibitor will set a good platform for probing the biology of FTO. This will also provide a foundation for therapeutic intervention. A better understanding of the interplay between FTO demethylation and m6A determined conformational changes will shed light in the physiological functions of FTO in affecting RNA transcripts.

CHAPTER 2

MATERIALS AND METHODS

2.1 Molecular cloning and small scale protein expression

The human full length FTO construct was kindly provided by Dr. Giles Yeo (University of Cambridge Metabolic Research Lab) for this study. Other cDNAs used in this thesis were kindly provided by Protein Production Platform (PPP), NTU.

Molecular cloning and small scale expression analysis of FTO, ΔN31-FTO, *E.coli* ΔN11-AlkB and ALKBH5 (residues 66-292): Truncations of proteins were selected based on previously reported enzymatically active constructs. The cloning vector pNIC28-Bsa4 (GenBank accession EF198106), developed by Opher Gileadi at SGC Oxford was constructed by the help of PPP, NTU [254, 255]. Briefly, ligase independent cloning (LIC) technology was employed with RecA- strain of *E.coli* for cloning and amplification of vector containing gene insert. The gene is fused with N-terminal hexa-histidine tag for downstream nickel affinity purification and TEV-protease cleavage site for optional tag removal. The construct is then re-transformed into *E.coli* expression strain (T1-phage resistant BL21-DE3 Rosetta) developed by PPP. The clone was put through a small scale expression screening by PPP to test for soluble expression. Briefly, cultures are grow in 1 mL of terrific broth at 37°C and induced by IPTG when the optical density at 600 nm (OD_{600}) reached 2-3. The temperature is lowered to 18°C upon induction. After overnight culture, the cells are harvested and lysed, then the supernatant clarified by centrifugation was applied for nickel affinity chromatography to assess the solubility and relative abundance of protein via SDS-PAGE.

Molecular cloning of ΔN31-FTO: The cloning of the truncated FTO was performed using the human full length FTO construct provided by Dr. Giles Yeo as cDNA template. The cDNA coding for ΔN31-FTO was amplified using polymerase chain reaction (PCR) with two primers (forward, 5'-GGAATTCCATATGACCCCCAAAGATGATGA-3' and reverse, 5'-CCCAAGCTTCAGGGTTTTGCTTCCAGAA-3') with NdeI and HindIII restriction sites. The PCR product and the vector pET28b are digested by the respective restriction enzymes and ligated by ligase. The ligated products were then transformed into DH5α and plated onto an agar plate overnight at 37°C. Colony PCR was then performed to check for insertion, subsequently sequencing to verify that there are no frameshift and mutation for the truncated gene insert. The FTO construct is fused with N-terminal hexa-histidine tag for downstream nickel affinity purification and thrombin protease cleavage site for optional tag removal. The constructed plasmid is then re-transformed into *E.coli* BL21 (DE3) for protein preparation.

Molecular cloning and small scale expression analysis of ALKBH2 (residues 56-258): The cloning of the truncated ALKBH2 was performed using cDNA template kindly provided by PPP, NTU. The PCR amplification for ALKBH2 (residues 56-258) was carried out using two primers (forward, 5'-GGAATTCCATATGAGCTGGCGGCACATTC-3' and reverse, 5'-CCCAAGCTTCAAAGCAAAATTTTACGAAAAGT-3') with NdeI and HindIII restriction sites. The amplicons and the vector pET28b are digested by the respective restriction enzymes and ligated by ligase. The ligated products were then transformed into DH5α and plated onto an agar plate overnight at 37°C. Colony PCR was then performed to check for insertion into the vector. The vector was then sequenced to verify that there are no frameshift and mutation for the truncated gene insert. The

ALKBH2 (residues 56-258) construct is fused with N-terminal hexa-histidine tag for downstream nickel affinity purification and thrombin protease cleavage site for optional tag removal. The plasmid is consequently re-transformed into *E.coli* BL21 (DE3) for protein expression. Small scale expression test was carried out. Briefly, 50 mL of 2XTY growth medium was inoculated with 500 μL of start culture at 37°C and 200 rpm. When OD_{600} reached around 0.6-0.8, protein expression was induced by 0.5 mM IPTG (Gold Biotechnology). The temperature is lowered to 16°C upon induction. After overnight culture, the cells are lysed, clarified and nickel affinity chromatography was performed to assess the solubility and relative abundance of protein via SDS-PAGE.

Molecular cloning and small scale expression analysis of YTHDF2 (residues 385-576): The cloning of the truncated YTHDF2 was performed using cDNA template kindly provided by PPP, NTU. The PCR amplification for YTHDF2 (residues 385-576) was carried out using two primers (forward, 5'-CGGGATCCATGCCCCACCCAGTGTTGG-3' and reverse, 5'-CCCTCGAG TCAACCTTGACGTTCCTTTTTAA-3') with BamHI and XhoI restriction sites. The amplicons and the vector pGEX-6P-1 is digested by the respective restriction enzymes and ligated by ligase. The ligated products were then transformed into DH5α and plated onto a agar plate overnight at 37°C. Colony PCR was then performed to check for insertion into the vector. The vector was then sequenced to verified that there are no frameshift and mutation for the truncated gene insert. The YTHDF2 (residues 285-576) is fused with N-terminal glutathione S-transferase (GST) tag for downstream gluthione sepharose affinity purification and PreScission™ protease cleavage site for optional GST tag removal. The plasmid is consequently re-transformed into *E.coli* BL21 (DE3) for protein expression. Small scale expression

test was carried out. Briefly, 50 mL of 2XTY growth medium was inoculated with 500 μL of start culture at 37°C and 200 rpm. When OD_{600} reached around 0.6-0.8, protein expression was induced by 0.5 mM IPTG (Gold Biotechnology). The temperature is lowered to 16°C upon induction. After overnight culture, the cells are lysed, clarified and glutathione sepharose affinity chromatography was performed to assess the solubility and relative abundance of protein via SDS-PAGE.

2.2 Large scale protein preparation

ALKBH3 (residues 1-286, fused to 39-amino acid histidine tag at the N-terminal) used for biochemical assays in this thesis was commercially purchased from ProSpec-Tany Technogene Ltd, Israel. We eventually managed to optimise the purification protocol and routinely obtain ALKBH3 with high yield and high purity. Full length PHD2 (fused to MYC/DKK tag at C-terminal) was purchased from Origene Technology, USA.

Large scale preparation of FTO: BL21 (DE3) Rosetta T1R transformed with pNIC28-Bsa4 FTO was grown in 6 L 2XTY media supplemented with chloramphenicol (33 μg/mL) and kanamycin (30 μg/mL) at 37°C and 200 rpm to an OD_{600} of ~0.6-0.8. Protein expression was induced by addition of 0.5 mM IPTG (Gold Biotechnology). Cell was continuously grown for 16 hours at 16°C before harvested by centrifugation. The cell pellet can be stored at -80°C for future purification or can be immediately suspended to homogeneity in 25 mM Tris, pH 7.5, 500 mM NaCl, 40 mM imidazole, pH 7.5 and 5 mM β-ME. Purification was performed at 4°C. Cells were disrupted by French press and the lysate clarified by centrifugation and filtration. FTO was purified from crude cell lysate by passing through the Ni-NTA affinity column (GE healthcare), then eluted with a gradient increase of imidazole up to 500

mM. Further purification was performed by size exclusion chromatography using HiLoad superdex 200 26/60 (GE Healthcare), simultaneously with buffer exchanged to 25 mM Tris, pH 7.5, 100 mM NaCl, 5% (v/v) glycerol and 5mM β-ME. Final protein purity and concentration was determined via SDS-PAGE and NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific), respectively.

Large scale preparation of ΔN31-FTO: BL21 (DE3) transformed with pET-28b FTO was grown 6 L 2XTY media supplemented with chloramphenicol (33 μg/mL) and kanamycin (30 μg/mL) at 37°C and 200 rpm to an OD_{600} of ~0.6-0.8. Protein expression was induced by addition of 0.5 mM IPTG (Gold Biotechnology). Cell was continuously grown for 16 hours at 16°C before harvested by centrifugation. The cell pellet can be stored at -80°C for future purification or can be immediately suspended to homogeneity in 25 mM Tris, pH7.5, 500 mM NaCl, 40 mM imidazole, pH 7.5 and 5 mM β-ME. Purification was performed at 4°C. Cells were disrupted by French press and the lysate clarified by centrifugation and filtration. FTO was purified from crude cell lysate by passing through the Ni-NTA affinity column (GE healthcare), then eluted with a gradient increase of imidazole up to 500 mM. It was followed by anion chromatography using a 5 mL HiTrap Q HP column (GE Healthcare), with gradient elution to 1M NaCl. Further purification was performed by size exclusion chromatography using HiLoad superdex 200 26/60 (GE Healthcare) buffer exchanged into 25 mM Tris, pH 7.5, 50 mM NaCl and 5 mM β-ME. Final protein purity and concentration was determined via SDS-PAGE and NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific), respectively.

Large scale preparation of ΔN11-AlkB: AlkB purification was performed by PPP [254, 255]. BL21 (DE3) Rosetta T1R transformed with pNIC28-Bsa4 ΔN11 AlkB were grown in a LEX system using 750 mL of Terrific Broth (TB) supplemented with 8 g/L of glycerol and appropriate antibiotics. The culture was incubated at 37°C in the LEX system with aeration and agitation through the bubbling of filtered air through the culture. When the OD600 reached ~2, the temperature was reduced to 18°C and the cultures were induced after 30 to 60 min by addition of 0.5 mM IPTG. Growth was continued for 16-20 h and harvested by centrifugation. The cell pellets were suspended in 1.5 times the cell pellet weight of lysis buffer containing 100 mM HEPES, 500 mM NaCl, 10 mM Imidazole, 10% glycerol, 0.5 mM TCEP, pH 8.0, Benzonase (4 uL per 750 mL cultivation) 250 U/uL from Merck, Protease Inhibitor Cocktail Set III, EDTA free (1000x dilution in lysis buffer) from Calbiochem and stored at -80°C. Cells were lysed on ice by sonication, and the lysate cleared by centrifugation and filtration. Filtered lysates were loaded onto 1 mL Ni-NTA Superflow (Qiagen) in IMAC Wash buffer 1 (20 mM HEPES, 500 mM NaCl, 10 mM imidazole, 10 % (v/v) glycerol, 0.5 mM TCEP, pH 7.5) and subsequently washed with IMAC Wash buffer 2 (20 mM HEPES, pH7.5, 500 mM NaCl, 25 mM imidazole, 10% (v/v) glycerol, 0.5 mM TCEP). Bound proteins were eluted with 500 mM imidazole and loaded onto a HiLoad 16/60 Superdex-200 column (GE Healthcare) pre-equilibrated with 20 mM HEPES, pH 7.5, 300 mM, NaCl, 10% (v/v) glycerol, 0.5 mM TCEP.

Large scale preparation of ALKBH2 (residues 56 -258): BL21 (DE3) transformed with pET-28b ALKBH2 (residues 56-258) was grown in 6 L 2XTY media supplemented with chloramphenicol (33 μg/mL) and kanamycin (30 μg/mL) at 37°C and 200 rpm to an OD_{600} of ~0.6-0.8. Protein expression was induced by addition of 0.5 mM IPTG (Gold Biotechnology). Cell was continuously grown for 16 hours at 16°C before harvested by centrifugation. The cell pellet can be stored at -80°C for future purification or can be immediately suspended to homogeneity in 50 mM sodium phosphate, pH8, 300 mM NaCl, 10% (v/v) glycerol and 5 mM β-ME. Purifcation was performed at 4°C. Cells were disrupted by French press and the lysate clarified by centrifugation and filtration. ALKBH2 was purified from crude cell lysate by passing through the Ni-NTA affinity column (GE healthcare), with gradient elution to 500 mM imidazole. It was followed by affinity chromatography using a 5 mL HiTrap Heparin HP column (GE Healthcare), with gradient elution to 1M NaCl. Further purification was performed by size exclusion chromatography using HiLoad superdex 75 16/60 (GE Healthcare) buffer exchanged into 10 mM Tris, pH 8, 100 mM NaCl and 5 mM β-ME. Final protein purity and concentration was determined via SDS-PAGE and NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific), respectively.

Large scale preparation of ALKBH5 (residues 66 -292): BL21 (DE3) Rosetta T1R transformed with pNIC28-Bsa4 ALKBH5 (residues 66-292) was grown in 6 L 2XTY media supplemented with chloramphenicol (33 μg/mL) and kanamycin (30 μg/mL) at 37°C and 200 rpm to an OD_{600} of ~0.6-0.8. Protein expression was induced by addition of 0.5 mM IPTG (Gold Biotechnology). Cell was continuously grown for 16 hours at 16°C before harvested by centrifugation. The cell pellet can be stored at -80°C for future purification or can be immediately suspended to homogeneity in 50

mM tris, pH8, 500 mM NaCl, 40 mM imidazole, pH8 and 5 mM β-ME. Purification was performed at 4°C. Cells were disrupted by French press and the lysate clarified by centrifugation and filtration. The protein was purified from crude cell lysate by passing through the Ni-NTA affinity column (GE healthcare), with gradient elution to 500 mM imidazole. It was followed by anion chromatography using a 5 mL HiTrap Q HP column (GE Healthcare), with gradient elution to 1M NaCl. Further purification was performed by size exclusion chromatography using HiLoad superdex 75 26/60 (GE Healthcare) buffer exchanged into 20 mM Tris, pH 8, 100 mM NaCl and 5mM β-ME. Final protein purity and concentration was determined via SDS-PAGE and NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific), respectively.

Large scale preparation of GST-YTHDF2 (residues 385-576): BL21 (DE3) transformed with pGEX-6P-1 YTHDf2 (residues 385-576) was grown in 6 L 2XTY media supplemented with chloramphenicol (33 μg/mL) and ampicillin (100 μg/mL) at 37°C and 200 rpm to an OD_{600} of around 0.6-0.8. Protein expression was induced by addition of 0.5 mM IPTG (Gold Biotechnology). Cell was continuously grown for 16 hours at 16°C before harvested by centrifugation. The cell pellet can be stored at -80°C for future purification or can be immediately suspended to homogeneity in 50 mM tris, pH 8, 500 mM NaCl and 5 mM β-ME. Purification was performed at 4°C. Cells were disrupted by French press and the lysate clarified by centrifugation and filtration. The protein was purified from crude cell lysate was load into 5 mL GST[TM] HP column (GE Healthcare). The protein was eluted with 50 mM tris, pH8, 300 mM NaCl, 10 mM glutathione and 5mM β-ME. Further purification was performed by size exclusion chromatography using HiLoad superdex 200 26/60 (GE Healthcare) and buffer exchanged into 20 mM Tris, pH 8, 150 mM NaCl and 5 mM β-ME. Final

protein purity and concentration was determined via SDS-PAGE and NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific), respectively.

2.3 Differential scanning fluorimetry (DSF)-based binding assay

DSF-based assay for FTO was performed by the lab of Dr Esther Woon [256]. DSF-based binding assays were performed using a MiniOpticon™ Real-Time PCR Detection System (Bio-Rad), monitoring protein unfolding using SYPRO orange (Invitrogen) according to the reported method. FAM (492 nm) and ROX (610 nm) filters were used for excitation and emission respectively. Reaction mixes contained 2 μM protein, 50 μM $MnCl_2$, 100 μM compounds, 5X SYPRO orange in a final volume of 50 μl. Reagents were prepared in 50 mM HEPES buffer, pH 7.5 excluding metals, which were dissolved as 100 mM stocks in 20 mM HCl, then further diluted in MilliQ water. Compounds tested were prepared in 100% DMSO and added such that the final concentration of DMSO was no more than 1% v/v of assay mix. Fluorescence readings were taken every 1°C in a range of 25-95°C, with the temperature increased linearly at a rate of 1 °C min^{-1}. The software provided was used to perform global minimum subtraction. The inflection point, representing T_m, was calculated by fitting the Boltzmann equation to the sigmoidal curves obtained; data were processed using GraphPad Prism 6.0™. The T_m shift caused by the addition of small molecules/ fragments was determined by subtraction of the "reference" T_m (protein incubated with metal and 1% v/v DMSO) from the T_m obtained in the presence of the compound. The assay was performed in triplicate for each inhibitor, with standard deviations typically less than 1°C.

2.4 High performance liquid chromatography (HPLC)-based inhibition assay

HPLC-based assay were performed by the lab of Dr Esther Woon with modifications from previously published protocols [91, 246, 257]. Briefly, the assays were performed in triplicate for each inhibitor, with a final reaction volume of 25 μL. In general, each reaction consisted of the 2OG oxygenase (1 - 4 μM), 2OG (50 - 150 μM), $(NH_4)_2Fe(SO_4)_2.6H_2O$ (150 μM), L-ascorbate (2 mM) and methylated substrate (10-100 μM) in 50 mM HEPES buffer, pH 7.5 and inhibitors at various concentrations (final DMSO concentration no more than 1% v/v of assay mix). The reactions were typically incubated at 4°C or room temperature for 30 min to 1 h. Demethylation of substrates were quantified by HPLC system separated by Zorbax C18 column (4.6 mm x 250 mm) or Dionex DNAPac PA200 anion-exchange column (8 μm, 4 mm x 25 mm). The UV detection wavelength was set at 266 nm. The percentage inhibition was estimated based on integrated peak areas of substrates and product. MALDI-TOF MS was also used to quantify the demethylation of some methylated oligonucleotides via estimation based on relative intensities observed in the mass spectra. Typically, eight different concentrations (0, 1, 3, 10, 30, 100, 300, 1000 μM) of inhibitors were used for IC_{50} determination, which is calculated using nonlinear regression, with normalised dose-response fit on GraphPad Prism 6.0^{TM}.

2.5 MTT cell viability assay

MTT cell viability assay was performed by the lab of Dr Esther Woon with some modifications [246]. Human HeLa cells were seeded in 96-well plates at a density of 1×10^4 cells/well, and allowed to grow for 24 h before being exposed to varying concentrations of compound **25** (0, 5, 10, 50 and 100 µM). Cells that were treated with 1% v/v DMSO served as negative control. After incubation at 37°C for 24 h, the supernatant was removed, and each well was washed twice with PBS. 20 µL of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) solution (5 mg/ml in PBS) and 100 µL of medium were then introduced, and the solution was incubated at 37°C for another 4 h. The resultant formazan crystals were dissolved in 100 µL DMSO and its concentration is determined by measuring its absorbance intensity at 570 nm using a microplate reader. A background wavelength of 630 nm was used. Cell viability was presented as a percentage relative to the untreated control cells. All experiments were performed in triplicate.

2.6 Cell-based inhibition assay

Cell-based inhibition assay was performed by the lab of Dr Esther Woon. The assay was performed as previously reported [246]. Human HeLa cells were treated with either 10 µM or 50 µM of compound **25** or with 1% v/v DMSO (negative control) and allowed to incubation at 37°C for 6 h. Then the cells were harvested and the total RNA were isolated with TRAZOL Reagent (Invitrogen). Subsequently, the mRNA was extracted using PolyATtract® mRNA Isolation Systems (Promega), and the rRNA contaminants were removed using RiboMinus Transcriptome Isolation Kit (Invitrogen). The concentration of the resultant mRNA was measured using NanoDrop.

To digest the mRNA to individual nucleotides, 1 μg of mRNA was treated at 37ºC for 1 hour with 2 Units of nuclease P1 in 40 μL buffer containing 25 mM of NaCl, and 2.5 mM of $ZnCl_2$, followed by the addition of 1 M NH_4HCO_3 (3 μL), and alkaline phosphatase (0.5 Unit). After incubation at 37ºC for 3 h, the solution was diluted 3 times, and analyzed by a HPLC system. The nucleosides were separated using a Zorbax C18 column (4.6 mm x 250 mm), eluted by solution mixture of solvent A (water + 0.1% TFA) and solvent B (MeOH), with a gradient decrease of solvent A (95% to 40%) over 8 min, at a flow rate of 1mL/min at room temperature. The UV detection wavelength was set at 266 nm. The concentrations of N6-methyladenine (m6A) and adenine (A) were determined based on their peak areas, using calibration plots obtained from pure nucleoside standards. The percentage of m6A/A was then obtained using the calculated concentrations.

2.7 Screening and optimisation for FTO crystallisation condition

Screening for FTO crystallisation condition: ΔN31-FTO at concentrations of 6 mg/mL, 8 mg/mL and 12 mg/mL were subjected to co-crystallisation screening. Co-crystallisation mixture consisted of ΔN31-FTO, nickel (II) chloride and NOG or 2,4-PDCA. Crystallisation screening was performed using an automated liquid handling robot, Phoenix (Art Robbins Instruments) dispensing sample drops into microplate (INTELLI-PLATE® 96) at three different protein sample to reservoir ratios (1:2, 1:1, and 2:1). Commercial screening kits from Hampton Research and Emerald BioSystems were used for initial high throughput screening. After dispensing, the microplates are sealed by a UV transmissible film, spin down and stored in Rock Imager (Formulatrix) at 4°C and 20°C for scheduled imaging. Initial protein crystal formation was judged by visible light photos and trans UV illumination.

Optimisation of FTO crystallisation condition: Grid screening was carried out using a 24 well VDX plate (Hampton Research). The optimisation was performed using hanging drop vapour diffusion at 20°C. The pH of crystallisation buffer, 100 mM sodium citrate was varied from 5.0 to 6.5 across the row while the percentages of the additive, iso-propanol were varied from 2% (v/v) to 20% (v/v) across the column. Additives (methanol, ethanol, n-butanol and t-butanol) that have structural resemblance to isopropanol were used at various percentages (v/v) during the optimisation screening in place of isopropanol. Different percentages (w/v) of polyethylene glycol (PEG) precipitant 3350 and 4000 were varied from 5 to 20% w/v in the optimisation drop. Each well consisted of 400 μL of the reservoir solution. The ratios of protein to reservoir solution were 2 μL : 1 μL, 1 μL : 1 μL and 1 μL : 2 μL, respectively.

2.8 FTO complex co-crystallisation

Crystals of ΔN31-FTO in complex with compounds **12**, **16**, and **21** were grown in hanging drops at 20°C using hanging drop vapour diffusion methods. The ratio of protein to reservoir solution was 2:1 (total drop volume: 2.25 μL). The sample contained 8 mg/mL protein and 1.5 mM ligand, in the buffer 25 mM Tris, pH 7.5, 50 mM NaCl, 5 mM β-ME, 1 mM $NiSO_4$. The reservoir solutions for ΔN31-FTO in complex with **12**, **16**, and **21** contained 100 mM sodium citrate, pH 5.6, 10-12% w/v polyethylene glycol (PEG) 3350, 10-15% t-butanol. After 1 week, the crystals were harvested, cryoprotected using well solution with 22% glycerol, and flash frozen in liquid nitrogen.

2.9 Data collection and structural refinement

Crystal screening and data collection was carried out at the Swiss Light Source (SLS), Switzerland. Crystal diffraction data of protein ΔN31-FTO complexed with **12** and **16** were screened and collected on PXI beamline of SLS. Crystal diffraction data of ΔN31-FTO protein complexed with **21** were screened and collected on PXIII beamline of SLS. All diffraction data were collected at 100K, then indexed, integrated, scaled and merged by X-ray Detector Software (XDS) [258]. The phases were determined and solved by molecular replacement using Phaser-MR subroutine in Phenix with 4IE5 (PDB ID) as a search model [259]. The correctness of the molecular replacement was determined by the translation function Z (TFZ) and log-likelihood gain (LLG) score [260]. Iterative rounds of refinement and model building were carried out using Phenix and COOT, respectively, until R_{work} and R_{free} no longer converge [261]. The PDB coordinate models of the inhibitors were generated by PRODRG [262]. The geometry and metal ion coordination restraint files for refinement were prepared by ReadySet subroutine in Phenix.

2.10 Computational tool for protein-oligonucleotide energy minimisation

PDB models were prepared using PyMOL prior to submission to the YASARA energy minimisation server [263]. The fitting of 5'-GGm6ACU-3' onto the domains of both FTO (PDB ID 3LFM) and ALKBH5 (PDB ID 4NRO) was by superimposing AlkB (PDB ID 4NID) using pymol. The m6A containing dsDNA was then edited by PyMOL leaving 5'-GGm6ACU-3' placed in the domain of FTO and ALKBH5. Both PDB models were then submitted to YASARA energy minimisation server.

2.11 Circular dichroism (CD) spectroscopy

CD spectra were obtained with a JASCO J810 spectro-polarimeter. Oligonucleotides were prepared in 10 mM sodium phosphate, pH 7.4 and 150 mM NaCl buffer. 5 μM of oligonucleotides (unless stated otherwise) were heated to 90°C for 5 min and re-annealed by slow cooling to 4°C at a decrease rate of 1 °C/min. CD spectra were recorded in quartz cuvettes (path length 10 mm, 400 μL) from 200 nm to 350 nm using a 10 nm/min scan speed, with a spectral band width of 1 nm and a time constant of 4 s. All the spectra were subtracted with the buffer blank and smoothed using the Savitsky-Golay algorithm (polynomial order 10).

2.12 Non-denaturing polyacrylamide gel electrophoresis (PAGE) analysis

Annealed oligonucleotides were loaded to 20% native polyacrylamide gel and electrophoresis was performed at 4°C in Tris/Borate/EDTA (TBE) running buffer made up of 90 mM Tris, pH 8.3, 90 mM boric acid and 5 mM EDTA. The gel were stained with SYBR® Gold Nucleic Acid Gel Stain and visualised by Gel Dock XR + (Bio-Rad) and Image Lab 4.0 software (Bio-Rad). The fraction of monomolecular hairpin structures was evaluated based on the assumption that the efficiency of the staining of the base pairs in a hairpin was similar to that in a biomolecular duplex.

2.13 Electrophoretic mobility shift assay (EMSA)

RNA probes that were labelled at their 3'-ends with biotin were purchased from Keck Foundation Biotechnology Resource Laboratory. Prior to electrophoresis, 2 μL of annealed RNA probes (4 nM final concentration) were incubated with 2 μL of FTO or ALKBH5 or YTHDF2 (0.02, 0.5, 1, 2, 5 μM, and other concentrations as indicated) at 4°C for 30 min in a binding buffer containing 10 mM HEPES, pH 7.3, 5% glycerol,

20 mM KCl, 1 mM $MgCl_2$, 1 mM DTT and 8 U RNasin[®] Ribonuclease Inhibitor (Promega). 20 uL of the RNA-protein mixture was loaded to 7.5% native polyacrylamide gel and electrophoresis was performed at 4°C for 90 min at 90 V using TBE running buffer. Electrophoresis for ALKBH5 and FTO were performed in buffer without EDTA. Visualisation was carried out using CL-Xposure[TM] Film (Thermo Scientific), Gel Dock XR+ (Bio-Rad) and Image Lab 4.0 software (Bio-Rad).

2.14 Microscale thermophoresis (MST) measurement

MST experiments were performed on a Monolith NT Label Free system (NanoTemper Technologies) at 25°C using 40% MST power, and 60% LED power for FTO and 40% LED power for ALKBH5 [264], respectively. Laser on and off times were set at 30 s and 5 s, respectively. Zero background standard treated capillaries (NanoTemper Technologies) were used. All experiments were performed in triplicates. Oligos for MST measurement were purchased from Dharmacon[TM]. For determination of binding affinity with FTO, twelve different concentrations of RNA ranging from 1 mM to 490 nM were used. The annealed RNAs were incubated with a mixture of FTO/$NiSO_4$ complex (100 nM/1 mM) and NOG (500 μM) in 50 mM Tris buffer, pH 7.5, 150 mM NaCl and 0.05% Tween 20 at 25°C for 20 minutes. Data derived from thermophoresis measurement and temperature-dependent change in fluorescence (T-Jump) were used to determine the binding affinities (K_D) of the RNAs. Curve fitting was performed using the NT Analysis software provided. For experiments with ALKBH5, ALKBH5/$MnCl_2$ complex (500 nM/1 mM) was used.

CHAPTER 3

RESULTS

3.1 Bacterial expression of proteins

Bacteria expression using engineered *E. coli* represents one of the fastest and easiest ways to prepare large amounts of recombinant proteins for structural and biochemical studies. Gratifyingly, most of the AlkB family of proteins we worked with in this thesis were able to get soluble expression using bacterial expression systems [265].

3.1.1 Small scale protein expression check

Small scale protein expression tests for some of the AlkB constructs fused with N-terminal hexa-histidine tag were performed by PPP, NTU, using Ni-NTA beads (Fig. 15A-D). Small scale protein expression check of ALKBH2 was done in-house (Fig. 15E). The soluble expression and purity of the proteins were judged by SDS-PAGE to ascertain whether they are suitable for large scale preparation.

Small scale protein expression check of the N-terminal GST-tagged YTHDF2 was performed using Glutathione Sepharose 4 Fast Flow (GE Healthcare) beads (Fig. 15F). The soluble expression and purity of the protein pull-down by glutathione was judged by SDS-PAGE to ascertain suitability for scaling up.

Gratifyingly, the constructs displayed significant overexpression upon pull-down and were suitable for scaling up.
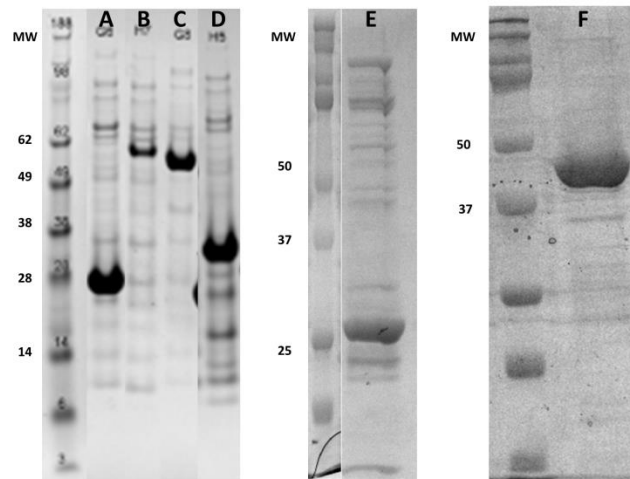
Figure 15. Small scale expression test. Lane **A** to **D** was performed by PPP. Lane **A**: ΔN11- AlkB (MW=25470 Da); Lane **B**: Full length FTO (MW=60966 Da); Lane **C**: ΔN31-FTO (MW=57139 Da); Lane **D**: ALKBH5 (AA66-292) (MW=28704 Da); Lane **E**: ALKBH2 (AA56-258) (MW=25257 Da); Lane **F**: GST-YTHDF2 (385-576) (MW=49345 Da).

3.1.2 Large scale protein preparation

After the constructs were shown to have good soluble expression, large scale purification was carried out (Fig.16). Due to stringent buffer requirement, only purification ΔN11-AlkB was performed by PPP (Fig. 16C). Proteins after final size-exclusion chromatography purification step were at least 95% pure by SDS-PAGE analysis and were collected for assays and crystallisation.
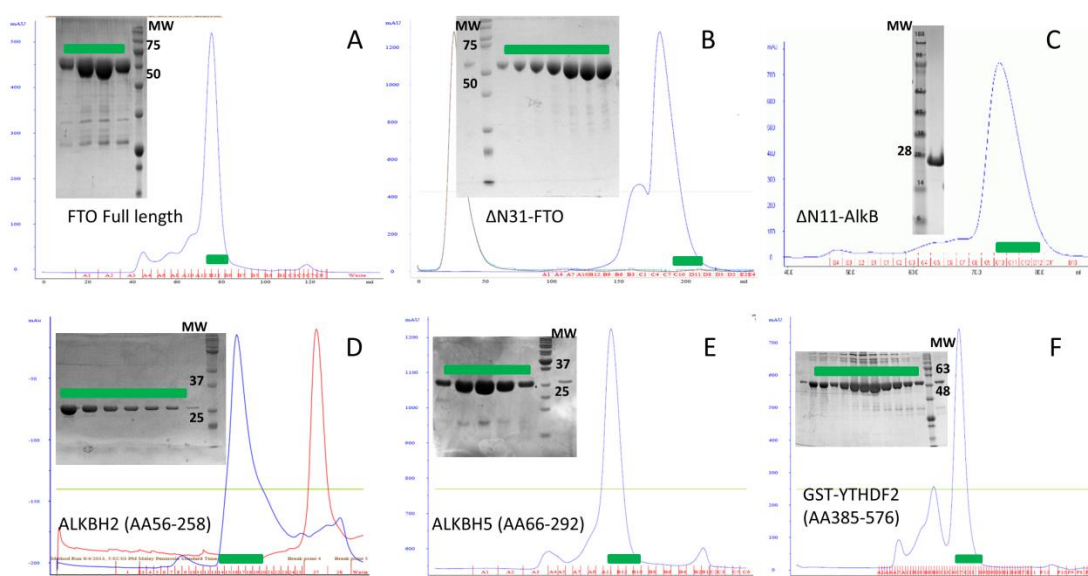
Figure 16. SDS-PAGE and final gel filtration profile of large scale protein purification. **A**: Full length FTO; **B**: ΔN31-FTO; **C**: ΔN11-AlkB; **D**: ALKBH2 (AA56-258); **E**: ALKBH5 (AA66-292); **F**: GST-YTHDF2 (AA385-576). Fractions highlighted green were collected for assays and/or crystallisation.

3.2 Substrate selectivity of FTO and other AlkB oxygenases

NOG binding: An analysis of substrate selectivity of FTO and other AlkB oxygenases was carried out prior to the commencement of the project. To do this, various cDNAs of AlkB oxygenases were cloned into constructs that were reported to result in active recombinant proteins. DSBH domain is a typical characteristic of AlkB oxygenases. Freshly purified recombinant proteins were probed for DSBH structural integrity via DSF using NOG (Table 1). In the presence of a metal (eg. Fe(II), Ni(II), Mn(II) or Zn(II)) and generic inhibitor NOG, thermal shifts (T_m shift) of more than 2°C relative to the apo proteins were observed, which was caused by additional stability afforded to the DSBH domains of the proteins. Note that a thermal shift (T_m shift) of more than 2°C is considered to be significant. *In vitro* enzymatic activity assays were also performed to test the activity of recombinant proteins. DSF and enzymatic assays were performed by the lab of Dr Esther Woon.

	NOG	
	IC_{50}/ µM	Tm shift/°C
FTO	36.5	4.1
AlkB	108.3	2.7
ALKBH2	124.7	2.3
ALKBH3	96.8	2.9
ALKBH5	145.2	2.1

Weak   Moderate   Strong

Table 1. T_m shift and IC_{50} data of AlkB oxygenases against NOG. NOG is a weak to moderate inhibitor to AlkB enzymes. AlkB oxygenases generally have a T_m shift of more than more than 2°C for NOG. (Data obtained from Dr Esther Woon.)

M3T binding: All the AlkB recombinant proteins reported in this thesis were active and able to demethylate their respective *N*-methylated nucleotide substrates [245]. Consistent with the previous reports, recombinant FTO was able to demethylate m3T nucleotide reasonably well [91]. In line with this, DSF of FTO with m3T showed a large T_m shift of 12.1°C (Table 2). Demethylation of m3T was observed at very low efficiency for other AlkB oxygenases, with weak T_m shift observed by DSF. Notably, FTO was unable to demethylate m6A nucleotide although m6A containing RNA oligonucleotides were reported to be its 'natural' substrate. Similar to FTO, ALKBH5 was unable to demethylate m6A mononucleoside although ALKBH5 bind strongly to m6A mononucleoside with an observed T_m shift of 9.8°C. The natural substrate for ALKBH5 was reported to be m6A. FTO and ALKBH5 were unable to bind strongly with oligonucleotides as observed from the T_m shift assay. Despite the weak binding, FTO and ALKBH5 were able to demethylate their respective m3T and m6A containing oligonucleotide substrates very well.

| | Tm shift/ °C | | | | |
	FTO	AlkB	ALKBH2	ALKBH3	ALKBH5
m3T	12.1*	1.4	1.1	1.3	1.1
GG(m3T)CT	2.9*	2.5	1.0	3.3	3.1
m6A	2.4	2.6	1.3	2.1	9.8
GG(m6A)CU	3.2*	3.0	0.8	3.1	2.9*

Weak   Moderate   Strong
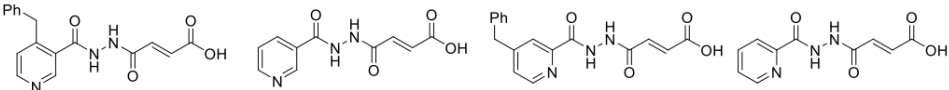
Table 2. T_m shift data of AlkB oxygenases to m3T and m6A containing nucleotides. Only FTO bind to m3T strongly while only ALKBH5 bind to m6A strongly. All AlkB oxygenases bind weakly to 5-mer oligonucleotides. Asterisks (*) indicate good *in vitro* demethylation activity of the enzymes toward the substrates. (Data obtained from Dr Esther Woon.)

3.3 Inhibitory activity of 12 and other analogs *in vitro*

Strategy employed for inhibitor design will be discussed further in chapter 4.2. Briefly, we have established that FTO and other human AlkB 2OG oxygenases have varied substrate preferences (chapter 3.2). Thus, we envisaged the design of inhibitors that simultaneously occupy the 2OG-binding and substrate binding sites to obtain potency and selectivity. Many two-component small molecules were designed, of which, inhibitor **12** demonstrated best profile (Table 3).

Inhibitor **12** demonstrated the highest potency and selectivity (IC$_{50}$ = 0.81 µM; T_{m} = 11.2°C) among all the inhibitors studied (Table 3). Inhibitory activities profiling of **12** and other analogs were performed by the lab of Dr Esther Woon. DSF and HPLC based demethylation assays were the two main assays used to identify and quantified the potency and selectivity of **12** and other analogs towards FTO. The table of data shown below are inhibitory profiles of selected small molecules bought or synthesised by the lab of Dr Esther Woon that are relevant to the discussion in this thesis.

In general, fumarate hydrazide compounds give a better binding and inhibitory profile towards FTO. Thus, fumarate hydrazide is preferred over succinate and maleate hydrazide for the 2OG-binding component when designing the inhibitors. For the substrate-binding component, 3-pyridyl compounds preferred to bind and inhibit FTO compared to their 2-pyridyl counterparts. DSF and HPLC demethylation assays were used to quantify the binding and inhibitory activities respectively. Generally, a higher T_{m} shift value gives a lower IC$_{50}$ value.

	12		21		22		23	
	IC$_{50}$/ μM	Tm shift/^{o}C	IC$_{50}$/ μM	Tm shift/^{o}C	IC$_{50}$/ μM	Tm shift/^{o}C	IC$_{50}$/ μM	Tm shift/^{o}C
FTO	0.81	11.2	11.8	7.4	86.8	3.0	109.4	2.7
AlkB	33.5	4.5	168.8	1.8	67.1	3.2	181.1	1.5
ALKBH2	25.9	4.9	152.3	2.0	39.9	3.7	127.1	2.3
ALKBH3	66.2	3.2	114.6	2.5	53.5	3.5	167.8	2.0
ALKBH5	108.1	2.7	167.9	2.1	104.6	2.8	156.7	2.1

Weak   Moderate   Strong

Table 3. Selectivity and inhibitory profile of selected small molecules. 12 shows the best potency and selectivity toward FTO. Position of nitrogen in the pyridyl ring is crucial for selectivity and potency. (Data obtained from Dr Esther Woon.)

3.4 Cellular toxicity and activity of 12

Cellular toxicity and efficacy profiling were performed by the lab of Dr Esther Woon (Fig. 17). The selection and discussion of data are accorded to the relevance of this thesis. The compound **25**, an esterified form of **12**, was used for cellular activity test to increase the penetrability and thus availability to the cells.

HeLa cells could tolerate the treatment of **25** well. Over a 24 hours period, more than 80% of the cell remained viable when treated with 100 µM of **25**. Quantification of cell viability was calculated with respect to a negative control containing 1% v/v DMSO.

Cells treated with 10 and 50 µM of **25** were found to have a notable increase of m6A in the mRNA compared to the control. The amount of m6A was quantified by UV absorbance using HPLC relative to known standards. It was found that there is an increase of 19% and 36% for cells treated with 10 and 50 µM of **25** respectively. Although more data points are desired, the inhibition by **25** appeared to be dose dependent.
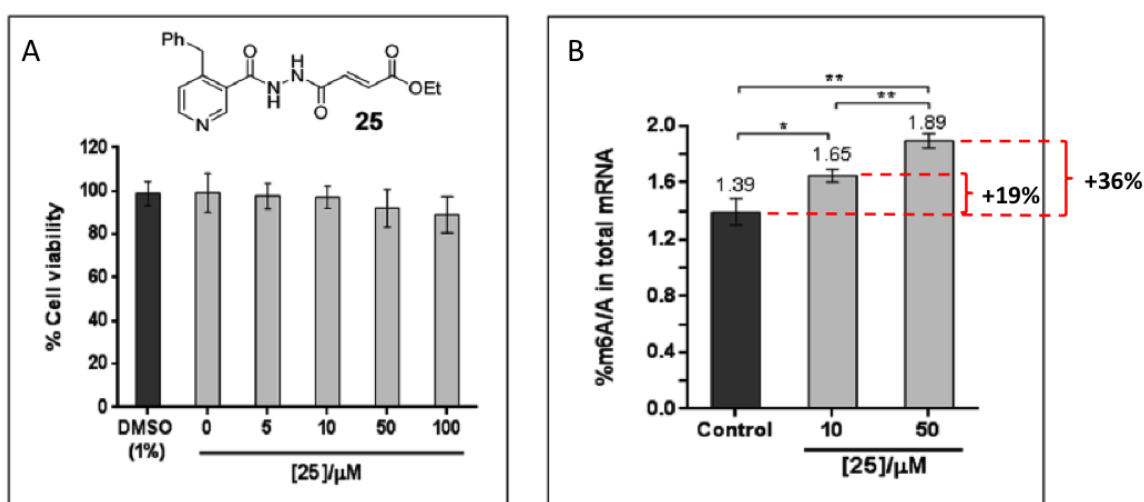
Figure 17. Cellular profile of 12. 25, an esterified from of **12** was used. **A.** Cellular toxicity. **B.** Cellular efficacy. (Data obtained from Dr Esther Woon.)

3.5 Crystallisation screening and optimisation

The crystal structure of ΔN31-FTO was already reported prior to the start of this project [91, 247]. Initial attempts to repeat the crystallisation proved to be futile despite following the published protocol strictly. Co-crystallisation was performed during the crystallisation attempts with the Fe(II) co-factor, generic inhibitor NOG and co-substrate m3T. Generic inhibitor 2,4-PDCA was also tried in place of NOG for co-crystallisation screens. In the very first few attempts for FTO crystallisation, ΔN31-FTO with N-terminal hexa-histidine tag and TEV linker were used. This was done as we believed that the relatively short linker would not interfere with the nucleation and crystal packing during crystallisation of FTO.

Unable to repeat the crystallisation experiment, ΔN31-FTO was cloned into pET28b to resemble the construct of the first reported crystal structure, ΔN31-FTO with N-terminal hexa-histidine tag with thrombin protease cleavage linker. However, numerous attempts of grid screening to explore the crystallisation space around the reported crystallisation (100 mM sodium citrate, pH 5.6 and 10% PEG 3350) yielded no sign of crystallisation. Additional variations around the reported condition were explored. For instance, temperature, final buffer, concentration of ΔN31-FTO and the ratio of protein to reservoir were varied. Despite numerous permutations around the reported condition, no crystallisation was observed.

Crystal screening: Next, we decided to screen for ΔN31-FTO crystallisation conditions from commercial screening kits. A few concentrations (6, 8, 12 mg/mL) were subjected to screening. Screening with apo protein and co-crystallisation was carried out. Co-factors were $NiCl_2$, $FeCl_2$, $ZnCl_2$ and $MnCl_2$. NOG and 2,4-PDCA were used during co-crystallisation in place for 2OG (will result in turnover) in hope

to stabilise the DSBH domain of FTO, thus promoting crystallisation. Gratifyingly, crystals were finally obtained in PEGRx (Hampton Research) screening kit. Crystalline formation was observed in well condition G4, 0.1 M sodium citrate pH 5.5, 20% (w/v) PEG 4000 and 18% (v/v) iso-propanol. The crystals were observed in well 1 and well 2 with protein to reservoir ratio of 2:1 and 1:1 respectively. UV illumination on the wells resulted in the crystals 'shining' against the dark background (Fig 18). This reduced the possibility of my crystals being a salt as only tryptophan residues in protein emit light at high intensity upon UV absorption. The crystals were small, highly nucleated and were unpick-able for X-ray diffraction test.
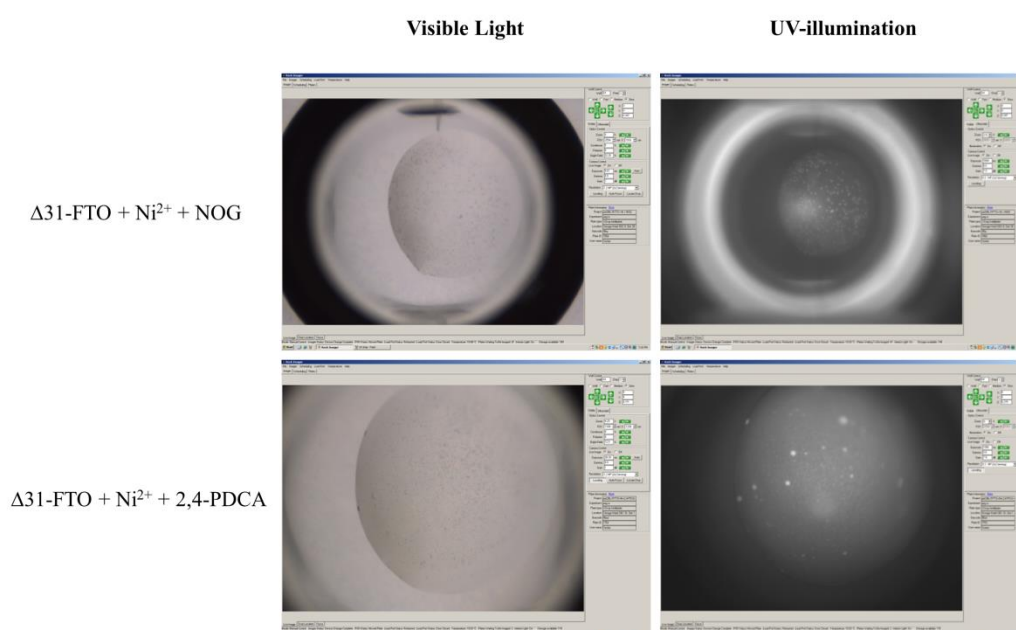
Figure 18. Visible light and UV-illumination images of Δ31-FTO-complex crystals. FTO crystals were small and highly nucleated. FTO crystals shone brightly against the background under UV-illumination.

Crystal optimisation: Grid screening of crystallisation space around the hit condition was carried out using hanging drop. Grid screening allowed two-dimensional exploration of crystallisation space. The pH of the buffer and percentage of precipitant were subjected to initial screening. Further investigation revealed pH does not affect the size and nucleation of the crystals. However, lowering the percentage of precipitant resulted in better nucleation (lesser crystal formation). Finally, changing the type of precipitant from 10% (w/v) PEG 4000 to 10% (w/v) PEG 3350 resulted in slightly bigger crystals which could be picked. The crystals were observed to adopt a cubic form of morphology (Fig. 19A). At this point, crystals were picked, cryo-protected and could diffract up to 3.5 Å at SLS. The diffraction data could also be processed by XDS and solved by molecular replacement using phases derived from previously solved structure (PDB ID: 4IE5). However, bigger crystals were needed to achieve higher resolution.

Before optimisation

After optimisation
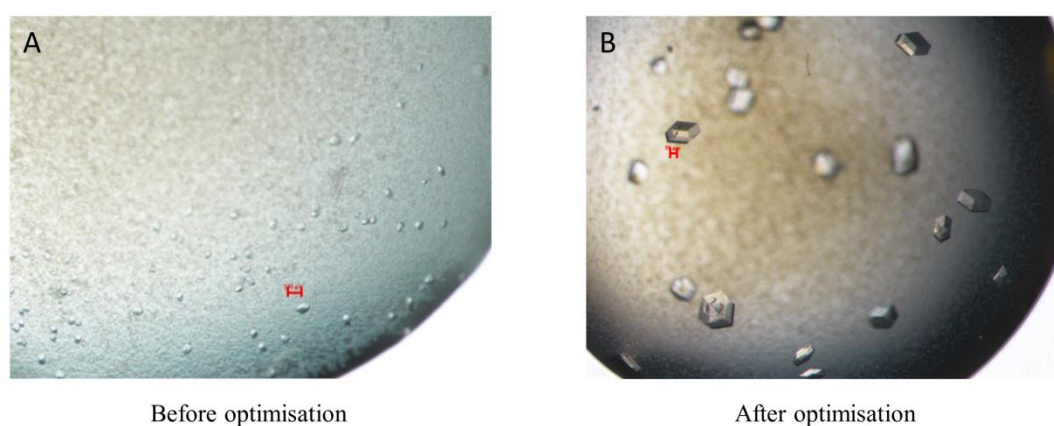
Figure 19. Δ31-FTO crystal hit optimisation. A. Small and highly nucleated crystals that grew in 100 mM sodium citrate, pH 5.6, 10% PEG 4000, 15% iso-propanol. **B.** Large cuboid crystals that grew in 100 mM sodium citrate, pH 5.6, 10% PEG 3350, 15% t-butanol.

During the period of optimising the condition of FTO, a paper on various co-crystal structures of FTO was published. 4% (v/v) t-butanol was used as additive in place of iso-propanol [247]. This prompted us to screen various organic\ solvents with similar structure as iso-propanol. Gratifyingly, at around 12% (v/v) t-butanol, crystals of FTO grew to a much bigger and pick-able size within a week (Fig. 19B). These FTO crystals were able to diffract up to 2.7 to 3.0 Å at SLS which allowed us to observe small molecular inhibitors binding to FTO at atomic resolution. It is noteworthy that no additive was used for crystallisation condition of the first reported crystal structure of FTO [91]. We were unable to reproduce crystallisation of FTO supplemented with 4% (v/v) t-butanol as reported [247].

3.6 Structural determination

Diffraction experiments were conducted in Swiss Light Source (SLS) at Paul Scherrer Institute (PSI) (Switzerland). Most of the crystals diffracted up to ~5 to 6 Å. In addition, some crystals which were able to diffract beyond 3 Å do not have visible electron density for the target ligands after data processing by XDS and structure solved by MR. We managed to collect three full datasets for FTO co-crystals which showed unambiguous occupancy of our inhibitors at atomic resolution. Template (PDB ID: 4IE5) was used to obtain the phases for MR [247]. The crystals crystallised as a monomer with one molecule per asymmetrical unit, adopted a trigonal crystal system with a H3 space group. Finally, the structures were refined with nickel (II) and their respective inhibitors until R_{work} and R_{free} no longer converge (Table 4).

PDB ID	4CXW	4CXX	4CXY
Data collection			
Resolution (Å)	50.00-3.05	49.18-2.71	50.00-2.60
Resolution[#] (Å)	3.10-3.05	2.76-2.71	2.65-2.60
Space group	R3	R3	R3
Cell dimension			
a, b, c (Å)	141.6, 141.6, 83.8	140.7, 140.7, 83.3	141.9, 141.9, 83.8
α°, β°, γ°	90, 90, 120	90, 90, 120	90, 90, 120
No. reflections	59141	67330	101120
Unique	22939 (3570)	16658 (2707)	38122 (5970)
Redundancy[#]	2.6 (2.6)	4.0 (3.9)	2.6 (2.6)
Completeness	95.2 (90.5)	99.5 (99.4)	98.6 (95.6)
$I/\sigma(I)$ [#]	11.3 (2.0)	18.0 (1.9)	13.5 (1.0)
R_{merge} (%)[#]	6.1 (50.5)	4.5 (70.3)	5.1 (98.4)
Refinement			
R_{work} (%)	18.3	21.0	21.2
R_{free} (%)	24.6	26.3	24.5
RMS Deviation			
Bond length (Å)	0.011	0.011	0.010
Bind angle (°)	1.385	1.315	1.398
No. of water	65	113	103

Table 4. X-ray diffraction data and refinement statistics. [#]Values in parentheses are for the highest resolution shell.

3.7 Binding mode of inhibitors with FTO

As expected, the FTO in complex with various inhibitors contained two domains, the N-terminal DSBH domain and the C-terminal helical bundles. Overall, they did not display any significant difference compared to the previously published structures [91, 247].

In the 2OG binding site of FTO, Ni(II) which substituted in place of natural co-factor Fe(II), was observed to coordinate to three highly conserved residues (His231, Asp233 and His307), which form the catalytic triad characteristic of AlkB oxygenases (Fig. 20 -23). A water molecule coordinating to Ni(II) was consistently observed *trans* to His231. In addition, the Fo-Fc omit map depicted unambiguous occupancy of the inhibitors could be observed in the 2OG and nucleotide binding sites of FTO. PDB coordinates of the inhibitors generated by PRODRG server were able to fit snugly into the Fo-Fc omit map and refined for analysis of the mechanism of inhibition. These inhibitors did not induce any noticeable conformational change to the residues lining the FTO active site.

Inhibitor 12 (PDB ID 4CXW): The Fo-Fc omit map demostrated the umambiguous occupancy of inhibitor 12 in the catalytic domain of FTO (Fig. 20). The inhibitor **12** consists of two components, namely the 2OG-binding component that resembles fumarate and the substrate binding component that occupies the substrate binding site. The two components are tethered by a hydrazide functional group. The acylhdryazine group of **12** forms a bidentate coordination to Ni(II), completing a hexadentate coordination around Ni(II). C-4 carbonyl oxygen of **12** binds to Ni(II) *trans* (2.1 Å) to Asp233 and hydrazine nitrogen of **12** binds to Ni(II) *trans* (2.1 Å) to His307. The carboxylate oxygens from C-1 of **12** form salt bridges to Arg316 (3.0 Å and 3.1 Å). In

addition, these carboxylate oxygens form hydrogen bonding interactions with nearby hydroxyl group of Tyr295 (2.7 Å) and Ser318 (2.9 Å). Superimposition of 4CXW with 3LFM (FTO-m3T-NOG) complex revealed that NOG and the 2OG binding component of **12** overlay very well, demonstrating similar NOG/2OG contacts in both structures (Fig. 21).

The 4-benzyl pyridine that makes up the substrate binding component of **12** occupies the nucleotide substrate binding site of FTO (Fig. 21). The 4-benzyl pyridine and the nucleobase of m3T occupy almost similar position (Fig. 21). They are positioned almost parallel to His231 forming a π-π stacking interaction. They are also positioned in an almost parallel fashion as Tyr108, similarly forming a π-π stacking interaction, albeit weaker compared to the former (His231). The 4-benzyl pyridine aligns well with m3T of 3LFM. The pyridyl nitrogen is positioned similarly to O^4-m3T, which points away from the substrate entrance forming a hydrogen bonding interaction with the amide oxygen of Glu234 (2.9 Å). The pyridyl carbonyl oxygen of **12** forms a hydrogen bonding interaction with Arg322 (2.8 Å). This interaction is absent in the corresponding m3T. The 4-benzyl substituent of **12** extend toward the opening of the catalytic domain, occupying a similar position with the ribose sugar ring of m3T. They participate in a similar hydrophobic interaction with nearby residues Leu109 and Val228.

Figure 20. Crystal structure of FTO (white sticks) bound to 12 (orange sticks).
The diagram on left shows the detailed interaction map of **12** with its surrounding residues. The diagram on the right shows the F_o-F_c omit map with **12** fitted into it.
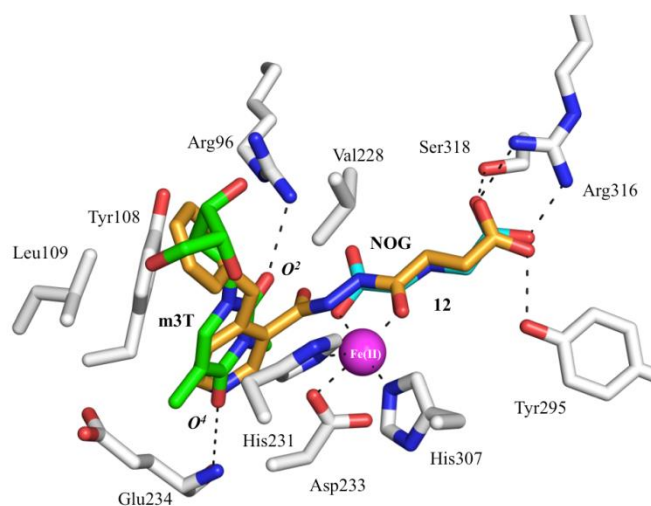
Figure 21. Crystal structure of m3T and NOG superposed with 12 (orange sticks). The 4-benzyl pyridine and the nucleobase of m3T occupy almost similar position. The 2OG-binding component of **12** aligns with NOG very well.

Inhibitor 16 (PDB ID 4CXX): The inhibitor **16** is an analog of **12**, with an additional t-butyl group at the *para* position of 4-benzyl substituent of **12** (Fig. 22). The t-butyl group points toward the opening of FTO catalytic domain and is observed to make hydrophobic interaction with the hydrocarbon ring of Pro93 and sidechain of Ile85.
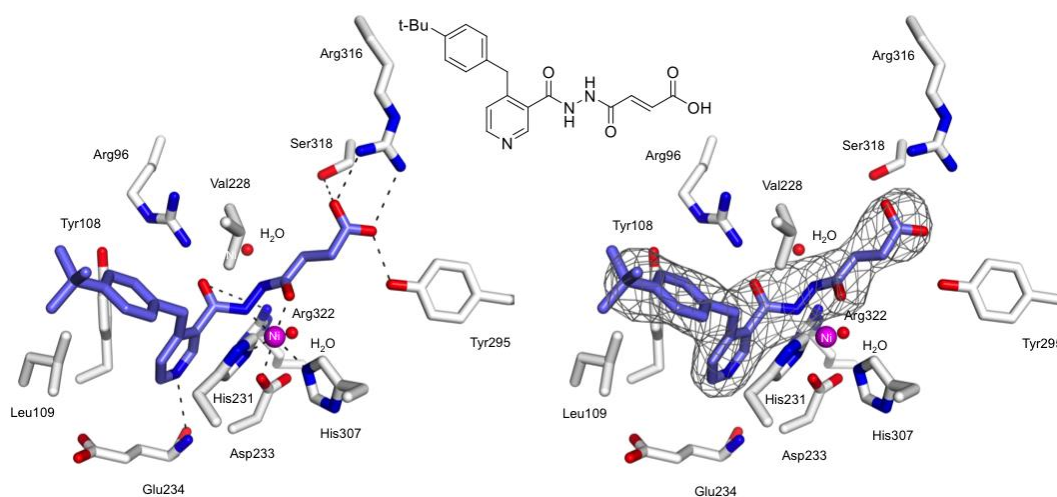
Figure 22. Crystal structure of FTO (white sticks) bound to 16 (blue sticks). The diagram on left shows the detailed interaction map of **16** with its surrounding residues. The diagram on the right shows the F_o-F_c omit map with **16** fitted into it.

Inhibitor 21 (PDB ID 4CXY): Inhibitor **21** is an analogue of **12**, with absence of the 4-benzyl substituent. **21** makes similar interactions as **12** (Fig. 23). Interestingly, an intense Fo-Fc omit map that is characteristic of metal (electron dense) was observed during the refinement of FTO-**21** crystal structure. The metal was presumed to be Ni(II) as it was present at high concentration during crystallisation. This observation was absent for inhibitor **12** and **16,** where the similar position was occupied by a water molecule. The nickel makes an interaction with the hydrazine nitrogen (2.9 Å) and oxygen (3.6 Å, nearer to the pyridien ring) that are not involved in the chelation of the active site Ni(II).
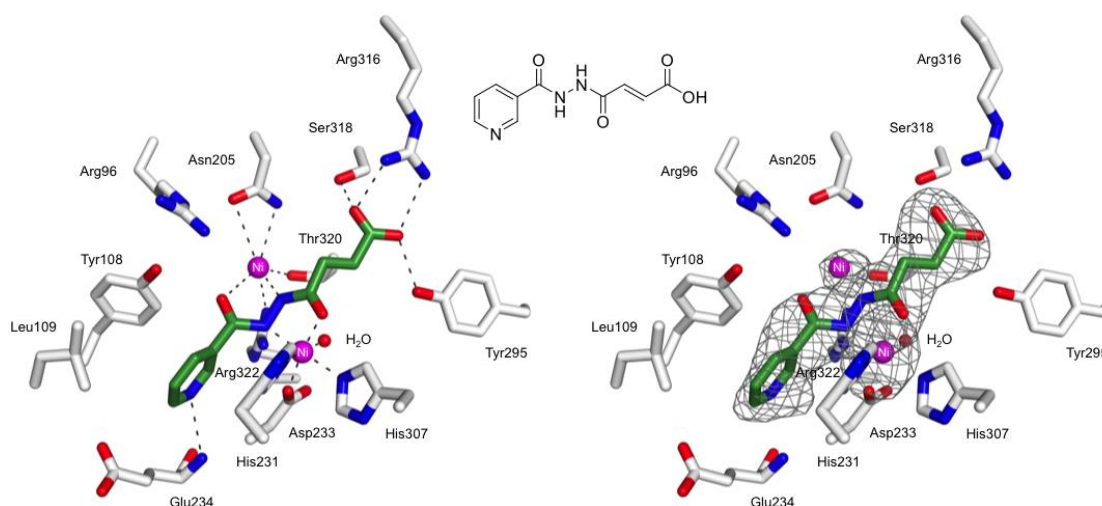
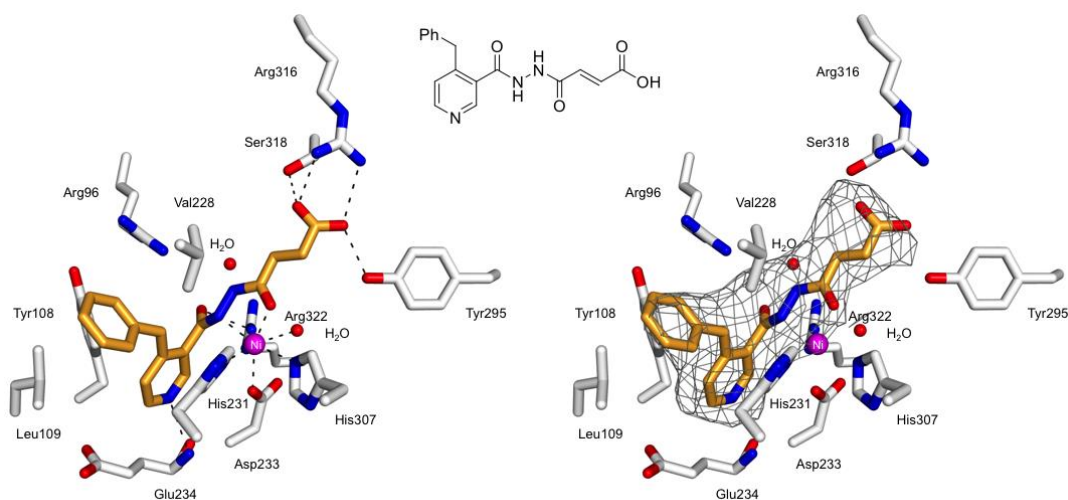**Figure 23. Crystal structure of FTO (white sticks) bound to 21 (green sticks).** The diagram on left shows the detailed interaction map of **21** with its surrounding residues. The diagram on the right shows the F_o-F_c omit map with **21** fitted into it.

3.8 Energy minimisation of FTO and ALKBH5 oligonucleotide complex

In silico modeling experiments were performed to understand if there are any conformational changes between enzyme inhibitor and substrate complexes. Not much conformational change was observed after the modelling experiments, and only slight rotation of Arg96 and Tyr108 in FTO has taken place (Fig. 24A). In ALKBH5, a flip of Tyr141 (indicated by red arrow) was observed when oligonucleotide was fitted into the domain (Fig. 24B). These residues are known to either make electrostatic interaction with phosphate backbone of oligonucleotide or form π-π stacking interaction with the nucleobases. Presumably, their conformational changes take place to make interaction with the oligonucleotide. FTO and ALKBH5 are most likely not to make a large conformation change upon substrate binding reflected by their inherent low affinity toward oligonucleotides. However, more experimental data such as structural studies are required to test this hypothesis.
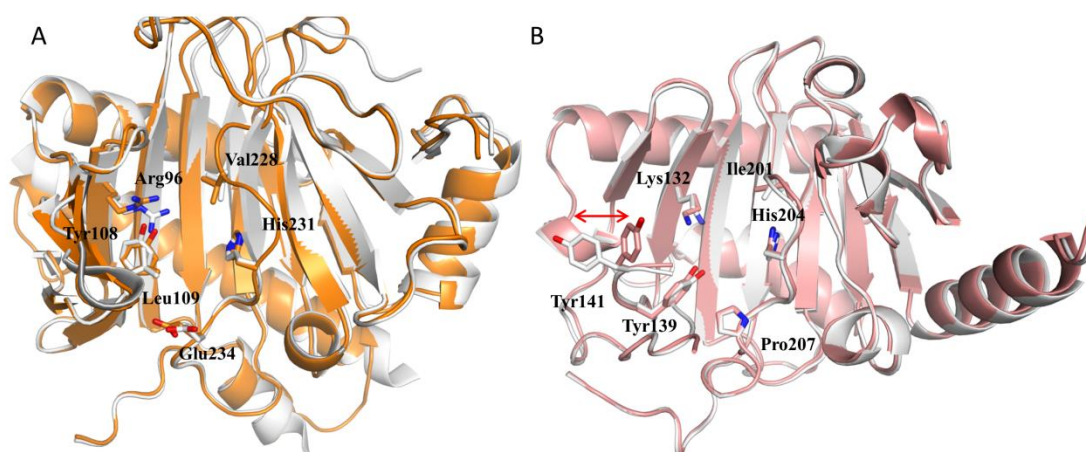
Figure 24. Energy minimised models. A. Unliganded FTO (white) superimposed with energy minimised FTO-GG(m6A)CU complex (orange) (RMSD of all atoms = 0.6 Å) . **B.** Unliganded ALKBH5 (white) superimposed with energy minimised ALKBH5- GG(m6A)CU complex (pink) (RMSD of all atoms = 0.3 Å). The C-terminal helical bundle of FTO and oligonucleotides were removed for clarity. A flip of Tyr141 is indicated by a double-headed red arrow.

3.9 Characterisation of RNA oligonucleotide substrates

Several purposefully designed RNA oligonucleotides (**RNA-1** to **6**) were investigated in order to shed light on the substrate preference of FTO (Fig. 25). The oligonucleotides **RNA-2**, **RNA-4** and **RNA-6** potentially could adopt different conformations, consequently their influence on preferential demethylation of FTO and ALKBH5 was investigated. The two conformations are:

1) Duplex conformation by intermolecular base pairing.

2) Hairpin conformation by intramolecular base pairing.
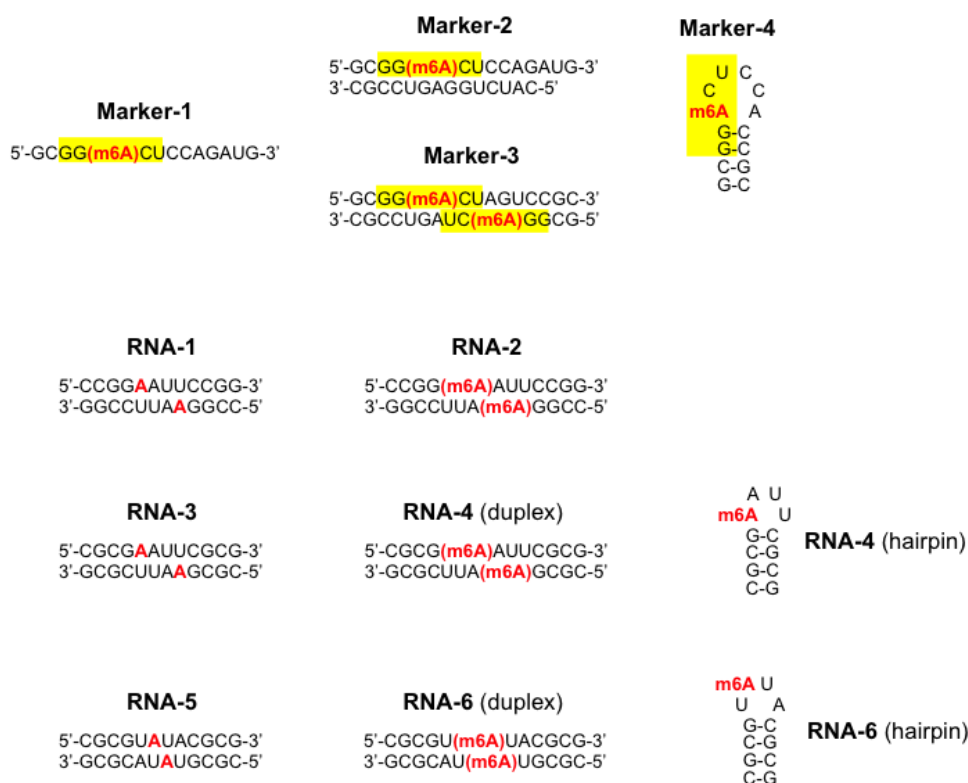
Figure 25. RNA sequences. Various standard markers were designed to aid in our investigation. **Marker-1** is a single-strand RNA standard. **Marker-2** and **3** are duplex standard. **Marker-4** is a hairpin standard. **RNA-1** to **6** are purposefully designed RNA oligonucleotides that we used for our studies.

3.9.1 CD spectroscopy

Unmethylated **RNA-1** (rCCGGAAUUCCGG) and methylated **RNA-2** (rCCGG(m6A)AUUCCGG) did not show significant conformation differences in their CD profiles. CD profiles of both oligonucleotides superimposed very well (Fig. 26A). Both exhibited a strong positive UV absorption band at 264 nm and a negative absorption band at the region around 212 nm, characteristic of A-form double helical structure [266]. **RNA 1** and **2** are believed to adopt a fully base-paired duplex structure based on the CD profiles.

Unmethylated **RNA-3** (rCGCGAAUUCGCG) and methylated **RNA-4** (rCGCG(m6A)AUUCGCG) showed significant conformation differences in their CD profiles (Fig. 26B). **RNA-3** has very similar CD profile with **RNA 1** and **2**. **RNA-3** has a strong positive UV absorption peak at ~264 nm and a negative absorption peak at ~212 nm, characteristic of A-form double helical structure. However, **RNA-4** has a distinctly different CD profile compared to its unmethylated analog **RNA-3**. The strong absorption peak was shifted from ~264 nm to ~280 nm. The CD profile of **RNA-4** still showed a negative absorption peak at ~212 nm. Based on the CD profile, **RNA-3** is believed to be fully base paired. However, **RNA-4** is believed to have a major change in its conformation due to introduction of m6A. It was later revealed that **RNA-4** existed as a mixture of duplex and hairpin conformation. The mixture of conformers resulted in the change of CD profile. Supported by other orthogonal experiments UV-melting experiments, Van't Hoff analyses and non-denaturing PAGE analyses, a mixture of conformers in **RNA-4** were indeed present.

Like **RNA-3** and **4**, unmethylated **RNA-5** (rCGCGUAUACGCG) and methylated **RNA-6** (rCGCGU(m6A)UACGCG) showed significant conformation differences in their CD profiles (Fig. 26C). **RNA-5** has a similar CD profile compared with that of **RNA-3**, demonstrating a certain characteristic of A-form double helical structure. **RNA-6** has a CD profile comparable to **RNA-4**, reflecting a mixture of duplex and hairpin conformation, in agreement with the results by other orthogonal experiments UV-melting experiments, Van't Hoff analyses and non-denaturing PAGE analyses [132].
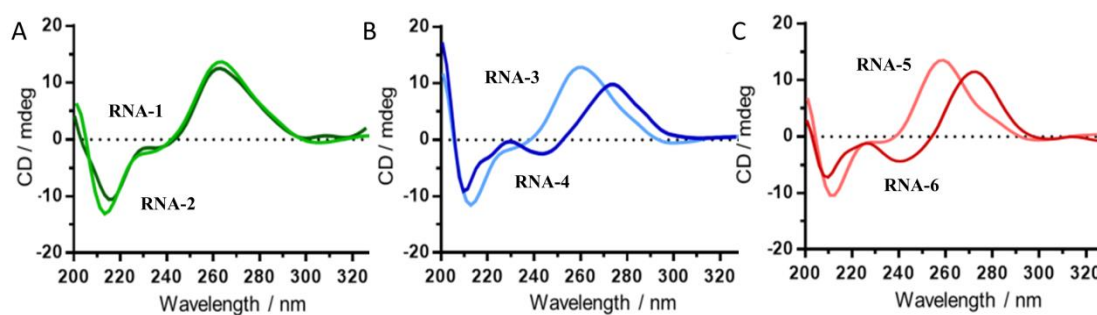
Figure 26. CD profiles of RNA oligonucleotides. A. RNA 1 and **2** showed similar CD profile. **B. RNA 3** and **4** showed different CD profile, indicating different conformation between methylated and non-methylated oligonucleotides. **C. RNA 5** and **6** showed different CD profile, indicating different conformation between methylated and non-methylated oligonucleotides.

3.9.2 Non-denaturing PAGE analysis

Non-denaturing PAGE was used to visualise the presence of different conformations of RNA oligonucleotides, given its structure/conformation directly linking to shift. As shown in Fig. 27, **Marker-1** (rGCGG(m6A)CUCCAGAUG) served as a single-stranded marker, and **Marker-2**, as a duplex conformation marker, is made up by annealing the same concentration of **Marker-1** and its complementary strand (rCAUCUGGAGUCCGC). Finally, **Marker-4** (rGCGG(m6A)CUCCACCGC) served as a hairpin marker. In a typical non-denaturing PAGE, **Marker-2** has the lowest mobility, whereas **Marker-4** has the highest mobility among the three markers. All the standard markers have different rate of migration, which could be deciphered unambiguously (Fig. 27).

Non-denaturing PAGE analysis of **RNA 1** to **6** at three concentrations (2.5, 5 and 10 μM) revealed the presence duplexes and hairpins, while compared with the standard markers. **RNA-1** to **3** and **5** have migration pattern similar to duplex standard marker. This observation is consistent with the CD profiles of **RNA-1** to **3** and **5** where they displayed clear A-form duplex structure.

However, **RNA-4** and **6** were demonstrated to form hairpin. Furthermore, high proportion of hairpin formation could be observed at relatively low concentration of **RNA-4** and **6** (2.5 μM and 5 μM). The hairpin band was inherently faint compared to the bands of duplex or single- strand at similar concentration. This is probably due to the current usage of SYBR® Gold Nucleic Acid Gel Stain not optimised for hairpin visualisation. Taking this into consideration, the percentage population of hairpin for **RNA-4** and **6** was at least 20% and 35% at 5 μM and 2.5 μM, respectively. The presence of both hairpin and duplex could be directly interpreted from the CD profiles

of **RNA-4** and **6**. Hence, we could infer that the hairpin formation is due to the presence of m6A.
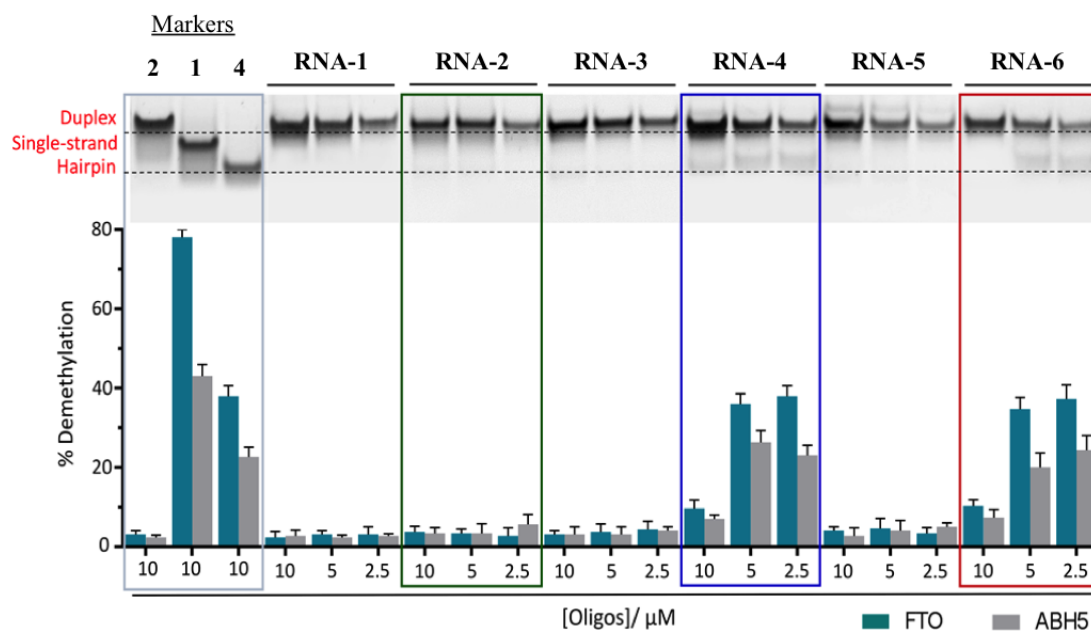
Figure 27. Native PAGE and *in vitro* demethylation of different conformations of RNA. FTO and ALKBH5 preferentially demethylate single strand and hairpin RNA oligonucleotides. (Demethylation data was from the lab of Dr Esther Woon.)

3.10 Characterisation of FTO and substrate interaction

Both m6A demethylases, FTO and ALKBH5 were noted to preferentially demethylate **marker-1** and **4**. However, there was no notable amount of demethylation activity with **marker-2** (based on the result from the lab of Dr Esther Woon).

3.10.1 Biotin-EMSA

The lack of demethylation activity of FTO and ALKBH5 toward duplex could be rationalised by the lack of binding affinity. There was no detectable binding of biotin-labelled **marker-3** to FTO and ALKBH5. When the proteins were incubated even at 1250 molar excess with respect 3'-biotinylated **marker-3** (4 nM), no retardation was observed (Fig. 28A and B).

On the contrary, demethylation activity of FTO and ALKBH5 towards single-stranded **marker-1** can be rationalised by the presence of binding affinity (Fig. 28A and B). Dose-dependent retardation of 3'-biotinylated **marker-1** could be observed in the biotin-EMSA experiment. Significant retardation could be observed when FTO and ALKBH5 was incubated at 500 molar excess with respect to 4 nM of 3'-biotinylated **marker-1**. Nevertheless, the affinity of FTO and ALKBH5 towards single-stranded oligonucleotide is considered to be weak based on the biotin-EMSA experiments.

No apparent binding of FTO and ALKBH5 towards hairpin **marker-4** was observed despite the ability of FTO and ALKBH5 to demethylate a hairpin (Fig. 28C). This may be due to 3'-biotin affecting the interaction of FTO and ALKBH5. Nevertheless, we manage to quantify the affinity of FTO and ALKBH5 via MST (describe in the next section).
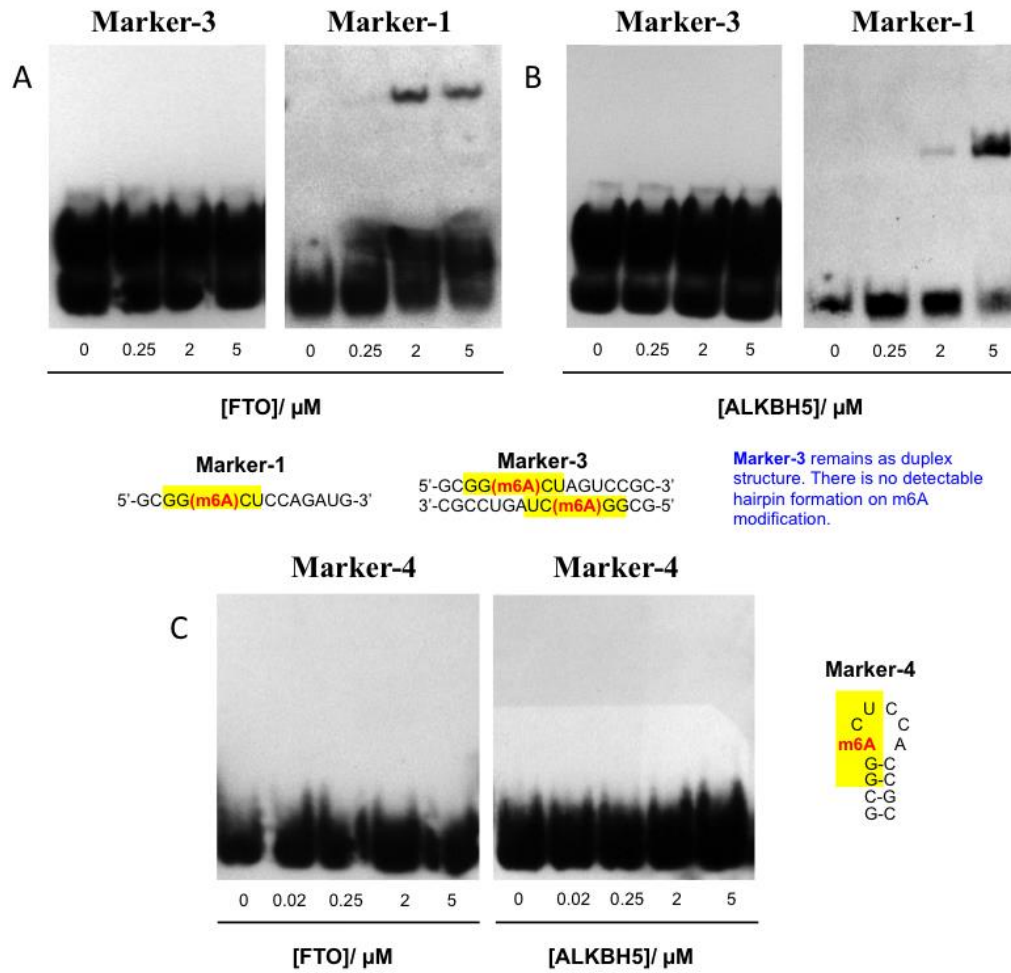
Figure 28. Biotin-EMSA of FTO and ALKBH5 with 3'-biotin labelled oligonucleotides. A. Biotin-EMSA of FTO with 3'-biotinylated **marker-1** and **3**. **B.** Biotin-EMSA of ALKBH5 with 3'-biotinylated **marker-1** and **3**. **C.** Biotin-EMSA of FTO and ALKBH5 with 3'-biotinylated **marker-4**.

3.10.2 MST (label-free) assay

MST allows an accurate determination of binding affinity of proteins to their binding partners (without labels) through quantification of dissociation constant (K_D). The quantification of the affinity between proteins and their partners is based on the population of proteins at equilibrium upon thermophoresis. The population of the proteins are quantified by the intrinsic fluorescence of the fluorophore, tryptophan residues. As such, protein without tryptophan is unable to be quantified via MST. In addition, binding partners of proteins should not have or possess at most a third of protein fluorescence such that interference of the protein fluorescence measurement is kept to the minimal.

Gratifyingly, the intrinsic fluorescence of **marker-1** (single-strand) and **marker-4** (hairpin) was less than three times of the fluorescence of FTO (100 nM) and ALKBH5 (500 nM). This allowed us to determine the binding affinity of **marker-1** and **marker-4** for the proteins (Fig. 29). As expected, FTO and ALKBH5 has a weak binding affinity towards **marker-1** (K_D (FTO) ~97.1 μM; K_D (ALKBH5) ~52.9 μM) and **marker-4** (K_D (FTO) ~91.3 μM; K_D (ALKBH5) ~75.3 μM). The binding data was unsurprising as AlkB oxygenases have weak and transient affinity towards oligonucleotides.
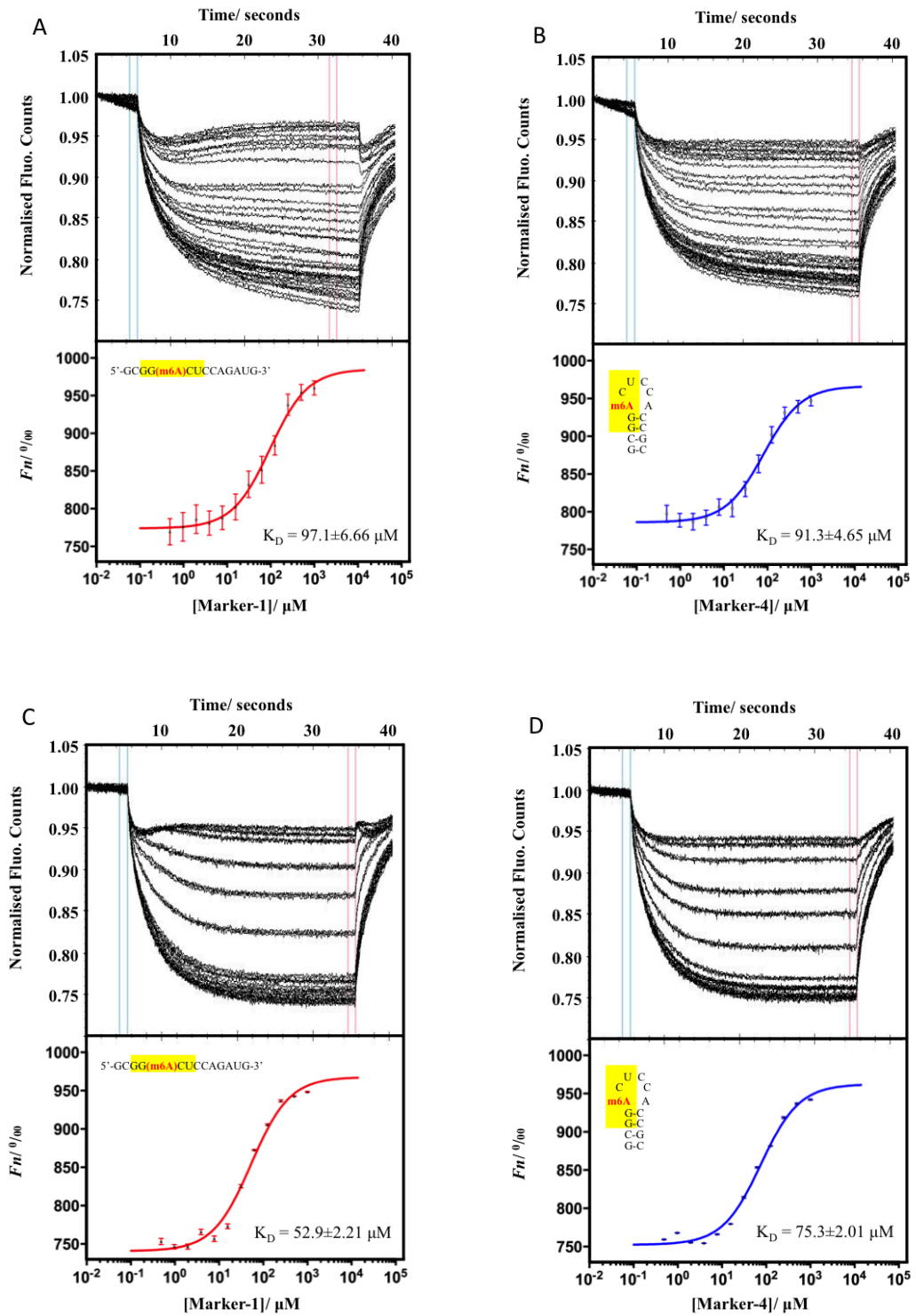
Figure 29. MST assay to quantify binding affinity of FTO and ALKBH5 to marker-1 and 4. FTO and ALKBH5 have weak binding to **marker-1** and **4**. **A.** MST assay of FTO with **marker-1**. **B.** MST assay of FTO with **marker-4**. **C.** MST assay of ALKBH5 with **marker-1**. **D.** MST assay of ALKBH5 with **marker-4**.

CHAPTER 4

DISCUSSION

4.1 Structural basis of FTO-substrate affinity

4.1.1 FTO and oligonucleotides

Most of the AlkB oxygenases in this thesis were able to selectively demethylate *N*-methylated nucleobases of single-stranded or duplex oligonucleotides, while crystallographic data aided towards a better understanding of the recognition between AlkB oxygenases and oligonucleotides. Sample preparation of highly concentrated, pure and homogenous AlkB oxygenase-oligonucleotide complexes is a prerequisite to crystallisation trial, which proved to be challenging. Prior to the commencement of our structural and mechanistic study of FTO, different groups of scientists had tried to crystallise *E. coli* AlkB, an FTO homolog, in complex with DNA. Among which, Dr Chuan He stated that the crystallisation of *E. coli* AlkB is almost impossible. *E. coli* AlkB has inherently weak affinity for DNA with and without alkylation damage [217]. In addition, the inability of AlkB to recognise a specific sequence of nucleotides make it increasingly difficult to form a homogenous complex. Consequently, reviewed earlier in this thesis, Dr. Chuan He employed a disulphide cross-linking strategy to trap the oligonucleotides in a fixed conformation in AlkB and ALKBH2. The crystal structures of AlkB and ALKBH2 with covalently trapped oligonucleotides represented the only known type of AlkB-DNA complexes thus far. This provided us with vital information of how FTO could possibly interact with its nucleotide substrates.

FTO-ssRNA complex: Prompted by the ability of recombinant FTO to demethylate some oligonucleotide substrates *in vitro*, we attempted to interrogate the structure of FTO complexed to m3T-containing ssDNA and m6A-containing RNA oligonucleotides. Although crystallisation trials yielded isomorphous crystals that

diffracted up to 3 Å, there was no sign of the occupancy of oligonucleotides upon solving the crystal structures. DSF analysis showed weak binding (T_m shift ~3°C) of FTO to various methylated or unmodified oligonucleotides. This observation was supported by biotin-EMSA, in which over a thousand fold excess of FTO was required for a small amount of retardation to be observed. It was further corroborated by MST assay, where the binding affinity was quantified (K_D (FTO) ~97.1 μM), indicating rather weak binding. Hence, these data suggested that the interaction between FTO and oligonucleotides is inherently weak.

DSF was performed on some recombinant AlkB N-demethylases and the binding data indicated that AlkB oxygenases also interact weakly with the oligonucleotides (Table 2). Further investigations on m6A demethylase ALKBH5 also showed weak binding via biotin-EMSA and MST assay (Fig. 28 and 29).

Observation from the crystal structure of FTO revealed various basic and aromatic residues present in the NRL1, NRL2 and L1 loop that were reasonably speculated to interact with oligonucleotides. To further rationalise the weak binding between FTO and its substrate oligonucleotide, a modelling study of a m6A-containing RNA docked into the domain and subjected to energy minimisation was carried out. Ideally, if the binding between the two entities is favourable, conformational changes of various protein residues to interact with the phosphate backbone or nucleobases should be observed. However, minimal conformational change was observed. Similar observations were noticed for ALKBH2. When oligonucleotide was fitted into the domain of ALKBH5, there was a flip of Tyr141, which presumably to make an interaction with a nucleobase (Fig. 24B).

TET as a close family to AlkB: TET1-3 belong to another 2OG oxygenase family that is known to perform oxidation on methylated nucleotides (5mC, 5hmC and 5fC) through a similar mechanism to that of AlkB oxygenases (Fig. 1). Superimposition of the DSBH domains of TET2 with *E.coli* AlkB and FTO showed highly aligned secondary structures. However, TET2 has much higher affinity for oligonucleotides, and the crystal structure of TET2:DNA complex (PDB ID 5DEU) was solved without cross-linking strategy. Although with highly similar DSBH domains, there were stark differences in the loops that preceded and in between various β-strands of the DSBH domain in how they interact with the oligonucleotides. The loops of TET2 potentially contributed to the high oligonucleotide affinity.

Together with Dr. Chuan He's and our observations, we could conclude that AlkB family of *N*-demethylases only weakly bind to their substrates. Strong interaction of proteins with their substrate oligonucleotides would hinder the folding or binding of important factors onto the DNA or RNA. It is reasonable to suggest that AlkB proteins are meant to perform the covalent modifications on the oligonucleotides and 'free' the oligonucleotides quickly to perform their downstream biological functions.

4.1.2 FTO and nucleotides

DSF and *in vitro* demethylation assays have indicated that recombinant FTO has stronger binding (T_m shift = 12.1°C) and demethylation profile (~20% conversion) for m3T compared to other recombinant AlkB oxygenases. Structural analysis of FTO complexed with m3T and NOG (PDB ID 3LFM) with other AlkB oxygenases structures could perhaps rationalize the above difference in affinity.

We superimposed FTO with crystal structures of AlkB (PDB ID 3I3M), ALKBH2 (PDB ID 3BUC), ALKBH3 (PDB ID 2IUW), and ALKBH5 (PDB ID 4NJ4) (Fig. 30). Residues lining the AlkB oxygenases' functional domains were in the active conformation, complexed to co-factors and some with substrates that stabilised the DSBH domains. Nucleotide binding sites of the AlkB oxygenases compared were largely similar to FTO (Fig. 30). The nucleotide binding sites were lined with some highly conserved residues or substituted for structurally equivalent residues that were able to make meaningful interactions with m3T. The loss of binding affinity towards m3T in other AlkB oxygenases is highly possible due to the substitution of residues that involves in strong interaction with m3T.

Substitution of $Arg96_{FTO}$ by $Met61_{AlkB}$ and $Gln112_{ALKBH2}$ led to the loss of interaction and decreased efficiency of binding with O^2-m3T, respectively (Fig. 30A and D). The interaction was further weakened, where amide backbones of $Asp132_{AlkB}$ and $Asp174_{ALKBH2}$ were positioned at a distance of ~4.6 Å (average), thus unable to make efficient contact with O^4-m3T in comparison to the corresponding $Glu234_{FTO}$. Crucially, corresponding $Tyr108_{FTO}$ residues in ALKBH3 and ALKBH5 were unable to make stabilising π-π interaction with the nucleobase of m3T (Fig. 30B and C). Both $Tyr143_{ALKBH3}$ and $Tyr139_{ALKBH5}$ were positioned perpendicular to the

nucleobase of m3T, weakening the π-π interaction. Nevertheless, this rationalisation might be inaccurate as the crystal structures of ALKBH3 and ALKBH5 used for comparison were not solved with substrate lining the nucleotide binding pocket. Potentially, $\text{Tyr143}_{\text{ALKBH3}}$ and $\text{Tyr139}_{\text{ALKBH5}}$ could 'switch' and form a π-π interaction with the nucleobase of their preferred substrate. Further analysis revealed that the amide backbones of $\text{Tyr143}_{\text{ALKBH3}}$ and $\text{Tyr139}_{\text{ALKBH5}}$ were unable to 'switch' to form a meaningful π-π interaction without interrupting and destabilising the alignment of nearby Cα backbone. In addition, energy minimisation of ALKBH5 with oligonucleotide revealed no significant conformational change of $\text{Tyr139}_{\text{ALKBH5}}$. The way C$\alpha$ backbone near $\text{Tyr139}_{\text{ALKBH5}}$ was positioned prevented such π-π interaction. Finally, $\text{Asp194}_{\text{ALKBH3}}$ and $\text{Pro207}_{\text{ALKBH5}}$ were unable to form a good interaction with O^4-m3T.
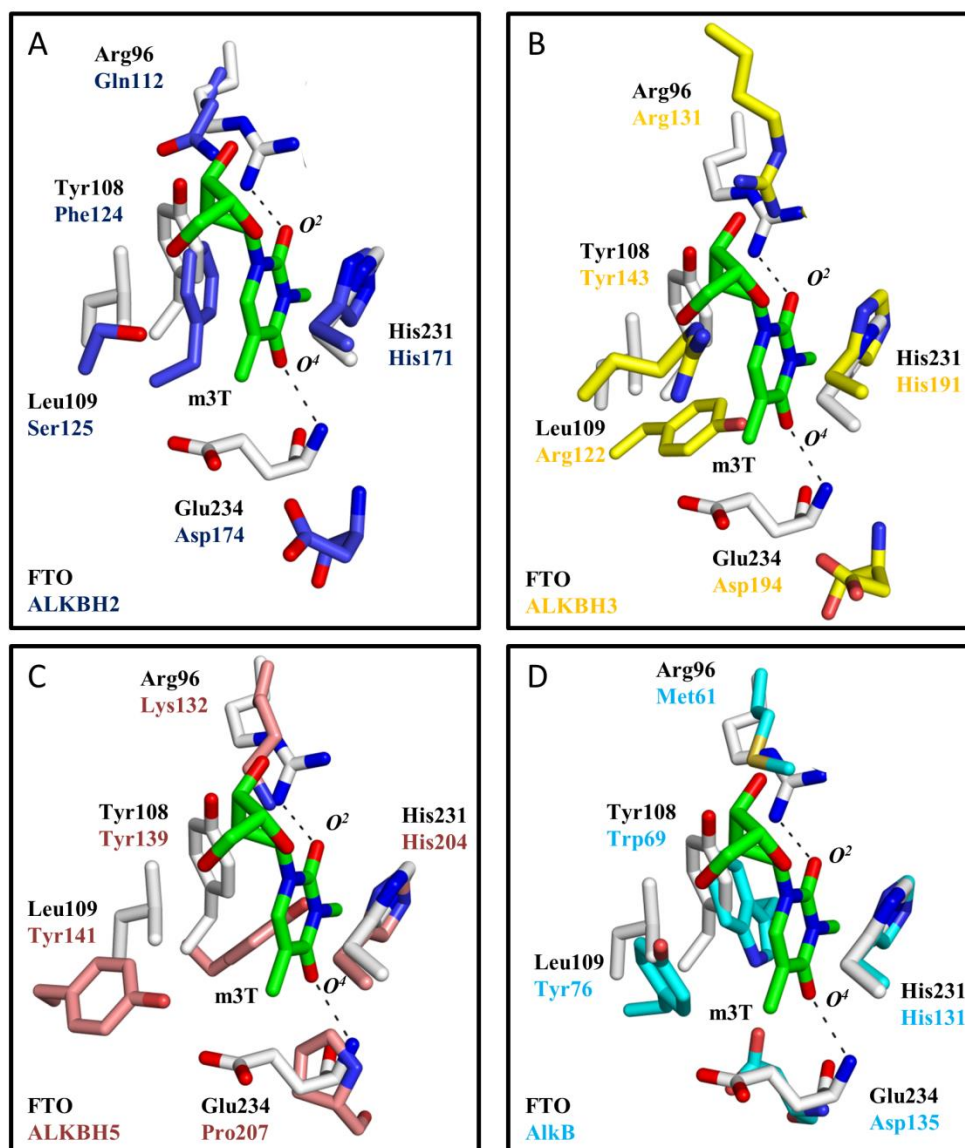
Figure 30. Comparison of various AlkB oxygenases substrate binding sites with FTO. A. Comparison of ALKBH2 with FTO. **B.** Comparison of ALKBH3 with FTO. **C.** Comparison of ALKBH5 with FTO. **D.** Comparison of AlkB with FTO.

4.2 Strategy for selective inhibitor design

No report of potent and selective inhibitor for AlkB oxygenases was present previously. Rhein was reported to inhibit FTO *in vitro* and in cells, however, rhein also inhibits other AlkB oxygenases [207, 246]. The inhibitors available (NOG and 2,4-PDCA) were generic and they bind indiscriminately to the 2OG-binding site of AlkB oxygenases, chelating to the Fe(II) co-factor. Analysis of the substrate binding sites of AlkB oxygenases crystal structures revealed important differences that accounted for their substrate selectivity. We took advantage of these structural differences for the design of our inhibitors for FTO. Given that the substrate binding site could be exploited for selective inhibition, designing an inhibitor that could simultaneously occupy the 2OG and substrate binding sites of FTO would fuel its potency and selectivity (Fig. 31). The 'two-component' inhibitor was made up of a '2OG-binding component' and a 'substrate-binding component' tethered together. Among all the inhibitors synthesised from the lab of Dr Esther Woon, **12** gave the best binding and inhibitory profile.
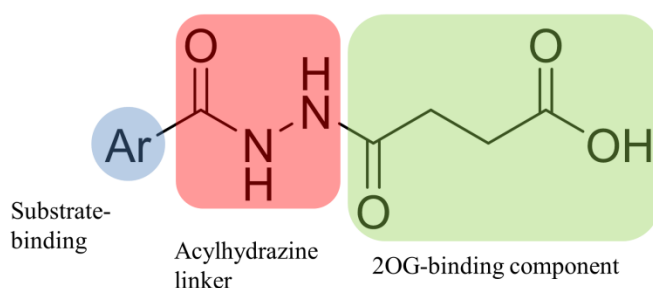
Figure 31. The 'two-component' inhibitor. The inhibitor is designed to simultaneously occupy the 2OG-binding and substrate-binding site.

4.3 Analysis of the binding mode of inhibitor 12

Kinetic studies were performed in Dr Esther Woon's lab. The analyses showed that **12** had a mixed mode of inhibition with m3T. Thus, **12** could bind to a free FTO and compete with FTO-m3T. Indeed, **12** demonstrated a competitive inhibition towards co-substrate 2OG. This made sense as **12** occupied the same site as 2OG.

A series of analogues of **12** were synthesised and tested for inhibitory activities against FTO and other AlkB oxygenases by DSF and HPLC *in vitro* demethylation assay (Table 3). **12** showed weak inhibition towards AlkB (IC_{50} = 33.5 μM; T_m = 4.5°C) and ALKBH2 (IC_{50} = 25.9 μM; T_m = 4.9°C), recpectively. However, **12** still maintained over more than 20 fold potency towards FTO (IC_{50} = 0.81 μM; T_m = 11.2°C) compared to AlkB and ALKBH2. Remarkably, **12** was more than 100 fold selective toward FTO compared to ALKBH5 (IC_{50} = 108.1 μM; T_m = 2.7°C). ALKBH5 and FTO have m6A as their major natural substrate. This is important as the ability of **12** to selectively inhibit FTO made it a potential functional probe to understand the biological roles of FTO and ALKBH5 in m6A-mediated epigenetic.

The benzyl group of **12** is most probably not crucial for **its** selectivity towards FTO. However, this could be seen in **21**, with the benzyl substituent removed. Although **21** inhibited FTO (IC_{50} = 11.8 μM; T_m = 7.4°C) to a lesser extent, it maintained at least a 100 fold selectivity towards ALKBH3 (IC_{50} = 114.6 μM; T_m = 2.5°C). The benzyl group most likely contributed to the potency of **12** towards FTO via hydrophobic interactions. Crucially a complete elimination of selectivity was observed when the pyridyl nitrogen was moved from the *meta* to *ortho* position. The position of pyridyl nitrogen is most likely a key determinant for selectivity. In the crystal structure of

FTO complexed to **12**, it showed that the position of the nitrogen is indeed important

for selectivity (Fig. 20).

4.4 Rationalisation of inhibitor 12 selectivity

The selective mechanism of **12** was revealed when we compared the structure of FTO-**12** complex with that of AlkB oxygenases. Inspection of the 2-OG binding sites of FTO and the AlkB oxygenases revealed some highly conserved residues. Particularly, the catalytic triad aligned very well across all the AlkB oxygenases (FTO, AlkB, ALKBH 2, 3 and 5) used for comparison. In addition, the arginine residues that form salt bridges to the carboxylate end of **12** were also conserved in all the AlkB oxygenases. Due to highly similar binding mechanism between the AlkB oxygenases and the 2OG-binding component of **12**, it could be concluded that this component conferred potency but not selectivity to the binding. Potency of **12** towards FTO could also be explained through the expanded hydrogen bonding network where the carboxylate end of **12** interacts with $Ser318_{FTO}$ and $Tyr295_{FTO}$. Corresponding residues in other AlkB oxygenases were substituted for residues that are less polar or unable to make direct contact due to unfavorable distance.

The selectivity of **12** for FTO could be rationalised by aligning the substrate-binding sites of the crystal structures of AlkB oxygenases (Fig. 32). There are some distinct differences in the residues lining the substrate-binding site that weaken binding affinity of **12**. The π-π interactions with the pyridyl ring of **12** were severely weakened for $Tyr143_{ALKBH3}$ and $Tyr139_{ALKBH5}$ as the positions of how the tyrosines were placed discouraged stacking interactions observed in $Tyr109_{FTO}$, $Trp169_{AlkB}$ and $Phe124_{ALKBH2}$ (Fig. 32 A and B).

It is worth noting that the position of the nitrogen in the pyridyl ring played a significant role in determining the potency and selectivity of **12**. The pyridyl nitrogen of **12** made a hydrogen bonding with amide backbone of $Glu234_{FTO}$, while the

corresponding residues of Glu234$_{FTO}$ in other AlkB oxygenases, Asp135$_{AlkB}$, Asp174$_{ALKBH2}$, Asp194$_{ALKBH3}$ and Pro207$_{ALKBH5}$ were positioned too far away from the pyridyl nitrogen to make any interactions (Fig. 32). That accounted for the severely diminished or abolished affinity of **12** for other AlkB oxygenases. In addition, the crystal structure of FTO-**12** complex offered an explanation regarding the loss of potency of **23** (pyridyl nitrogen was moved from the *meta* to *ortho* position) towards FTO.

Prior to the development of inhibitor **12**, there were no potent and selective binders to FTO [246, 247]. Notably, natural product rhein bind promiscuously to FTO and other AlkB enzymes inhibiting their catalytic activities (Fig. 33A). Overlying both crystal structures of FTO complexed to inhibitor **12** and rhein revealed both inhibitors occupied similar nucleotide binding region and made similar interactions with the surrounding residues (Fig. 33A). However, rhein lacks the pyridine nitrogen that offered **12** the selectivity toward FTO. Following the report of inhibitor **12**, there were two notable reports of selective inhibitors for FTO [249, 251]. The reported inhibitors (meclofanamic acid and a fluorescein derivatised compound) revealed a novel binding mechanism (Fig. 33A & B). Both inhibitors contained a phenyl ring that interacted hydrophobically with Ile85, Leu90 and Pro93 of the NRL1 situated near the entrance of FTO catalytic domain. NRL1 is longer and extended in FTO. On the other hand, the much shorter NRL1 in ALKBH5 is unable to make similar interaction (Fig. 33D). This rationalised the selectivity toward FTO over ALKBH5.
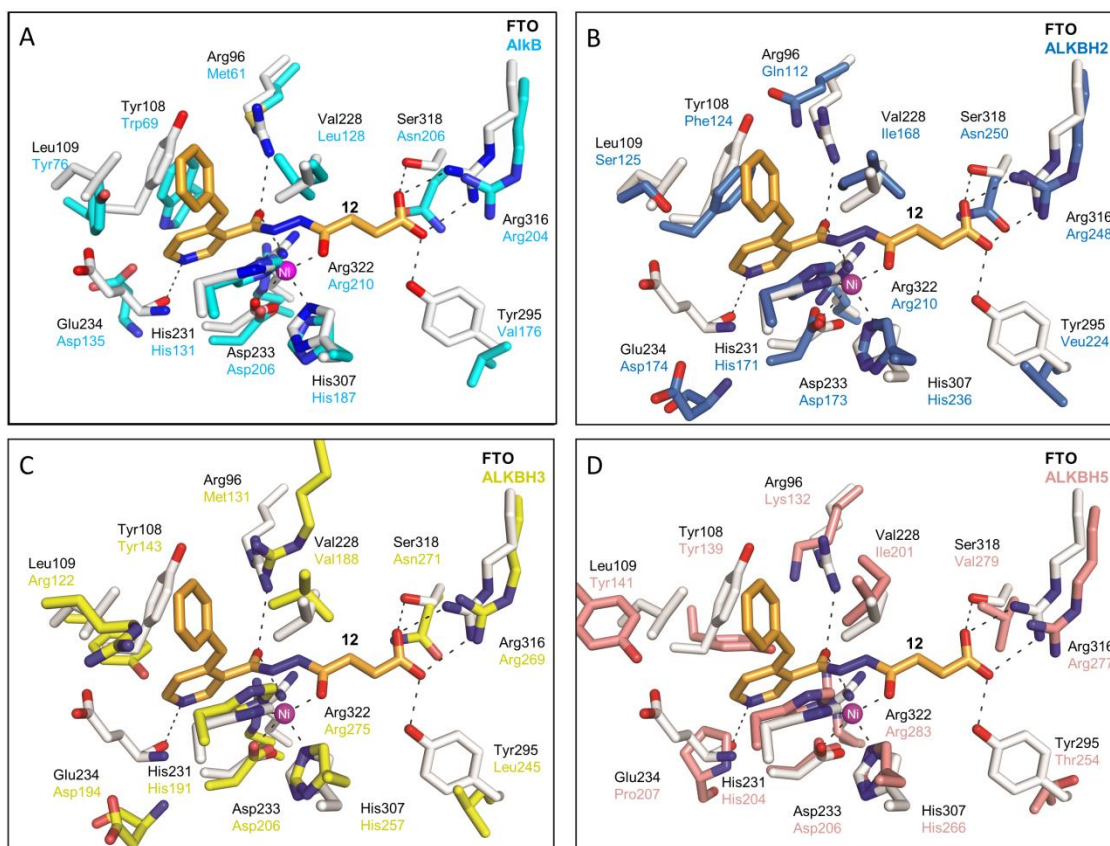
Figure 32. Rationalisation of selectivity of 12 by superimposing available crystal structures. Particularly, **12** was unable to form good contact with residues of other AlkB oxygenases corresponding to Glu234$_{FTO}$ **A.** Superimposition of FTO and AlkB . **B.** Superimposition of FTO and ALKBH2. **C.** Superimposition of FTO and ALKBH3. **D.** Superimposition of FTO and ALKBH5.
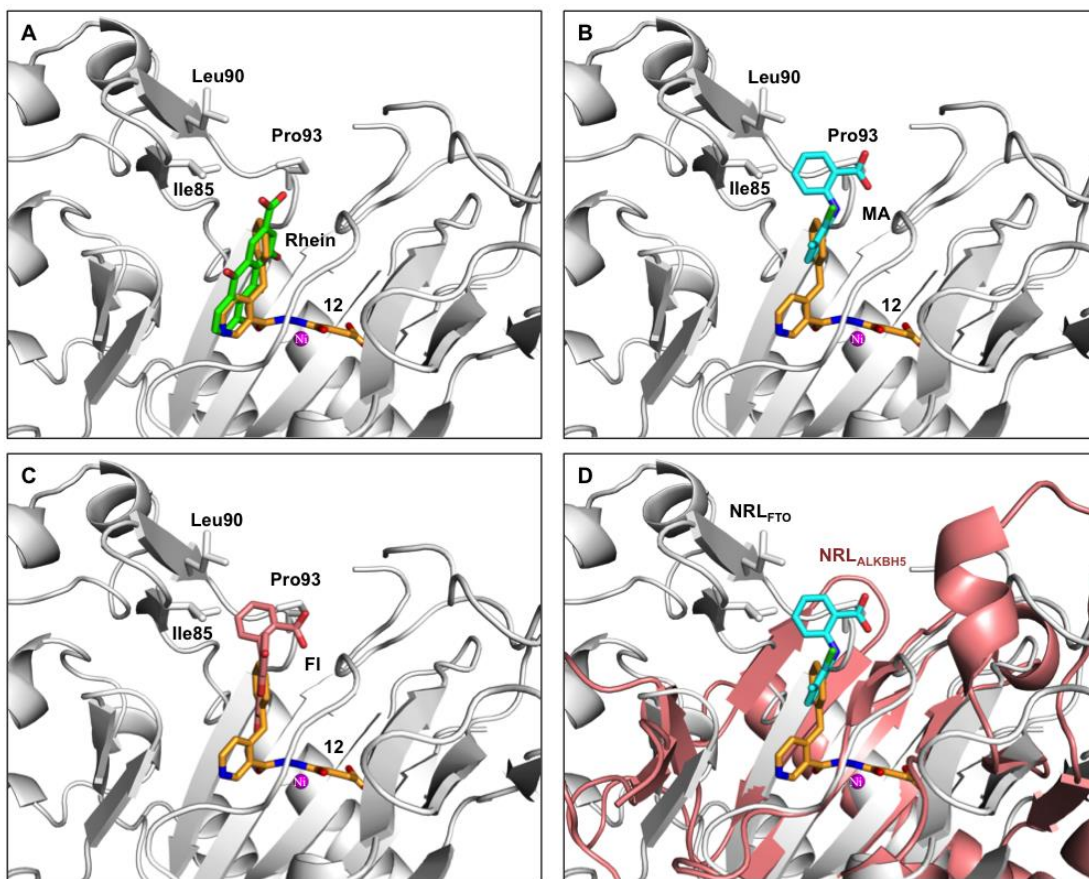
Figure 33. Binding modes of different inhibitors. A. Overlay of **12** with natural product rhein. **B.** Overlay of **12** with meclofanamic acid (MA). **C.** Structural comparison of NRL1 of FTO (white) and ALKBH5 (salmon). **D.** Overlay of **12** with fluorescein derivatised inhibitor (Fl).

4.5 Potential of 12 as functional probe

The compound **25**, an ethyl ester derivative of **12**, was able to penetrate the cell membrane better than **12**, and subsequently, hydrolysed to **12** in the cells. Preliminary studies on **25** showed that it is non-toxic to cells up to 100 μM, and is able to inhibit the demethylation activity of FTO.

There are established criterion that classify compounds as drug-like if they fall largely within an acceptable quantified categories. On the other hand, classifying compounds as probe-like is a relatively new concept. There are various literatures and commentaries that suggested what constitute to be good probes [267-269]. It is important to note that drugs and probes may have overlapping profiles, but they serve distinct purposes and need not be a subset of each other. The Structural Genomic Consortium (SGC) classifies a probe that possesses good *in vitro* potency (less than 100 nM), high selectivity within a family of structurally related proteins (more than 30 fold) and a good on-target in cells (at less than 1 μM) [267-269].

12 showed promising results in *in vitro* studies and preliminary cellular profiling. However, further verification is required with regards to whether FTO is indeed specifically targeted and whether there is minimal or insignificant off-target effects in cells. Only then it could be possible to perturb and study the mechanistic roles FTO play in various cell lines.

4.6 M6A can influence conformations of oligonucleotides

M6A is the most common epigenetic modification found within highly conserved consensus sequences (DRACH; where D = A, G or U; R = G or A and H = A, C or U) in eukaryotic mRNA and is able to dictate the fate of mRNA. Reviewed earlier in this thesis, m6A did not prevent Watson-Crick base pairing, however it destabilised the base pairing by 0.5-1.7 kcal/mol, depending on sequence context. Various secondary structures are adopted by mRNA and there have been reports of m6A modifications that influence secondary structures of RNA oligonucleotides [109, 131, 270]. Given that FTO and ALKBH5 are able to recognise and demethylate their major substrate, m6A, it is of significant interest to know whether they are able to affect the change of RNA secondary structures upon demethylation.

Several RNA oligonucleotides were designed to adopt duplex or hairpin conformation, involving Watson-Crick base pairing. As shown in chapter 3.9, not all m6A-containing RNA oligonucleotides were able to elicit a change in the secondary structures. In the RNA oligonucleotides studied (**RNA-2**, **4** and **6**), despite having the same base composition, two bases in the immediate proximity to the m6A modification could provoke a change in the overall conformation. Particularly, this becomes even evident when the concentration of oligonucleotides (**RNA-2**, **4** and **6**) was below 5 μM. Observations of the behavior of these methylated oligonucleotides were confirmed by various experimental methods such as CD and non-denaturing PAGE, as described in chapter 3.9. Potentially, if FTO and ALKBH5 is able to recognise and demethylate the m6A, a conformational change should be observed.

4.7 M6A-induced RNA conformational change influences substrate specificity of FTO

Given that m6A was practically found within a set of consensus sequence (DRACH), we asked how the consensus sequence is important for FTO and ALKBH5 demethylation. HPLC-based *in vitro* demethylation on various DNA and RNA oligonucleotides containing m6A showed that the consensus sequence is not crucial for demethylation [132]. However, a minimal of four nucleotides is needed for m6A demethylation to occur for FTO and ALKBH5 (Table 5). It is noteworthy to mention that FTO could demethylate a single m3T nucleotide efficiently and not m6A [132, 245]. On the other hand, such consensus sequence is crucial for other m6A 'writer' and 'reader' proteins, for instance, METTL3 and YTHDF2, where the consensus sequence motifs are crucial for recognition and effect of functions [85, 118, 120].

The demethylation of FTO and ALKBH5 against highly similar sequences, **RNA-2**, **4** and **6** were profiled (Fig. 27). The assay results in Fig. 27 showed that the demethylases preferred to catalyse **RNA-4** and **6** at 5 μM and below. However no activity was observed for **RNA-2**. Notably, for **RNA-4** and **6** at 5 μM, substantial hairpin conformation was observed but not for **RNA-2**. This suggested the demethylases preferentially demethylate certain m6A-induced conformation, hairpin in this context. When **RNA-4** and **6** were at 10 μM, both existed as duplexes and no demethylation was observed. Upon demethylation, **RNA-4** and **6** became **RNA-3** and **5**, respectively. **RNA-3** and **5** are duplexes and it could be speculated that *in vivo*, this drastic conformation change might possibly deter binding of certain regulatory proteins, thus affecting downstream biological events. However, more *in vivo*

investigations are required to prove that FTO and ALKBH5 are able to partake such

physiological function.

Sequence	% Demethylation	
	FTO	ALKBH5
m6A	2	2
G(m6A)C	5	1
GG(m6a)C	31	21
GG(m6A)CU	69	37
CCCC(m6A)CCCCCCCC	54	30

Table 5. Percentage demethylation of different sequences by FTO and ALKBH5. FTO and ALKBH5 displayed similar demethylation profile. Substantial demethylation observed with at least four-nucleotide long sequence. Consensus sequence is not essential for demethylation to occur. Recombinant proteins were purified in-house and demethylation studies were performed by Dr Esther Woon.

CHAPTER 5

CONCLUSION AND PERSPECTIVE

Initially, m6A was thought to be a permanent mark in RNA. *FTO* gene is the first GWAS identified gene that has direct link to obesity, as well as its coding protein FTO is the first of the two m6A demethylases known so far that is able to remove the m6A modifications in RNA. The discovery of FTO therefore revealed that m6A modifications are dynamic and epigenetic regulation can occur in the layer of RNA transcripts. ALKBH5, the second m6A demethylase, was later discovered. Both FTO and ALKBH5 belong to AlkB 2OG oxygenases family of proteins and share highly similar domain architecture. FTO and ALKBH5 have stark physiological functions, however, the mechanisms of how FTO and ALKBH5 affect their functions were largely unknown.

It was of interest to develop a potent and cell-active probe that is selective for FTO to help us understand the mechanism of action of FTO in cells. This may set a good platform for therapeutic intervention for obesity in the long run. We cloned and purified various FTO and AlkB oxygenases and employed X-ray crystallography, DSF, *in vitro* demethylation assays and modelling studies to develop an inhibitor that is highly selective and active towards FTO. It was the first selective and active inhibitor for a member of human AlkB oxygenases family. We have shown that despite highly similar homologies among human AlkB members, the substrate-binding site could be exploited for enhancing selectivity of inhibitor.

RNA transcripts were known to adopt various conformations *in vivo*. Those conformations might play a part in various physiological functions, for instance, binding of regulation factors and translation regulation. We employed CD analyses, gel-shift technique, thermophoresis measurements and biochemical assays to show that, in principle, FTO can demethylate m6A-containing RNA *in vitro* and evoke a conformation change of the RNA oligonucleotide. With these *in vitro* data, we may

ask how the molecular mechanism of FTO in carrying out its physiological role entails conformational changes to RNA transcripts.

Many have reported FTO and m6A in isolation. We understand the physiological effect of over and under expression of FTO. We are also beginning to understand the physiological effect of the presence and absence of m6A in different parts of mRNA. However, we do not understand when and how FTO expression is regulated. We also do not know the mechanism of FTO and ALKBH5 recognition and discrimination of any mRNA transcripts given that both demethylases have distinct physiological roles. Arguably, understanding the mechanism behind the interplay between FTO and its substrate, m6A has only taken its very first step.

References

1.   Sono, M., Roach, M.P., Coulter, E.D., and Dawson, J.H. (1996). Heme-Containing Oxygenases. Chem Rev *96*, 2841-2888.

2.   Meunier, B., de Visser, S.P., and Shaik, S. (2004). Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. Chem Rev *104*, 3947-3980.

3.   Que, L., Jr., and Ho, R.Y. (1996). Dioxygen Activation by Enzymes with Mononuclear Non-Heme Iron Active Sites. Chem Rev *96*, 2607-2624.

4.   Schofield, C.J., and Zhang, Z. (1999). Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. Curr Opin Struct Biol *9*, 722-731.

5.   Purpero, V., and Moran, G.R. (2007). The diverse and pervasive chemistries of the alpha-keto acid dependent enzymes. J Biol Inorg Chem *12*, 587-601.

6.   Kawai, Y., Ono, E., and Mizutani, M. (2014). Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. Plant J *78*, 328-343.

7.   Prescott, A.G., and John, P. (1996). DIOXYGENASES: Molecular Structure and Role in Plant Metabolism. Annu Rev Plant Physiol Plant Mol Biol *47*, 245-271.

8.   Aravind, L., and Koonin, E.V. (2001). The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol *2*, RESEARCH0007.

9.   Flashman, E., and Schofield, C.J. (2007). The most versatile of all reactive intermediates? Nat Chem Biol *3*, 86-87.

10.   Johansson, C., Tumber, A., Che, K., Cain, P., Nowak, R., Gileadi, C., and Oppermann, U. (2014). The roles of Jumonji-type oxygenases in human disease. Epigenomics *6*, 89-120.

11.   Hutton, J.J., Jr., Trappel, A.L., and Udenfriend, S. (1966). Requirements for alpha-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase. Biochem Biophys Res Commun *24*, 179-184.

12.   Stetten, M.R. (1949). Some aspects of the metabolism of hydroxyproline, studied with the aid of isotopic nitrogen. J Biol Chem *181*, 31-37.

13.   Loenarz, C., and Schofield, C.J. (2008). Expanding chemical biology of 2-oxoglutarate oxygenases. Nat Chem Biol *4*, 152-156.

14.   Loenarz, C., and Schofield, C.J. (2011). Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. Trends Biochem Sci *36*, 7-18.

15.   Tarhonskaya, H., Chowdhury, R., Leung, I.K., Loik, N.D., McCullagh, J.S., Claridge, T.D., Schofield, C.J., and Flashman, E. (2014). Investigating the contribution of the active site environment to the slow reaction of hypoxia-inducible factor prolyl hydroxylase domain 2 with oxygen. Biochem J *463*, 363-372.

16.   Majmundar, A.J., Wong, W.J., and Simon, M.C. (2010). Hypoxia-inducible factors and the response to hypoxic stress. Mol Cell *40*, 294-309.

17.   Markolovic, S., Leissing, T.M., Chowdhury, R., Wilkins, S.E., Lu, X., and Schofield, C.J. (2016). Structure-function relationships of human JmjC oxygenases-demethylases versus hydroxylases. Curr Opin Struct Biol *41*, 62-72.

18. Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

19. Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

20. Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol *14*, 341-356.

21. Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300-1303.

22. Fu, L., Guerrero, C.R., Zhong, N., Amato, N.J., Liu, Y., Liu, S., Cai, Q., Ji, D., Jin, S.G., Niedernhofer, L.J., et al. (2014). Tet-mediated formation of 5-hydroxymethylcytosine in RNA. J Am Chem Soc *136*, 11582-11585.

23. Aas, P.A., Otterlei, M., Falnes, P.O., Vagbo, C.B., Skorpen, F., Akbari, M., Sundheim, O., Bjoras, M., Slupphaug, G., Seeberg, E., et al. (2003). Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature *421*, 859-863.

24. Drablos, F., Feyzi, E., Aas, P.A., Vaagbo, C.B., Kavli, B., Bratlie, M.S., Pena-Diaz, J., Otterlei, M., Slupphaug, G., and Krokan, H.E. (2004). Alkylation damage in DNA and RNA--repair mechanisms and medical significance. DNA Repair (Amst) *3*, 1389-1407.

25. Ougland, R., Rognes, T., Klungland, A., and Larsen, E. (2015). Non-homologous functions of the AlkB homologs. J Mol Cell Biol *7*, 494-504.

26. Gerken, T., Girard, C.A., Tung, Y.C., Webby, C.J., Saudek, V., Hewitson, K.S., Yeo, G.S., McDonough, M.A., Cunliffe, S., McNeill, L.A., et al. (2007). The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science *318*, 1469-1472.

27. Clifton, I.J., McDonough, M.A., Ehrismann, D., Kershaw, N.J., Granatino, N., and Schofield, C.J. (2006). Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. J Inorg Biochem *100*, 644-669.

28. McDonough, M.A., Loenarz, C., Chowdhury, R., Clifton, I.J., and Schofield, C.J. (2010). Structural studies on human 2-oxoglutarate dependent oxygenases. Curr Opin Struct Biol *20*, 659-672.

29. Aik, W., McDonough, M.A., Thalhammer, A., Chowdhury, R., and Schofield, C.J. (2012). Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr Opin Struct Biol *22*, 691-700.

30. Zhang, Z., Ren, J., Stammers, D.K., Baldwin, J.E., Harlos, K., and Schofield, C.J. (2000). Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. Nat Struct Biol *7*, 127-133.

31. Elkins, J.M., Hewitson, K.S., McNeill, L.A., Seibel, J.F., Schlemminger, I., Pugh, C.W., Ratcliffe, P.J., and Schofield, C.J. (2003). Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha. J Biol Chem *278*, 1802-1806.

32. Elkins, J.M., Ryle, M.J., Clifton, I.J., Dunning Hotopp, J.C., Lloyd, J.S., Burzlaff, N.I., Baldwin, J.E., Hausinger, R.P., and Roach, P.L. (2002). X-ray crystal structure of Escherichia coli taurine/alpha-ketoglutarate dioxygenase complexed to ferrous iron and substrates. Biochemistry *41*, 5185-5192.

33. Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., Wang, P., and Xu, Y. (2013). Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell *155*, 1545-1555.

34. Koehntop, K.D., Emerson, J.P., and Que, L., Jr. (2005). The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. J Biol Inorg Chem *10*, 87-93.

35. Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

36. Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res *42*, W320-324.

37. Zhu, Q., Hillwig, M.L., Doi, Y., and Liu, X. (2016). Aliphatic Halogenase Enables Late-Stage C-H Functionalization: Selective Synthesis of a Brominated Fischerindole Alkaloid with Enhanced Antibacterial Activity. Chembiochem *17*, 466-470.

38. Mitchell, A.J., Zhu, Q., Maggiolo, A.O., Ananth, N.R., Hillwig, M.L., Liu, X., and Boal, A.K. (2016). Structural basis for halogenation by iron- and 2-oxo-glutarate-dependent enzyme WelO5. Nat Chem Biol *12*, 636-640.

39. 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.

40. Yu, B., Edstrom, W.C., Benach, J., Hamuro, Y., Weber, P.C., Gibney, B.R., and Hunt, J.F. (2006). Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. Nature *439*, 879-884.

41. Zheng, G., Fu, Y., and He, C. (2014). Nucleic acid oxidation in DNA damage repair and epigenetics. Chem Rev *114*, 4602-4620.

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

43. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol *7*, 885-887.

44. Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vagbo, C.B., Shi, Y., Wang, W.L., Song, S.H., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell *49*, 18-29.

45. Shen, L., Song, C.X., He, C., and Zhang, Y. (2014). Mechanism and function of oxidative reversal of DNA and RNA methylation. Annu Rev Biochem *83*, 585-614.

46. Abu-Omar, M.M., Loaiza, A., and Hontzeas, N. (2005). Reaction mechanisms of mononuclear non-heme iron oxygenases. Chem Rev *105*, 2227-2252.

47. Yi, C., Yang, C.G., and He, C. (2009). A non-heme iron-mediated chemical demethylation in DNA and RNA. Acc Chem Res *42*, 519-529.

48. Barrows, L.R., and Magee, P.N. (1982). Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. Carcinogenesis *3*, 349-351.

49. Falnes, P.O., Klungland, A., and Alseth, I. (2007). Repair of methyl lesions in DNA and RNA by oxidative demethylation. Neuroscience *145*, 1222-1232.

50. Sedgwick, B. (2004). Repairing DNA-methylation damage. Nat Rev Mol Cell Biol *5*, 148-157.

51. Fedeles, B.I., Singh, V., Delaney, J.C., Li, D., and Essigmann, J.M. (2015). The AlkB Family of Fe(II)/alpha-Ketoglutarate-dependent Dioxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond. J Biol Chem *290*, 20734-20742.

52. Mishina, Y., and He, C. (2006). Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins. J Inorg Biochem *100*, 670-678.

53. Sedgwick, B., Robins, P., and Lindahl, T. (2006). Direct removal of alkylation damage from DNA by AlkB and related DNA dioxygenases. Methods Enzymol *408*, 108-120.

54. Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes Dev *23*, 781-783.

55. Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet *33 Suppl*, 245-254.

56. Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

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

58. Volkel, P., and Angrand, P.O. (2007). The control of histone lysine methylation in epigenetic regulation. Biochimie *89*, 1-20.

59. Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

60. Ehrlich, M., and Wang, R.Y. (1981). 5-Methylcytosine in eukaryotic DNA. Science *212*, 1350-1357.

61. Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. Cell *128*, 669-681.

62. Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev *16*, 6-21.

63. De Carvalho, D.D., You, J.S., and Jones, P.A. (2010). DNA methylation and cellular reprogramming. Trends Cell Biol *20*, 609-617.

64. Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev *25*, 1010-1022.

65. 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.

66. Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129-1133.

67. Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun *2*, 241.

68. Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell *42*, 451-464.

69.	Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606-610.

70.	Song, C.X., Szulwach, K.E., Dai, Q., Fu, Y., Mao, S.Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H., et al. (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell *153*, 678-691.

71.	He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303-1307.

72.	Inoue, A., Shen, L., Dai, Q., He, C., and Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res *21*, 1670-1676.

73.	Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., et al. (2016). DNA methylation on N(6)-adenine in mammalian embryonic stem cells. Nature *532*, 329-333.

74.	Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell *149*, 1635-1646.

75.	Meyer, K.D., and Jaffrey, S.R. (2014). The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol *15*, 313-326.

76.	Jia, G., Fu, Y., and He, C. (2013). Reversible RNA adenosine methylation in biological regulation. Trends Genet *29*, 108-115.

77.	Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci U S A *71*, 3971-3975.

78.	Wei, C., Gershowitz, A., and Moss, B. (1975). N6, O2'-dimethyladenosine a novel methylated ribonucleoside next to the 5' terminal of animal cell and virus mRNAs. Nature *257*, 251-253.

79.	Krug, R.M., Morgan, M.A., and Shatkin, A.J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. J Virol *20*, 45-53.

80.	Stoltzfus, C.M., and Dimock, K. (1976). Evidence of methylation of B77 avian sarcoma virus genome RNA subunits. J Virol *18*, 586-595.

81.	Bodi, Z., Zhong, S., Mehra, S., Song, J., Graham, N., Li, H., May, S., and Fray, R.G. (2012). Adenosine Methylation in Arabidopsis mRNA is Associated with the 3' End and Reduced Levels Cause Developmental Defects. Front Plant Sci *3*, 48.

82.	Clancy, M.J., Shambaugh, M.E., Timpte, C.S., and Bokar, J.A. (2002). Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. Nucleic Acids Res *30*, 4509-4518.

83.	Keith, J.M., Ensinger, M.J., and Mose, B. (1978). HeLa cell RNA (2'-O-methyladenosine-N6-)-methyltransferase specific for the capped 5'-end of messenger RNA. J Biol Chem *253*, 5033-5039.

84.	Csepany, T., Lin, A., Baldick, C.J., Jr., and Beemon, K. (1990). Sequence specificity of mRNA N6-adenosine methyltransferase. J Biol Chem *265*, 20117-20122.

85.   Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol *10*, 93-95.

86.   Ping, X.L., Sun, B.F., Wang, L., Xiao, W., Yang, X., Wang, W.J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.S., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res *24*, 177-189.

87.   Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., et al. (2016). Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. Nature *534*, 575-578.

88.   Wang, P., Doxtader, K.A., and Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. Mol Cell *63*, 306-317.

89.   Zhou, K.I., and Pan, T. (2016). Structures of the m(6)A Methyltransferase Complex: Two Subunits with Distinct but Coordinated Roles. Mol Cell *63*, 183-185.

90.   Jia, G., Yang, C.G., Yang, S., Jian, X., Yi, C., Zhou, Z., and He, C. (2008). Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. FEBS Lett *582*, 3313-3319.

91.   Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., Wang, Q., Cheng, W., Wang, J., Feng, Y., and Chai, J. (2010). Crystal structure of the FTO protein reveals basis for its substrate specificity. Nature *464*, 1205-1209.

92.   Thalhammer, A., Bencokova, Z., Poole, R., Loenarz, C., Adam, J., O'Flaherty, L., Schodel, J., Mole, D., Giaslakiotis, K., Schofield, C.J., et al. (2011). Human AlkB homologue 5 is a nuclear 2-oxoglutarate dependent oxygenase and a direct target of hypoxia-inducible factor 1alpha (HIF-1alpha). PLoS One *6*, e16210.

93.   Fu, Y., Jia, G., Pang, X., Wang, R.N., Wang, X., Li, C.J., Smemo, S., Dai, Q., Bailey, K.A., Nobrega, M.A., et al. (2013). FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. Nat Commun *4*, 1798.

94.   Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. Nat Protoc *8*, 176-189.

95.   Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., and Jaffrey, S.R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods *12*, 767-772.

96.   Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m(6)A RNA methylation. Nature Reviews Genetics *15*, 293-306.

97.   Munns, T.W., Sims, H.F., and Liszewski, M.K. (1977). Immunospecific retention of oligonucleotides possessing N6-methyladenosine and 7-methylguanosine. J Biol Chem *252*, 3102-3104.

98.   Munns, T.W., Liszewski, M.K., and Sims, H.F. (1977). Characterization of antibodies specific for N6-methyladenosine and for 7-methylguanosine. Biochemistry *16*, 2163-2168.

99.   Wei, C.M., Gershowitz, A., and Moss, B. (1976). 5'-Terminal and internal methylated nucleotide sequences in HeLa cell mRNA. Biochemistry *15*, 397-401.

100.   Wei, C.M., and Moss, B. (1977). Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. Biochemistry *16*, 1672-1676.

101.   Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA *19*, 1848-1856.

102.   Berulava, T., Rahmann, S., Rademacher, K., Klein-Hitpass, L., and Horsthemke, B. (2015). N6-adenosine methylation in MiRNAs. PLoS One *10*, e0118438.

103.   Yang, L., Ma, Y., Han, W., Li, W., Cui, L., Zhao, X., Tian, Y., Zhou, Z., Wang, W., and Wang, H. (2015). Proteinase-activated receptor 2 promotes cancer cell migration through RNA methylation-mediated repression of miR-125b. J Biol Chem *290*, 26627-26637.

104.   Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature *518*, 560-564.

105.   Zhou, K.I., Parisien, M., Dai, Q., Liu, N., Diatchenko, L., Sachleben, J.R., and Pan, T. (2016). N(6)-Methyladenosine Modification in a Long Noncoding RNA Hairpin Predisposes Its Conformation to Protein Binding. J Mol Biol *428*, 822-833.

106.   Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature *505*, 117-120.

107.   Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell *161*, 1388-1399.

108.   Xiao, W., Adhikari, S., Dahal, U., Chen, Y.S., Hao, Y.J., Sun, B.F., Sun, H.Y., Li, A., Ping, X.L., Lai, W.Y., et al. (2016). Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell *61*, 507-519.

109.   Alarcon, C.R., Goodarzi, H., Lee, H., Liu, X., Tavazoie, S., and Tavazoie, S.F. (2015). HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. Cell *162*, 1299-1308.

110.   Chen, T., Hao, Y.J., Zhang, Y., Li, M.M., Wang, M., Han, W., Wu, Y., Lv, Y., Hao, J., Wang, L., et al. (2015). m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. Cell Stem Cell *16*, 289-301.

111.   Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell *15*, 707-719.

112.   Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature *526*, 591-594.

113.   Zhao, X., Yang, Y., Sun, B.F., Shi, Y., Yang, X., Xiao, W., Hao, Y.J., Ping, X.L., Chen, Y.S., Wang, W.J., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res *24*, 1403-1419.

114. Zhang, M., Zhang, Y., Ma, J., Guo, F., Cao, Q., Zhou, B., Chai, J., Zhao, W., and Zhao, R. (2015). The Demethylase Activity of FTO (Fat Mass and Obesity Associated Protein) Is Required for Preadipocyte Differentiation. PLoS One *10*, e0133788.

115. Wang, C.Y., Shie, S.S., Wen, M.S., Hung, K.C., Hsieh, I.C., Yeh, T.S., and Wu, D. (2015). Loss of FTO in adipose tissue decreases Angptl4 translation and alters triglyceride metabolism. Sci Signal *8*, ra127.

116. Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. Science *347*, 1002-1006.

117. Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. Cell *163*, 999-1010.

118. Zhu, T., Roundtree, I.A., Wang, P., Wang, X., Wang, L., Sun, C., Tian, Y., Li, J., He, C., and Xu, Y. (2014). Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. Cell Res *24*, 1493-1496.

119. Luo, S., and Tong, L. (2014). Molecular basis for the recognition of methylated adenines in RNA by the eukaryotic YTH domain. Proc Natl Acad Sci U S A *111*, 13834-13839.

120. Xu, C., Wang, X., Liu, K., Roundtree, I.A., Tempel, W., Li, Y., Lu, Z., He, C., and Min, J. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat Chem Biol *10*, 927-929.

121. Fu, Y., and He, C. (2012). Nucleic acid modifications with epigenetic significance. Curr Opin Chem Biol *16*, 516-524.

122. Niu, Y., Zhao, X., Wu, Y.S., Li, M.M., Wang, X.J., and Yang, Y.G. (2013). N6-methyl-adenosine (m6A) in RNA: an old modification with a novel epigenetic function. Genomics Proteomics Bioinformatics *11*, 8-17.

123. Lee, M., Kim, B., and Kim, V.N. (2014). Emerging roles of RNA modification: m(6)A and U-tail. Cell *158*, 980-987.

124. Roundtree, I.A., and He, C. (2016). RNA epigenetics--chemical messages for posttranscriptional gene regulation. Curr Opin Chem Biol *30*, 46-51.

125. Schwartz, S. (2016). Cracking the epitranscriptome. RNA *22*, 169-174.

126. Liu, N., and Pan, T. (2016). N6-methyladenosine-encoded epitranscriptomics. Nat Struct Mol Biol *23*, 98-102.

127. Gilbert, W.V., Bell, T.A., and Schaening, C. (2016). Messenger RNA modifications: Form, distribution, and function. Science *352*, 1408-1412.

128. Wang, Y., and Zhao, J.C. (2016). Update: Mechanisms Underlying N6-Methyladenosine Modification of Eukaryotic mRNA. Trends Genet.

129. Sternglanz, H., and Bugg, C.E. (1973). Conformation of N6-methyladenine, a base involved in DNA modification: restriction processes. Science *182*, 833-834.

130. Kierzek, E., and Kierzek, R. (2003). The thermodynamic stability of RNA duplexes and hairpins containing N6-alkyladenosines and 2-methylthio-N6-alkyladenosines. Nucleic Acids Res *31*, 4472-4480.

131. Roost, C., Lynch, S.R., Batista, P.J., Qu, K., Chang, H.Y., and Kool, E.T. (2015). Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. J Am Chem Soc *137*, 2107-2115.

132. Zou, S., Toh, J.D., Wong, K.H., Gao, Y.G., Hong, W., and Woon, E.C. (2016). N(6)-Methyladenosine: a conformational marker that regulates the substrate specificity of human demethylases FTO and ALKBH5. Sci Rep *6*, 25677.
133. Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T., et al. (2015). Structural imprints in vivo decode RNA regulatory mechanisms. Nature *519*, 486-490.
134. Loos, R.J., and Bouchard, C. (2008). FTO: the first gene contributing to common forms of human obesity. Obes Rev *9*, 246-250.
135. Frayling, T.M. (2007). Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet *8*, 657-662.
136. Larder, R., Cheung, M.K., Tung, Y.C., Yeo, G.S., and Coll, A.P. (2011). Where to go with FTO? Trends Endocrinol Metab *22*, 53-59.
137. Fawcett, K.A., and Barroso, I. (2010). The genetics of obesity: FTO leads the way. Trends Genet *26*, 266-274.
138. Farooqi, I.S. (2011). FTO and obesity: the missing link. Cell Metab *13*, 7-8.
139. Yeo, G.S., and O'Rahilly, S. (2012). Uncovering the biology of FTO. Mol Metab *1*, 32-36.
140. Meyre, D. (2012). Is FTO a type 2 diabetes susceptibility gene? Diabetologia *55*, 873-876.
141. Tung, Y.C., Yeo, G.S., O'Rahilly, S., and Coll, A.P. (2014). Obesity and FTO: Changing Focus at a Complex Locus. Cell Metab *20*, 710-718.
142. Hernandez-Caballero, M.E., and Sierra-Ramirez, J.A. (2015). Single nucleotide polymorphisms of the FTO gene and cancer risk: an overview. Mol Biol Rep *42*, 699-704.
143. Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., Rayner, N.W., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science *316*, 889-894.
144. Dina, C., Meyre, D., Gallina, S., Durand, E., Korner, A., Jacobson, P., Carlsson, L.M., Kiess, W., Vatin, V., Lecoeur, C., et al. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet *39*, 724-726.
145. Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagaraja, R., Orru, M., Usala, G., et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. PLoS Genet *3*, e115.
146. Peeters, A., Beckers, S., Verrijken, A., Roevens, P., Peeters, P., Van Gaal, L., and Van Hul, W. (2008). Variants in the FTO gene are associated with common obesity in the Belgian population. Mol Genet Metab *93*, 481-484.
147. Ohashi, J., Naka, I., Kimura, R., Natsuhara, K., Yamauchi, T., Furusawa, T., Nakazawa, M., Ataka, Y., Patarapotikul, J., Nuchnoi, P., et al. (2007). FTO polymorphisms in oceanic populations. J Hum Genet *52*, 1031-1035.
148. Hotta, K., Nakata, Y., Matsuo, T., Kamohara, S., Kotani, K., Komatsu, R., Itoh, N., Mineo, I., Wada, J., Masuzaki, H., et al. (2008). Variations in the FTO gene are associated with severe obesity in the Japanese. J Hum Genet *53*, 546-553.
149. Chang, Y.C., Liu, P.H., Lee, W.J., Chang, T.J., Jiang, Y.D., Li, H.Y., Kuo, S.S., Lee, K.C., and Chuang, L.M. (2008). Common variation in the fat mass

and obesity-associated (FTO) gene confers risk of obesity and modulates BMI in the Chinese population. Diabetes *57*, 2245-2252.

150. Villalobos-Comparan, M., Teresa Flores-Dorantes, M., Teresa Villarreal-Molina, M., Rodriguez-Cruz, M., Garcia-Ulloa, A.C., Robles, L., Huertas-Vazquez, A., Saucedo-Villarreal, N., Lopez-Alarcon, M., Sanchez-Munoz, F., et al. (2008). The FTO gene is associated with adulthood obesity in the Mexican population. Obesity (Silver Spring) *16*, 2296-2301.

151. Zhang, F., Xu, L., Jin, L., and Wang, X.F. (2008). A common variant in the FTO gene is associated with obesity in the Uyghur population. J Endocrinol Invest *31*, 1043.

152. Yajnik, C.S., Janipalli, C.S., Bhaskar, S., Kulkarni, S.R., Freathy, R.M., Prakash, S., Mani, K.R., Weedon, M.N., Kale, S.D., Deshpande, J., et al. (2009). FTO gene variants are strongly associated with type 2 diabetes in South Asian Indians. Diabetologia *52*, 247-252.

153. Gonzalez-Sanchez, J.L., Zabena, C., Martinez-Larrad, M.T., Martinez-Calatrava, M.J., Perez-Barba, M., and Serrano-Rios, M. (2009). Variant rs9939609 in the FTO gene is associated with obesity in an adult population from Spain. Clin Endocrinol (Oxf) *70*, 390-393.

154. Hennig, B.J., Fulford, A.J., Sirugo, G., Rayco-Solon, P., Hattersley, A.T., Frayling, T.M., and Prentice, A.M. (2009). FTO gene variation and measures of body mass in an African population. BMC Med Genet *10*, 21.

155. Legry, V., Cottel, D., Ferrieres, J., Arveiler, D., Andrieux, N., Bingham, A., Wagner, A., Ruidavets, J.B., Ducimetiere, P., Amouyel, P., et al. (2009). Effect of an FTO polymorphism on fat mass, obesity, and type 2 diabetes mellitus in the French MONICA Study. Metabolism *58*, 971-975.

156. Peters, T., Ausmeier, K., and Ruther, U. (1999). Cloning of Fatso (Fto), a novel gene deleted by the Fused toes (Ft) mouse mutation. Mamm Genome *10*, 983-986.

157. van der Hoeven, F., Schimmang, T., Volkmann, A., Mattei, M.G., Kyewski, B., and Ruther, U. (1994). Programmed cell death is affected in the novel mouse mutant Fused toes (Ft). Development *120*, 2601-2607.

158. Anselme, I., Laclef, C., Lanaud, M., Ruther, U., and Schneider-Maunoury, S. (2007). Defects in brain patterning and head morphogenesis in the mouse mutant Fused toes. Dev Biol *304*, 208-220.

159. Gotz, K., Briscoe, J., and Ruther, U. (2005). Homozygous Ft embryos are affected in floor plate maintenance and ventral neural tube patterning. Dev Dyn *233*, 623-630.

160. Fredriksson, R., Hagglund, M., Olszewski, P.K., Stephansson, O., Jacobsson, J.A., Olszewska, A.M., Levine, A.S., Lindblom, J., and Schioth, H.B. (2008). The obesity gene, FTO, is of ancient origin, up-regulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain. Endocrinology *149*, 2062-2071.

161. Robbens, S., Rouze, P., Cock, J.M., Spring, J., Worden, A.Z., and Van de Peer, Y. (2008). The FTO gene, implicated in human obesity, is found only in vertebrates and marine algae. J Mol Evol *66*, 80-84.

162. Sanchez-Pulido, L., and Andrade-Navarro, M.A. (2007). The FTO (fat mass and obesity associated) gene codes for a novel member of the non-heme dioxygenase superfamily. BMC Biochem *8*, 23.

163. Wahlen, K., Sjolin, E., and Hoffstedt, J. (2008). The common rs9939609 gene variant of the fat mass- and obesity-associated gene FTO is related to fat cell lipolysis. J Lipid Res *49*, 607-611.

164. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature *445*, 168-176.

165. Olszewski, P.K., Fredriksson, R., Olszewska, A.M., Stephansson, O., Alsio, J., Radomska, K.J., Levine, A.S., and Schioth, H.B. (2009). Hypothalamic FTO is associated with the regulation of energy intake not feeding reward. BMC Neurosci *10*, 129.

166. Hubacek, J.A., Stanek, V., Gebauerova, M., Pilipcincova, A., Dlouha, D., Poledne, R., Aschermann, M., Skalicka, H., Matouskova, J., Kruger, A., et al. (2010). A FTO variant and risk of acute coronary syndrome. Clin Chim Acta *411*, 1069-1072.

167. Pausova, Z., Syme, C., Abrahamowicz, M., Xiao, Y., Leonard, G.T., Perron, M., Richer, L., Veillette, S., Smith, G.D., Seda, O., et al. (2009). A common variant of the FTO gene is associated with not only increased adiposity but also elevated blood pressure in French Canadians. Circ Cardiovasc Genet *2*, 260-269.

168. Wehr, E., Schweighofer, N., Moller, R., Giuliani, A., Pieber, T.R., and Obermayer-Pietsch, B. (2010). Association of FTO gene with hyperandrogenemia and metabolic parameters in women with polycystic ovary syndrome. Metabolism *59*, 575-580.

169. Fischer, J., Koch, L., Emmerling, C., Vierkotten, J., Peters, T., Bruning, J.C., and Ruther, U. (2009). Inactivation of the Fto gene protects from obesity. Nature *458*, 894-898.

170. Church, C., Moir, L., McMurray, F., Girard, C., Banks, G.T., Teboul, L., Wells, S., Bruning, J.C., Nolan, P.M., Ashcroft, F.M., et al. (2010). Overexpression of Fto leads to increased food intake and results in obesity. Nat Genet *42*, 1086-1092.

171. Tung, Y.C., Ayuso, E., Shan, X., Bosch, F., O'Rahilly, S., Coll, A.P., and Yeo, G.S. (2010). Hypothalamic-specific manipulation of Fto, the ortholog of the human obesity gene FTO, affects food intake in rats. PLoS One *5*, e8771.

172. Haupt, A., Thamer, C., Staiger, H., Tschritter, O., Kirchhoff, K., Machicao, F., Haring, H.U., Stefan, N., and Fritsche, A. (2009). Variation in the FTO gene influences food intake but not energy expenditure. Exp Clin Endocrinol Diabetes *117*, 194-197.

173. Speakman, J.R., Rance, K.A., and Johnstone, A.M. (2008). Polymorphisms of the FTO gene are associated with variation in energy intake, but not energy expenditure. Obesity (Silver Spring) *16*, 1961-1965.

174. Tanofsky-Kraff, M., Han, J.C., Anandalingam, K., Shomaker, L.B., Columbo, K.M., Wolkoff, L.E., Kozlosky, M., Elliott, C., Ranzenhofer, L.M., Roza, C.A., et al. (2009). The FTO gene rs9939609 obesity-risk allele and loss of control over eating. Am J Clin Nutr *90*, 1483-1488.

175. Timpson, N.J., Emmett, P.M., Frayling, T.M., Rogers, I., Hattersley, A.T., McCarthy, M.I., and Davey Smith, G. (2008). The fat mass- and obesity-associated locus and dietary intake in children. Am J Clin Nutr *88*, 971-978.

176. Church, C., Lee, S., Bagg, E.A., McTaggart, J.S., Deacon, R., Gerken, T., Lee, A., Moir, L., Mecinovic, J., Quwailid, M.M., et al. (2009). A mouse model for

the metabolic effects of the human fat mass and obesity associated FTO gene. PLoS Genet *5*, e1000599.

177. Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G.S., Meyre, D., Golzio, C., Molinari, F., Kadhom, N., et al. (2009). Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. Am J Hum Genet *85*, 106-111.
178. Caglayan, A.O., Tuysuz, B., Coskun, S., Quon, J., Harmanci, A.S., Baranoski, J.F., Baran, B., Erson-Omay, E.Z., Henegariu, O., Mane, S.M., et al. (2016). A patient with a novel homozygous missense mutation in FTO and concomitant nonsense mutation in CETP. J Hum Genet *61*, 395-403.
179. Smemo, S., Tena, J.J., Kim, K.H., Gamazon, E.R., Sakabe, N.J., Gomez-Marin, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature *507*, 371-375.
180. Merkestein, M., and Sellayah, D. (2015). Role of FTO in Adipocyte Development and Function: Recent Insights. Int J Endocrinol *2015*, 521381.
181. Ben-Haim, M.S., Moshitch-Moshkovitz, S., and Rechavi, G. (2015). FTO: linking m6A demethylation to adipogenesis. Cell Res *25*, 3-4.
182. Jiao, Y., Zhang, J., Lu, L., Xu, J., and Qin, L. (2016). The Fto Gene Regulates the Proliferation and Differentiation of Pre-Adipocytes in Vitro. Nutrients *8*, 102.
183. Chen, X., Zhou, B., Luo, Y., Huang, Z., Jia, G., Liu, G., and Zhao, H. (2016). Tissue Distribution of Porcine FTO and Its Effect on Porcine Intramuscular Preadipocytes Proliferation and Differentiation. PLoS One *11*, e0151056.
184. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nat Methods *12*, 7-8.
185. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc *10*, 845-858.
186. Aik, W., Scotti, J.S., Choi, H., Gong, L., Demetriades, M., Schofield, C.J., and McDonough, M.A. (2014). Structure of human RNA N(6)-methyladenine demethylase ALKBH5 provides insights into its mechanisms of nucleic acid recognition and demethylation. Nucleic Acids Res *42*, 4741-4754.
187. Kataoka, H., Yamamoto, Y., and Sekiguchi, M. (1983). A new gene (alkB) of Escherichia coli that controls sensitivity to methyl methane sulfonate. J Bacteriol *153*, 1301-1307.
188. Samson, L., Derfler, B., and Waldstein, E.A. (1986). Suppression of human DNA alkylation-repair defects by Escherichia coli DNA-repair genes. Proc Natl Acad Sci U S A *83*, 5607-5610.
189. Volkert, M.R. (1988). Adaptive response of Escherichia coli to alkylation damage. Environ Mol Mutagen *11*, 241-255.
190. Chen, B.J., Carroll, P., and Samson, L. (1994). The Escherichia coli AlkB protein protects human cells against alkylation-induced toxicity. J Bacteriol *176*, 6255-6261.
191. Dinglay, S., Gold, B., and Sedgwick, B. (1998). Repair in Escherichia coli alkB mutants of abasic sites and 3-methyladenine residues in DNA. Mutat Res *407*, 109-116.
192. Dinglay, S., Trewick, S.C., Lindahl, T., and Sedgwick, B. (2000). Defective processing of methylated single-stranded DNA by E. coli AlkB mutants. Genes Dev *14*, 2097-2105.

193. Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature *419*, 174-178.
194. Falnes, P.O., Johansen, R.F., and Seeberg, E. (2002). AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature *419*, 178-182.
195. Begley, T.J., and Samson, L.D. (2003). AlkB mystery solved: oxidative demethylation of N1-methyladenine and N3-methylcytosine adducts by a direct reversal mechanism. Trends Biochem Sci *28*, 2-5.
196. Ougland, R., Zhang, C.M., Liiv, A., Johansen, R.F., Seeberg, E., Hou, Y.M., Remme, J., and Falnes, P.O. (2004). AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. Mol Cell *16*, 107-116.
197. Falnes, P.O., Bjoras, M., Aas, P.A., Sundheim, O., and Seeberg, E. (2004). Substrate specificities of bacterial and human AlkB proteins. Nucleic Acids Res *32*, 3456-3461.
198. Koivisto, P., Robins, P., Lindahl, T., and Sedgwick, B. (2004). Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. J Biol Chem *279*, 40470-40474.
199. Delaney, J.C., and Essigmann, J.M. (2004). Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. Proc Natl Acad Sci U S A *101*, 14051-14056.
200. Falnes, P.O. (2004). Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. Nucleic Acids Res *32*, 6260-6267.
201. Yu, B., and Hunt, J.F. (2009). Enzymological and structural studies of the mechanism of promiscuous substrate recognition by the oxidative DNA repair enzyme AlkB. Proc Natl Acad Sci U S A *106*, 14315-14320.
202. Sundheim, O., Talstad, V.A., Vagbo, C.B., Slupphaug, G., and Krokan, H.E. (2008). AlkB demethylases flip out in different ways. DNA Repair (Amst) *7*, 1916-1923.
203. Zhu, C., and Yi, C. (2014). Switching demethylation activities between AlkB family RNA/DNA demethylases through exchange of active-site residues. Angew Chem Int Ed Engl *53*, 3659-3662.
204. Welford, R.W., Schlemminger, I., McNeill, L.A., Hewitson, K.S., and Schofield, C.J. (2003). The selectivity and inhibition of AlkB. J Biol Chem *278*, 10157-10161.
205. Woon, E.C., Demetriades, M., Bagg, E.A., Aik, W., Krylova, S.M., Ma, J.H., Chan, M., Walport, L.J., Wegman, D.W., Dack, K.N., et al. (2012). Dynamic combinatorial mass spectrometry leads to inhibitors of a 2-oxoglutarate-dependent nucleic acid demethylase. J Med Chem *55*, 2173-2184.
206. Krylova, S.M., Koshkin, V., Bagg, E., Schofield, C.J., and Krylov, S.N. (2012). Mechanistic studies on the application of DNA aptamers as inhibitors of 2-oxoglutarate-dependent oxygenases. J Med Chem *55*, 3546-3552.
207. Li, Q., Huang, Y., Liu, X., Gan, J., Chen, H., and Yang, C.G. (2016). Rhein Inhibits AlkB Repair Enzymes and Sensitizes Cells to Methylated DNA Damage. J Biol Chem *291*, 11083-11093.
208. Wei, Y.F., Carter, K.C., Wang, R.P., and Shell, B.K. (1996). Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair. Nucleic Acids Res *24*, 931-937.

209. Westbye, M.P., Feyzi, E., Aas, P.A., Vagbo, C.B., Talstad, V.A., Kavli, B., Hagen, L., Sundheim, O., Akbari, M., Liabakk, N.B., et al. (2008). Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. J Biol Chem *283*, 25046-25056.

210. Ougland, R., Lando, D., Jonson, I., Dahl, J.A., Moen, M.N., Nordstrand, L.M., Rognes, T., Lee, J.T., Klungland, A., Kouzarides, T., et al. (2012). ALKBH1 is a histone H2A dioxygenase involved in neural differentiation. Stem Cells *30*, 2672-2682.

211. Haag, S., Sloan, K.E., Ranjan, N., Warda, A.S., Kretschmer, J., Blessing, C., Hubner, B., Seikowski, J., Dennerlein, S., Rehling, P., et al. (2016). NSUN3 and ABH1 modify the wobble position of mt-tRNAMet to expand codon recognition in mitochondrial translation. EMBO J.

212. Ougland, R., Jonson, I., Moen, M.N., Nesse, G., Asker, G., Klungland, A., and Larsen, E. (2016). Role of ALKBH1 in the Core Transcriptional Network of Embryonic Stem Cells. Cell Physiol Biochem *38*, 173-184.

213. Pan, Z., Sikandar, S., Witherspoon, M., Dizon, D., Nguyen, T., Benirschke, K., Wiley, C., Vrana, P., and Lipkin, S.M. (2008). Impaired placental trophoblast lineage differentiation in Alkbh1(-/-) mice. Dev Dyn *237*, 316-327.

214. Nordstrand, L.M., Svard, J., Larsen, E., Nilsen, A., Ougland, R., Furu, K., Lien, G.F., Rognes, T., Namekawa, S.H., Lee, J.T., et al. (2010). Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects. PLoS One *5*, e13827.

215. Duncan, T., Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T., and Sedgwick, B. (2002). Reversal of DNA alkylation damage by two human dioxygenases. Proc Natl Acad Sci U S A *99*, 16660-16665.

216. Koivisto, P., Duncan, T., Lindahl, T., and Sedgwick, B. (2003). Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. J Biol Chem *278*, 44348-44354.

217. Mishina, Y., Lee, C.H., and He, C. (2004). Interaction of human and bacterial AlkB proteins with DNA as probed through chemical cross-linking studies. Nucleic Acids Res *32*, 1548-1554.

218. Lee, D.H., Jin, S.G., Cai, S., Chen, Y., Pfeifer, G.P., and O'Connor, T.R. (2005). Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. J Biol Chem *280*, 39448-39459.

219. Ringvoll, J., Nordstrand, L.M., Vagbo, C.B., Talstad, V., Reite, K., Aas, P.A., Lauritzen, K.H., Liabakk, N.B., Bjork, A., Doughty, R.W., et al. (2006). Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. EMBO J *25*, 2189-2198.

220. Ringvoll, J., Moen, M.N., Nordstrand, L.M., Meira, L.B., Pang, B., Bekkelund, A., Dedon, P.C., Bjelland, S., Samson, L.D., Falnes, P.O., et al. (2008). AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA. Cancer Res *68*, 4142-4149.

221. Nay, S.L., Lee, D.H., Bates, S.E., and O'Connor, T.R. (2012). Alkbh2 protects against lethality and mutation in primary mouse embryonic fibroblasts. DNA Repair (Amst) *11*, 502-510.

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

223. Lee, S.Y., Luk, S.K., Chuang, C.P., Yip, S.P., To, S.S., and Yung, Y.M. (2010). TP53 regulates human AlkB homologue 2 expression in glioma resistance to Photofrin-mediated photodynamic therapy. Br J Cancer *103*, 362-369.

224. Johannessen, T.C., Prestegarden, L., Grudic, A., Hegi, M.E., Tysnes, B.B., and Bjerkvig, R. (2013). The DNA repair protein ALKBH2 mediates temozolomide resistance in human glioblastoma cells. Neuro Oncol *15*, 269-278.

225. Fujii, T., Shimada, K., Anai, S., Fujimoto, K., and Konishi, N. (2013). ALKBH2, a novel AlkB homologue, contributes to human bladder cancer progression by regulating MUC1 expression. Cancer Sci *104*, 321-327.

226. Monsen, V.T., Sundheim, O., Aas, P.A., Westbye, M.P., Sousa, M.M., Slupphaug, G., and Krokan, H.E. (2010). Divergent ss-hairpins determine double-strand versus single-strand substrate recognition of human AlkB-homologues 2 and 3. Nucleic Acids Res *38*, 6447-6455.

227. Chen, B., Liu, H., Sun, X., and Yang, C.G. (2010). Mechanistic insight into the recognition of single-stranded and double-stranded DNA substrates by ABH2 and ABH3. Mol Biosyst *6*, 2143-2149.

228. Konishi, N., Nakamura, M., Ishida, E., Shimada, K., Mitsui, E., Yoshikawa, R., Yamamoto, H., and Tsujikawa, K. (2005). High expression of a new marker PCA-1 in human prostate carcinoma. Clin Cancer Res *11*, 5090-5097.

229. Tasaki, M., Shimada, K., Kimura, H., Tsujikawa, K., and Konishi, N. (2011). ALKBH3, a human AlkB homologue, contributes to cell survival in human non-small-cell lung cancer. Br J Cancer *104*, 700-706.

230. Yamato, I., Sho, M., Shimada, K., Hotta, K., Ueda, Y., Yasuda, S., Shigi, N., Konishi, N., Tsujikawa, K., and Nakajima, Y. (2012). PCA-1/ALKBH3 contributes to pancreatic cancer by supporting apoptotic resistance and angiogenesis. Cancer Res *72*, 4829-4839.

231. Nakao, S., Mabuchi, M., Shimizu, T., Itoh, Y., Takeuchi, Y., Ueda, M., Mizuno, H., Shigi, N., Ohshio, I., Jinguji, K., et al. (2014). Design and synthesis of prostate cancer antigen-1 (PCA-1/ALKBH3) inhibitors as anti-prostate cancer drugs. Bioorg Med Chem Lett *24*, 1071-1074.

232. Mabuchi, M., Shimizu, T., Ueda, M., Sasakawa, Y., Nakao, S., Ueda, Y., Kawamura, A., Tsujikawa, K., and Tanaka, A. (2015). Improving the bioavailability and anticancer effect of the PCA-1/ALKBH3 inhibitor HUHS015 using sodium salt. In Vivo *29*, 39-43.

233. Neta, G., Brenner, A.V., Sturgis, E.M., Pfeiffer, R.M., Hutchinson, A.A., Aschebrook-Kilfoy, B., Yeager, M., Xu, L., Wheeler, W., Abend, M., et al. (2011). Common genetic variants related to genomic integrity and risk of papillary thyroid cancer. Carcinogenesis *32*, 1231-1237.

234. Choi, S.Y., Jang, J.H., and Kim, K.R. (2011). Analysis of differentially expressed genes in human rectal carcinoma using suppression subtractive hybridization. Clin Exp Med *11*, 219-226.

235. Dango, S., Mosammaparast, N., Sowa, M.E., Xiong, L.J., Wu, F., Park, K., Rubin, M., Gygi, S., Harper, J.W., and Shi, Y. (2011). DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. Mol Cell *44*, 373-384.

236. Zheng, G., Dahl, J.A., Niu, Y., Fu, Y., Klungland, A., Yang, Y.G., and He, C. (2013). Sprouts of RNA epigenetics: the discovery of mammalian RNA demethylases. RNA Biol *10*, 915-918.

237. Zhou, B., and Han, Z. (2013). Crystallization and preliminary X-ray diffraction of the RNA demethylase ALKBH5. Acta Crystallogr Sect F Struct Biol Cryst Commun *69*, 1231-1234.

238. Chen, W., Zhang, L., Zheng, G., Fu, Y., Ji, Q., Liu, F., Chen, H., and He, C. (2014). Crystal structure of the RNA demethylase ALKBH5 from zebrafish. FEBS Lett *588*, 892-898.

239. Feng, C., Liu, Y., Wang, G., Deng, Z., Zhang, Q., Wu, W., Tong, Y., Cheng, C., and Chen, Z. (2014). Crystal structures of the human RNA demethylase Alkbh5 reveal basis for substrate recognition. J Biol Chem *289*, 11571-11583.

240. Xu, C., Liu, K., Tempel, W., Demetriades, M., Aik, W., Schofield, C.J., and Min, J. (2014). Structures of human ALKBH5 demethylase reveal a unique binding mode for specific single-stranded N6-methyladenosine RNA demethylation. J Biol Chem *289*, 17299-17311.

241. Zhang, C., Samanta, D., Lu, H., Bullen, J.W., Zhang, H., Chen, I., He, X., and Semenza, G.L. (2016). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m(6)A-demethylation of NANOG mRNA. Proc Natl Acad Sci U S A *113*, E2047-2056.

242. Chan, M.C., Holt-Martyn, J.P., Schofield, C.J., and Ratcliffe, P.J. (2016). Pharmacological targeting of the HIF hydroxylases--A new field in medicine development. Mol Aspects Med *47-48*, 54-75.

243. McAllister, T.E., England, K.S., Hopkinson, R.J., Brennan, P.E., Kawamura, A., and Schofield, C.J. (2016). Recent Progress in Histone Demethylase Inhibitors. J Med Chem *59*, 1308-1329.

244. Mecinovic, J., Loenarz, C., Chowdhury, R., and Schofield, C.J. (2009). 2-Oxoglutarate analogue inhibitors of prolyl hydroxylase domain 2. Bioorg Med Chem Lett *19*, 6192-6195.

245. Toh, J.D.W., Sun, L.Y., Lau, L.Z.M., Tan, J., Low, J.J.A., Tang, C.W.Q., Cheong, E.J.Y., Tan, M.J.H., Chen, Y., Hong, W.J., et al. (2015). A strategy based on nucleotide specificity leads to a subfamily-selective and cell-active inhibitor of N-6-methyladenosine demethylase FTO. Chem Sci *6*, 112-122.

246. Chen, B., Ye, F., Yu, L., Jia, G., Huang, X., Zhang, X., Peng, S., Chen, K., Wang, M., Gong, S., et al. (2012). Development of cell-active N6-methyladenosine RNA demethylase FTO inhibitor. J Am Chem Soc *134*, 17963-17971.

247. Aik, W., Demetriades, M., Hamdan, M.K., Bagg, E.A., Yeoh, K.K., Lejeune, C., Zhang, Z., McDonough, M.A., and Schofield, C.J. (2013). Structural basis for inhibition of the fat mass and obesity associated protein (FTO). J Med Chem *56*, 3680-3688.

248. McMurray, F., Demetriades, M., Aik, W., Merkestein, M., Kramer, H., Andrew, D.S., Scudamore, C.L., Hough, T.A., Wells, S., Ashcroft, F.M., et al. (2015). Pharmacological inhibition of FTO. PLoS One *10*, e0121829.

249. Huang, Y., Yan, J., Li, Q., Li, J., Gong, S., Zhou, H., Gan, J., Jiang, H., Jia, G.F., Luo, C., et al. (2015). Meclofenamic acid selectively inhibits FTO demethylation of m6A over ALKBH5. Nucleic Acids Res *43*, 373-384.

250. He, W., Zhou, B., Liu, W., Zhang, M., Shen, Z., Han, Z., Jiang, Q., Yang, Q., Song, C., Wang, R., et al. (2015). Identification of A Novel Small-Molecule Binding Site of the Fat Mass and Obesity Associated Protein (FTO). J Med Chem *58*, 7341-7348.

251. Wang, T., Hong, T., Huang, Y., Su, H., Wu, F., Chen, Y., Wei, L., Huang, W., Hua, X., Xia, Y., et al. (2015). Fluorescein Derivatives as Bifunctional

Molecules for the Simultaneous Inhibiting and Labeling of FTO Protein. J Am Chem Soc *137*, 13736-13739.

252. Qiao, Y., Zhou, B., Zhang, M., Liu, W., Han, Z., Song, C., Yu, W., Yang, Q., Wang, R., Wang, S., et al. (2016). A Novel Inhibitor of the Obesity-Related Protein FTO. Biochemistry *55*, 1516-1522.

253. Svensen, N., and Jaffrey, S.R. (2016). Fluorescent RNA Aptamers as a Tool to Study RNA-Modifying Enzymes. Cell Chem Biol *23*, 415-425.

254. Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., et al. (2008). Protein production and purification. Nat Methods *5*, 135-146.

255. Savitsky, P., Bray, J., Cooper, C.D., Marsden, B.D., Mahajan, P., Burgess-Brown, N.A., and Gileadi, O. (2010). High-throughput production of human proteins for crystallization: the SGC experience. J Struct Biol *172*, 3-13.

256. Niesen, F.H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc *2*, 2212-2221.

257. Rose, N.R., Woon, E.C., Kingham, G.L., King, O.N., Mecinovic, J., Clifton, I.J., Ng, S.S., Talib-Hardy, J., Oppermann, U., McDonough, M.A., et al. (2010). Selective inhibitors of the JMJD2 histone demethylases: combined nondenaturing mass spectrometric screening and crystallographic approaches. J Med Chem *53*, 1810-1818.

258. Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr *66*, 133-144.

259. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

260. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr *61*, 458-464.

261. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

262. Schuttelkopf, A.W., and van Aalten, D.M. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr *60*, 1355-1363.

263. Krieger, E., Joo, K., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins *77 Suppl 9*, 114-122.

264. Wienken, C.J., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. Nat Commun *1*, 100.

265. Baneyx, F. (1999). Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol *10*, 411-421.

266. Kypr, J., Kejnovska, I., Renciuk, D., and Vorlickova, M. (2009). Circular dichroism and conformational polymorphism of DNA. Nucleic Acids Res *37*, 1713-1725.

267. Workman, P., and Collins, I. (2010). Probing the probes: fitness factors for small molecule tools. Chem Biol *17*, 561-577.

268. Bunnage, M.E., Chekler, E.L., and Jones, L.H. (2013). Target validation using chemical probes. Nat Chem Biol *9*, 195-199.

269. Arrowsmith, C.H., Audia, J.E., Austin, C., Baell, J., Bennett, J., Blagg, J., Bountra, C., Brennan, P.E., Brown, P.J., Bunnage, M.E., et al. (2015). The promise and peril of chemical probes. Nat Chem Biol *11*, 536-541.

270. Brion, P., and Westhof, E. (1997). Hierarchy and dynamics of RNA folding. Annu Rev Biophys Biomol Struct *26*, 113-137.