

# Mechanisms of substrate recognition and *N*<sup>6</sup>-methyladenosine demethylation revealed by crystal structures of ALKBH5–RNA complexes

Simranjeet Kaur<sup>1</sup>, Nok Yin Tam<sup>1</sup>, Michael A. McDonough<sup>1,2,\*</sup>, Christopher J. Schofield<sup>1,2,\*</sup> and Wei Shen Aik<sup>1,\*</sup>

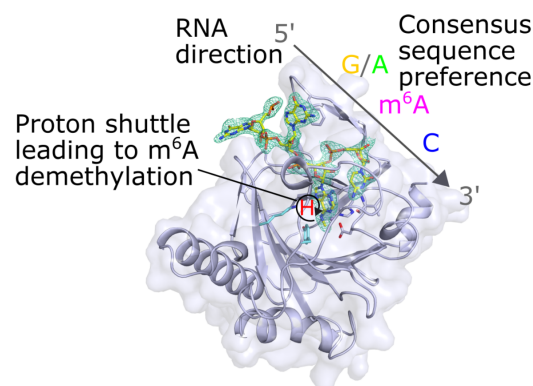
<sup>1</sup>Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Hong Kong SAR, China and <sup>2</sup>The Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, UK

Received January 24, 2022; Revised March 10, 2022; Editorial Decision March 11, 2022; Accepted March 14, 2022

## ABSTRACT

AlkB homologue 5 (ALKBH5) is a ferrous iron and 2-oxoglutarate dependent oxygenase that demethylates RNA *N*<sup>6</sup>-methyladenosine (m<sup>6</sup>A), a post-transcriptional RNA modification with an emerging set of regulatory roles. Along with the fat mass and obesity-associated protein (FTO), ALKBH5 is one of only two identified human m<sup>6</sup>A RNA oxidizing enzymes and is a potential target for cancer treatment. Unlike FTO, ALKBH5 efficiently catalyzes fragmentation of its proposed nascent hemiaminal intermediate to give formaldehyde and a demethylated nucleoside. A detailed analysis of the molecular mechanisms used by ALKBH5 for substrate recognition and m<sup>6</sup>A demethylation is lacking. We report three crystal structures of ALKBH5 in complex with an m<sup>6</sup>A-ssRNA 8-mer substrate and supporting biochemical analyses. Strikingly, the single-stranded RNA substrate binds to the active site of ALKBH5 in a 5'-3' orientation that is opposite to single-stranded or double-stranded DNA substrates observed for other AlkB subfamily members, including single-stranded DNA bound to FTO. The combined structural and biochemical results provide insight into the preference of ALKBH5 for substrates containing a (A/G)m<sup>6</sup>AC consensus sequence motif. The results support a mechanism involving formation of an m<sup>6</sup>A hemiaminal intermediate, followed by efficient ALKBH5 catalyzed demethylation, enabled by a proton shuttle network involving Lys132 and Tyr139.

## GRAPHICAL ABSTRACT



ALKBH5-RNA Structural Complex

## INTRODUCTION

*N*<sup>6</sup>-Methyladenosine (m<sup>6</sup>A) is the most abundant stable internal RNA modification and has diverse cellular functions including in regulation of mRNA splicing, RNA processing, nuclear transport, translation and RNA stability (1). ALKBH5 is upregulated in hypoxia and is a hypoxia inducible factor (HIF) target gene (2), suggesting it may play a role in the hypoxic response. Along with the fat mass and obesity-associated protein (FTO), human AlkB homologue 5 (ALKBH5) is one of only two identified human enzymes catalyzing RNA m<sup>6</sup>A methyl group oxidation (3). The m<sup>6</sup>A RNA modification is reported to have important roles in cancer and viral infection (4,5). The m<sup>6</sup>A demethylase activity of ALKBH5 plays roles in the promotion of certain cancers, including glioblastoma, hepatocarcinoma, and acute myeloid leukemia (6–8). ALKBH5 is emerging as a potential drug-target for cancer, however, *ALKBH5* over-

\*To whom correspondence should be addressed. Tel: +852 3411 6682; Fax: +852 3411 7348; Email: aikweishen@hkbu.edu.hk  
Correspondence may also be addressed to Michael A. McDonough. Tel: +44 1865 275 629; Fax: +44 1865 285 002;  
Email: michael.mcdonough@chem.ox.ac.uk  
Correspondence may also be addressed to Christopher J. Schofield. Tel: +44 1865 275 625; Fax: +44 1865 285 002;  
Email: christopher.schofield@chem.ox.ac.uk

expression is reported to inhibit several cancer cell types, including bladder and pancreatic cancers (9,10). Dissecting the apparently conflicting context-dependent roles of m<sup>6</sup>A demethylation in different types of cancer with the use of small molecule probes is of current interest (11).

ALKBH5 is a member of the Fe(II) and 2-oxoglutarate (2OG)-dependent oxygenase (2OG oxygenase) superfamily (2), which catalyzes oxidations of diverse substrates, including nucleic acids, lipids, proteins, and small molecule metabolites (12,13). Within the 2OG oxygenase superfamily, ALKBH5 belongs to the AlkB subfamily, which includes the bacterial *Escherichia coli* nucleic acid repair enzyme AlkB and the human AlkB homologues ALKBH1-8 and FTO.

ALKBH5 is proposed to demethylate m<sup>6</sup>A in RNA by oxidizing the N<sup>6</sup>-methyl group, producing a transient N<sup>6</sup>-hydroxymethyladenosine (hm<sup>6</sup>A) intermediate that efficiently fragments to give adenosine and formaldehyde (Figure 1) (12,13). By contrast, enzyme free hm<sup>6</sup>A is a relatively stable species at physiological pH (14,15). The oxidation of m<sup>6</sup>A as catalyzed by FTO is reported to produce a mixture of adenosine, hm<sup>6</sup>A, and N<sup>6</sup>-formyladenosine (f<sup>6</sup>A) (Figure 1) (15–17), with hm<sup>6</sup>A being the major product, at least under some assay conditions (14). By contrast with FTO, only the demethylated adenosine product has been observed following ALKBH5 catalysis (17). Recent studies report a role for FTO in specifically demethylating the N<sup>6</sup>-methyl of N<sup>6</sup>,2'-O-dimethyladenosine (m<sup>6</sup>Am) adjacent to the 5'-cap of small nuclear RNA (snRNA) (18). The apparently different substrates and products of FTO and ALKBH5 catalysis may relate to their different physiological roles. Defining detailed structural and mechanistic differences that underlie the different substrate and product selectivities of ALKBH5 and FTO is thus of biological importance with respect to their different roles in RNA regulation and will be useful for the development of selective inhibitors for each.

Structural studies of nucleic acid oxygenases (NAOX) have revealed distinct nucleic acid substrate binding modes. For instance, crystallographic analyses of an FTO variant, AlkB, ALKBH2 and the ten-eleven translocation enzymes (TETs) in complex with their respective ssDNA or dsDNA substrates reveal differences in how they bind their substrates (19–24). We, and others, have determined structures of ALKBH5 in the absence of an RNA substrate (25–28). However, a structure of ALKBH5 bound to an m<sup>6</sup>A-RNA substrate has been unavailable to date. Here, we report crystal structures of ALKBH5 in complex with an m<sup>6</sup>A-containing ssRNA substrate to address this shortfall. The results reveal an unprecedented substrate-binding mode for ALKBH5 relative to other AlkB subfamily oxygenases and enable proposal of a structure-based mechanism rationalizing why, in contrast to FTO, ALKBH5 catalyzes efficient demethylation of m<sup>6</sup>A to A.

## MATERIALS AND METHODS

### Molecular cloning and site-directed mutagenesis

A gene encoding for N-terminally hexahistidine-tagged ALKBH5 residues 74–292 (ALKBH5<sub>74–292</sub>) was cloned into the pNIC28-Bsa4 vector (29) using Transfer-PCR (30).

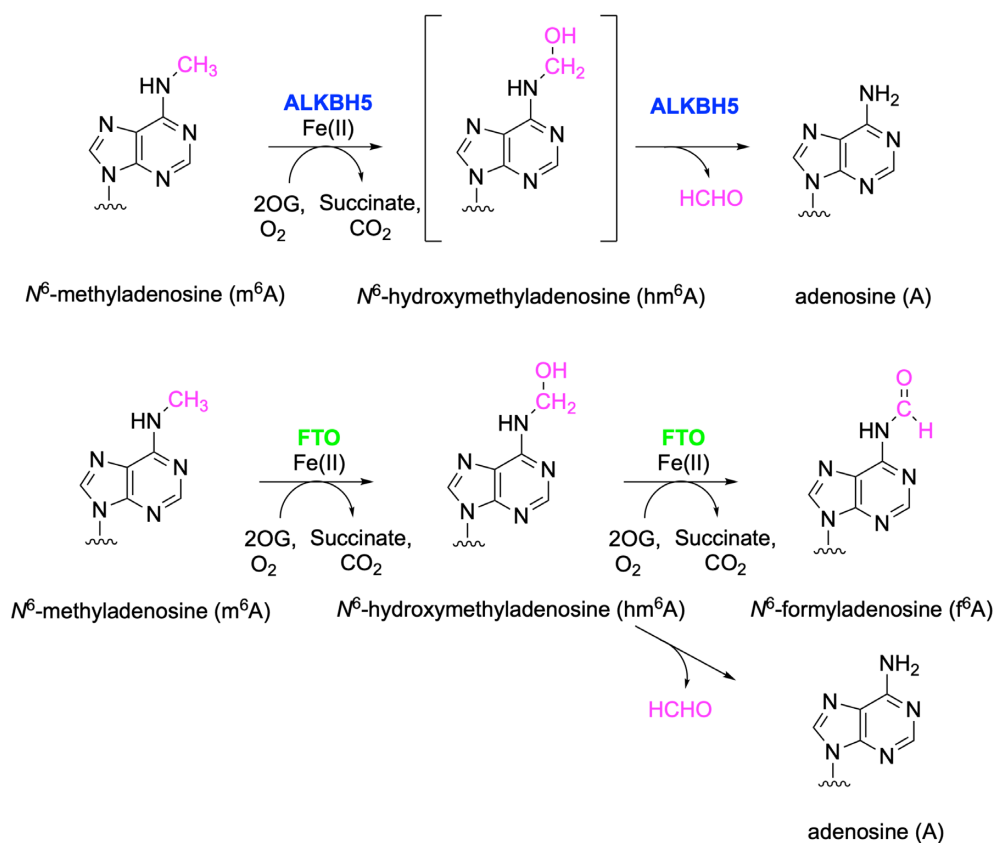
ALKBH5<sub>74–292</sub> variants were generated by site-directed mutagenesis using Transfer-PCR (30), using the same donor and acceptor plasmid pNIC28-Bsa4-ALKBH5<sub>74–292</sub> (primers are given in Supplementary Table S2).

### Production and purification of wild type and variant ALKBH5<sub>74–292</sub> proteins

The plasmid pNIC28-Bsa4-ALKBH5<sub>74–292</sub> was transformed into *Escherichia coli* Rosetta (DE3) pLysS competent cells and grown in 2YT media supplemented with 50 µg/ml kanamycin and 34 µg/ml chloramphenicol at 37°C. When the cell culture OD<sub>600</sub> reached 0.8–1.0; expression was induced using a final concentration of 0.8 mM isopropyl β-D-1-thiogalactopyranoside; cells were then grown at 18°C overnight. They were harvested by centrifugation; the cell pellets were resuspended in lysis buffer (25 mM Tris, pH 7.2, 500 mM NaCl, 10 mM imidazole, and 17.4 µg/ml phenylmethanesulfonyl fluoride), and lysed by sonication. The lysates were separated from the insoluble fractions by high-speed centrifugation (Beckman Avanti JA 25.5) at 69673 g for 30 min at 4°C. The lysate supernatant was filtered using a syringe filter with pore size of 0.45 µm and purified by nickel affinity chromatography. For structural studies, the lysate supernatant was loaded onto a 5 ml His-trap column (Cytiva, USA) pre-equilibrated with lysis buffer. For biochemical studies, the lysate supernatant was incubated with 0.5 ml Ni-NTA agarose beads (Macherey-Nagel, Germany) at 4°C for 1 h before being loaded onto a gravity column. The columns were washed with 10 column volumes of wash buffer (25 mM Tris, pH 7.2, 500 mM NaCl, 40 mM imidazole) and protein was eluted with 3–4 column volumes of elution buffer (25 mM Tris, pH 7.2, 500 mM NaCl, 375 mM imidazole). The eluate was incubated for 30 min with a final concentration of 100 mM ethylenediaminetetraacetic acid (EDTA) and the His-tag was cleaved by TEV protease (protease to target protein ratio (w/w) of 1:20) overnight while being dialyzed into a buffer containing 25 mM Tris, pH 7.2, 100 mM NaCl and 10 mM β-mercaptoethanol (BME). The eluted proteins were individually purified using a 5 ml HiTrap Heparin column (Cytiva, USA) pre-equilibrated with Buffer A (25 mM Tris, pH 7.2) using a linear gradient of Buffer A and Buffer B (25 mM Tris, pH 7.2, 1 M NaCl). For structural studies, the wild type protein was further purified by size exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column (Cytiva, USA) pre-equilibrated with buffer containing 20 mM MES, pH 6.5, 150 mM NaCl, 1 mM dithiothreitol (DTT). The relevant protein fractions were pooled, concentrated to 47.5 mg/ml, flash frozen in liquid nitrogen and stored at –80°C.

### Production and purification of FTOΔ31

FTOΔ31 was produced and purified as reported (31). In brief, FTOΔ31 was produced in *E. coli* Rosetta pLysS cells and purified by nickel affinity chromatography, treated with a final concentration of 200 mM EDTA, and further purified by heparin affinity, followed by MonoQ ion exchange chromatography. The purified protein was buffered exchanged into 25 mM Tris, pH 7.5, snap frozen in liquid nitrogen and stored at –80°C.



**Figure 1.** Schematic representation of the ALKBH5 and FTO catalyzed reactions.

### Crystallization of ALKBH5<sub>74-292</sub>-m<sup>6</sup>A-containing RNA complex

To prepare a near homogeneous ALKBH5<sub>74-292</sub>-ssRNA complex for crystallisation; ALKBH5<sub>74-292</sub> in 25 mM Tris pH 7.5 buffer, MnCl<sub>2</sub>, *N*-oxalylglycine (NOG), and an 8-mer single-stranded RNA substrate (5'-UGG(m<sup>6</sup>A)CUGC-3', Horizon Discovery Ltd, UK) were mixed in a 1.2:6:6:1 (ALKBH5<sub>74-292</sub>:MnCl<sub>2</sub>:NOG:RNA) molar ratio in the above order, with a final volume of 1 ml with concentrations of 400 μM ALKBH5<sub>74-292</sub>, 2 mM MnCl<sub>2</sub>, 2 mM NOG and 333 μM RNA substrate. Since NOG was not observed in the active site of the resultant structure, we increased the concentrations of NOG or 2OG 5-fold to increase ligand occupancy. To prepare near homogeneous ALKBH5<sub>74-292</sub>-2OG-ssRNA or ALKBH5<sub>74-292</sub>-NOG-ssRNA for crystallization; ALKBH5<sub>74-292</sub> in 25 mM Tris, pH 7.5 buffer, MnCl<sub>2</sub>, 2OG or NOG, and an 8-mer single-stranded RNA substrate (5'-UGG(m<sup>6</sup>A)CUGC-3', Horizon Discovery Ltd, UK) were mixed in a 1.2:6:30:1 (ALKBH5<sub>74-292</sub>:MnCl<sub>2</sub>:2OG/NOG:RNA) molar ratio in the above order, with a final volume of 1 ml with concentrations of 400 μM ALKBH5<sub>74-292</sub>, 2 mM MnCl<sub>2</sub>, 10 mM 2OG or NOG and 333 μM RNA substrate. The mixtures were incubated on ice for 30 min. Each complex was purified by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva, USA), pre-equilibrated with 25 mM Tris pH 7.5 buffer. The fractions corresponding to protein-RNA complex, as

determined by UV-vis 260/280 nm ratio (>1), were pooled and concentrated to 17.6 mg/ml (ALKBH5<sub>74-292</sub>-RNA), 11.0 mg/ml (ALKBH5<sub>74-292</sub>-2OG-RNA), and 19.2 mg/ml (ALKBH5<sub>74-292</sub>-NOG-RNA). The protein-RNA complex concentrations were estimated using the summed extinction coefficient (63561 M<sup>-1</sup> cm<sup>-1</sup>) at 280 nm for both ALKBH5<sub>74-292</sub> (15930 M<sup>-1</sup> cm<sup>-1</sup>) and 8-mer m<sup>6</sup>A-containing ssRNA (47631 M<sup>-1</sup> cm<sup>-1</sup>), which was taken to be the extinction coefficient at 260 nm (74400 M<sup>-1</sup> cm<sup>-1</sup>) divided by the 260/280 ratio of its UV absorption (1.562).

Crystals of the ALKBH5<sub>74-292</sub>-RNA complex and the ALKBH5<sub>74-292</sub>-NOG-RNA complex were grown at 293 K by the sitting drop vapor diffusion method in CrystalMation Intelli-plate 96-3 High Profile plates (Hampton Research, USA). Crystallization plates were prepared using a Crystal Gryphon dispenser (Art Robbins Instruments, USA). 0.2 μl of 12 mg/ml ALKBH5-RNA complex was mixed with an equal volume of reservoir solution containing 0.2 M sodium sulfate decahydrate, 20% (w/v) polyethylene glycol (PEG) 3350 and equilibrated against 60 μl of reservoir solution. Crystals of the ALKBH5<sub>74-292</sub>-RNA complex grew as rods after two days with typical dimensions 280 μm × 10 μm × 7 μm. 0.2 μl of solution containing final concentrations of 12 mg/ml ALKBH5<sub>74-292</sub>-NOG-RNA complex, 2 mM MnCl<sub>2</sub> and 10 mM NOG, was mixed with an equal volume of reservoir solution containing 0.2 M sodium formate, 20% (w/v) PEG 3350 and equilibrated against 60 μl of reservoir solution. Crystals of the ALKBH5<sub>74-292</sub>-NOG-RNA complex grew as rods af-



ter two days with typical dimensions  $170\ \mu\text{m} \times 15\ \mu\text{m} \times 7\ \mu\text{m}$ .

To obtain the crystals of the ALKBH5<sub>74–292</sub>–2OG–RNA complex using the hanging drop vapour diffusion method in 24-well VDX plates (Hampton Research, USA) a 1  $\mu\text{l}$  solution containing final concentrations of 10 mg/ml ALKBH5<sub>74–292</sub>–2OG–RNA complex, 2 mM MnCl<sub>2</sub> and 10 mM 2OG was mixed with an equal volume of reservoir solution containing 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES) monohydrate, pH 6.5, 16% w/v PEG 6000 and 5% (v/v) 2-methyl-2,4-pentanediol (MPD); the mixture was equilibrated against 500  $\mu\text{l}$  of reservoir solution at 293 K. Crystals of the ALKBH5<sub>74–292</sub>–2OG–RNA complex grew as rods and appeared after two days with typical dimensions  $1000\ \mu\text{m} \times 50\ \mu\text{m} \times 40\ \mu\text{m}$ . The reservoir solution was diluted with 100% (v/v) glycerol to a final concentration of 20% (v/v) glycerol for use as cryo-protectant. All crystals were harvested and transferred to the cryo-protectant solution using a nylon loop (Hampton Research, USA), then flash cooled in liquid nitrogen and stored under liquid nitrogen until data collection.

#### Data collection and structure determination

X-ray diffraction data were collected from a single crystal of the ALKBH5<sub>74–292</sub>–m<sup>6</sup>A RNA complex at Diamond Light Source Beamline I03 equipped with an Eiger2 XE 16M detector using X-rays tuned to wavelength 0.9763 Å. The autoprocessed data were indexed as high symmetry space group *P*<sub>31</sub>12, but initial attempts at molecular replacement (MR) failed to provide a reasonable structure solution. Analysis using PHENIX.XTRIAGE (32) indicated merohedral twinning may be present. The data were then re-processed (integrated, merged and scaled) with XIA2 (33) in the lower symmetry space group *P*<sub>3</sub><sub>2</sub>. The structure was solved in the *P*<sub>3</sub><sub>2</sub> space group by molecular replacement using PHASER (34) and a structure of ALKBH5<sub>66–292</sub> (PDB ID 4NJ4) (25) as the search model. Iterative rounds of model building and refinement were performed using COOT (35) and PHENIX.REFINE (32) by including the twin law *-k, -h, -l* in the refinement parameters until the decreasing *R*<sub>work</sub> and *R*<sub>free</sub> no longer converged. X-ray diffraction data for single crystals of either the ALKBH5<sub>74–292</sub>–2OG–m<sup>6</sup>A RNA or ALKBH5<sub>74–292</sub>–NOG–m<sup>6</sup>A RNA complexes were collected at Diamond Light Source Beamline I24 equipped with a Pilatus3 6M detector using X-rays tuned to wavelength 0.9999 Å. The data for the ALKBH5<sub>74–292</sub>–2OG–m<sup>6</sup>A RNA complex were processed (integrated, merged and scaled) using XIA2 in space group *P*<sub>3</sub><sub>2</sub>. The structure was solved in the *P*<sub>3</sub><sub>2</sub> space group by molecular replacement using PHASER and the ALKBH5<sub>74–292</sub>–m<sup>6</sup>A RNA complex (PDB ID 7V4G) as the search model. Crystallographic refinement proceeded as above using PHENIX.REFINE by including the twin law *-k, -h, -l*. For the ALKBH5<sub>74–292</sub>–NOG–m<sup>6</sup>A RNA complex data were collected separately in three sweeps from two rod-shaped crystals (0.1° oscillation per image). Two of three data sets were collected from the same crystal but exposed in different regions so as to avoid radiation damage, the other was collected from a different crystal that was

harvested from the same crystallization drop. Autoprocessing using XIA2 indexed each individual data set in the *P*<sub>2</sub><sub>1</sub> space group and twinning was not detected. To improve the completeness and *CC*<sub>1/2</sub> value for the high resolution bin, the three datasets were successfully merged. The structure was solved by molecular replacement using PHASER and a structure of the ALKBH5<sub>74–292</sub>–m<sup>6</sup>A RNA complex (PDB ID: 7V4G) as the search model. The model was iteratively fitted and refined using COOT and PHENIX.REFINE, respectively. Crystallographic data processing and refinement statistics can be found in Supplementary Table S1.

#### m<sup>6</sup>A demethylation assays

To investigate the activities of ALKBH5<sub>74–292</sub> and FTOΔ31 (FTO residues 32–505) using m<sup>6</sup>A-containing 8-mer RNA substrates containing varied sequences (Horizon Discovery, UK) (Supplementary Table S3), 50  $\mu\text{l}$  reactions containing 0.5  $\mu\text{M}$  ALKBH5<sub>74–292</sub> or FTOΔ31, 300  $\mu\text{M}$  2OG sodium salt, 2 mM sodium L-ascorbate, 150  $\mu\text{M}$  diammonium iron(II) sulfate, and 4  $\mu\text{M}$  m<sup>6</sup>A-containing 8-mer ssRNA in 25 mM Tris pH 7.5 were performed in triplicate. Reactions were incubated at 37°C for 3 min (ALKBH5<sub>74–292</sub>) or 60 min (FTOΔ31) and quenched by adding an equivalent volume of 2 mM EDTA.

To investigate the activities of ALKBH5<sub>74–292</sub> variants relative to wild type enzyme, 50  $\mu\text{l}$  reactions containing 1  $\mu\text{M}$  ALKBH5<sub>74–292</sub> (wild type or variants), 300  $\mu\text{M}$  2OG sodium salt, 2 mM sodium L-ascorbate, 150  $\mu\text{M}$  diammonium iron(II) sulfate, and 8  $\mu\text{M}$  m<sup>6</sup>A-containing ssRNA with the sequence 5'-UGG(m<sup>6</sup>A)CUGC-3' in 25 mM Tris pH 7.5 were performed in triplicate. The reactions were carried out at 37°C for 7 min and quenched by an equivalent volume of 2 mM EDTA.

The quenched samples (2  $\mu\text{l}$ ) were mixed with 1  $\mu\text{l}$  of Matrix Assisted Laser Desorption Ionization (MALDI) matrix (prepared by adding 2 parts of 2,4,6-trihydroxyacetophenone and one part of ammonium citrate) and analyzed using MALDI-Time of Flight (TOF) mass spectrometry (Bruker autoflex<sup>®</sup> maX) and flexAnalysis software.

## RESULTS AND DISCUSSION

In our original structural studies on ALKBH5, in which we reported a structure of ALKBH5<sub>66–292</sub> (residues 66–292) (25), we attempted co-crystallisation of ALKBH5<sub>66–292</sub> (residues 66–292) with a 5-mer ssRNA substrate (GGm<sup>6</sup>ACU) by mixing ALKBH5<sub>66–292</sub>, Mn(II), NOG, and 5-mer ssRNA prior to performing crystallization trials; however, this procedure did not lead to crystals. In this study, our initial attempts to prepare a complex of ALKBH5<sub>66–292</sub>–8-mer RNA containing m<sup>6</sup>A using size exclusion chromatography provided clustered crystals unsuitable for diffraction studies. We modified the construct to produce N-terminally truncated ALKBH5<sub>74–292</sub> (residues 74–292); we were able to co-purify ALKBH5<sub>74–292</sub> in complex with Mn (substituting for Fe) and the m<sup>6</sup>A-containing 8-mer ssRNA substrate (U<sub>-3</sub>G<sub>-2</sub>G<sub>-1</sub>(m<sup>6</sup>A)<sub>0</sub>C<sub>+1</sub>U<sub>+2</sub>G<sub>+3</sub>C<sub>+4</sub>). Following optimization, crystals were obtained that diffracted to 2.1 Å

resolution (Figure 2D) in the  $P3_2$  space group and a structure of ALKBH5<sub>74–292</sub> in complex with m<sup>6</sup>A-containing RNA was solved by molecular replacement using a structure of human ALKBH5<sub>66–292</sub> (PDB ID: 4NJ4) (25) as the search model. Three copies of the ALKBH5<sub>74–292</sub>–ssRNA complex are present in the asymmetric unit. Clear electron density is present for four ( $G_{-2}G_{-1}(m^6A)_0C_{+1}$ ) of the eight nucleotides of the ssRNA at the ALKBH5 active site region in all three copies in the asymmetric unit and the observable ssRNA does not appear to be involved in crystal contacts (Supplementary Figure S1). Interestingly, in this initial substrate complex structure the ALKBH5 active site did not contain the cosubstrate 2OG analogue NOG, although NOG had been added to the sample mixture prior to purification by size exclusion chromatography.

Using higher ligand concentrations in the sample mixture and the crystallization drops, we subsequently determined structures of (i) ALKBH5<sub>74–292</sub> in complex with the same m<sup>6</sup>A-containing 8-mer ssRNA and the 2OG cosubstrate in the  $P3_2$  crystal form to 2.1 Å resolution (Figure 2B), and (ii) a structure of ALKBH5<sub>74–292</sub> in complex with NOG and the m<sup>6</sup>A-containing 8-mer ssRNA in the  $P2_1$  crystal form to 2.5 Å resolution (Figure 2C). The structure of the ALKBH5<sub>74–292</sub>–NOG–RNA complex in space group  $P2_1$  has five copies of the ALKBH5<sub>74–292</sub>–NOG–ssRNA complex in the asymmetric unit, two of which manifest electron density for all eight nucleotides of the ssRNA substrate (Figure 2C), likely in part due to stabilization by crystal packing contacts with neighboring molecules (Supplementary Figure S2). Four ( $G_{-2}G_{-1}(m^6A)_0C_{+1}$ ) of the eight nucleotides are visible in the other three molecules in the asymmetric unit, in a similar manner for all the ssRNA substrate molecules in both the  $P3_2$  crystal form structures. The three structures from both crystal forms manifest similar overall protein conformations among molecules in the asymmetric unit of each structure, and when compared among the three structures using Chain A from each structure as representatives (RMSDs  $C_\alpha$  0.2 Å (Chain A from the ALKBH5<sub>74–292</sub>–NOG–RNA complex) and 0.2 Å (Chain A from the ALKBH5<sub>74–292</sub>–RNA complex) relative to Chain A from the ALKBH5<sub>74–292</sub>–2OG–RNA complex). Aside from the four extra nucleotides observed in the  $P2_1$  structure, the active site ssRNA binding modes in all the chains in the asymmetric units of the three structures are similar (Figure 2B–D, Supplementary Figure S3).

To investigate the importance of selected ALKBH5–substrate interactions revealed by the structures and to probe the ALKBH5 mechanisms, we assayed recombinant wild type ALKBH5<sub>74–292</sub>, ALKBH5 active site variants, and FTOΔ31 (FTO residues 32–505) using a set of 8-mer m<sup>6</sup>A-containing RNA substrates with varied sequences (Supplementary Table S3) and determined their relative activities using matrix assisted laser desorption ionization mass spectrometry (MALDI-MS) based assays (25) (Supplementary Figure S4). Bases upstream and downstream of the m<sup>6</sup>A were varied individually and in combination in one instance (Supplementary Table S3) to investigate their roles in substrate recognition.

## Overall structure of ALKBH5 in complex with ssRNA

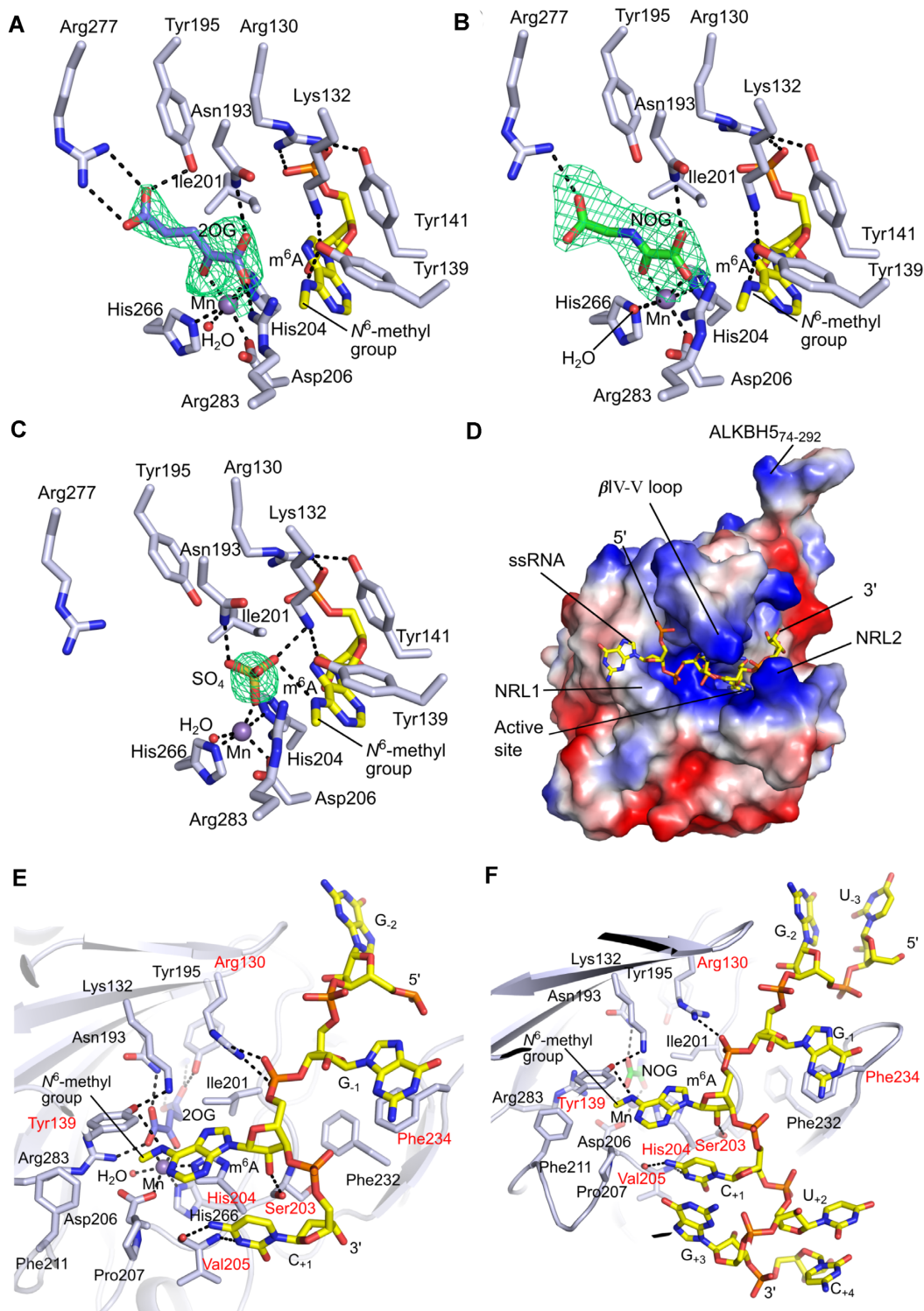
Structures of ALKBH5 in the absence of substrate (25–28) have revealed the 2OG oxygenase superfamily characteristic distorted double-stranded  $\beta$ -helix (DSBH) core fold comprising eight  $\beta$ -strands; in the case of ALKBH5, this has an extended  $\beta$ IV–V loop (residues 229–242) containing a  $3_{10}$  helix and has an N-terminal extension, comprised of three  $\alpha$ -helices and five  $\beta$ -strands (Figure 2A). Two nucleotide recognition lids (NRL1 (residues 124–137) and NRL2 (residues 138–155)) are common features of the AlkB family of 2OG oxygenases and are positioned within the N-terminal extension of ALKBH5: NRL1 is formed by  $\beta$ -strands 2 and 3 and NRL2 by  $\beta$ -strands 4 and 5. In the reported structures of ALKBH5 (25–28), NRL1 is shorter than in other AlkB subfamily members, while NRL2 is partially disordered.

The overall structure of the ALKBH5<sub>74–292</sub>–2OG–ssRNA complex is similar to the human ALKBH5 structures without substrate (25,26,28) (e.g. RMSD  $C_\alpha$  = 0.3 Å between Chain A from ALKBH5<sub>74–292</sub>–2OG–RNA complex (PDB ID: 7WKV) and Chain A from ALKBH5<sub>66–292</sub> without substrate (PDB ID: 4NJ4)), with the exception of conformational changes in the loops (NRL2 and the  $\beta$ IV–V loop) surrounding the active site (Supplementary Figure S5). Clear electron density is apparent for four of the eight ssRNA nucleotides (Figure 2B), which bind in a large positively charged groove extending from the active site (Figure 3D). Compared to the structures without substrate, the  $\beta$ IV–V loop and NRL2 are folded to enclose the substrate at the active site (Figure 3D and Supplementary Figure S5).

At the ALKBH5 active site in the structures reported here, a catalytically inert Mn ion (substituting for catalytically active Fe(II)) is coordinated by the highly conserved metal binding triad His204, Asp206, and His266, a water molecule, and either 2OG, NOG, or a sulfate ion (Figure 3A–C, Supplementary Figure S6A and B). In the structure of the ALKBH5<sub>74–292</sub>–RNA complex without 2OG or NOG, a sulfate ion (likely derived from the crystallization buffer) was observed to ligate to the Mn ion at the 2OG cosubstrate C-1 carboxylate binding site via one of its oxygens; the sulfate ion is also positioned to interact with Asn193 and Lys132 (Figure 3C). The observation of the Mn-ligating sulfate was unexpected because NOG was present in the initial complexation buffer. In the structure of the ALKBH5<sub>74–292</sub>–2OG–RNA complex, the 2OG binding mode is the same as previously reported (PDB IDs: 4OCT and 4NRO) (26,28) (Supplementary Figure S6C). The C-5 carboxylate group of 2OG is positioned to form a salt bridge with the side chain of Arg277. The 2OG ketone oxygen and one of the 2OG C-1 carboxylate oxygens chelate to the metal in a bidentate manner and form a hydrogen bond and electrostatic interactions with the side chains of Asn193 and Arg283, respectively (Figure 3A and Supplementary Figure S7). In the structure of the ALKBH5<sub>74–292</sub>–NOG–RNA complex, the NOG adopts a highly similar binding mode to 2OG, although with slight variations in the conformation of its C-5 carboxylate group (Figures 3B and Supplementary Figure S8).







**Figure 3.** m<sup>6</sup>A ssRNA substrate recognition and 2OG cosubstrate binding by ALKBH5. Views of the active site of ALKBH5<sub>74-292</sub> with (A) 2OG (PDB ID: 7WKV), (B) NOG (PDB ID: 7WL0), and (C) sulfate ion (PDB ID: 7V4G) bound to the active site metal (Mn substituting for Fe). The OMIT electron density maps for (A) 2OG, (B) NOG and (C) sulfate ion ( $mF_o - DF_c$ , contoured to  $3.0 \sigma$ ) are in green mesh. (D) Electrostatic surface potential of the ALKBH5<sub>74-292</sub>-2OG-ssRNA complex (PDB ID: 7WKV) generated using PyMOL (The PyMOL Molecular Graphics System, Version 2.3.0, Schrödinger, LLC). RNA (sticks) is bound into a basic/positively charged cleft (blue). Acidic/negatively charged surface (red), neutral (white). Views of the RNA substrate binding site of structures of (E) the ALKBH5<sub>74-292</sub>-2OG-ssRNA complex (PDB ID: 7WKV) and (F) the ALKBH5<sub>74-292</sub>-NOG-ssRNA complex (PDB ID: 7WL0). The (A/G)m<sup>6</sup>AC consensus motif is recognized by Phe234, Tyr139, His204 and Val205. RNA-interacting residues are in a red font. Colors: ALKBH5 secondary structure and residues, light blue cartoons/sticks; RNA C, yellow sticks; 2OG C, slate sticks; NOG C, green sticks; O, red; N, blue; P, orange; Mn, purple sphere; water molecule, red sphere; hydrogen bonding/electrostatic interactions, black dashes. Note:  $mF_o - DF_c$  electron density map is a Sigma-A weighted variant of the  $F_o - F_c$  difference electron density map.

Recent NMR studies of the ALKBH5 2OG binding mode to an active site Zn(II) ion, coupled with molecular dynamics (MD) studies suggest a mixture of ‘off-line’ and ‘on-line’ 2OG C-1 carboxylate binding modes (36). The crystallographically observed 2OG and NOG binding modes in ALKBH5 when complexed with Mn and an m<sup>6</sup>A substrate are of the ‘off-line’ mode. However, our crystallographic data may represent a snapshot of the complex in the catalytic process, where rearrangement of the 2OG C-1 carboxylate could occur. Another difference between our crystallographic studies and the reported molecular dynamics studies (36) is that the modeled binding mode of the ssDNA substrate to ALKBH5 in the MD simulations was based on a structure of the ALKBH2–dsDNA complex (PDB ID: 3RZJ) (37), and thus the 5′ to 3′ direction of the modeled nucleic acid is reversed relative to that of the ssRNA substrate when bound to ALKBH5 as observed in our crystallographic studies (see later).

### Conformational changes of ALKBH5 upon substrate binding

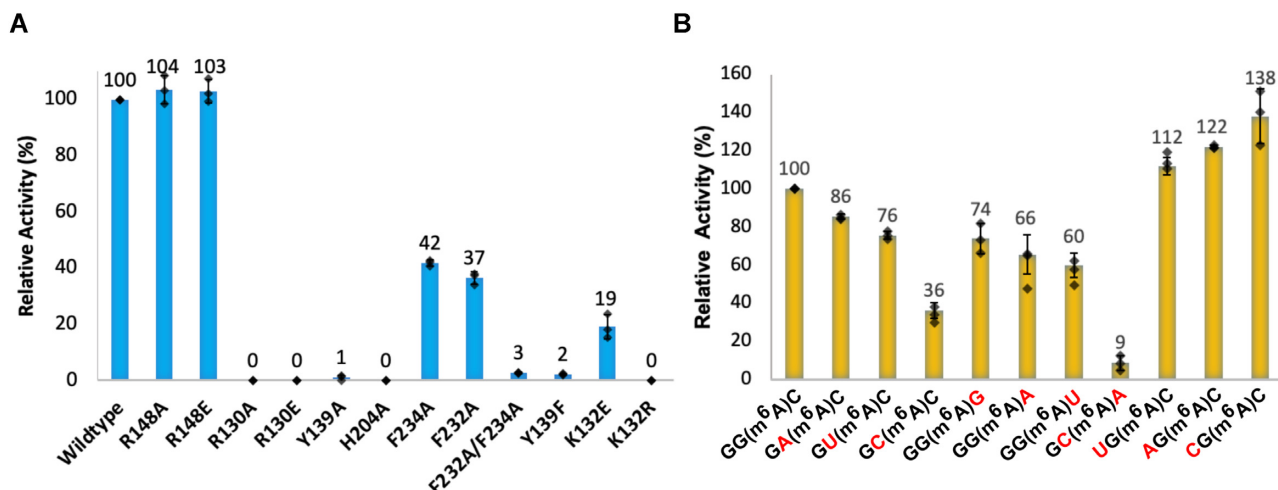
Comparison of the ALKBH5 structures with and without ssRNA reveals major conformational changes, implying induced fit upon substrate binding. The conformational changes primarily occur in the NRL2 and the βIV–V loop bordering the active site (Supplementary Figure S5). In contrast with structures of ALKBH5 without substrate, where NRL2 is partially disordered (25–28), when a substrate is bound the entire NRL2 loop is well ordered. NRL2 residues interact with the ssRNA substrate on the 3′ side of the m<sup>6</sup>A. The β-hairpin of NRL2 observed in the absence of substrate (25) is distorted upon substrate binding, forming a <sub>310</sub> helix followed by random coil (Figure 2B and Supplementary Figure S5). The βIV–V loop makes contact with the RNA substrate oligomer on the 5′ side of the m<sup>6</sup>A; the <sub>310</sub> helix observed in the βIV–V loop without substrate is transformed to a β-hairpin on substrate binding (Figures 2B and Supplementary Figure S5) (25,26,28). Secondary structure changes of NRL2 and the βIV–V loop as well as rigidification of NRL2 was also observed in the recently reported NMR studies by Purslow *et al.* (36), consistent with our crystallographic observations. The m<sup>6</sup>A-containing ssRNA binds across the active site opening and is enclosed by NRL1, NRL2 and the βIV–V loop (Figure 2B and Supplementary Figure S5). Our structural analyses support the reported chemical shift perturbations by NMR for residues on NRL1, NRL2 and the βIV–V loop (36). These three structural elements are shaped like ‘two-fingers and a thumb’, that collectively grip the ssRNA. The conformational changes resulting in the substrate being near entirely enclosed at the active site region are reminiscent of analogous changes observed on substrate binding with other 2OG dependent oxygenases, i.e. human HIF prolyl hydroxylase domain 2 (PHD2), *Pseudomonas* PHD (PPHD), *Trichoplax adhaerens* PHD (TaPHD), and human γ-butyrobetaine hydroxylase 1 (BBOX1) (38–41). Electrostatic surface potential analysis of ALKBH5<sub>74–292</sub> shows that most of its substrate-binding cleft is positively charged, so enabling ionic interactions with negatively charged oligonucleotide substrates (Figure 3D).

### ALKBH5 prefers m<sup>6</sup>A consensus motif (A/G)m<sup>6</sup>AC

To explore how ALKBH5 selectively binds its proposed (A/G)m<sup>6</sup>AC substrate consensus RNA motif (42,43), we systematically investigated interactions between each observed nucleotide base and the protein in the ALKBH5-substrate structural complexes. The m<sup>6</sup>A inserts into the active site with the 5′-phosphate of (m<sup>6</sup>A)<sub>0</sub> interacting with Arg130 from NRL1 (average distance 3.0 Å) (Figures 3E and F). The side chain of Arg148 is generally disordered, but in a number of protein molecules in the ALKBH5<sub>74–292</sub>–2OG–RNA (Chains A and C), ALKBH5<sub>74–292</sub>–NOG–RNA (Chain C) and ALKBH5<sub>74–292</sub>–RNA (Chains A, B and C) complexes, the observed side chain electron density for Arg148 suggests that it may interact with the C<sub>+1</sub> 5′-phosphate group (average distance 3.3 Å) in solution (Supplementary Figure S9). The 2′-OH group of the m<sup>6</sup>A ribose forms a hydrogen bond with the main chain carbonyl of Ser203 (average distance 2.8 Å) (Figure 3E and F). The m<sup>6</sup>A nucleobase likely forms a π–π interaction with the imidazole ring of His204, and the N<sup>6</sup>-amine is positioned within hydrogen bonding distance to the side chain OH of Tyr139 (average distance 3.2 Å) although the geometry in the crystal structure is not ideal for strong hydrogen bonding; we cannot rule out small adjustments of this interaction during catalysis in solution (Figure 3E and F). The N<sup>6</sup>-methyl group of m<sup>6</sup>A is positioned towards a hydrophobic region formed by the side chains of Phe211, Tyr139, and, to a lesser extent, Pro207; these residues also project towards the Watson–Crick edge of the m<sup>6</sup>A base (Figure 3E and F). This arrangement appears to help position the m<sup>6</sup>A methyl group close to the metal ion (Figure 3E and F). The R130A, R130E and H204A variants were inactive while the Y139A variant had little or no activity (~1% of wild type) (Figure 4A), validating the proposals based on structural observations that these are key residues in substrate binding and/or catalysis. However, the NRL2 loop variants R148A and R148E show similar activities to the wild type (Figure 4A), implying the potential interaction between the Arg148 side chain and the C<sub>+1</sub> phosphoryl group is not critical.

G<sub>–1</sub> is positioned ~180° about the phosphoribosyl backbone relative to m<sup>6</sup>A, and one face of its six-membered ring is positioned to make π–π interactions with the side chain of Phe234 on the βIV–V loop (Figure 3E and F). This observation supports the proposal that ALKBH5 selectively binds the larger bicyclic purine nucleobases (G or A) over smaller monocyclic pyrimidines (C or U) at its –1 binding position, potentially due to more effective π–π overlap with the Phe234 side chain. To test this hypothesis, we investigated the activities of purified ALKBH5<sub>74–292</sub> with different sequence substitutions (either A, C or U) at position –1. The results show that ALKBH5<sub>74–292</sub> has higher activity with G<sub>–1</sub> (100%) and A<sub>–1</sub> (86 ± 1%) compared to U<sub>–1</sub> (76 ± 2%) and C<sub>–1</sub> (36 ± 4%) (Figure 4B), in support of the proposed preference for purines at the –1 position. The F234A and F232A variants both have reduced, but substantial activity (42 ± 1 and 37 ± 2%, respectively of wild type). However, the double F232A/F234A variant manifests severely impaired activity (<3% of wild type) (Figure 4A). Thus, F234 is important in substrate recognition while F232 may also play a role in substrate recognition.





**Figure 4.** Activities of wild type and variant ALKBH5<sub>74-292</sub> and with different ssRNA substrates. (A) Relative activities of wild type ALKBH5<sub>74-292</sub> and variants of ALKBH5<sub>74-292</sub>. (B) Relative activities of wild type ALKBH5<sub>74-292</sub> with m<sup>6</sup>A-containing 8-mer ssRNA sequence variants at positions -2, -1 and +1; m<sup>6</sup>A is at position 0. Nucleotides that are varied from original sequence used in crystallographic studies (...G<sub>-2</sub>G<sub>-1</sub>(m<sup>6</sup>A)<sub>0</sub>C<sub>+1</sub>...) are in red. Heights of the bar graphs indicate the mean ( $n = 3$ ); black diamond markers indicate individual data points; error bars indicate standard deviation of the mean ( $n = 3$ ).

The cytosine base at the +1 position is adjacent to the m<sup>6</sup>A nucleobase at the active site opening. Importantly, a pair of hydrogen bonds are formed between the cytosine N3 and N<sup>4</sup>-amine atoms and the main chain amide NH and carbonyl O atoms of Val205, respectively (average distances 3.1 and 3.0 Å, respectively) (Figure 3E and F). The interactions that C<sub>+1</sub> makes with ALKBH5 imply that a bulkier purine base (A or G) at this position might require alteration of the RNA backbone conformation due to steric hindrance with the main chain at Val205. A uracil base at the +1 position will result in the loss of a hydrogen bond, because the 4-carbonyl group of uracil is unable to form a hydrogen bond with the main chain carbonyl of Val205, so explaining why a cytosine is preferred at the +1 position (consistent with the (A/G)m<sup>6</sup>AC consensus motif for ALKBH5) (42,43). Indeed, our *in vitro* activity assays show that ALKBH5 has reduced activity with RNA sequences with G, A or U at the +1 position (74 ± 8%, 66 ± 10% and 60 ± 7% respectively) relative to sequences containing C at the +1 position (100%) (Figure 4B), in accord with structural observations suggesting a preference for the (A/G)m<sup>6</sup>AC consensus motif by ALKBH5. Strikingly, when we tested the activity of an 8-mer ssRNA substrate containing both a pyrimidine (C<sub>-1</sub>) at -1 position and a purine (A<sub>+1</sub>) at +1 position (...C<sub>-1</sub>(m<sup>6</sup>A)<sub>0</sub>A<sub>+1</sub>...), the activity of ALKBH5 was severely reduced (9 ± 4%) (Figure 4B), supporting the proposal that the identity of the nucleobases directly adjacent to m<sup>6</sup>A can have a dramatic effect on activity.

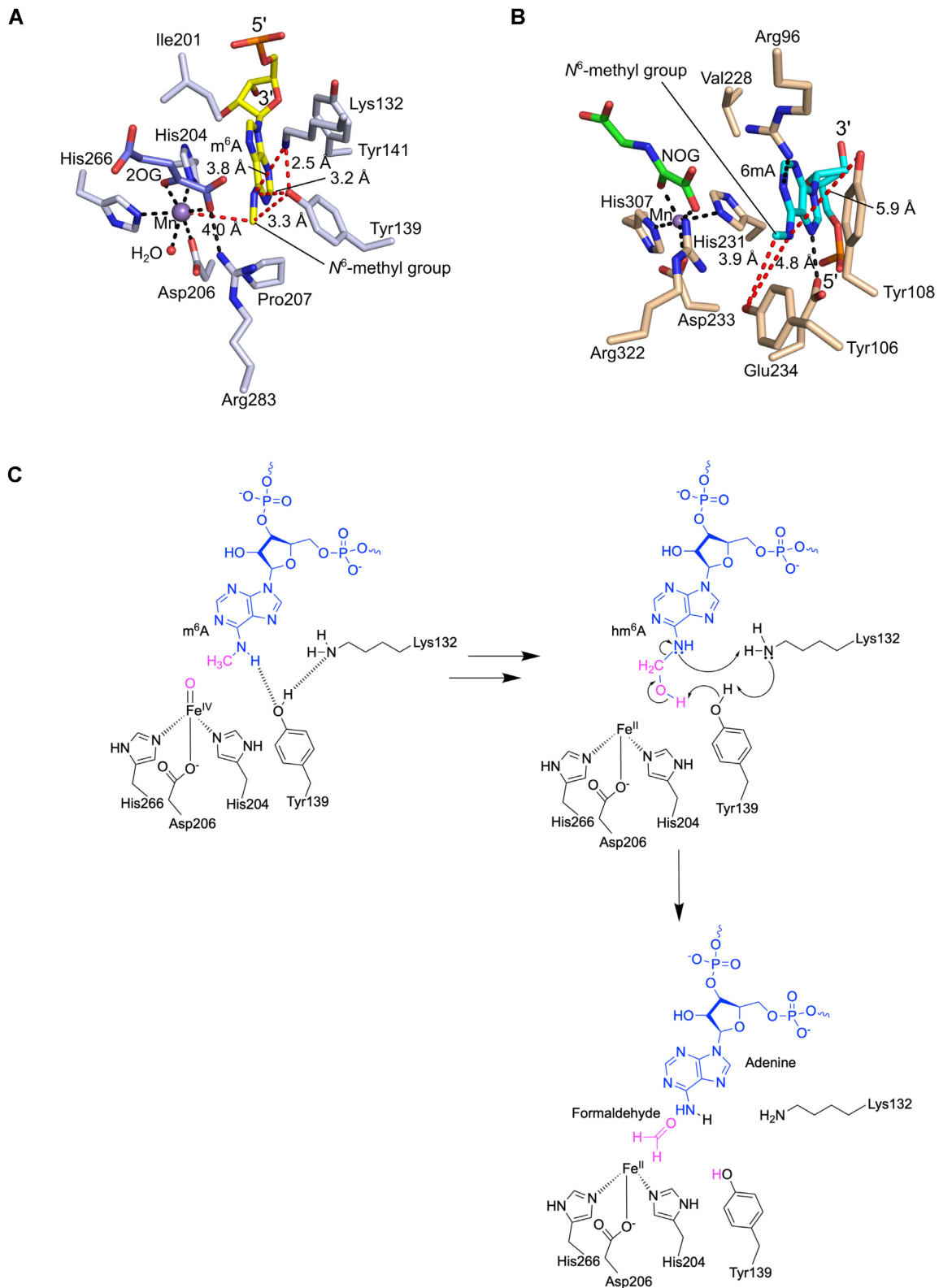
Recognition of the G nucleobase at the -2 position appears to be less selective, as it does not form direct interactions with ALKBH5 (Supplementary Figure S10), consistent with a m<sup>6</sup>A consensus motif (A/G/U)(A/G)m<sup>6</sup>AC(42, 43). We tested ALKBH5 activity using RNA substrate sequences with U, A or C at the -2 position; these sequences all resulted in higher activity (112 ± 4%, 122 ± 1% and 138 ± 14%, respectively) than the substrate with a G at -2.

Analysis of the protein structures does not provide an obvious rationale for these higher activities with nucleobases other than G at the -2 position, which might in part reflect intramolecular ssRNA interactions. Further studies employing modeling are of interest to explore this observation.

Overall, our structural and biochemical data suggest that ALKBH5 prefers substrate sequences with the (A/G)m<sup>6</sup>AC motif, consistent with the prevalence of m<sup>6</sup>A in biological contexts, including DRACH motifs, where A is the potentially N6-methylated adenosine (D is A, G or U; R is A or G; and H is A, C or U) (42,43). In cells there is evidence that DRACH motifs are substrates for the RNA m<sup>6</sup>A methyltransferases METLL3/METTL14 complex and subsequently ALKBH5 (44). Our work supports the proposed action of ALKBH5 on DRACH motifs. Interestingly, however, with FTO no discernible consensus sequence preference for the tested substrates was identified (Supplementary Figure S11). Although further detailed *in vitro* and *in vivo* studies are required, these observations suggest that ALKBH5 and FTO might target m<sup>6</sup>A in different sequence (or RNA secondary structure) contexts, consistent with their different biological roles.

#### The RNA phosphoribosyl backbone binding direction is reversed in ALKBH5 compared to substrates of several AlkB subfamily members

To further investigate how ALKBH5 confers substrate selectivity, we compared the ALKBH5-ssRNA structures with the available structures of nucleic acid oxygenases (NAOX) in complex with their substrates. The results reveal a striking difference in the 5'-3' phosphoribosyl backbone binding direction for ALKBH5 compared to other NAOX substrate structural complexes, *i.e.* that the phosphoribosyl backbone of the ssRNA binds to ALKBH5 in an opposite direction to that observed in structures of: (i) an



**Figure 5.** Proposed role for a proton shuttle in ALKBH5-catalyzed m<sup>6</sup>A hydroxylation followed by demethylation. **(A)** View of the active site of ALKBH5<sub>574-292</sub>-2OG-m<sup>6</sup>A RNA complex (PDB ID: 7WKV; Chain A). Colors: ALKBH5 residue C (light blue sticks), m<sup>6</sup>A C (yellow), 2OG C (slate), O (red), N (blue), P (orange), Mn (purple sphere), water molecule (red sphere), interactions (black dashes), distances (red dashes). **(B)** View of the active site of FTO variant-6mA ssDNA complex (PDB ID: 5ZMD) (22). Colors: FTO residue C (light brown), 6mA C (cyan), NOG C (green), O (red), N (blue), P (orange), Mn (purple sphere), interactions (black dashes), distances (red dashes). Note: the binding orientation of the face of 6mA nucleotide is inverted in FTO relative to that of m<sup>6</sup>A in ALKBH5 reflecting their opposing 5' to 3' nucleotide directions. **(C)** Proposed outline mechanism for demethylation of m<sup>6</sup>A catalyzed by ALKBH5. Note the proposed roles of Lys132 and Tyr139 in ALKBH5 catalysed m<sup>6</sup>A demethylation (17) and that for FTO directly analogous residues appear to be absent explaining why it does not (efficiently) catalyze demethylation.

FTO variant in complex with ssDNA (PDB ID: 5ZMD); (ii) ALKBH2 in complex with dsDNA (PDB ID: 3BTY); (iii) AlkB in complex with ssDNA or dsDNA (PDB IDs: 2FD8 and 3BIE) (Supplementary Figures S12 and S13) (21–23). Furthermore, the facial orientation of the m<sup>6</sup>A nucleobase in ALKBH5 is opposite to that observed in FTO (Supplementary Figure S13), an observation that likely impacts on catalysis and substrate selectivity/type of product (see below).

### Demethylation vs hydroxylation reaction mechanisms

We next focused on comparing the ALKBH5 and FTO active sites as both oxidize m<sup>6</sup>A in a manner ultimately leading to demethylation, though recent studies have shown that FTO generates hm<sup>6</sup>A as its major nascent product, at least in the tested sequence contexts (14,17). We were particularly interested in investigating the potential consequences on demethylation of the inverted 5' to 3' ssRNA substrate direction through the ALKBH5 active site cleft compared with that of FTO. In the case of FTO, in addition to a structure with a N<sup>3</sup>-methylthymidine nucleoside in the active site (45), a structure of an FTO variant with N<sup>6</sup>-methyldeoxyadenosine (6mA) bearing ssDNA is reported (Figures 5A and B, and Supplementary Figure S9) (22). Wild type FTO and the variant used for structural studies have been shown to be active against 6mA in ssDNA, although the available evidence is that mRNA is likely to be a more physiologically relevant substrate (16,22).

In the ALKBH5<sub>74–292</sub>-2OG-substrate complex, the N<sup>6</sup>-methyl group of m<sup>6</sup>A is positioned ~4.0 Å away from the Mn (Figure 5A), a substrate-metal distance consistent with a (near) catalytically productive orientation as observed in other structures of 2OG oxygenase enzyme-substrate complexes (46). The phenolic hydroxyl of Tyr139 is positioned to form a hydrogen bond with N6 of m<sup>6</sup>A, and with Lys132. As proposed (17), the proximity of N<sup>ε</sup> of Lys132 N to Tyr139 OH (average distance: 2.9 Å) and the m<sup>6</sup>A N6 (average distance 3.9 Å), and Tyr139 OH to the m<sup>6</sup>A N<sup>6</sup>-methyl (average distance: 3.3 Å) enables efficient demethylation following initial hydroxylation to give hm<sup>6</sup>A (Figure 5C). Upon hydroxylation of m<sup>6</sup>A, the side chain hydroxyl group of Tyr139 may participate in a proton shuttle by donating a proton to the side chain amine of Lys132 while receiving a proton from the hydroxyl group of hm<sup>6</sup>A. The deprotonated hm<sup>6</sup>A can then undergo fragmentation forming formaldehyde and an adenosine. During fragmentation, the N6 of adenosine accepts a proton from the protonated amine of Lys132, completing the proton transfer cycle. This mechanism allows m<sup>6</sup>A to be enzymatically hydroxylated and demethylated in quick succession, i.e. it accounts for the lack of hm<sup>6</sup>A detected in ALKBH5 assays.

Substitutions of the ALKBH5 active site residues Lys132 and Tyr139 were carried out to investigate their roles in hm<sup>6</sup>A fragmentation. Our results support the importance of these residues in overall catalysis as identified by Toh *et al.* (17). The K132R (to mimic the arginine present in FTO (Arg96<sub>FTO</sub>) – see below) and Y139A ALKBH5 variants were inactive and the K132E and Y139F variants manifested only 19 ± 4% and 2 ± 0% demethylation activities relative to wild type ALKBH5 (Figure 4A).

Comparison of the ALKBH5-ssRNA complex structure with that of an FTO variant in complex with N<sup>6</sup>-methyldeoxyadenosine (6mA) in ssDNA implies that residues in the active site of FTO are not positioned to participate in a proton shuttle in an analogous manner to that we propose for ALKBH5 (Figure 5B). While in FTO there are two tyrosine residues (Tyr106 and Tyr108) near the 6mA base, these residues are significantly farther away from the N6 of 6mA (4.8 Å for Tyr106 OH and 5.9 Å for Tyr108 OH) compared to the Tyr139 OH to N6 distance (average distance: 3.2 Å) in ALKBH5. Importantly, the inverted orientation of the face of the m<sup>6</sup>A nucleobase in ALKBH5 compared to FTO allows Arg96<sub>FTO</sub>, which occupies a similar position as Lys132 of ALKBH5, to form a hydrogen bond with N1 of m<sup>6</sup>A (3.0 Å), thus its ability to supply a proton to N6 of m<sup>6</sup>A during potential fragmentation to give formaldehyde is less likely (the arginine guanidino group is also less likely to act as a base due its potentially higher intrinsic pK<sub>a</sub> compared to the N<sup>ε</sup>-amino group of lysine). Consequently, FTO-catalyzed m<sup>6</sup>A oxidation produces hm<sup>6</sup>A-containing RNA as a major product, which is a relatively stable species at physiological pH (14,15,17).

### CONCLUSIONS

Three crystal structures of ALKBH5 in complex with its ssRNA substrate and co-substrate/co-substrate mimic reveal an unprecedented substrate binding mode for a 2OG oxygenase. Comparison of FTO and ALKBH5 substrate structural complexes provides a structural basis for the different type of products made by each of the two enzymes with m<sup>6</sup>A substrates, i.e. demethylation with formaldehyde formation (ALKBH5) or N-methyl hydroxylation (FTO). This knowledge should be useful in the development of selective ALKBH5 and FTO inhibitors, including those exploiting the different substrate binding modes of the two oxygenases. The results further illustrate the remarkably flexible nature of substrate (nucleic acid) binding by and reactivity of 2OG oxygenases, despite the presence of a conserved DSBH fold harboring closely related Fe(II) and 2OG binding elements. The results also provide structural information on why ALKBH5 prefers substrates with m<sup>6</sup>A located within the consensus (A/G)m<sup>6</sup>AC motif.

### DATA AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession numbers 7V4G, 7WKV, and 7WL0.

### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

### ACKNOWLEDGEMENTS

W.S.A. thanks the Research Grants Council of Hong Kong and Hong Kong Baptist University for funding. C.J.S. thanks the Biotechnology and Biological Research Council, the Wellcome Trust, and Cancer Research UK for funding. We thank the Diamond Light Source and staff for the



allocation of beam time and support, Judy Tsz-Shan Lum for technical support on MALDI-TOF MS, and Yuen Kit Cheng and Chi Sing Lee for helpful discussion. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

## FUNDING

Research Grants Council of Hong Kong Early Career Scheme 2019/20 [22301719]; Hong Kong Baptist University Tier 2 Start-Up Grant [RC-SGT2/18-19/SCI/003]; Wellcome Trust [106244/Z/14/Z]. Funding for open access charge: Hong Kong Baptist University.

*Conflict of interest statement.* None declared.

## REFERENCES

- Yang, Y., Hsu, P.J., Chen, Y.-S. and Yang, Y.-G. (2018) Dynamic transcriptomic m<sup>6</sup>A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res.*, **28**, 616–624.
- Thalhammer, A., Bencokova, Z., Poole, R., Loenarz, C., Adam, J., O'Flaherty, L., Schödel, J., Mole, D., Giaslaktiotis, 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 1 $\alpha$  (HIF-1 $\alpha$ ). *PLoS One*, **6**, e16210.
- Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C.J., Vågbo, 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.
- Wang, T., Kong, S., Tao, M. and Ju, S. (2020) The potential role of RNA N<sup>6</sup>-methyladenosine in cancer progression. *Mol. Cancer*, **19**, 88.
- Wu, F., Cheng, W., Zhao, F., Tang, M., Diao, Y. and Xu, R. (2019) Association of N<sup>6</sup>-methyladenosine with viruses and related diseases. *Viral J.*, **16**, 133.
- Zhang, S., Zhao, B.S., Zhou, A., Lin, K., Zheng, S., Lu, Z., Chen, Y., Sulman, E.P., Xie, K., Bögl, O. *et al.* (2017) m<sup>6</sup>A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell*, **31**, 591–606.
- Shen, C., Sheng, Y., Zhu, A.C., Robinson, S., Jiang, X., Dong, L., Chen, H., Su, R., Yin, Z., Li, W. *et al.* (2020) RNA demethylase ALKBH5 selectively promotes tumorigenesis and cancer stem cell self-renewal in acute myeloid leukemia. *Cell Stem Cell*, **27**, 64–80.
- Chen, Y., Zhao, Y., Chen, J., Peng, C., Zhang, Y., Tong, R., Cheng, Q., Yang, B., Feng, X., Lu, Y. *et al.* (2020) ALKBH5 suppresses malignancy of hepatocellular carcinoma via m<sup>6</sup>A-guided epigenetic inhibition of LYPD1. *Mol. Cancer*, **19**, 123.
- Yu, H., Yang, X., Tang, J., Si, S., Zhou, Z., Lu, J., Han, J., Yuan, B., Wu, Q., Lu, Q. *et al.* (2021) ALKBH5 inhibited cell proliferation and sensitized bladder cancer cells to cisplatin by m<sup>6</sup>A-CK2 $\alpha$ -Mediated glycolysis. *Mol. Ther. - Nucleic Acids*, **23**, 27–41.
- Guo, X., Li, K., Jiang, W., Hu, Y., Xiao, W., Huang, Y., Feng, Y., Pan, Q. and Wan, R. (2020) RNA demethylase ALKBH5 prevents pancreatic cancer progression by posttranscriptional activation of PER1 in an m<sup>6</sup>A-YTHDF2-dependent manner. *Mol. Cancer*, **19**, 91.
- Selberg, S., Seli, N., Kankuri, E. and Karelson, M. (2021) Rational design of novel anticancer small-molecule RNA m<sup>6</sup>A demethylase ALKBH5 inhibitors. *ACS Omega*, **6**, 13310–13320.
- Islam, M.S., Leissing, T.M., Chowdhury, R., Hopkinson, R.J. and Schofield, C.J. (2018) 2-Oxoglutarate-dependent oxygenases. *Annu. Rev. Biochem.*, **87**, 585–620.
- Hausinger, R.P. (2004) Fe(II)/ $\alpha$ -ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.*, **39**, 21–68.
- Shishodia, S., Zhang, D., El-Sagheer, A.H., Brown, T., Claridge, T.D.W., Schofield, C.J. and Hopkinson, R.J. (2018) NMR analyses on N-hydroxymethylated nucleobases – implications for formaldehyde toxicity and nucleic acid demethylases. *Org. Biomol. Chem.*, **16**, 4021–4032.
- 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 N<sup>6</sup>-hydroxymethyladenosine and N<sup>6</sup>-formyladenosine in mammalian RNA. *Nat. Commun.*, **4**, 1798.
- Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G. *et al.* (2011) N<sup>6</sup>-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.*, **7**, 885–887.
- Toh, J.D.W., Crossley, S.W.M., Bruemmer, K.J., Ge, E.J., He, D., Iovan, D.A. and Chang, C.J. (2020) Distinct RNA N-demethylation pathways catalyzed by nonheme iron ALKBH5 and FTO enzymes enable regulation of formaldehyde release rates. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 25284–25292.
- Mauer, J., Sindelar, M., Despic, V., Guez, T., Hawley, B.R., Vasseur, J.-J., Rentmeister, A., Gross, S.S., Pellizzoni, L., Debart, F. *et al.* (2019) FTO controls reversible m<sup>6</sup>Am RNA methylation during snRNA biogenesis. *Nat. Chem. Biol.*, **15**, 340–347.
- Hashimoto, H., Pais, J.E., Zhang, X., Saleh, L., Fu, Z.-Q., Dai, N., Corrêa, I.R., Zheng, Y. and Cheng, X. (2014) Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature*, **506**, 391–395.
- 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.
- 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.
- Zhang, X., Wei, L.-H., Wang, Y., Xiao, Y., Liu, J., Zhang, W., Yan, N., Amu, G., Tang, X., Zhang, L. *et al.* (2019) Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 2919–2924.
- 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.
- Holland, P.J. and Hollis, T. (2010) Structural and mutational analysis of escherichia coli AlkB provides insight into substrate specificity and DNA damage searching. *PLoS One*, **5**, e8680.
- Aik, W., Scotti, J.S., Choi, H., Gong, L., Demetriades, M., Schofield, C.J. and McDonough, M.A. (2014) Structure of human RNA N<sup>6</sup>-methyladenine demethylase ALKBH5 provides insights into its mechanisms of nucleic acid recognition and demethylation. *Nucleic Acids Res.*, **42**, 4741–4754.
- 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.
- 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.
- 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 N<sup>6</sup>-methyladenosine RNA demethylation. *J. Biol. Chem.*, **289**, 17299–17311.
- Savitsky, P., Bray, J., Cooper, C.D.O., 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.
- Erijman, A., Dantes, A., Bernheim, R., Shifman, J.M. and Peleg, Y. (2011) Transfer-PCR (TPCR): a highway for DNA cloning and protein engineering. *J. Struct. Biol.*, **175**, 171–177.
- Aik, W., Demetriades, M., Hamdan, M.K.K., Bagg, E.A.L., 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.
- Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J. *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. Sect. D*, **75**, 861–877.
- Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.*, **43**, 186–190.

34. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
35. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D*, **66**, 486–501.
36. Purslow, J.A., Nguyen, T.T., Balabhadra, K., Aayushi, S. and Vincenzo, V. (2022) N6-methyladenosine binding induces a metal-centered rearrangement that activates the human RNA demethylase alkhh5. *Sci. Adv.*, **7**, eabi8215.
37. Yi, C., Chen, B., Qi, B., Zhang, W., Jia, G., Zhang, L., Li, C.J., Dinner, A.R., Yang, C.-G. and He, C. (2012) Duplex interrogation by a direct DNA repair protein in search of base damage. *Nat. Struct. Mol. Biol.*, **19**, 671–676.
38. Lippl, K., Boleininger, A., McDonough, M.A., Abboud, M.I., Tarhonskaya, H., Chowdhury, R., Loenarz, C. and Schofield, C.J. (2018) Born to sense: biophysical analyses of the oxygen sensing prolyl hydroxylase from the simplest animal *Trichoplax adhaerens*. *Hypoxia (Auckland, N.Z.)*, **6**, 57–71.
39. Leung, I.K.H., Krojer, T.J., Kochan, G.T., Henry, L., von Delft, F., Claridge, T.D.W., Oppermann, U., McDonough, M.A. and Schofield, C.J. (2010) Structural and mechanistic studies on  $\gamma$ -butyrobetaine hydroxylase. *Chem. Biol.*, **17**, 1316–1324.
40. Chowdhury, R., McDonough, M.A., Mecinović, J., Loenarz, C., Flashman, E., Hewitson, K.S., Domene, C. and Schofield, C.J. (2009) Structural basis for binding of hypoxia-inducible factor to the oxygen-sensing prolyl hydroxylases. *Structure*, **17**, 981–989.
41. Scotti, J.S., Leung, I.K.H., Ge, W., Bentley, M.A., Paps, J., Kramer, H.B., Lee, J., Aik, W., Choi, H., Paulsen, S.M. *et al.* (2014) Human oxygen sensing may have origins in prokaryotic elongation factor Tu prolyl-hydroxylation. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 13331–13336.
42. 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.
43. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M. *et al.* (2012) Topology of the human and mouse m<sup>6</sup>A RNA methylomes revealed by m<sup>6</sup>A-seq. *Nature*, **485**, 201–206.
44. Lorenz, D.A., Sathe, S., Einstein, J.M. and Yeo, G.W. (2020) Direct RNA sequencing enables m<sup>6</sup>A detection in endogenous transcript isoforms at base-specific resolution. *RNA*, **26**, 19–28.
45. 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.
46. Aik, W.S., Chowdhury, R., Clifton, I.J., Hopkinson, R.J., Leissing, T., McDonough, M.A., Nowak, R., Schofield, C.J. and Walport, L.J. (2015) Introduction to structural studies on 2-oxoglutarate-dependent oxygenases and related enzymes. In: *RSC Metallobiology*. pp. 59–94.
47. Gille, C. and Frömmel, C. (2001) STRAP: editor for STRuctural alignments of proteins. *Bioinformatics*, **17**, 377–378.
48. Gille, C., Föhling, M., Weyand, B., Wieland, T. and Gille, A. (2014) Alignment-Annotator web server: rendering and annotating sequence alignments. *Nucleic Acids Res.*, **42**, W3–W6.
49. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320–W324.