THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 17, pp. 11571–11583, April 25, 2014


# Crystal Structures of the Human RNA Demethylase Alkbh5 Reveal Basis for Substrate Recognition*[S]

Received for publication, January 1, 2014, and in revised form, March 7, 2014 Published, JBC Papers in Press, March 10, 2014, DOI 10.1074/jbc.M113.546168

**Chong Feng[‡], Yang Liu[‡], Guoqiang Wang[‡], Zengqin Deng[‡], Qi Zhang[‡], Wei Wu[‡], Yufeng Tong[§¶], Changmei Cheng[‖1], and Zhongzhou Chen[‡2]**

*From the [‡]State Key Laboratory of Agrobiotechnology, China Agricultural University, Beijing 100193, China , the [‖]Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology, Ministry of Education, Department of Chemistry, Tsinghua University, Beijing 100084, China, the [§]Structural Genomics Consortium and the [¶]Department of Pharmacology and Toxicology, University of Toronto, 101 College Street, Toronto, Ontario M5G 1L7, Canada*

**Background:** Human AlkB homolog 5 (Alkbh5) is an RNA demethylase that erases m$^6$A modification.
**Results:** Crystal structures of an enzymatically active Alkbh5 construct in complex with cofactors or small molecules were determined.
**Conclusion:** Structure and activity analyses showed that Alkbh5 strongly prefers single-stranded oligos and small molecule inhibitors.
**Significance:** The Alkbh5 structure reveals potential for structure-based design of selective inhibitors.

**$N^6$-Methylation of adenosine is the most ubiquitous and abundant modification of nucleoside in eukaryotic mRNA and long non-coding RNA. This modification plays an essential role in the regulation of mRNA translation and RNA metabolism. Recently, human AlkB homolog 5 (Alkbh5) and fat mass- and obesity-associated protein (FTO) were shown to erase this methyl modification on mRNA. Here, we report five high resolution crystal structures of the catalytic core of Alkbh5 in complex with different ligands. Compared with other AlkB proteins, Alkbh5 displays several unique structural features on top of the conserved double-stranded $\beta$-helix fold typical of this protein family. Among the unique features, a distinct "lid" region of Alkbh5 plays a vital role in substrate recognition and catalysis. An unexpected disulfide bond between Cys-230 and Cys-267 is crucial for the selective binding of Alkbh5 to single-stranded RNA/DNA by bringing a "flipping" motif toward the central $\beta$-helix fold. We generated a substrate binding model of Alkbh5 based on a demethylation activity assay of several structure-guided site-directed mutants. Crystallographic and biochemical studies using various analogs of $\alpha$-ketoglutarate revealed that the active site cavity of Alkbh5 is much smaller than that of FTO and preferentially binds small molecule inhibitors. Taken together, our findings provide a structural basis for understanding the substrate recognition specificity of Alkbh5 and offer a foundation for selective drug design against AlkB members.**

Over 100 post-transcriptional modifications have been identified in different types of RNA, including rRNA, tRNA, mRNA, and snRNA (1). $N^6$-Methyladenosine (m$^6$A)[3] is the most prevalent and abundant internal methylated nucleoside in mammalian mRNA and is also found in the RNA of plants and viruses (2, 3). Mapping of m$^6$A in human and mouse RNA identified that m$^6$A modification happens mainly within the consensus sequence (G/A)(G/A)m$^6$AC(A/C/U) in a non-stoichiometric manner with only three to five m$^6$A sites observed per mRNA molecule (4–6). m$^6$A sites are enriched near stop codons and in the 3′-UTR of mRNA (5, 6). As a result, scientific interest in the m$^6$A modification has increased. The m$^6$A modification is predicted to be involved in various pathways of RNA metabolism and to play a vital role in the regulation of gene expression (5, 6). Similar to DNA and protein modifications, the m$^6$A methylation is reversible and can be temporally and spatially regulated by methyltransferases and demethylases (7, 8).

Alkbh5 belongs to the conserved AlkB family of non-heme Fe(II)/$\alpha$-KG-dependent dioxygenases (9, 10) that repair *N*-alkylated nucleobases by oxidative demethylation. The *Escherichia coli* AlkB has been shown to restore 3-meC, 1-meA, 3-meT, 1-meG, and other lesions in DNA and RNA (11, 12). The human genome encodes nine AlkB homologs (Alkbh1–8 and FTO). Alkbh1 is a mitochondrial protein that demethylates 3-meC in both RNA and DNA (13). Alkbh2 and Alkbh3 share the same repair activities as AlkB, but Alkbh3 and AlkB prefer ssDNA and ssRNA substrates, whereas Alkbh2 prefers dsDNA (14–16). Alkbh8 catalyzes the hydroxylation of hypermodified wobble uridines in tRNA that are associated with bladder cancer invasion (17). FTO was initially shown to demethylate 3-methylthymine and 3-methyluracil in synthetic ssDNA and

* This work was supported by the National Basic Research Program of China (973 Program Grants 2011CB965304 and 2009CB825501), National Natural Science Foundation of China (Grants 31370720, 31222032, 90919043, and 31070664), Specialized Research Fund for the Doctoral Program of Higher Education (Grant 20100008110009), and National Laboratory of Medical Molecular Biology (Peking Union Medical College) (to Z. C.).
[S] This article contains supplemental Fig. S1 and Movies S1 and S2.
*The atomic coordinates and structure factors (codes 4NRM, 4NRO, 4NRP, 4NRQ, and 4O7X) have been deposited in the Protein Data Bank (http://wwpdb.org/).*
[1] To whom correspondence may be addressed. Tel.: 86-10-62784642; Fax: 86-10-62784642; E-mail: chengcm@mail.tsinghua.edu.cn.
[2] To whom correspondence may be addressed. Tel.: 86-10-6273-4078; Fax: 86-10-6273-4078; E-mail: chenzhongzhou@cau.edu.cn.

[3] The abbreviations used are: m$^6$A, $N^6$-methyladenosine; Alkbh5, human AlkB homolog 5; FTO, fat mass- and obesity-associated protein; $\alpha$-KG, $\alpha$-ketoglutarate; ssRNA, single-stranded RNA; ssDNA, single-stranded DNA; NOG, *N*-oxalylglycine; PDCA, pyridine 2,4-dicarboxylate; ITC, isothermal titration calorimetry; DSBH, double-stranded $\beta$-helix; dsRNA, double-stranded RNA.

ssRNA (18, 19). Structural studies have confirmed that FTO selects against double-stranded nucleic acids (20). Furthermore, FTO was recently shown to exhibit high demethylation activity toward m$^6$A, confirming that ssRNA-containing m$^6$A is the preferred substrate of this enzyme (21).

*Alkbh5* is a direct transcriptional target of hypoxia inducible factor-1 and is induced by hypoxia in a range of cell types (22). The expression of Alkbh5 can be negatively regulated by PRMT7 (23). A proteome study revealed that Alkbh5 binds preferentially toward the distal 5′ regions of the coding sequences (24). Although Alkbh5 has been known for many years, the substrate and function of this enzyme remained elusive until recently. Zheng *et al.* (25) demonstrated that Alkbh5 is the second mammalian RNA demethylase that removes the m$^6$A modification both *in vitro* and *in vivo*. This study showed that Alkbh5 co-localizes with nuclear speckles, which are involved in the assembly of mRNA-processing factors. Alkbh5-mediated demethylation of mRNA significantly affects nuclear mRNA export and RNA metabolism. Alkbh5 demethylation activity was also shown to regulate spermatogenesis and apoptosis in mouse testes. The study also thoroughly characterized the enzyme activity of Alkbh5, which shows a preference for single-stranded oligos with a selective bias toward m$^6$A within the consensus sequence (25).

Given the striking biological role of Alkbh5 as an m$^6$A RNA demethylase, it is of considerable scientific interest to elucidate the molecular mechanism underlying Alkbh5 demethylation activity. To this end, we determined five crystal structures of the human Alkbh5 catalytic core in complex with citrate and acetate, $Mn^{2+}$, and $Mn^{2+}$ with ligand $\alpha$-KG, *N*-oxalylglycine (NOG), or pyridine 2,4-dicarboxylate (PDCA). Structure comparison with AlkB homologs and activity analysis of site-directed mutants were also carried out and provided insights into the distinct substrate preference of Alkbh5 and selective inhibitor design.

## EXPERIMENTAL PROCEDURES

*Protein Expression and Purification*—DNAs encoding the wild-type and various mutants of human *Alkbh5* were amplified by PCR and subcloned into a modified pET-28a (Novagen) vector encoding a tobacco etch virus protease recognition site. The final clones were verified by DNA sequencing. All of the recombinant plasmids were transformed into *E. coli* strain BL21(DE3). Cells were grown in LB at 37 °C until $A_{600}$ reached 0.8–1.0, and protein overexpression was induced overnight at 18 °C by addition of 0.2 mM isopropyl 1-thio-$\beta$-D-galactopyranoside. The cultures were harvested by centrifuging at 4000 × *g* for 10 min. Cell pellets were then resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 500 mM NaCl) replenished with Triton X-100 and PMSF (Invitrogen). The cells were lysed by sonication, and lysates were clarified by centrifuging at 20,000 × *g* for 15 min. The supernatant was then filtrated through a 0.45-$\mu$m filter membrane to remove cell debris before being applied to a His affinity column (GE Healthcare). After the sample was loaded, the column was washed with buffer B (buffer A containing 20 mM imidazole), and the target protein was eluted with buffer C (buffer A containing 300 mM imidazole). Tobacco etch virus protease was added to the elutant at a 1:10 (w/w, protease/protein) ratio for 3 h at 4 °C to remove the His tag. Uncleaved proteins and tobacco etch virus protease were removed by passing the elutant through a second $Ni^{2+}$-chelating column. The Alkbh5 sample was then loaded onto an ion exchange column (Q-Sepharose, GE Healthcare) eluted with buffer containing 20 mM Tris-HCl, pH 8.0 and 50 mM NaCl. Further purification was performed by size exclusion chromatography using a Superdex 200 column (GE Healthcare) with GF buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl). Peak fractions were collected and examined using Coomassie Blue-stained SDS-polyacrylamide gels. Fractions containing the target protein were pooled and concentrated to 8–40 mg/ml. All the protein purification procedures were performed at 4 °C.

*Crystallization and Data Collection*—Crystallizations were performed at 24 and 4 °C using both the hanging drop and sitting drop vapor diffusion methods. Initial crystals of Alkbh5 were grown in a reservoir solution of 30% PEG 4000, 200 mM $CH_3CO_2NH_4$, 100 mM sodium citrate, pH 5.6 at 24 °C. High quality crystals were finally obtained after 2 weeks at 24 °C in a hanging drop containing 1 $\mu$l of protein solution (150 mM NaCl, 20 mM Tris-HCl, pH 8.0, 5 mM $MnCl_2$, 10 mM $\beta$-mercaptoethanol, 10 mM $\alpha$-KG) mixed with an equal amount of optimized well solution (28% PEG 4000, 200 mM $CH_3CO_2NH_4$, 100 mM sodium citrate, pH 5.4, 2% (v/v) glycerol). The complexes of Alkbh5·$Mn^{2+}$, Alkbh5·$Mn^{2+}$·$\alpha$-KG, Alkbh5·$Mn^{2+}$·NOG, and Alkbh5·$Mn^{2+}$·PDCA were prepared by soaking the native crystals in well solution containing 5 mM $MnCl_2$ or additionally with 50 mM $\alpha$-KG, 5 mM NOG, or 5 mM PDCA, respectively. All the crystals were transferred to a cryobuffer (reservoir buffer supplemented with 20% ethylene glycol) and flash frozen in liquid nitrogen before data collection.

Native data were collected at beamline NE3A at Photon Factory. The data for complex crystals containing $Mn^{2+}$ and different ligands ($\alpha$-KG, NOG, and PDCA) were collected at beamline NE3A at Photon Factory, beamline BL17U1 at the Shanghai Synchrotron Radiation Facility, and beamline 1W2B at the Beijing Synchrotron Radiation Facility, respectively. Data were indexed, integrated, and scaled with the HKL2000 suite of programs (26). Data collection details and processing statistics are shown in Table 1.

*Structure Determination and Refinement*—An initial molecular replacement solution was obtained from the BALBES (27) server using the structure of Alkbh2·dsDNA (Protein Data Bank code 3BTZ; 19% identity to Alkbh5) as a model after extensive trials of high resolution ranges between 2.0 and 3.0 Å. The solution has a molecular replacement score of 3.45 and an $R_{work}/R_{free}$ value of 0.497/0.574. The primitive density map has a relatively clear outline with several recognizable $\alpha$-helices and $\beta$-strand bundles. After removing the inappropriate main and side chains in the model using the program Coot (28), we iteratively refined and built the model with CNS (29), Phaser (30), AutoBuild (31), and REFAMC5 (32). The best solution stopped at an $R_{work}/R_{free}$ value of 0.395/0.452. At this point, we attempted to use the IPCAS package from the CCP4 suite (33). The running mode was set as partial model extension without single wavelength anomalous diffraction/single isomorphous replacement information with 10 interactive cycles of phasing by OASIS (34), density modification by DM (35), and model

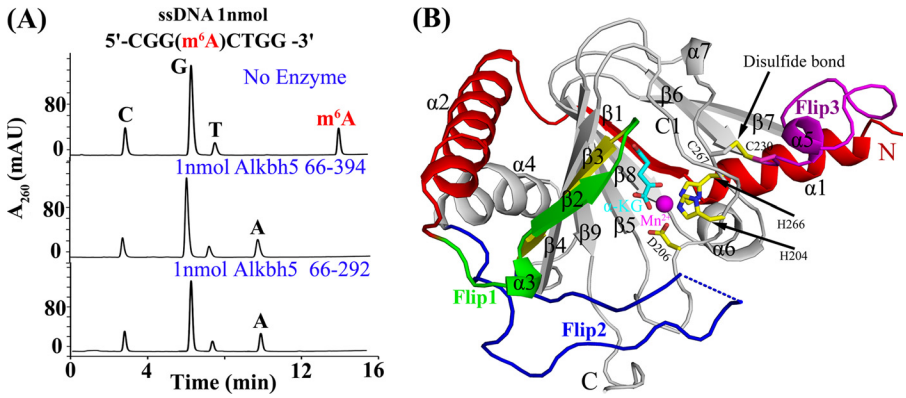


FIGURE 1. **Overall structure of human Alkbh5·α-KG·Mn$^{2+}$.** *A*, Alkbh5 *in vitro* enzymatic activity assays. As shown by the HPLC analysis of digested substrates, Alkbh5 66–292 used for crystallization shows the same repair activity as the longer fragment Alkbh5 66–394. *B*, the overall structures of Alkbh5 66–292 in the presence of α-KG and $Mn^{2+}$. The $Mn^{2+}$ ion is shown as a *magenta sphere*. The disulfide bond between Cys-230 and Cys-267 is highlighted in *yellow* and indicated by the *black arrow*. Key residues as well as α-KG are shown as *sticks*, and secondary structures are labeled. Flip1, Flip2, and Flip3 are shown in *green*, *blue*, and *magenta*, respectively. *mAU*, milli-absorbance units.

building by Buccaneer (36), all of which improved the $R_{work}/R_{free}$ value. The final cycle of refinement gave a qualified model with the $R_{work}/R_{free}$ values of 0.266/0.308. After five more cycles of manual rebuilding by Coot and refinement with REFAMC5, the structure was refined to 2.17 Å with an $R_{work}$ of 20.3% and $R_{free}$ of 23.2%. The crystal of Alkbh5 belongs to space group $P4_32_12$ with one molecule in the asymmetric unit. The complex structures of $Mn^{2+}$-bound Alkbh5 or additionally with different ligands (α-KG, NOG, and PDCA) were all determined by molecular replacement using the native structure as a model. All structural figures were prepared using the PyMOL program (37).

*Preparation of Substrates*—Two m$^6$A-containing single-stranded DNA oligonucleotides (8-mer, CGG(m$^6$A)CTGG; 10-mer, GTCA(m$^6$A)CAGCC) were chemically synthesized (GenScript). A hemimethylated 10-mer dsDNA was prepared by annealing a slight excess of complementary oligos to the above 10-mer m$^6$A-containing ssDNA in 10 mM Tris-HCl, pH 8.0, 100 mM NaCl buffer. Annealing was performed first by heating the mixture at 96 °C for 5 min and then slowly cooling it to room temperature. Additionally, the 6-methyldeoxyadenosine was synthesized according to a previous report (38).

*In Vitro Demethylation Activity Assays*—The demethylation activity assay was carried out in a 100-μl reaction mixture containing 10 μM m$^6$A-containing oligos, 10 μM human Alkbh5 wild type or mutant proteins, 50 mM HEPES, pH 7.2, 150 μM $(NH_4)_2Fe(SO_4)_2$, 300 μM α-KG, and 2 mM L-ascorbic acid. The sample was incubated at room temperature for 30 min before the reaction was quenched by heating at 96 °C for 5 min. We then digested the single-stranded oligonucleotides into single nucleosides using nuclease P1 (Sigma N8630) and alkaline phosphatase (Takara 2250A). The treated solution was analyzed in an HPLC system equipped with an Agilent Eclipse XDB-C18 analysis column (150 × 4.6 mm) eluted with buffer A ($H_2O$) and buffer B (methanol) with a flow rate of 0.5 ml/min at room temperature. The detection wavelength was set at 260 nm. When the dsDNA was used as substrate, the reaction mixture was heated at 100 °C for 5 min and then chilled rapidly in an ice bath. The denatured DNA was digested and analyzed using the same procedure as that for ssDNA.

*Inhibition Assays for Alkbh5*—For the inhibition assay, 5 μM Alkbh5 was incubated in a reaction buffer containing 50 mM HEPES, pH 7.2, 150 μM $(NH_4)_2Fe(SO_4)_2$, 2 mM L-ascorbic acid, and inhibitors at various concentrations for 30 min at 24 °C prior to the addition of 10 μM 8-mer m$^6$A-ssDNA and 160 μM α-KG. After incubation for 3 h, the reactions were quenched by heating at 96 °C for 5 min. Other procedures were the same as those in the demethylation activity assays. Most of the assays were repeated in triplicate. All inhibitors were assayed at five different concentrations. $IC_{50}$ values were calculated as the inhibitor concentration that reduces Alkbh5 activity to half its maximal level using GraphPad Prism 5.0 software.

*Isothermal Titration Calorimetry (ITC) Assays*—Alkbh5 protein was subjected to size exclusion chromatography with buffer D (20 mM Tris, pH 8.0, 150 mM NaCl). Both α-KG and succinate were dissolved in buffer D at 100 mM, and the pH of the stock solutions was adjusted to 8.0 using sodium hydroxide. The solutions were then diluted to a final concentration of 4 mM using buffer D. α-KG and succinate in the syringe were titrated into 0.22 and 0.38 mM Alkbh5 protein in the cell, respectively. The experiments were performed using a Nano-2G-ITC instrument (TA Corp.) at 4 °C, and the ITC data were processed with NanoAnalyze software.

## RESULTS

*Structure of the Enzymatically Active Alkbh5 Core Exhibits a Conserved Double-stranded β-Helix Fold*—Full-length human Alkbh5 and a series of truncated variants were overexpressed and purified from *E. coli*. After extensive crystallization trials, diffracting crystals could only be obtained for a truncated form of Alkbh5 (residues 66–292) lacking the N-terminal 65 residues and the C-terminal 103 residues. To verify whether this truncated fragment, which contains the entire predicted double-stranded β-helix (DSBH) fold, was enzymatically active, we performed an *in vitro* enzyme assay to compare the activity of various truncation variants of Alkbh5. In agreement with previously published results (25), fragment 66–394 displayed full demethylation activity toward ssDNA containing m$^6$A (Fig. 1*A*). The crystallizable fragment 66–292 also retained full demethylation activity. For simplicity, we hereafter refer to the

**TABLE 1**

**Data collection and refinement statistics of *h*Alkbh5 complexes***

There were three crystal experiments for each structure. r.m.s., root mean square.

| | Alkbh5·citrate·acetate | Alkbh5·$Mn^{2+}$ | Alkbh5·$Mn^{2+}$·α-KG | Alkbh5·$Mn^{2+}$·NOG | Alkbh5·$Mn^{2+}$·PDCA |
|---|---|---|---|---|---|
| **Data collection** | | | | | |
| Wavelength | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Space group | $P4_32_12$ | $P4_32_12$ | $P4_32_12$ | $P4_32_12$ | $P4_32_12$ |
| Cell dimensions | | | | | |
| *a*, *b*, *c* (Å) | 56.73, 56.73, 146.32 | 57.09, 57.09, 145.73 | 56.18, 56.18, 144.94 | 56.55, 56.55, 143.87 | 56.77, 56.77, 144.71 |
| α, β, γ (°) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution (Å)[a] | 50-2.17 (2.23-2.17) | 50-1.78 (1.83-1.78) | 50-2.30 (2.36-2.30) | 50-1.80 (1.83-1.80) | 50-2.50 (2.57-2.50) |
| $R_{merge}$ (%) | 7.7 (74.4) | 6.7 (70.1) | 9.9 (62.3) | 6.1 (67.9) | 9.7 (50.3) |
| $I/\sigma$ | 50.3 (2.4) | 54.5 (2.3) | 38.2 (4.9) | 53.8 (2.2) | 19.9 (3.1) |
| Completeness (%) | 99.3 (91.0) | 99.3 (100) | 90 (92.1) | 99.7 (99.9) | 99.4 (99.8) |
| Total no. of reflections | 264,645 | 258,635 | 89,316 | 466,373 | 61,371 |
| Unique reflections | 13,441 | 24,002 | 11,010 | 22,530 | 8,780 |
| Redundancy | 19.8 (6.2) | 10.9 (11.4) | 9.0 (9.3) | 20.8 (9.8) | 7.0 (7.2) |
| **Refinement** | | | | | |
| Resolution (Å) | 50-2.17 (2.22-2.17) | 50-1.78 (1.83-1.78) | 50-2.30 (2.36-2.30) | 50-1.8 (1.83-1.80) | 50-2.50 (2.57-2.50) |
| No. of reflections | 12,575 (756) | 22,569 (1,453) | 9,350 (486) | 20,719 (983) | 8,215 (477) |
| $R_{work}/R_{free}$ (%) | 20.3/23.2 | 21.4/23.7 | 23.7/25.6 | 21.8/19.6 | 22.3/24.9 |
| No. of atoms | | | | | |
| Protein | 1,636 | 1,639 | 1,572 | 1,664 | 1,608 |
| Ligand/ion | 17 | 1 | 11 | 11 | 13 |
| Water | 66 | 77 | 110 | 106 | 4 |
| *B*-factors (Å$^2$) | | | | | |
| Protein | 39.04 | 32.98 | 53.59 | 23.23 | 32.17 |
| Ligand/ion | 45.87 | 27.24 | 60.45 | 21.57 | 30.23 |
| Water | 43.6 | 35.71 | 61.76 | 28.95 | 31.71 |
| r.m.s. deviations | | | | | |
| Bond lengths (Å) | 0.005 | 0.006 | 0.007 | 0.005 | 0.005 |
| Bond angles (°) | 1.34 | 1.057 | 1.166 | 0.98 | 1.00 |
| Ramachandran plot ( %)[b] | 93.4/6.6/0/0 | 92.8/7.2/0/0 | 90.2/9.8/0/0 | 93.4/6.6/0/0 | 92.3/7.7/0/0 |

[a] Statistics for highest resolution shell.
[b] Residues in most favored, additional allowed, generously allowed, and disallowed regions of the Ramachandran plot.

latter fragment as Alkbh5 throughout the remainder of this study unless otherwise specified.

Blast searches of Alkbh5 sequence against the solved structures in the Protein Data Bank returned hits with low sequence homology of 16–26% identity. The Alkbh2 structure (Protein Data Bank code 3BTZ) was used for molecular replacement to get the initial model of Alkbh5. After extensively iterative manual model building and refinement, the native structure of Alkbh5 was refined to a final $R_{work}$ of 20.3% and $R_{free}$ of 23.2% at 2.17 Å (Table 1). Unexpectedly, the native structure contains citrate and acetate molecules from the reservoir solution (see below for details). The structures of $Mn^{2+}$-bound Alkbh5 ($Mn^{2+}$ is a substitute for $Fe^{2+}$) or additionally with cofactor α-KG or inhibitors NOG and PDCA were obtained at 1.78, 2.3, 1.8, and 2.5 Å, respectively (Table 1), using the structure of Alkhb5 as a model.

The Alkbh5 structure contains seven α-helices and nine β-sheets (Fig. 1*B* and supplemental Movie S1). The N-terminal seven residues could not be modeled due to a lack of electron density. Residues 143–149 in the loop between β3 and α4 also had no visible electron density and therefore could not be built (Fig. 1*B*, shown as *blue dashed lines*). The catalytic core of Alkbh5 also contains the DSBH fold (also known as the jelly roll fold) characteristic of the α-KG-dependent dioxygenase superfamily but without the canonical eight antiparallel β-strands: only six β-strands (β4 to β9) were observed in the fold with β4, β5, β8, and β9 forming the major sheet, whereas β6, β7, and a short α-helix (α7) plus a long loop (C1) formed the minor sheet (39). Helix α1 lies in the N terminus of the DSBH fold followed by β1 located on the same side of the major sheet, which is further buttressed by two helices, α2 and α4. In addition, the active site contains three residues, His-204, Asp-206 and His-266, that form the conserved H*X*(D/E)$X_n$H motif that coordinates the metal ion (Fig. 1*B*).

*The Unique Structural Features of Alkbh5*—As recognized extensively in the literature (39), it is the outer secondary structural elements of the common DSBH fold that contribute most toward the substrate selectivity of the AlkB family. The activity of Alkbh5 has been confined during evolution to demethylate m$^6$A exclusively without any activity toward other base lesions. The latter are targeted by the rest of the AlkB family members (22). Therefore, we compared the structure of Alkbh5 with those of other AlkB family members to uncover the basis of this selectivity.

Structural comparison revealed that the most significant structural difference between Alkbh5 and other AlkB proteins rests on the so-called "nucleotide recognition lid" outside the DSBH fold, which has been further divided into two sections named "Flip1" and "Flip2" (39) (Fig. 2*A*). Here, we took Alkbh2 and FTO for example to illustrate the unique structural elements of Alkbh5. The Flip1 regions of both Alkbh2 and FTO contain two antiparallel β-strands and one α-helix. However, the Flip1 region of Alkbh5 (residues 117–129) contains one α-helix (α3) and one β-strand (β2) (supplemental Fig. S1), exposing an uncovered and relatively large space over the active site (Fig. 2*A*). In addition, the Flip2 region of Alkbh5 contains a long loop, whereas those of Alkbh2 and FTO contain a shorter loop and two additional antiparallel β-sheets (Fig. 2*A*). The Alkbh5 Flip2 region (residues 136–165) is also highly flexible with higher *B*-factors (supplemental Movie S2), especially

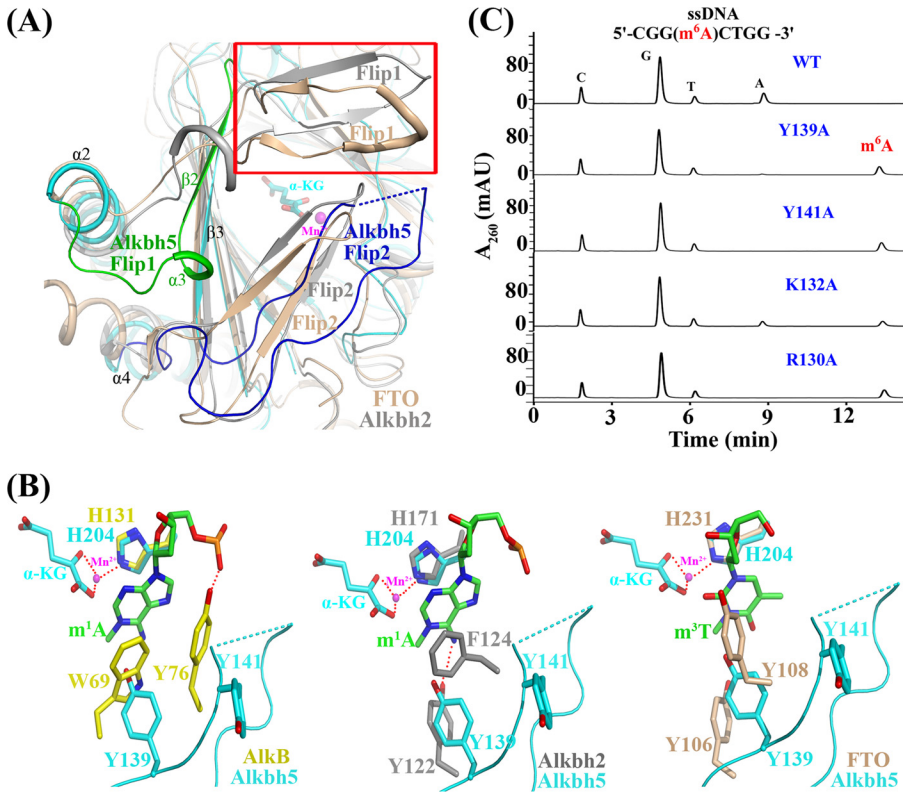


FIGURE 2. **The distinct lid region of Alkbh5 is vital for its substrate recognition and catalysis.** *A*, structural comparison of Alkbh5 with FTO (*tinted*; Protein Data Bank code 3LFM) and Alkbh2 (*gray*; Protein Data Bank code 3BTZ) around the lid region. The most distinct regions, Flip1 (residues 117–129) and Flip2 (residues 136–165), of Alkbh5 are shown in *green* and *blue*, respectively. The discrete region is shown in *blue dashed lines*. The uncovered and larger space over the active site of Alkbh5 is labeled in the *red frame*. *B*, superposition of the partial key residues involved in nucleoside recognition. Alkbh5 (*cyan*), AlkB (Protein Data Bank code 3BIE; *yellow*; *left panel*), Alkbh2 (Protein Data Bank code 3BTZ; *gray*; *middle panel*), and FTO (Protein Data Bank code 3LFM; *tinted*; *right panel*) are shown. The nucleoside is colored *green*. Hydrogen bonds are indicated as *red dashed lines*. α-KG and $Mn^{2+}$ from Alkbh5 are labeled. *C*, mutations of the key residues in the lid region (shown in supplemental Movie S1) greatly impair Alkbh5 demethylation activity. *mAU*, milli-absorbance units.

within the sections adjacent to the missing residues 143–149 (red dashed lines shown in supplemental Movie S2). According to previous enzymological and structural studies of the AlkB family members (39–41), the lid region is a unique structural feature of the AlkB family that determines substrate binding and specificity. The distinctive composition and conformation of Flip1 and Flip2 in Alkbh5 are likely to confer the substrate selectivity characteristic of this enzyme.

Another significant structural feature of Alkbh5 is the formation of an unexpected disulfide bond between residues Cys-230 and Cys-267. The disulfide bond restrains a region defined as Flip3 (residues 229–242) between β6 and β7 toward the minor sheet of the DSBH fold, presenting it as a noticeable overhang vertical to the plane of the minor sheet (Fig. 3*A* and supplemental Fig. S1). As a direct consequence, when the double-stranded DNA from the AlkB or Alkbh2 complex structures was modeled onto the catalytic site of Alkbh5, the Flip3 region of Alkbh5 was well accommodated by the modified strand, whereas it greatly interfered with the binding of the unmethylated strand (Fig. 3*B*). Therefore, we propose that Flip3 impedes the access of dsDNA and dsRNA to the active site of Alkbh5 and that this is the basis for the selectivity of Alkbh5 toward single-stranded substrates.

*Key Residues in the Lid Region Are Potential Determinants of m$^6$A Recognition and Demethylation by Alkbh5*—A combination of hydrogen bonds, hydrophobic interactions, and electrostatic interactions are believed to hold protein and substrate complexes together (15, 20). Accordingly, identifying residues that participate in the substrate-enzyme interaction network shall suggest the mechanisms underlying the substrate specificity and catalytic activity of the protein of interest. With this in mind, we compared the primary sequence and the three-dimensional structure of Alkbh5 with those of other AlkB family proteins that have been comprehensively studied. Several crucial residues of Alkbh5 were selected for mutagenesis to investigate the possible mechanisms of m$^6$A recognition and catalysis (supplemental Fig. S1).

In the structure of the AlkB·Tm$^1$AT complex, the nucleobase ring of m$^1$A is sandwiched between the side chains of Trp-69 from Flip2 and the invariant $Fe^{2+}$-ligating residue, His-131, from the active site (42). This well known aromatic stacking interaction is highly conserved in the AlkB family. The aromatic ring corresponding to Trp-69 is presented by Phe-124 in Alkbh2 (15) and Tyr-108 in FTO (20) (Fig. 2*B*). In the Flip2 region of Alkbh5, Tyr-141 is the candidate residue for the same role (supplemental Fig. S1); the side chain of this residue is directed toward the exterior of Alkbh5 in the native structure (supplemental Movie S1 and Fig. 2*B*), but upon binding to the substrate, it may undergo a conformational change to stack with the nucleobase (43). In addition to Tyr-141, neighboring aromatic residue Tyr-139 is also well conserved (supplemental Movie S1 and Fig. 2*B*). The counterparts of Tyr-139 in AlkB

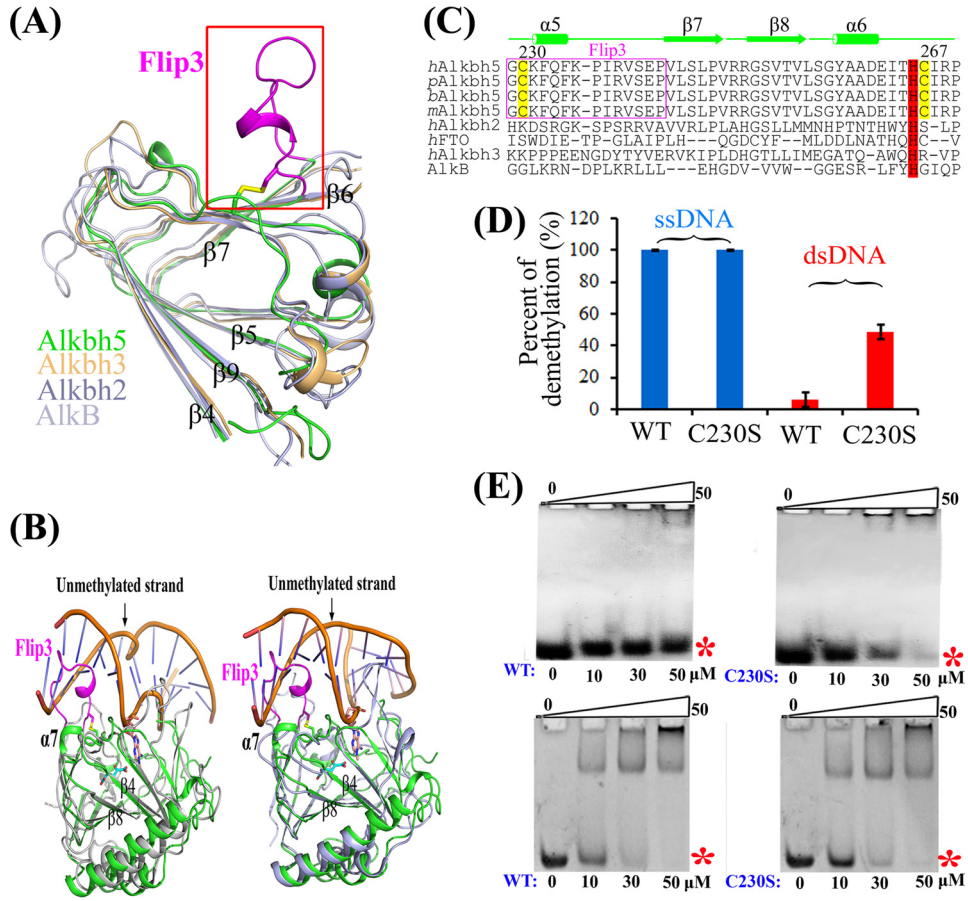


FIGURE 3. **The unique disulfide bond of Alkbh5 decides its binding preference for single-stranded nucleic acids.** *A*, structural comparison of Alkbh5 with Alkbh3 (*tinted*; Protein Data Bank code 2IUW), Alkbh2 (*blue*; Protein Data Bank code 3BTZ), and AlkB (*gray*; Protein Data Bank code 3BIE) around the jelly roll motif. The unique Flip3 (*magenta*; residues 229–242) of Alkbh5 is highlighted within the *red frame*. *B*, structural alignment of Alkbh5 with AlkB·dsDNA complex (*left*) and Alkbh2·dsDNA complex (*right*). The unmethylated strand of dsDNA would sterically clash with "Flip3" of Alkbh5. 1-meA (*purple*) from the methylated strand of dsDNA and α-KG (*cyan*) from Alkbh5 are shown in *sticks*. *C*, structure-based sequence alignment of Alkbh5 of different species and its family members within the disulfide bond-forming region. The conserved $Fe^{2+}$-binding residue His-266 is colored in *red*. *h*, *Homo sapiens*; *p*, *Pan troglodytes*; *b*, *Bos taurus*; *m*, *Mus musculus*. *D*, comparison of the demethylation activity of WT and C230S mutant Alkbh5 for the single-stranded and double-stranded $m^6A$-containing oligonucleotides. All experiments were repeated three times. *Error bars* indicate ±S.D. *E*, EMSA binding assays were performed for dsDNA (*upper panel*) and partial duplex DNA with a 5′ eight-nucleotide ss-DNA (*lower panel*) with increasing amounts (indicated at the *bottom*) of Alkbh5 WT (*left*) and C230S mutant (*right*). The sequence of the partial duplex DNA with a 5′ eight-nucleotide ssDNA was as follows: 5′-CGGACTGGCGGCAGCACTGC-3′ and 5′-GCAGTGCTGCCG-3′. Total DNA substrate used was 10 μM. The position of free DNA is indicated by a *red asterisk*.

(Tyr-76) and in Alkbh2 (Tyr-122) (supplemental Fig. S1) are closely engaged in the recognition of $m^1A$ (15, 43) (Fig. 2*B*); thus, it is reasonable to speculate that Tyr-139 in Alkbh5 may form a critical hydrogen bond with $m^6A$ to reinforce binding. The importance of these two neighboring aromatic residues in the Flip2 region of Alkbh5 was confirmed by the fact that mutant Y139A displayed less than 2% activity of the wide type, whereas the variant Y141A abolished Alkbh5 catalytic activity (Fig. 2*C*).

In all solved structures of the AlkB family members, Flip1 and Flip2 were separated by a β-strand (β3 in Alkbh5) that contains key positively charged residues that are involved in interactions with the nucleoside (supplemental Fig. S1). Arg-131 in Alkbh3 and Arg-110 in Alkbh2 were reported to interact with the 1-meA N-3 atom (15, 16). In FTO, a hydrogen bond between Arg-96 and the O-2 atom of 3-meT is vital for its repair activity (20). In Alkbh5, Arg-130, which lies on the β3 strand (supplemental Movie S1), may rotate to make contact with $m^6A$, and a nearby charged residue, Lys-132, may also have interaction with $m^6A$ (supplemental Fig. S1 and Movie S1). In support of this hypothesis, mutant R130A was catalytically inactive, whereas mutant K132A severely impaired Alkbh5 (Fig. 2*C*).

*The Active Center Pocket Plays a Key Role in Substrate Discrimination and Oxidative Demethylation Activity*—In the active center of Alkbh5, many conserved interactions are observed. The manganese ion is chelated in an octahedral coordination geometry by a water molecule and residues His-204, Asp-206, and His-266 and bidentately by the C-1 carboxylate and C-2 ketone groups of α-KG cofactor (Fig. 4*B*). Apart from its chelation with $Mn^{2+}$, α-KG is further stabilized by a series of interactions with residues conserved throughout the AlkB family. These include two hydrogen bonds formed by the side chains of Asn-193 and Tyr-195 as well as three salt bridges formed by Lys-132, Arg-283, and Arg-277 (Fig. 4*C*). Echoing the significance of the interactions involving the $Mn^{2+}$ or α-KG, any of the three mutants H204A, Y195A, and R277A/R283A showed completely abolished repair activity (Fig. 4*D*), confirming that Alkbh5 implements $m^6A$ oxidative demethylation in an iron- and α-KG-dependent fashion.

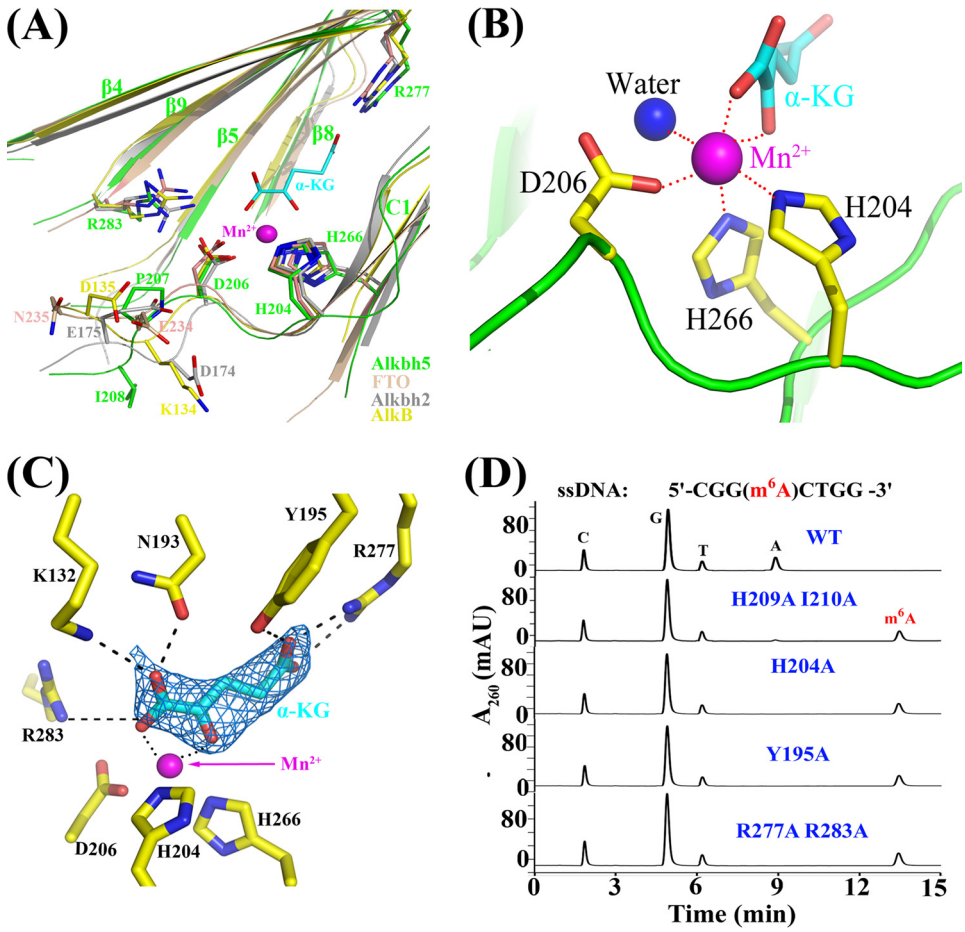


FIGURE 4. **The detailed interaction of the active center of Alkbh5.** *A*, the superposition of Alkbh5 (*green*) with AlkB (*yellow*), Alkbh2 (*gray*), and FTO (*tinted*) at the active site. All the critical residues are shown as *sticks*. The two residues following the conserved H*X*(D/E) motif from the four proteins, α-KG, and $Mn^{2+}$ from Alkbh5 are labeled. *B*, the interaction network around $Mn^{2+}$. $Mn^{2+}$, water, α-KG, and the side chains of Alkbh5 are colored in *magenta*, *blue*, *cyan*, and *yellow*, respectively. Interaction between $Mn^{2+}$ and its coordinating atoms are shown in *red dashed lines*. *C*, the interaction network around α-KG and the involved residues are shown as *sticks*. The $F_o - F_c$ differential electron density map (contour level, $\sigma = 3.0$) is indicated as a *marine blue mesh*. *D*, demethylation activity to m$^6$A-containing ssDNA of the wide-type Alkbh5 and its mutants involved in the $Mn^{2+}$ binding, α-KG binding, and the residues around the entrance loop of the jelly roll fold. *mAU*, milli-absorbance units.

The loop linking the canonical second and third strands of the DSBH fold is directly involved in substrate binding and catalysis (40, 41). In the loops of most AlkB family members, besides the conserved H*X*(D/E) motif, there are two polar amino acids and at least one aspartic acid or glutamic acid (Asp-135 in AlkB, Glu-175 in Alkbh2, and Glu-234 in FTO) that form an important hydrogen bond with the nucleobase that contributes to the selection of different methylated nucleobases within the target proteins (15, 16, 20) (Fig. 4*A* and supplemental Fig. S1). However, the same region in Alkbh5 is more hydrophobic and lacks a polar side chain (Fig. 4*A*), which may contribute to the selectivity of Alkbh5 toward m$^6$A. We constructed a double mutant, H209A/I210A (supplemental Fig. S1), which exhibited dramatically reduced repair performance of less than 1% activity in comparison with the WT protein (Fig. 4*D*). These two residues are located in the entrance loop of the active pocket and probably interact with the phosphate backbone of the RNA substrate (Fig. 5*E*) and may also help maintain the stability of the active site structure of Alkbh5.

*A Unique Disulfide Bond Assists Alkbh5 to Differentiate between Single-stranded and Double-stranded Substrates*—As demonstrated in Fig. 3*A* and supplemental Movie S1, Alkbh5 harbors a unique disulfide bond between Cys-230 and Cys-267 that may confer Alkbh5 the ability to discriminate against double-stranded nucleic acids. To test this hypothesis, a C230S mutant was constructed.

An electrophoretic mobility shift assay (EMSA) was carried out wherein double-stranded oligodeoxynucleotides containing m$^6$A were incubated with the recombinant wide-type or the mutant C230S protein. Although the protein·DNA complexes in several lanes did not enter the acrylamide gel and the complex bands were not always apparent, it was clear that with increasing protein concentration the amount of free dsDNA detected on the gel decreased quickly when mixed with the mutant C230S, whereas the amount of free dsDNA did not show remarkable change when mixed with the wild-type protein (Fig. 3*E*, *upper panel*). Conversely, when a partial duplex DNA with a 5′ eight-nucleotide ss-DNA was used, the WT and mutant C230S proteins displayed similar binding affinities (Fig. 3*E*, *lower panel*).

We also used an HPLC-based demethylation assay to evaluate the activity of the mutant C230S. As expected, the mutant C230S retained full activity toward single-stranded nucleic acids. Meanwhile, its repair capacity toward double-stranded

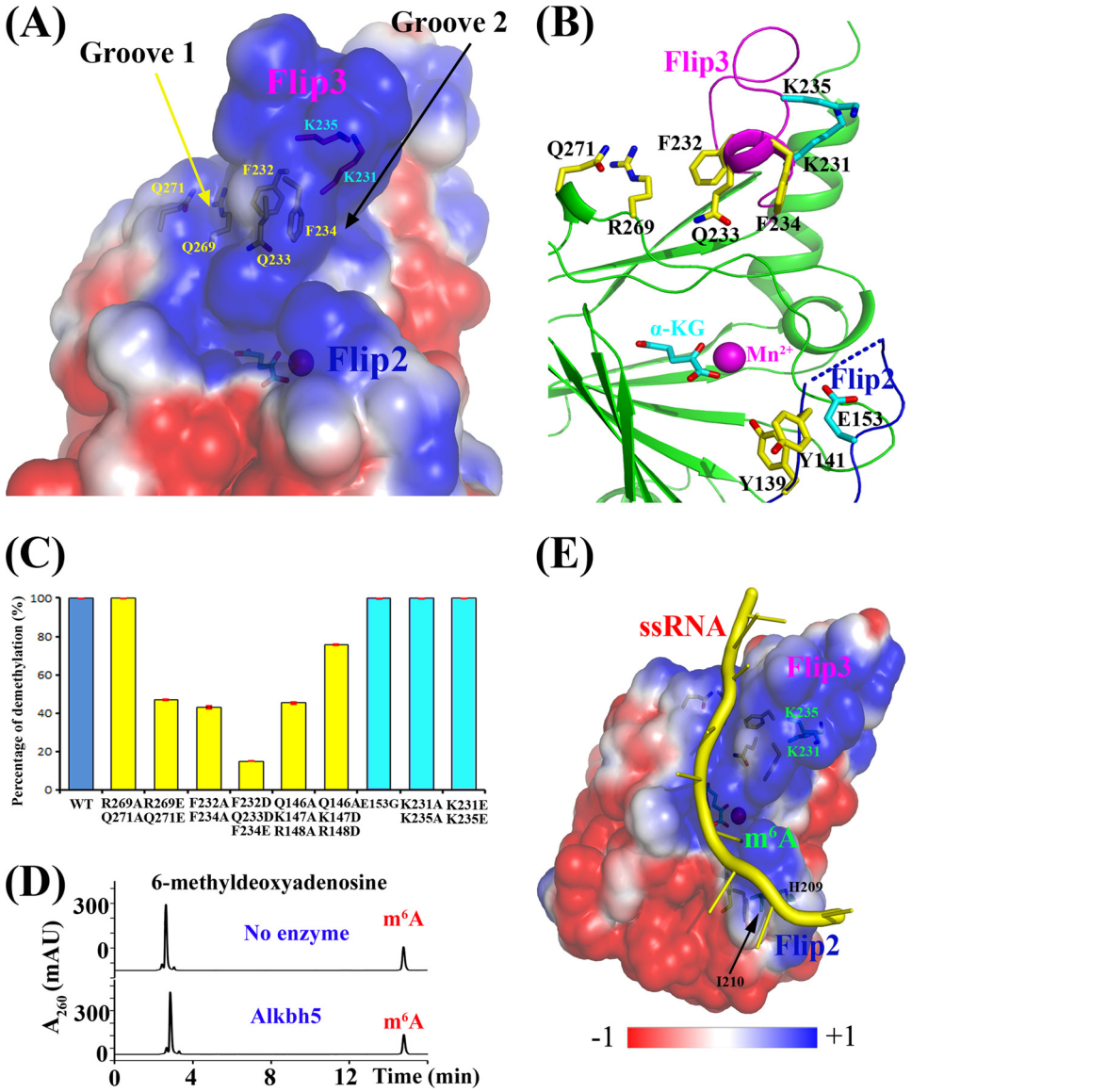


FIGURE 5. **The substrate recognition and potential binding model of Alkbh5.** α-KG and $Mn^{2+}$ are shown. *A*, the electrostatic surface potential of Alkbh5. The positively charged surface is colored *blue*, and the negatively charged surface is colored *red*. The two positively charged grooves around Flip3 are highlighted and labeled. The key residues are shown as *sticks* and labeled. *B*, the critical residues located at the positively charged surface of Alkbh5 are presented. The view orientation is the same as *A*. The residues in the missing part of Flip2 are shown in *blue dashed lines*. *C*, repair activity of Alkbh5 surface residue variants on the 8-mer m$^6$A-containing ssDNA. *Error bars* indicate ±S.D. for triplicate experiments. *D*, Alkbh5 shows no repair activity for 6-methyldeoxyadenosine as detected by HPLC. *E*, the proposed model of Alkbh5 binding to ssRNA (*yellow*). The m$^6$A nucleobase inserted into the catalytic pocket is indicated, and Flip2 and Flip3 of Alkbh5 are also labeled. The key residues are shown as *sticks*; residues Lys-231, Lys-235, His-209, and Ile-210 are also labeled. *mAU*, milli-absorbance units.

nucleic acids increased markedly compared with the almost undetectable activity of WT (Fig. 3*D*).

To examine whether disruption of the disulfide bridge affects the overall secondary structure of Alkbh5, we performed a circular dichroism (CD) measurement. Far-UV CD spectra were recorded in the 180–260-nm range, and the results indicated that there were no significant conformational differences between the WT protein and the mutant C230S (data not shown).

This unique disulfide bond is highly conserved among Alkbh5 proteins from different species but is not shared by other AlkB family members (Fig. 3*C*). It can be predicted that the disulfide bond could have been selected during evolution to ensure a proper conformation of Alkbh5 and its selectivity toward single-stranded nucleic acids.

*A Substrate Model Based on Mutagenesis Data*—We made every effort to crystalize the m$^6$A-containing ssDNA·Alkbh5 complex, but all attempts were unsuccessful. Eventually, we settled for exploring protein-DNA interactions of the complex. To this end, we carried out site-directed mutagenesis for selected surface residues of Alkbh5 along with a demethylation assay.

Initially, the electrostatic potential map of Alkbh5 was calculated (Fig. 5*A*). This map shows that a positively charged area is distributed around the active site and extends from the cavity of the catalytic core toward the minor β-sheet where it is separated into two possible substrate-binding grooves (groove1 and groove2) by the Flip3 region, which shows up as a positively charged protrusion. Additionally, the discrete portion of the Flip2 loop is also positively charged, further suggesting that the area might bind with the substrate (Fig. 5*A*).

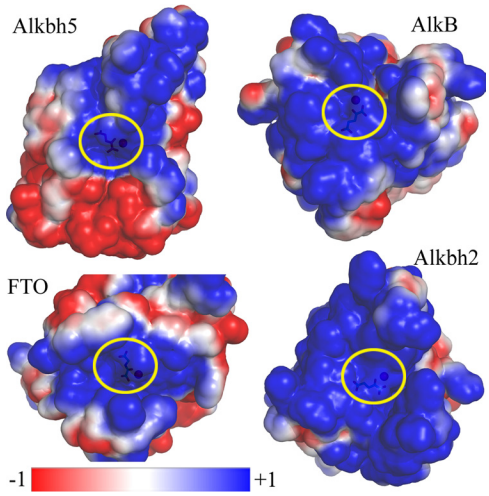


FIGURE 6. **Comparison of electrostatic surfaces among Alkbh5, AlkB, FTO, and Alkbh2.** All the structures are shown in the same scale. Positively charged surface is colored in *blue*, negatively charged surface is colored in *red*, and neutral surface is colored in *white*. Alkbh5 is less positively charged compared with AlkB (Protein Data Bank code 3BIE), FTO (Protein Data Bank code 3LFM), and Alkbh2 (Protein Data Bank code 3BTZ). α-KG and $Mn^{2+}$ are shown. The active center is shown in a *yellow ellipse*.

On the whole, the distribution of the surface positive charges around the active site represents the possible binding region of the single-stranded nucleic acids. This region is much narrower compared with those of other AlkB members (Fig. 6). To further probe the substrate-binding mode, we mutated several residues located within different areas of the positively charged surface and measured demethylation activity (Fig. 5*B*).

Although residues Gln-146, Lys-147, and Arg-148 are located in the middle of Flip2 (supplemental Fig. S1) and are not visible in the solved structure of Alkbh5 (Fig. 1*B*), their positive charges and polar side chains make them good candidates for substrate binding. Two triple mutants were generated. When residues Gln-Lys-Arg were mutated to Ala-Asp-Asp, this triple mutant displayed 73.3% activity toward ssDNA, whereas another triple mutant, Q146A/K147A/R148A, only retained 44% of the WT activity (Fig. 5*C*). Combining the data from the Y141A and Y139A mutants mentioned above (Figs. 5*B* and 2*C*) suggests that the long Flip2 loop of Alkbh5 consolidates its interactions with one end of the single-stranded nucleic acids and confers recognition of the specific nucleotide base similarly to what occurs for the other AlkB members (15, 16, 20, 43).

Next, we designed a series of mutants of the residues located at the two grooves. Residues Arg-269, Gln-271, Phe-232, Gln-233, and Phe-234 were selected from groove1, and Lys-231 and Lys-235 were selected from groove2 (Fig. 5, *A* and *B*). Double mutation R269A/Q271A did not affect the repair efficiency of Alkbh5; however, the R269E/Q271E mutant greatly reduced catalytic activity to ~50% (Fig. 5*C*), which suggests their interactions with the phosphate group of the substrate. Similarly, the F232A/F234A mutant exhibited 41% activity toward m$^6$A-containing ssDNA. When three negatively charged residues (Asp, Asp, and Glu) replaced residues Phe-232, Gln-233, and Phe-234, the variant displayed a severe loss of activity, demonstrating only 13.5% of WT activity (Fig. 5*C*). These mutagenesis results suggested that the other end of the single-stranded nucleic acids may stretch along the minor β-sheet or specifically along the positively charged groove1 on the surface, making contacts with the residues mentioned above. On the contrary, neither the K231A/K235A nor the K231E/K235E double mutation in groove2 had an impact on the repair capacity of Alkbh5 (Fig. 5*C*). Thus, based on these data, we believe groove1 and Flip2, but not groove2, are involved in substrate interaction. A model of ssRNA binding to Alkbh5 is shown in Fig. 5*E*.

*The Structures of Alkbh5 with Different Ligands Shed Light on Potential Selective Drug Design*—Although cofactors $Mn^{2+}$ and α-KG were premixed with the purified Alkbh5 before setting up crystallization, the actual binding site in the resolved structure was occupied by the citrate and acetate molecules from the reservoir buffer. Soaking was used to obtain $Mn^{2+}$- and α-KG-bound Alkbh5 crystals. An overlay of Alkbh5·citrate·acetate and Alkbh5·$Mn^{2+}$·α-KG structures revealed little difference in the overall conformation (root mean square deviation of 0.4 Å) except for the side chain of Asp-206, which rotates by 86.8° and binds $Mn^{2+}$ in the α-KG complex structure (Fig. 7*B*).

Interestingly, although the citrate and acetate molecules compete with α-KG for binding and ultimately block α-KG from entering the active center, their positions do not superimpose with the actual α-KG binding site. Instead, they are located at the two ends of α-KG, respectively (Fig. 7*B*). To elaborate, the acetate molecule resides in the position proximate to the C-5 carboxylate of α-KG and makes tight contacts with Arg-277 and two crystallographic water molecules (Fig. 7, *A* and *B*). The citrate molecule is close to the entrance of the active pocket with one of its carboxylates pointing toward the same direction as the C-1 carboxylate oxygen of α-KG, forming extensive hydrogen bonds with Asn-193, Lys-132, and three water molecules. The central carboxylate of the citrate was involved in electrostatic and hydrogen-bonding interactions with a water molecule and the side chains of Arg-283, His-204, Asp-206, and His-266. The third carboxylate is orientated toward the deep interior of the active site, forming a hydrogen bond with a water molecule in the crystal structure (Fig. 7*A*). We aligned the structure of Alkbh5·α-KG to that of FTO·citrate in which the citrate molecule occupies the α-KG binding site. The superimposition uncovered that there would be considerable steric hindrance between the citrate molecule of FTO and residues Ile-281 and Tyr-195 of Alkbh5, explaining why the citrate molecule in the Alkbh5 structure cannot locate to the position predicted from the FTO structure (Fig. 7*D*).

The naturally grown crystal structure of Alkbh5 in complex with the citrate and acetate molecules prompted us to investigate selective inhibitors of Alkbh5. In this regard, we tested two well characterized α-KG oxygenase inhibitors, NOG and PDCA, both of which are non-reactive analogs of α-KG. In addition, citrate and succinate (the product of α-KG decarboxylation) were chosen to explore their inhibition of Alkbh5 activity. To identify a suitable substrate concentration for these inhibition assays, we first performed an ITC experiment to determine the binding affinity of α-KG toward Alkbh5. The equilibrium dissociation constant ($K_d$) between Alkbh5 and α-KG was ~9.6 μM (Fig. 7*E*). By increasing the inhibitor concentrations relative to α-KG in a proportional manner and then monitoring the demethylation capacity of Alkbh5 by HPLC, we determined the $IC_{50}$ values of the four candidate inhibitors (Fig.

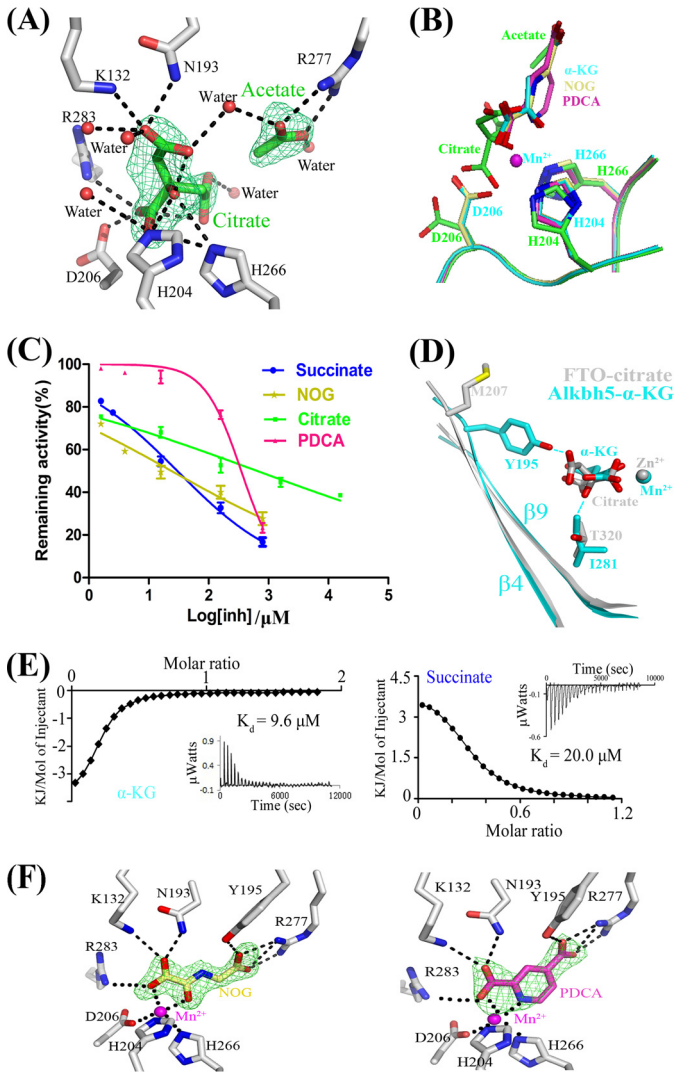


FIGURE 7. **The binding of different ligands to Alkbh5.** Citrate/acetate, α-KG, NOG, and PDCA are shown in *green*, *cyan*, *magenta*, and *pale yellow*, respectively. *A*, the interaction network around the citrate and acetate molecules in the naturally grown crystal structure of Alkbh5. Also the $F_o - F_c$ differential electron density map of the bound citrate and acetate at 3 $\sigma$ is presented. Hydrogen bonds are indicated as *black dashes*. *B*, the superimposition of the different ligands at the active sites of Alkbh5·cofactor complexes. The greatest difference lies in residue Asp-206 of the Alkbh5·citrate complex; it rotates by 86.8° in the other three structures. Only the key residues from the Alkbh5·citrate complex and Alkbh5·α-KG complex are labeled. *C*, $IC_{50}$ curves for inhibitors (*inh*), including succinate, PDCA, NOG, and citrate. Succinate is colored in *blue*. *Error bars* indicate ±S.D. *D*, superimposition of Alkbh5 (*cyan*) and FTO (*gray*; Protein Data Bank code 4IE7) around the binding pocket. Considerable steric hindrance (shown in *cyan lines*) between residues Ile-281 and Tyr-195 of Alkbh5 with the citrate would be introduced if citrate in Alkbh5 is bound like that in FTO·citrate. *E*, quantification of the binding affinity of α-KG (*left*) and succinate (*right*) to Alkbh5 by ITC. The integrated heat is plotted against the molar ratio of ligand added to Alkbh5 in the cell. The dissociation constant ($K_d$) is indicated. *F*, the interaction network around NOG (*left*) and PDCA (*right*). The residues involved are shown as *sticks*. The $F_o - F_c$ differential electron density maps (contour level, $\sigma = 3.0$) of NOG and PDCA are both indicated as a *lime green mesh*.

7*C*). PDCA was found to be a relatively moderate inhibitor of Alkbh5 with an $IC_{50}$ value of 347.2 μM, whereas citrate was a much less effective inhibitor with an $IC_{50}$ of 627.9 μM. NOG and succinate inhibited Alkbh5 with $IC_{50}$ values of 25.85 and 30.00 μM, respectively. This result is different from that by FTO (44) for which PDCA is the strongest inhibitor and succinate is the weakest inhibitor among the four inhibitors. To further confirm this result, we measured the binding affinity of succinate toward Alkbh5 by ITC. The ITC data revealed a $K_d$ value of ~20 μM (Fig. 7*E*), which is consistent with the $IC_{50}$ value. These results imply that these small molecules have very different inhibitory specificities toward FTO and Alkbh5, which may provide a good foundation for the future design of selective inhibitors toward the AlkB family. Previous studies have shown that FTO can demethylate several methylated nucleotides (18, 38), whereas Alkbh5 can only demethylate m$^6$A. A possible explanation is that the active site cavity of Alkbh5 is relatively small. Calculations reveal that the cavity volume is 490.2 Å$^3$ for Alkbh5 and 817.5 Å$^3$ for FTO. This might also explain the more potent binding affinity of smaller inhibitors toward Alkbh5 than larger inhibitors.

To examine the binding mode of the Alkbh5 inhibitors, we determined the structures of Alkbh5 in complex with NOG at 1.8 Å and with PDCA at 2.5 Å. As expected, both inhibitors are located in the α-KG binding site (Fig. 7*B*), with the C-5 carboxylate of NOG and C-4 carboxylate of PDCA forming hydrogen bonds and electrostatic interactions with Tyr-195 and Arg-277, respectively. A slight difference between the two inhibitors is that NOG bound to $Mn^{2+}$ via its C-1 carboxylate and C-2 ketone groups, whereas PDCA chelated $Mn^{2+}$ through its C-2 carboxylate and pyridyl nitrogen atom (Fig. 7*F*).

## DISCUSSION

In this study, we presented five high resolution crystal structures of the Alkbh5 catalytic core without $Mn^{2+}$, with $Mn^{2+}$ alone, or with $Mn^{2+}$ in addition to another ligands, NOG, α-KG, or PDCA. Alkbh5 was shown to possess several unique structural features compared with other AlkB family members. First of all, the nucleotide recognition lid of Alkbh5 emerges as a distinct region. On the one hand, Flip1 of Alkbh5 leaves a larger space open over the active site. On the other hand, the long and discrete Flip2 loop is much more flexible (Fig. 2*A* and supplemental Movie S2). Based on our mutagenesis analyses (Figs. 2*C* and 5*C*) and the known substrate recognition mechanisms of other AlkB members (15, 16, 20, 43), the long Flip2 loop of Alkbh5 was proposed to play a significant role in substrate recognition and binding through interaction with one end of the single-stranded nucleic acid substrate (Fig. 5*E*). Once in contact with the substrate, the Flip2 loop may undergo a large conformational change and turn upward to the Flip1-exposed open area to accommodate the substrate effectively. Moreover, the significance of the Flip2 region (residues 136–165) in Alkbh5 can also be inferred from the data in COSMIC (Catalogue of Somatic Mutations in Cancer) (45). The E153G mutation of Alkbh5 has been observed in human breast carcinoma, which led us to test the E153G mutant (Fig. 5*B*). However, the E153G mutation did not affect the demethylation activity of Alkbh5 *in vitro* (Fig. 5*C*). It was possible that E153G affected the co-regulation of Alkbh5 by endogenous partners *in vivo*.

Moreover, our structure-guided mutagenesis analysis further suggested the significance of the Alkbh5 lid region in m$^6$A recognition and catalysis. Among the lid residues, besides the relatively conserved residues in Flip2, the residues located in

$\beta$3, which separates Flip1 and Flip2, are also vital. Here, it is particularly worth mentioning that Alkbh5 is the only member that simultaneously bears two critical residues, Arg-130 and Lys-132, in the conserved $\beta$-strand (supplemental Fig. S1) among the AlkB family. This may explain Alkbh5 substrate specificity on m$^6$A. It is possible that both the residues Arg-130 and Lys-132 (supplemental Fig. S1 and Movie S1) make strong interactions only with m$^6$A instead of other modified nucleotides. Additionally, compared with FTO, Alkbh5 might have the extra interactions contributed by Arg-130, which may explain its higher m$^6$A binding affinity and specificity (18, 25). Furthermore, it is noteworthy that Lys-132 acetylation of Alkbh5 was identified by high resolution mass spectrometry in the presence of the histone deacetylase inhibitor suberoylanilide hydroxamic acid (46). Lysine acetylation has been proven to play a key role in diverse cellular processes such as RNA splicing, transcription, and nuclear transport. Acetylation of Lys-132 of Alkbh5 regulates the activity of Alkbh5 as shown by the diminished activity of the K132A mutant.

The structure of Alkbh5 presents a unique disulfide bond between residues Cys-230 and Cys-267. Our work has demonstrated that this novel disulfide bond enables Alkbh5 to distinguish single-stranded from double-stranded oligos. Both the EMSA and the HPLC-based demethylation assay supported the stronger dsDNA binding affinity and the higher dsDNA catalytic activity in the C230S mutant. It is likely that disruption of the disulfide bridge allows the Flip3 region of Alkbh5 to move away from the minor sheet to leave sufficient space to accommodate the unmethylated strand of double-stranded nucleic acids without steric hindrance (Fig. 3*B*). To our knowledge, this disulfide bond is a novel structural element that determines the substrate preference of AlkB family proteins besides the Flip1 hairpin in Alkbh2 (15) (Fig. 2*A*) and the L1 loop in FTO (20). It is well known that mRNA may exist as a partially double-stranded structure *in vivo*. Therefore, it is expected that reducing metabolites at certain points may be high enough to break the disulfide bond of Alkbh5. Under this physiological condition, Alkbh5 could effectively demethylate m$^6$A-containing dsRNA by changing its Flip3 conformation, which in turn would affect many biological processes, including gene expression and RNA metabolism. The substrate selectivity of Alkbh5 might in this way be dynamically regulated, further indicating the critical role of Alkbh5 in mammalian cells.

Recently, it was reported that FTO catalyzes the formation of $N^6$-hydroxymethyladenosine and $N^6$-formyladenosine in a stepwise manner during m$^6$A demethylation (38), whereas these two intermediates were not detected in the Alkbh5-catalyzed m$^6$A demethylation process (47). This may be structurally explained by the fact that the residues adjacent to the key motif H*X*(D/E) in the substrate catalysis pocket of Alkbh5 are more hydrophobic (Fig. 4*A*) and therefore unlikely to confer affinity toward the hydrophilic groups of $N^6$-hydroxymethyladenosine or $N^6$-formyladenosine. These residues might prevent Alkbh5 from carrying out further oxidation of these intermediates.

We proposed an ssRNA binding model of Alkbh5 based on the results of site-directed mutagenesis of the surface residues together with demethylation assays. However, we still cannot explain the specific preference of Alkbh5 for the consensus sequence PuPum$^6$AC(A/C/U) (where Pu represents purine) in single-stranded nucleic acids (25). In this study, we also found that Alkbh5 did not have any activity toward 6-methyldeoxyadenosine. No converted adenine peak was detected by HPLC (Fig. 5*D*), and no formaldehyde was measured using the Nash assay (data not shown) when Alkbh5 was incubated with 6-methyldeoxyadenosine. Presumably, the methyl group in the 6-methyldeoxyadenosine cannot be positioned correctly toward the Alkbh5 active site for effective demethylation. Hence, it is very likely that the interaction between Alkbh5 and the nucleotide backbones, especially the nucleobase portions flanking m$^6$A, is essential for the optimal orientation of m$^6$A, the productive binding of single-stranded nucleic acids, and the final Alkbh5-mediated m$^6$A oxidative demethylation. A structure of Alkbh5 in complex with nucleic acid substrate containing the consensus sequence is needed to fully account for this.

Our work also investigated the regulation of Alkbh5 repair activity by $\alpha$-KG competitors. Through crystallographic and biochemical studies, we found that, in contrast to FTO, Alkbh5 has potent binding preference toward smaller molecule inhibitors that is likely caused by the small active site cavity of Alkbh5. Although the sample size of Alkbh5 inhibitors presented here is small, our results offer a basis for the rational design of specific AlkB inhibitors and activators based on the distinct binding pocket of this protein family and prompt the exploration of the physiological functions of Alkbh5-mediated m$^6$A demethylation.

Finally, the recent findings have shown the significant biological, clinical, and therapeutic implications of Alkbh5. Human *Alkbh5* is one of the disease-related gene candidates located on chromosome 17p11. The genes along this chromosome are known to be involved in a wide range of human genetic diseases caused by chromosomal deletions or duplication (48, 49). In addition, ONCOMINE cancer database (Compendia Bioscience, Ann Arbor, MI) searches showed that human Alkbh5 levels dramatically decrease in Korkola seminoma and testicular yolk sac tumors as well as in several types of breast cancer. Thus, abrogated expression or enzymatic activity of Alkbh5 has been strongly implicated in human diseases. The study of Alkbh5 in transcription regulation will shed light on the mechanisms of RNA epigenetic regulation in tumorigenesis, which will provide the molecular basis for designing demethylase inhibitor-guided anticancer drugs in the future.

Taken together, our results further the structural knowledge of the AlkB family and provide insights into their diverse substrate recognitions. Additionally, our study of Alkbh5 as the m$^6$A demethylase will avail the investigation of the modulation of m$^6$A modification in diverse fundamental processes and cast light on the realm of RNA epigenetics.

*Acknowledgments—We thank Prof. Pinchao Mei and Prof. Jiemin Wong for cDNAs and discussion.*

## REFERENCES

1. Cantara, W. A., Crain, P. F., Rozenski, J., McCloskey, J. A., Harris, K. A., Zhang, X., Vendeix, F. A., Fabris, D., and Agris, P. F. (2011) The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.* **39,** D195–D201

2. 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
3. Canaani, D., Kahana, C., Lavi, S., and Groner, Y. (1979) Identification and mapping of N6-methyladenosine containing sequences in simian virus 40 RNA. *Nucleic Acids Res.* **6,** 2879–2899
4. Harper, J. E., Miceli, S. M., Roberts, R. J., and Manley, J. L. (1990) Sequence specificity of the human mRNA N6-adenosine methylase *in vitro*. *Nucleic Acids Res.* **18,** 5735–5741
5. 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
6. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., and Rechavi, G. (2012) Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485,** 201–206
7. Pan, T. (2013) N6-methyl-adenosine modification in messenger and long non-coding RNA. *Trends Biochem. Sci.* **38,** 204–209
8. 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
9. 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
10. Tsujikawa, K., Koike, K., Kitae, K., Shinkawa, A., Arima, H., Suzuki, T., Tsuchiya, M., Makino, Y., Furukawa, T., Konishi, N., and Yamamoto, H. (2007) Expression and sub-cellular localization of human ABH family molecules. *J. Cell. Mol. Med.* **11,** 1105–1116
11. Falnes, P. Ø., Johansen, R. F., and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **419,** 178–182
12. 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
13. Westbye, M. P., Feyzi, E., Aas, P. A., Vågbø, C. B., Talstad, V. A., Kavli, B., Hagen, L., Sundheim, O., Akbari, M., Liabakk, N. B., Slupphaug, G., Otterlei, M., and Krokan, H. E. (2008) Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. *J. Biol. Chem.* **283,** 25046–25056
14. 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
15. 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
16. Sundheim, O., Vågbø, C. B., Bjørås, M., Sousa, M. M., Talstad, V., Aas, P. A., Drabløs, 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
17. Pastore, C., Topalidou, I., Forouhar, F., Yan, A. C., Levy, M., and Hunt, J. F. (2012) Crystal structure and RNA binding properties of the RNA recognition motif (RRM) and AlkB domains in human AlkB homolog 8 (ABH8), an enzyme catalyzing tRNA hypermodification. *J. Biol. Chem.* **287,** 2130–2143
18. Gerken, T., Girard, C. A., Tung, Y.-C. L., Webby, C. J., Saudek, V., Hewitson, K. S., Yeo, G. S., McDonough, M. A., Cunliffe, S., McNeill, L. A., Galvanovskis, J., Rorsman, P., Robins, P., Prieur, X., Coll, A. P., Ma, M., Jovanovic, Z., Farooqi, I. S., Sedgwick, B., Barroso, I., Lindahl, T., Ponting, C. P., Ashcroft, F. M., O'Rahilly, S., and Schofield, C. J. (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* **318,** 1469–1472
19. 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
20. 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
21. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., and He, C. (2011) N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7,** 885–887
22. Thalhammer, A., Bencokova, Z., Poole, R., Loenarz, C., Adam, J., O'Flaherty, L., Schödel, J., Mole, D., Giaslakiotis, K., Schofield, C. J., Hammond, E. M., Ratcliffe, P. J., and Pollard, P. J. (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
23. Karkhanis, V., Wang, L., Tae, S., Hu, Y. J., Imbalzano, A. N., and Sif, S. (2012) Protein arginine methyltransferase 7 regulates cellular response to DNA damage by methylating promoter histones H2A and H4 of the polymerase $\delta$ catalytic subunit gene, POLD1. *J. Biol. Chem.* **287,** 29801–29814
24. Baltz, A. G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., and Landthaler, M. (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46,** 674–690
25. Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., Li, C. J., Vågbø, C. B., Shi, Y., Wang, W. L., Song, S. H., Lu, Z., Bosmans, R. P., Dai, Q., Hao, Y. J., Yang, X., Zhao, W. M., Tong, W. M., Wang, X. J., Bogdan, F., Furu, K., Fu, Y., Jia, G., Zhao, X., Liu, J., Krokan, H. E., Klungland, A., Yang, Y. G., and He, C. (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49,** 18–29
26. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol. Macromolecular Crystallography, part A* (C. W. Carter, Jr. & R. M. Sweet, eds.) Vol. 276, pp. 307–326, Academic Press, New York
27. Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* **64,** 125–132
28. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
29. Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2,** 2728–2733
30. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 32–41
31. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* **64,** 61–69
32. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 355–367
33. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242
34. Wu, L. J., Zhang, T., Gu, Y. X., Zheng, C. D., and Fan, H. F. (2009) Direct-method SAD phasing of proteins enhanced by the use of intrinsic bimodal phase distributions in the subsequent phase-improvement process. *Acta Crystallogr. D Biol. Crystallogr.* **65,** 1213–1216
35. Collaborative. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
36. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62,** 1002–1011
37. DeLano, W. L. (2010) *The PyMOL Molecular Graphics System*, version 1.3, Schrödinger, LLC, New York
38. Fu, Y., Jia, G., Pang, X., Wang, R. N., Wang, X., Li, C. J., Smemo, S., Dai, Q., Bailey, K. A., Nobrega, M. A., Han, K.-L., Cui, Q., and He, C. (2013) FTO-mediated formation of N6-hydroxymethyladenosine and N6-formylad-

enosine in mammalian RNA. *Nat. Commun.* **4,** 1798
39. Sundheim, O., Talstad, V. A., Vågbø, C. B., Slupphaug, G., and Krokan, H. E. (2008) AlkB demethylases flip out in different ways. *DNA Repair* **7,** 1916–1923
40. 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
41. 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
42. 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
43. 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
44. 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
45. Forbes, S. A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C. Y., Jia, M., Ewing, R., Menzies, A., Teague, J. W., Stratton, M. R., and Futreal, P. A. (2010) COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Res.* **38,** D652–D657
46. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840
47. 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
48. Zody, M. C., Garber, M., Adams, D. J., Sharpe, T., Harrow, J., Lupski, J. R., Nicholson, C., Searle, S. M., Wilming, L., Young, S. K., Abouelleil, A., Allen, N. R., Bi, W., Bloom, T., Borowsky, M. L., Bugalter, B. E., Butler, J., Chang, J. L., Chen, C.-K., Cook, A., Corum, B., Cuomo, C. A., de Jong, P. J., DeCaprio, D., Dewar, K., FitzGerald, M., Gilbert, J., Gibson, R., Gnerre, S., Goldstein, S., Grafham, D. V., Grocock, R., Hafez, N., Hagopian, D. S., Hart, E., Norman, C. H., Humphray, S., Jaffe, D. B., Jones, M., Kamal, M., Khodiyar, V. K., LaButti, K., Laird, G., Lehoczky, J., Liu, X., Lokyitsang, T., Loveland, J., Lui, A., Macdonald, P., Major, J. E., Matthews, L., Mauceli, E., McCarroll, S. A., Mihalev, A. H., Mudge, J., Nguyen, C., Nicol, R., O'Leary, S. B., Osoegawa, K., Schwartz, D. C., Shaw-Smith, C., Stankiewicz, P., Steward, C., Swarbreck, D., Venkataraman, V., Whittaker, C. A., Yang, X., Zimmer, A. R., Bradley, A., Hubbard, T., Birren, B. W., Rogers, J., Lander, E. S., and Nusbaum, C. (2006) DNA sequence of human chromosome 17 and analysis of rearrangement in the human lineage. *Nature* **440,** 1045–1049
49. Nickerson, M. L., Warren, M. B., Toro, J. R., Matrosova, V., Glenn, G., Turner, M. L., Duray, P., Merino, M., Choyke, P., Pavlovich, C. P., Sharma, N., Walther, M., Munroe, D., Hill, R., Maher, E., Greenberg, C., Lerman, M. I., Linehan, W. M., Zbar, B., and Schmidt, L. S. (2002) Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dubé syndrome. *Cancer Cell* **2,** 157–164