Substrate Recognition Hot Paper

DOI: 10.1002/ange.201310050

Switching Demethylation Activities between AlkB Family RNA/DNA Demethylases through Exchange of Active-Site Residues**

*Chenxu Zhu and Chengqi Yi**

Abstract: *The AlkB family demethylases AlkB, FTO, and ALKBH5 recognize differentially methylated RNA/DNA substrates, which results in their distinct biological roles. Here we identify key active-site residues that contribute to their substrate specificity. Swapping such active-site residues between the demethylases leads to partially switched demethylation activities. Combined evidence from X-ray structures and enzyme kinetics suggests a role of the active-site residues in substrate recognition. Such a divergent active-site sequence may aid the design of selective inhibitors that can discriminate these homologue RNA/DNA demethylases.*

Fe^{II} and 2-ketoglutarate (2KG) dependent dioxygenases perform a wide variety of oxidation functions in biology, including the oxidative demethylation of methylated DNA and RNA.[1] The first such dioxygenase identified was AlkB from *E. coli*, which oxidatively repairs damaged bases including N^1-methyladenine in DNA and RNA (m^1dA and m^1A, respectively).[2] ALKBH2 and ALKBH3, two AlkB human homologue proteins, can also demethylate m^1dA, and thus are repair proteins that protect our genome from methylation damage.[3] In contrast, FTO and ALKBH5, two other AlkB human homologues, display only negligible activity towards m^1dA, and are considered to play regulatory roles rather than DNA repair. The fat-mass and obesity-associated protein FTO, which has been demonstrated to influence human obesity and energy utilization in up to half of the world's population,[4] was first shown to demethylate N^3-methylthymidine and N^3-methyluridine,[5] but more recently also the abundant mRNA modification N^6-methyladenosine (m^6A) with much higher efficiency.[6] ALKBH5 is the second mammalian demethylase discovered to oxidatively reverse m^6A and it has an impact on RNA metabolism and mouse fertility.[7] Therefore, despite the common ground of all being nucleic acid demethylases, these AlkB family proteins optimally recognize differentially methylated substrates and also play different biological roles.

N^1-Methyladenine and N^6-methyladenine are both N-methylated nucleobases but differ in their position of modification and also biological consequences (Figure 1a). m^1dA can be generated in single-stranded DNA (ssDNA) by methylating agents and the adduct prevents formation of Watson–Crick base pairs, and is thus cytotoxic if unrepaired;[8] chemically induced m^1A could also impair the function of RNA molecules.[2c,d] On the other hand, m^6A is formed enzymatically in eukaryotic mRNAs, pairs to T/U regularly, and is believed to play important regulatory roles in mRNA processing and metabolism.[9] Thus, the biological roles of these demethylases are determined—to a large extent—by the differentially methylated RNA/DNA substrates they recognize. However, the mechanism of such selective recognition remains poorly understood. Although the crystal structure of FTO has been solved,[10] it does not contain the most preferred substrate m^6A. The structure of ALKBH3 has also been solved,[11] but it does not have a bound substrate. Although substrate recognition by AlkB and ALKBH2 has been relatively well-studied through crystallography,[12a–f] how the active site of AlkB can discriminate N^6-methyladenine remains unknown. The three-dimensional folding of the catalytic domain of these proteins looks similar in the existing

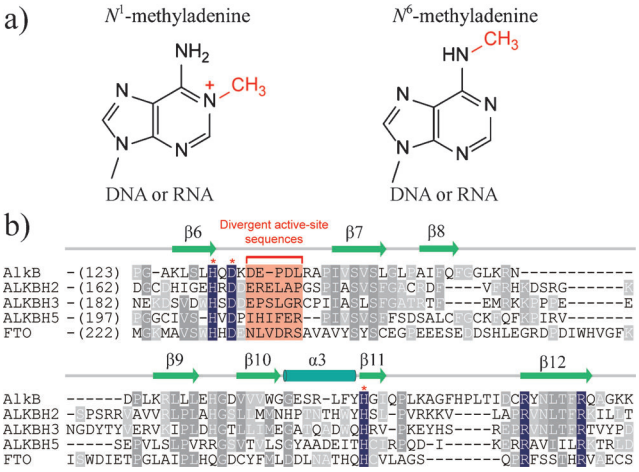

Figure 1. The AlkB family proteins recognize differentially methylated adenine bases. a) Chemical structures of N^1-methyladenine and N^6-methyladenine. b) Sequence alignment of five AlkB family proteins. Secondary structures are indicated on top of the aligned sequences. Letters highlighted in red represent the divergent active-site sequences, while those in blue denote the characteristic residues of the AlkB family proteins. Columns with "*" at the top are ligands for Fe^{II} ions and the numbers of the first amino acids within the red bracket are: D135 for AlkB, E175 for ALKBH2, E195 for ALKBH3, I209 for ALKBH5, and N235 for FTO.

[*] C. Zhu, Prof. C. Yi
State Key Laboratory of Protein and Plant Gene Research
School of Life Sciences, Synthetic and Functional Biomolecules Center, and Peking-Tsinghua Center for Life Sciences
Peking University, Beijing 100871 (China)
E-mail: chengqi.yi@pku.edu.cn

[**] This work was supported by the National Natural Science Foundation of China (No. 31270838) and the National Basic Research Foundation of China (No. 2014CB964900). We also thank the Shanghai Synchrotron Radiation Facility (BL17U) and Dr. F. Yu for assistance with X-ray data collection.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201310050.

structures. Such limited knowledge makes it difficult to dissect the molecular basis of their substrate recognition and also presents challenges to efficiently discriminate these closely related demethylases for functional studies and therapeutic efforts.

We present here a detailed study to understand the molecular basis of the demethylation specificity of AlkB, FTO, and ALKBH5 by using both biochemical approaches and X-ray crystallography. Inspired by the observations that AlkB and ALKBH2 use acidic residues (D135 in AlkB and E175 in ALKBH2) in the active site to recognize the exocyclic amino group of the positively charged m^1dA (see Figure S1 in the Supporting Information),[12a,c,g] we focused our attention on a loop region (between β6 and β7) immediately C-terminal to the invariant HXD motif that serves as ligands for Fe^{II} ions (Figure 1b, and see Figure S2 in the Supporting Information). A multiple sequence alignment shows that the amino acid composition varies significantly within this loop region. Particularly, for demethylases that work efficiently for the positively charged m^1dA/m^1A (AlkB, ALKBH2, and ALKBH3), negatively charged residues are found at the equivalent positions of AlkB D135, while amino acids with neutral side chains are often present for demethylases that act on the neutral N^6-methyladenine (Figure 1b). Thus, we envisioned that such divergent amino acid sequences may play a role in determining the optimal substrate recognized by these demethylases.

We chose methylated ssDNA and ssRNA for evaluation of the activity since AlkB, FTO, and ALKBH5 all prefer single-stranded substrates. We first generated single or double mutants of AlkB at the D135 and/or E136 positions by using amino acids at the equivalent positions of FTO or ALKBH5. AlkB had been previously shown to exhibit in vitro activity towards m^6dA,[13] although the efficiency is much lower than to m^1dA. We envisioned that such mutations could allow better accommodation of an N^6-methyladenine base in the active site, thereby enhancing the demethylation efficiency of AlkB towards m^6dA. To our delight, many of these mutant AlkBs indeed display better activity to m^6dA compared to wild-type AlkB (wtAlkB), with the two mutant AlkBs D135I and E136L showing the greatest improvement (Figure 2a,b). A similar trend of demethylation enhancement is also observed for N^6-methyladenine in the context of a 15 mer ssRNA with the same sequence (see Figure S3 in the Supporting Information). It is noticeable that the mutation of D135 to leucine (which is also neutral, but differs from the corresponding sequence of FTO and ALKBH5) shows no noticeable change in the m^6dA demethylation activity compared to wtAlkB (see Figure S4 in the Supporting Information), thereby arguing against the notion that any mutation introduced into the loop region can enhance the m^6dA demethylation activity of AlkB. All the AlkB mutants we examined have decreased activity towards m^1dA, as anticipated (see Figure S5 in the Supporting Information).

We next tested the possibility of enhancing the demethylation activity to m^1dA by replacing amino acids in the loop region of FTO and ALKBH5. Both FTO and ALKBH5 were reported to exhibit only negligible m^1dA demethylation activity in their wild-type forms.[5] As we had hoped, substituting just one or two amino acids in the loop region of FTO and ALKBH5 with sequences from either AlkB or ALKBH2 can significantly increase the demethylation activity towards m^1dA (Figure 2c,d), and meanwhile decrease the activity to m^6dA (see Figure S6 in the Supporting Information). The most noticeable enhancement was found with the FTO N235D/L236E double mutant and L236R single mutant, as well as the ALKBH5 I209D and I209E single mutants. Swapping active-site residues between these demethylases thus results in partially switched demethylation activities.

To elucidate the roles of these active-site residues at the molecular level, we solved the crystal structure of AlkB bound to m^6dA-containing duplex DNA and compared this

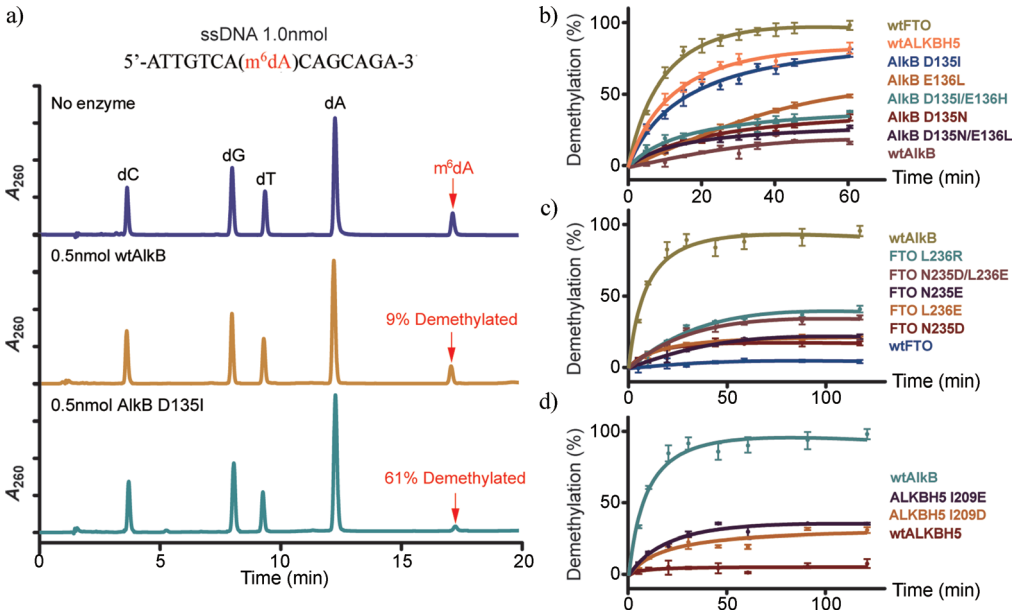

Figure 2. Swapping active-site residues between AlkB, FTO, and ALKBH5 results in switched demethylation activities. a) Typical HPLC traces of digested m^6dA-containing substrates, showing the enhanced m^6dA demethylation activity of AlkB D135I. Repair curves of b) mutant AlkBs with increased m^6dA demethylation activity as well as c) mutant FTOs and d) mutant ALKBH5s with increased m^1dA demethylation activity. The activities of wild-type enzymes to the cognate substrate are plotted as well for comparison; specific enzymes and their corresponding activity curves are shown in the same color. All the demethylation experiments shown were carried out in triplicate.

structure to previously reported structures of AlkB with a cognate m^1dA substrate (Figure 3a,b). Overlay of the m^6dA-wtAlkB structure with the m^1dA-wtAlkB structure (PDB ID: 3BIE) shows that the two structures are almost identical (rmsd of ca. 0.28 Å), with a major conformational change found in a flexible loop that caps the flipped base (Figure 3b). A closer examination reveals that this flexible loop (K134 to L139) coincides with the active-site region highlighted in Figure 1b. In the m^1dA structure, D135 forms a crucial hydrogen bond with the amino group of m^1dA, and E136 also projects towards m^1dA (see Figure S1 in the Supporting Information). In the case when m^6dA is bound, both residues point away from m^6dA (Figure 3b): the side chain of E136 is now fully exposed to solvent and, very interestingly, D135 forms a salt bridge with R183 (see Figure S7 in the Supporting Information). Additionally, R210, a characteristic residue of the AlkB family dioxygenases, now occupies the original position of E136 and is, together with Y78, within van der Waals contact with the methyl group of m^6dA (Figure 3c). The stacking interaction between W69 and H131 is preserved for both m^1dA and m^6dA. Thus, the flipped m^6dA base can still be recognized by AlkB; however, compared to the optimal binding of m^1dA, the recognition of m^6dA is less tight.

We further solved the m^6dA-bound structures of several AlkB mutants (D135I, E136L, and D135I/E136H) which have the swapped sequences from FTO or ALKBH5 and hence improved demethylation activity towards m^6dA and m^6A. The exact conformations of the mutated active-site loop vary between these structures, and they also differ from those seen in wtAlkB/m^6dA and wtAlkB/m^1dA structures (see Figure S8 in the Supporting Information). Although direct interactions between the mutated amino acids and the N^6-methyladenine base are not found, a common feature is that the introduced mutations all changed the original conformation of the loop in the wtAlkB/m^6dA structure, which appears unfavorable for m^6dA recognition. Therefore, we postulate that in addition to discriminating between m^1dA and m^6dA, the active-site loop of AlkB, when equipped with one or two mutations, could also allow better accommodation of the noncognate substrate m^6dA.

Finally we performed a detailed kinetic analysis to characterize the m^6A demethylation reactions catalyzed by wild-type and mutant AlkBs (see Figures S9 and S10 in the Supporting Information). For a fair comparison, we chose the exact m^6A-containing RNA sequence used previously to evaluate the demethylation kinetics of FTO and ALKBH5.[6,7] Under the conditions used, FTO and ALKBH5 are approximately 90-fold and 15-fold more efficient, respectively, than wtAlkB in terms of catalyzing m^6A demethylation (Table 1). By making one single mutation (D135I or E136L), we increased the catalytic efficiency of AlkB to about one third of that of ALKBH5. More interestingly, the turnover numbers of wild-type and mutant AlkBs towards m^6A are very similar (see Table 1). It appears to us that the increased demethylation efficiency results fully from the K_M value, an inverse measure of the substrate's affinity for the enzyme. Together with our crystallographic observations, we conclude that the improved recognition of N^6-methyladenine by the mutant AlkBs are mainly responsible for the observed enhancement in demethylation efficiency.

In summary, we have identified an active-site region in AlkB (and potentially ALKBH2), FTO, and ALKBH5 which is critical for substrate recognition and demethylation specificity. Swapping the active-site sequences between the demethylases results in partially exchanged demethylation activities. Combined evidence from crystallographic observations

Table 1: Kinetic constants of AlkBs, FTO, and ALKBH5.

Enzyme	Substrate[a]	K_M [μM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$min^{-1}\,\mu M^{-1}$]
AlkB	m^1dA	2.00 ± 0.35	7.41 ± 0.47	3.71
AlkB	m^1A	2.32 ± 0.31	3.72 ± 0.19	1.60
FTO	m^6A	0.60 ± 0.12	0.381 ± 0.114	0.634
ALKBH5	m^6A	1.66 ± 0.16	0.174 ± 0.008	0.105
AlkB	m^6A	14.93 ± 2.46	0.107 ± 0.013	0.007
AlkB D135N	m^6A	5.77 ± 0.84	0.103 ± 0.008	0.018
AlkB E136L	m^6A	3.47 ± 0.37	0.104 ± 0.005	0.030
AlkB D135I	m^6A	2.76 ± 0.34	0.097 ± 0.005	0.035

[a] All methylated adenine bases are part of a 15 mer ssDNA/ssRNA sequence.

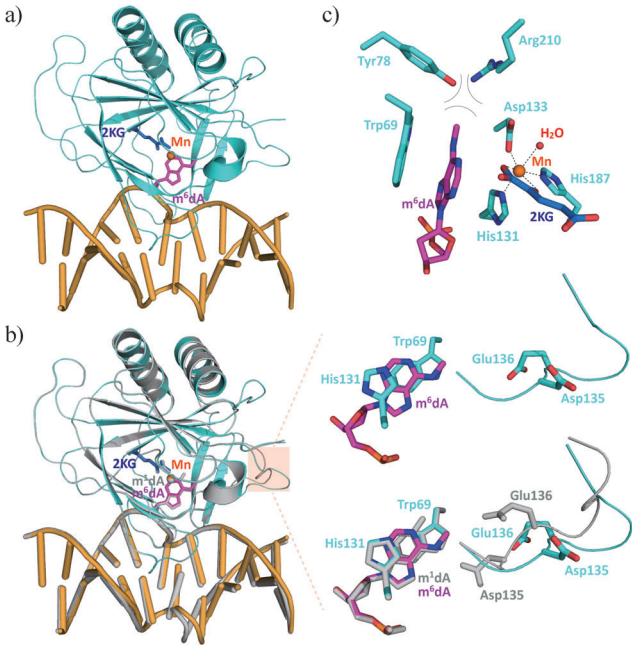

Figure 3. m^6dA recognition by AlkB. Protein is colored in cyan, m^6dA in magenta, and DNA in orange. a) Overall structure of m^6dA-wtAlkB. b) Overlay of m^6dA-wtAlkB and m^1dA-wtAlkB (3BIE). The pink box highlights the conformational change of the active-site loop of AlkB to recognize m^6dA, with a zoom-in view shown on the right. c) Detailed interactions in the active-site of AlkB to accommodate m^6dA. The black curves represent van der Waals contacts and dotted lines represent metal coordination.

and kinetic studies indicates a role of the active-site residues in substrate recognition. As selective inhibition of the AlkB family demethylases is of particular interest (in terms of developing both functional probes for research and potential lead compounds for therapeutics),[14] the divergent active-site sequences identified here provide opportunities for the design of selective inhibitors that are capable of discriminating between these closely related demethylases.

Received: November 19, 2013
Revised: January 4, 2014
Published online: March 5, 2014

Keywords: protein structures · RNA recognition · RNA/DNA demethylases · selective inhibition · substrate specificity

[1] a) W. A. Pastor, L. Aravind, A. Rao, *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 341 – 356; b) Y. Fu, C. He, *Curr. Opin. Chem. Biol.* **2012**, *16*, 516 – 524.

[2] a) P. O. Falnes, R. F. Johansen, E. Seeberg, *Nature* **2002**, *419*, 178 – 182; b) S. C. Trewick, T. F. Henshaw, R. P. Hausinger, T. Lindahl, B. Sedgwick, *Nature* **2002**, *419*, 174 – 178; c) P. A. Aas et al., *Nature* **2003**, *421*, 859 – 863; d) R. Ougland, C. M. Zhang, A. Liiv, R. F. Johansen, E. Seeberg, Y. M. Hou, J. Remme, P. O. Falnes, *Mol. Cell* **2004**, *16*, 107 – 116; e) P. Falnes, M. Bjørås, P. A. Aas, O. Sundheim, E. Seeberg, *Nucleic Acids Res.* **2004**, *32*, 3456 – 3461.

[3] a) J. Ringvoll, L. M. Nordstrand et al., *EMBO J.* **2006**, *25*, 2189 – 2198; b) T. Duncan, S. C. Trewick, P. Koivisto, P. A. Bates, T. Lindahl, B. Sedgwick, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16660 – 16665; c) S. Dango, N. Mosammaparast, M. E. Sowa, L. J. Xiong, F. Wu, K. Park, M. Rubin, S. Gygi, J. W. Harper, Y. Shi, *Mol. Cell* **2011**, *44*, 373 – 384.

[4] a) T. M. Frayling et al., *Science* **2007**, *316*, 889 – 894; b) C. Dina et al., *Nat. Genet.* **2007**, *39*, 724 – 726; c) A. Scuteri et al., *PLoS Genet.* **2007**, *3*, e115; d) J. Fischer, L. Koch, C. Emmerling, J. Vierkotten, T. Peters, J. C. Brüning, U. Rüther, *Nature* **2009**, *458*, 894 – 898; e) C. Church et al., *Nat. Genet.* **2010**, *42*, 1086 – 1092.

[5] a) T. Gerken et al., *Science* **2007**, *318*, 1469 – 1472; b) G. Jia, C. G. Yang, S. Yang, X. Jian, C. Yi, Z. Zhou, C. He, *FEBS Lett.* **2008**, *582*, 3313 – 3319.

[6] G. Jia, Y. Fu, X. Zhao, Q. Dai, G. Zheng, Y. Yang, C. Yi, T. Lindahl, T. Pan, Y. G. Yang, C. He, *Nat. Chem. Biol.* **2011**, *7*, 885 – 887.

[7] G. Zheng, J. A. Dahl et al., *Mol. Cell* **2013**, *49*, 18 – 29.

[8] a) B. Sedgwick, *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 148 – 157; b) F. Drabløs, E. Feyzi, P. A. Aas, C. B. Vaagbo, B. Kavli, M. S. Bratlie, J. Pena-Diaz, M. Otterlei, G. Slupphaug, H. E. Krokan, *DNA Repair* **2004**, *3*, 1389 – 1407; c) Y. Mishina, E. M. Duguid, C. He, *Chem. Rev.* **2006**, *106*, 215 – 232.

[9] a) K. D. Meyer, Y. Saletore, P. Zumbo, O. Elemento, C. E. Mason, S. R. Jaffrey, *Cell* **2012**, *149*, 1635 – 1646; b) G. Jia, Y. Fu, C. He, *Trends Genet.* **2013**, *29*, 108 – 115; c) G. Zheng, J. A. Dahl, Y. Niu, Y. Fu, A. Klungland, Y. G. Yang, C. He, *RNA Biol.* **2013**, *10*, 915 – 918; d) D. Dominissini et al., *Nature* **2012**, *485*, 201 – 206.

[10] Z. Han, T. Niu, J. Chang, X. Lei, M. Zhao, Q. Wang, W. Cheng, J. Wang, Y. Feng, J. Chai, *Nature* **2010**, *464*, 1205 – 1209.

[11] O. Sundheim, C. B. Vagbo, M. Bjoras, M. M. Sousa, V. Talstad, P. A. Aas, F. Drablos, H. E. Krokan, J. A. Tainer, G. Slupphaug, *EMBO J.* **2006**, *25*, 3389 – 3397.

[12] a) B. Yu, W. C. Edstrom, J. Benach, Y. Hamuro, P. C. Weber, B. R. Gibney, J. F. Hunt, *Nature* **2006**, *439*, 879 – 884; b) C. G. Yang, C. Yi, E. M. Duguid, C. T. Sullivan, X. Jian, P. A. Rice, C. He, *Nature* **2008**, *452*, 961 – 965; c) C. Yi, G. Jia, G. Hou, Q. Dai, W. Zhang, G. Zheng, X. Jian, C. G. Yang, Q. Cui, C. He, *Nature* **2010**, *468*, 330 – 333; d) C. Yi et al., *Nat. Struct. Mol. Biol.* **2012**, *19*, 671 – 676; e) P. J. Holland, T. Hollis, *PLoS One* **2010**, *5*, e8680; f) B. Yu, J. F. Hunt, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 14315 – 14320; g) A. M. Maciejewska et al., *J. Biol. Chem.* **2013**, *288*, 432 – 441.

[13] D. Li, J. C. Delaney, C. M. Page, X. Yang, A. S. Chen, C. Wong, C. L. Drennan, J. M. Essigmann, *J. Am. Chem. Soc.* **2012**, *134*, 8896 – 8901.

[14] a) B. Chen et al., *J. Am. Chem. Soc.* **2012**, *134*, 17963 – 17971; b) W. Aik, M. Demetriades, M. K. Hamdan, E. A. Bagg, K. K. Yeoh, C. Lejeune, Z. Zhang, M. A. McDonough, C. J. Schofield, *J. Med. Chem.* **2013**, *56*, 3680 – 3688.