

SUPPLEMENTARY INFORMATION FOR

Thymine DNA glycosylase recognizes the geometry alternation of minor groove induced by 5-formylcytosine and 5-carboxylcytosine

Tianran Fu^{1,2§}, Liping Liu^{3,4§}, Qing-Lin Yang^{5§}, Yuxin Wang^{1,2}, Pan Xu^{3,4}, Lin Zhang^{1,2}, Shien Liu^{3,4}, Qing Dai⁶, Qianjiang Ji⁷, Guo-Liang Xu⁵, Chuan He⁶, Cheng Luo^{3,4*} and Liang Zhang^{1,2*}

¹Department of Pharmacology and Chemical Biology, Shanghai Jiao Tong University School of Medicine, Shanghai, P. R. China; ²Shanghai Universities Collaborative Innovation Center for Translational Medicine, Shanghai, P. R. China; ³CAS Key Laboratory of Receptor Research, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China; ⁴University of Chinese Academy of Sciences, Beijing 100049, China; ⁵State Key Laboratory of Molecular Biology, Chinese Academy of Sciences Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai, China. ⁶Department of Chemistry, Department of Biochemistry and Molecular Biology, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois, USA. ⁷School of Physical Science and Technology, ShanghaiTech University, Shanghai, China.

§These authors contributed equally to this work.

*Correspondence: Liang Zhang: liangzhang2014@sjtu.edu.cn; Cheng Luo: cluo@simmm.ac.cn;

This PDF file includes:

Supplementary Table S1 and S2; Supplementary Figures S1 to S3;

Table S1. Data Collection and Refinement Statistics

	5mC-dsDNA	5hmC-dsDNA
Data collection		
Space group	C2	C2
Cell dimensions		
a, b, c(Å)	31.606, 25.788, 34.215	63.907, 23.818, 85.556
α , β , γ (°)	90.00, 115.934, 90.00	90.00, 111.117, 90.00
Wavelength (Å)	0.9795	0.9795
Resolution (Å)*	50-1.40 (1.45-1.40)	50-2.85 (2.95-2.85)
R _{merge} (%)	5.7 (7.9)	10.1 (52.2)
I/ σ I	29.5 (24.3)	16.7 (3.3)
Completeness (%)	99.1 (100)	99.9 (99.3)
Redundancy	6.6 (6.7)	6.2(5.6)
Refinement		
Resolution (Å)	30.7-1.40	26.6-2.85
No.reflections	4932	2611
R _{work} /R _{free}	0.182/0.217	0.231/0.282
No.atoms		
Nucleic acids	203	816

Water	86	N/A
Ligand/ion	N/A	N/A
B-factors		
Nucleic acids	15.676	41.246
Water	27.308	N/A
Ligand/ion	N/A	N/A
R.m.s deviations		
Bond lengths (Å)	0.005	0.005
Bond angles (°)	0.933	1.423

*Highest-resolution shell is shown in parentheses.

Table S2. K_{\max} of wild-type TDG and TDG R275A in excision of uracil containing substrate.

Enzyme	Substrate	k_{\max} (min^{-1})	$K_{\max(\text{WT})} / K_{\max}(\text{Mut})$
TDG	U·G	0.419 ± 0.013	/
TDG S200A	U·G	0.120 ± 0.004	3.49
TDG K201A	U·G	0.092 ± 0.005	4.55
TDG D202A	U·G	0.126 ± 0.011	3.33
TDG R275A	U·G	0.010 ± 0.002	41.9

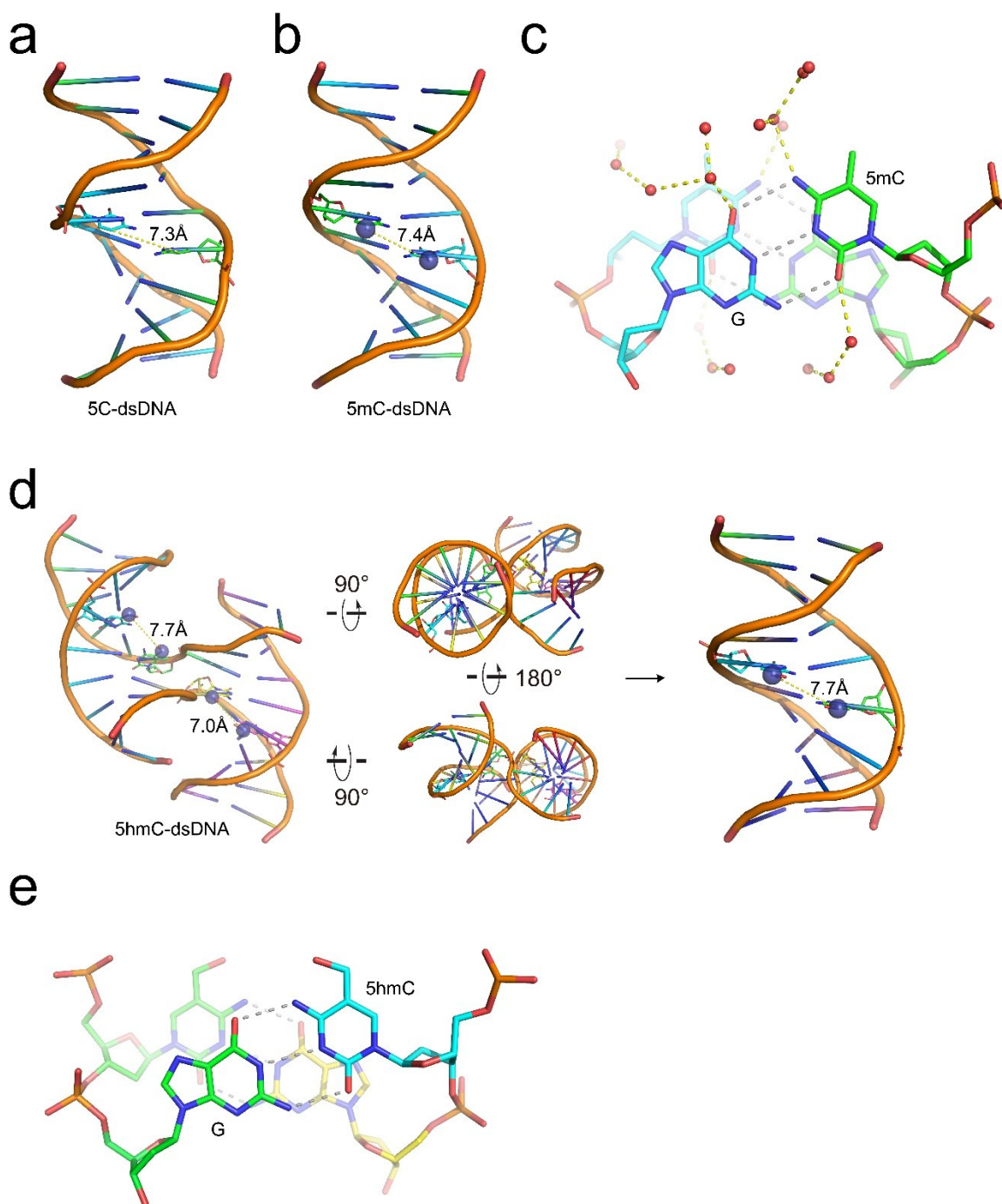


Figure S1. Figure 1 | Schematic diagram of 5C-dsDNA, 5mC-dsDNA and 5hmC-dsDNA crystal structures. (a) Overall structure of 5C-dsDNA (pdb code: 1EN9). Two complementary strands are colored in green and cyan, respectively. (b) Overall structure of 5mC-dsDNA. Two complementary strands are colored in green and cyan, respectively.

The fully symmetric 5mC bases are shown as sticks and their C5m atoms are shown as purple spheres. The distance between two C5m atoms is shown as yellow dashes. **(c)** The hydrogen bond networks around 5mC base pairing. The waters are shown in red dots. The hydrogen bonds are shown as yellow dashes. **(d)** Overall structure of 5hmC-dsDNA. Four complementary strands are colored in green, cyan, yellow and magenta, respectively. The overall structure is rotated to be viewed from the top or from the bottom. The regular B-formed dsDNA generated from the structure is also shown. The fully symmetric 5hmC bases are shown as sticks and their C5m atoms are shown as purple spheres. The distance between two C5m atoms is shown as yellow dashes. **(e)** The hydrogen bond networks within the 5hmC base pairing.

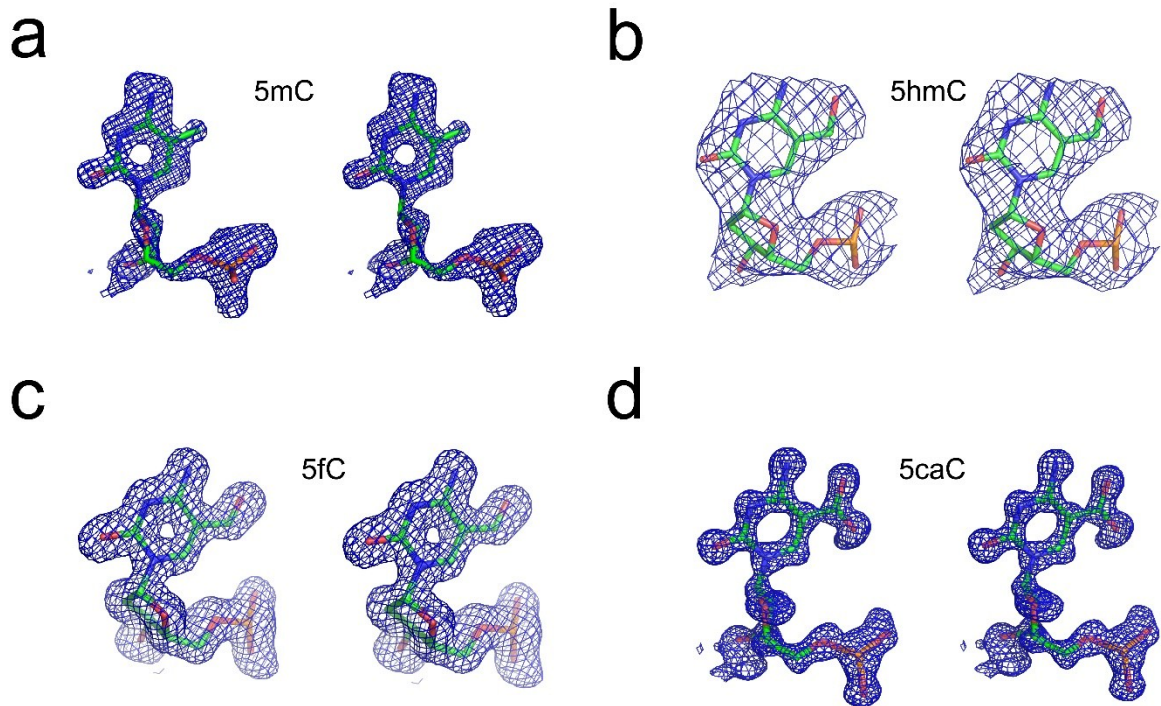


Figure S2. Stereo view of fofc electron density map contoured at 2.0 σ around 5mC (a), 5hmC (b), 5fC (c) and 5caC (d).

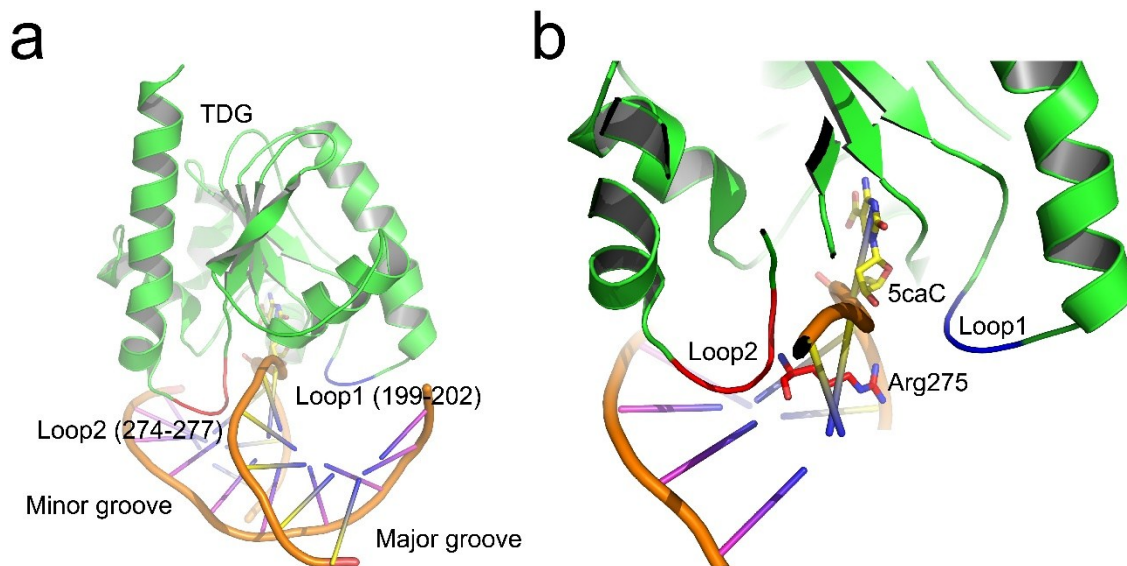


Figure S3. Schematic diagram of TDG-DNA structure. TDG and two DNA strands are colored in green, yellow and magenta respectively. 5caC base and residue R275 are shown as sticks. Two loops involved in DNA recognition are colored in blue and red respectively (PDB entry 3UO7).