

## SUPPLEMENTARY INFORMATION FOR

# **A TET homolog protein from *Coprinopsis cinerea* (CcTET) that biochemically converts 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine, and 5-carboxylcytosine**

Liang Zhang<sup>†‡</sup>, Weizhong Chen<sup>‡</sup>, Lakshminarayan M. Iyer<sup>§</sup>, Jennifer Hu<sup>†</sup>, Gloria Wang<sup>†</sup>, Ye Fu<sup>†</sup>, Miao Yu<sup>†</sup>, Qing Dai<sup>†</sup>, L. Aravind<sup>§</sup> and Chuan He<sup>\*†</sup>.

<sup>†</sup>Department of Chemistry and Institute for Biophysical Dynamics, The University of Chicago, 929 E. 57th Street, Chicago, Illinois 60637, USA.

<sup>§</sup>Department of Chemical Physics, University of Science and Technology of China, Hefei, Anhui 230026, China.

<sup>‡</sup>National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA.

*\*Corresponding author: Chuan He, [chuanhe@uchicago.edu](mailto:chuanhe@uchicago.edu), Tel: 1-773-702-5061*

This PDF file includes:

Experimental Method

Supplementary Results (Supplementary Table S1, Figure S1 to Figure S16)

Supplementary References

## Experimental Method

*Sequence analysis.* TET/JBP homologs were retrieved using sequence profile searches with the PSI-BLAST program<sup>1</sup> run against the non-redundant protein database of National Center for Biotechnology Information (NCBI), using previously described TET/JBP sequences as starting queries<sup>2,3</sup>. A comprehensive multiple sequence alignment of both prokaryotic and eukaryotic homologs was generated using the Kalign2 program<sup>4</sup>. The JPred program was used to predict secondary structures and this was used for further manual adjustment of the multiple sequence alignment<sup>5</sup>. A sequence profile generated from this alignment was used as seed to search the proteome of *Coprinopsis cinerea* strain Okayama 7#130 using the command-line PSI-BLAST program with the -B option. The retrieved sequences were manually inspected for completeness and active site residues based on which a set of 12 homologs with intact active sites were identified for further study.

*CcTET subcloning, expression and purification.* The 12 TET homolog candidate genes from *Coprinopsis cinerea* strain Okayama 7 (#130) was subcloned to PMCGS19 plasmid and expressed in BL21(DE3) cells containing vector pRK1037<sup>6</sup>. The strain grew in LB media supplemented with 100 ug/ml of ampicillin and 50 µg/ml kanamycin at 37 °C. When OD600 reached 0.6, 1 mM isopropyl-b-D-thiogalactopyranoside (IPTG) was added into the culture to induce the protein expression under room temperature for overnight. The cells were harvested by centrifugation at 6000 g for 10 min at 4 °C, and suspended in Ni-NTA buffer A (20 mM HEPES, pH 8.0, 500 mM NaCl, 10 mM imidazole, and 1 mM DTT). After sonication and high speed centrifuge (16,000 g, 30 min, 4 °C), the supernatant was loaded into 5ml Ni-NTA column (Qiagen) which has been pre-equilibrated with Ni-NTA buffer A. The column was washed with eight column volumes of buffer A, and eluted with three column volumes of Ni-NTA buffer B (20 mM HEPES, pH 8.0, 500 mM NaCl, 500 mM imidazole, and 1 mM DTT). The elution fraction was pooled and concentrated to 300 µl, and then diluted to 2 ml with monoS buffer A (20 mM HEPES, pH 8.0 and 1 mM DTT). The 2 ml fraction was load to the monoS column which has been pre-equilibrated with 6% monoS buffer B (20 mM HEPES, pH 8.0, 1 M NaCl and 1 mM DTT), and eluted with gradient percentage of monoS buffer B. the fractions were pooled and concentrated to 2 ml again and load to Superdex75 column with GF buffer (20 mM HEPES, pH 8.0, 150 mM NaCl and 1 mM DTT). The fractions were pooled and concentrated.

The purity of the proteins was determined on 13% SDS-PAGE. The final purified protein was quantified by Bradford reagent purchased from Bio-Rad, flash-frozen with 30% glycerol and stored at -80 °C.

*mTET1 subcloning, expression, and purification.* The catalytic domain of mouse *Tet1* gene (GU079948, residues 1367-2039) was cloned into *Bss*HIII and *Not*I sites of the N-terminal Flag-tagged pFastBac Dual vector (Invitrogen) and expressed in Bac-to-Bac baculovirus insect cell expression system. The recombinant Flag-mTET1 was purified according to the published procedure<sup>7,8</sup>. In general, the mTET1 protein was purified with anti-Flag M2 antibody agarose affinity gel purchased from Sigma-Aldrich, and further purified with GE Healthcare Superdex200 with running buffer containing 20 mM HEPES (pH 8.0), 150 mM NaCl, and 1 mM DTT. The fractions were pooled, concentrated, quantified by Bradford reagent purchased from Bio-Rad, and flash-frozen with 30% glycerol and stored at -80 °C.

*Oligonucleotide synthesis.* The oligonucleotides used in the biochemistry study were synthesized using solid-phase synthesis<sup>9</sup>. The reagents and phosphoramidites (including 5mC, 5hmC, 5fC and 5caC) were purchased from Glen Research and used in the solid-phase synthesis. All synthetic oligonucleotides were further purified with reverse-phase high-performance liquid chromatography. The oligonucleotides containing normal bases were purchased from Operon.

*CcTET and mouse TET1 (mTET1) oxidation activity assay (in the absence of ATP).* The activity assays were performed according to the published procedure at various pH (5.8, 7.0 and 8.0)<sup>10-12</sup>. A 9/11 mer duplex (or single) DNA containing one 5mCpG site on the 9mer was employed as the substrate. 20 pmol dsDNA (or ssDNA) substrate (final concentration 1 μM) and 40 pmol of recombinant protein (CcTET or mTET1, final concentration 2 μM) was added into 20 μl reaction mixture containing 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75 μM Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM α-KG, and 2.5 mM DTT for 30 min at 37 °C in triplicates (for CcTET activity assay, the reactions were processed at 25 °C and 37 °C). The reactions were quenched by heating to 100 °C for 5 min and immediately cooled in an ice bath. In reactions with ATP, 1.2 mM ATP was added to the reaction buffer.

*CcTET oxidation activity assay on A·T containing duplex DNA (in the absence of ATP).* The activity assays were performed according to the published procedure at various pH (5.8, 7.0 and 8.0)<sup>10-12</sup>. A 9/11 mer duplex DNA containing three A·T sites was employed as the substrate (Sequences: 5'-GACCGGAGT-3', 5'-GACTCCGGTCT-3'). 20 pmol dsDNA substrate (final concentration 1  $\mu$ M) and 40 pmol of CcTET (final concentration 2  $\mu$ M) was added into 20  $\mu$ l reaction mixture containing 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75  $\mu$ M Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT for 30 min at 25 °C in triplicates. The reactions were quenched by heating to 100 °C for 5 min and immediately cooled in an ice bath. The reaction products were digested and monitored by quantitative mass spectrometry (HPLC-MS/MS).

*Electrophoretic mobility shift assay (EMSA).* 50 nM 16mer dsDNA substrates (Sequences: 5'-TCTGGAACGGAATTCT-3', 5'-AAGAATTCXGTTCCAG-3'. X position indicates desired cytosine derived base including C, 5mC, 5hmC, 5fC and 5caC) and the various concentrations of CcTET were incubated at room temperature for 30 min, and then subjected to 8% polyacrylamide native PAGE (30 min pre-run at 100 V, and 80 min for sample separation at 100 V). The gels were stained with Sybr Green.

*Reaction progress curve of substrate and fraction products versus incubation time (20 min).* The reactions were conducted according to the CcTET oxidation activity assay (without ATP) at 37 °C. 20 pmol 5mC-containing dsDNA substrate (final concentration 1  $\mu$ M) and 4 pmol of recombinant CcTET protein (final concentration 0.2  $\mu$ M) were added into 20  $\mu$ l reaction mixture containing 50 mM HEPES, pH 7.0, 75  $\mu$ M Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT for various incubation time (0, 30, 60, 90, 120, 240, 360, 480, 600, 900 and 1200 seconds) at 37 °C. The reactions were quenched by heating to 100 °C for 5 min and immediately cooled in an ice bath.

*Reaction progress curve of substrate and fraction products versus incubation time (2.5 min).* The reactions were conducted using the CcTET oxidation activity assay (without ATP) at 37 °C. A 9/11 mer duplex DNA containing either one 5mCpG or 5hmCpG or 5fCpG site on the 9mer was employed as substrate. Various concentrations of dsDNA substrate (1, 2, 4 and 6  $\mu$ M for 5mC

and 5hmC; 0.5, 1, 2 and 6  $\mu\text{M}$  for 5fC) and recombinant CcTET protein (4 pmol CcTET for 5mC-containing substrate, final concentration of 0.2  $\mu\text{M}$ ; 5 pmol CcTET for 5hmC- and 5fC-containing substrate, final concentration of 0.25  $\mu\text{M}$ ) was added into 20  $\mu\text{l}$  reaction mixture containing 50 mM HEPES, pH 7.0, 75  $\mu\text{M}$   $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ , 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT for various incubation time (0, 30, 60, 90, 120, and 150 seconds) at 37  $^\circ\text{C}$  in triplicates. The reactions were quenched by heating to 100  $^\circ\text{C}$  for 5 min and immediately cooled in an ice bath.

*Kinetics of 5mC to 5hmC, 5hmC to 5fC, and 5fC to 5caC oxidation by CcTET in dsDNA.* The kinetics data was determined according to the previous published method<sup>13</sup>. In general, the  $C_{1/2V_{\max}}$  and  $V_{\max}$  values were made by keeping a constant CcTET enzyme concentration (0.2  $\mu\text{M}$  for 5mC, 0.25  $\mu\text{M}$  for 5hmC and 5fC respectively) and varying the substrate (5mC, 5hmC or 5fC containing 9/11mer dsDNA) concentration (1, 2, 4 and 6  $\mu\text{M}$  for 5mC and 5hmC; 0.5, 1, 2 and 6  $\mu\text{M}$  for 5fC). All reactions were performed at 37  $^\circ\text{C}$  within 2.5 min (0, 30, 60, 90, 120, and 150 seconds) in triplicates. The reactions were quenched by heating to 100  $^\circ\text{C}$  for 5 min and immediately cooled in an ice bath.

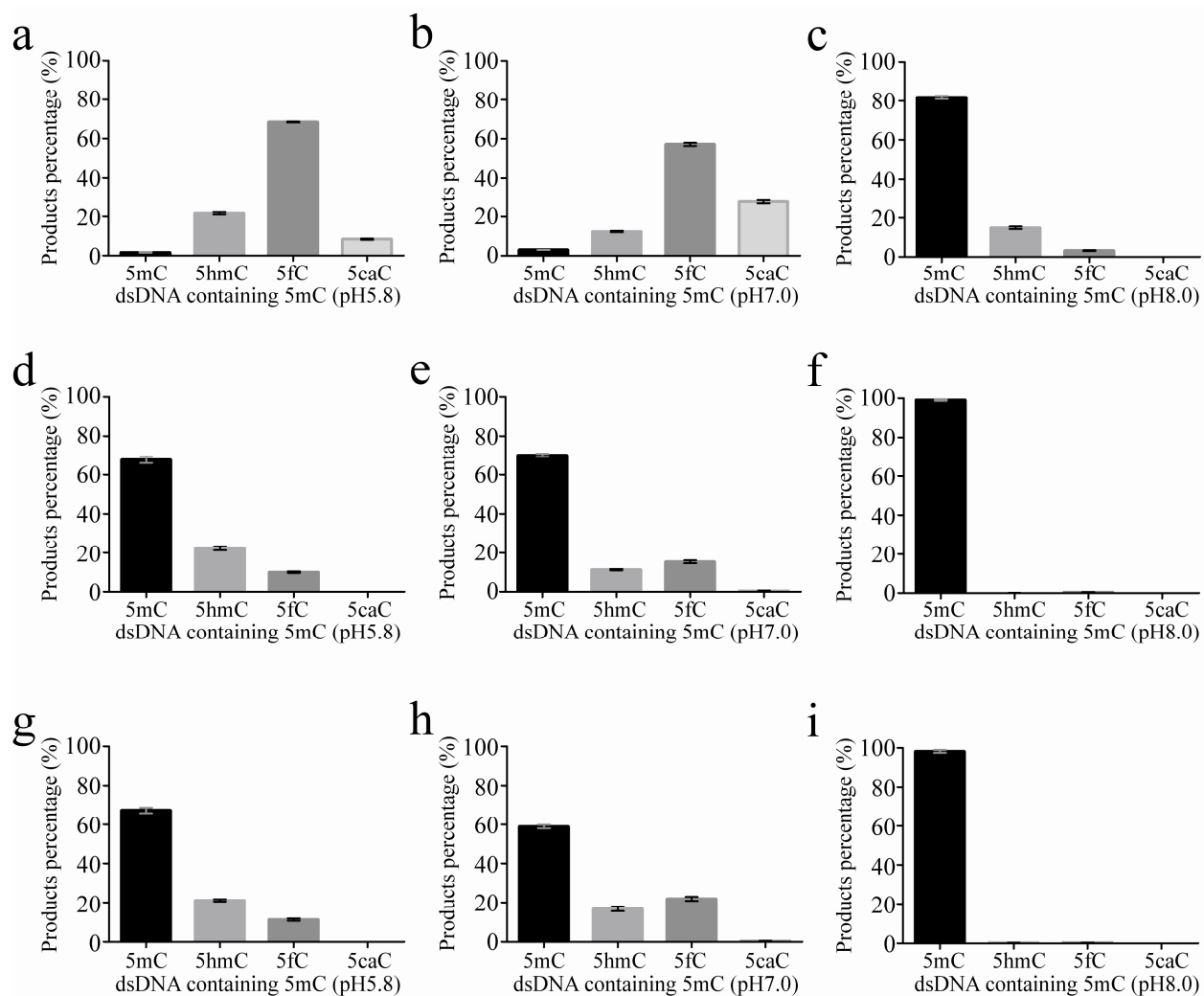
*LC-MS/MS detection and analysis.* The remaining substrate and oxidation products in the reactions were digested and monitored and validated by Agilent 6410 QQQ triple-quadrupole LC mass spectrometer (LC-MS/MS) in positive electrospray ionization mode. The quantitative analysis of the desired products level was processed according to the previous published methods<sup>13,14</sup>.

## Supplementary Results

Supplementary Table S1. kinetic constants of 5mC to 5hmC, 5hmC to 5fC, and 5fC to 5caC conversion catalyzed by CcTET on a dsDNA at 37 °C (Sequences: 5'-GACXGGAGT-3', 5'-GACTCCGGTCT-3', X position indicates desired cytosine derived base including 5mC, 5hmC, and 5fC).  $C_{1/2V_{max}}$  refers to the substrate concentration at which the reaction rate is at half-maximum;  $V_{max}$  refers to reaction maximum rate.

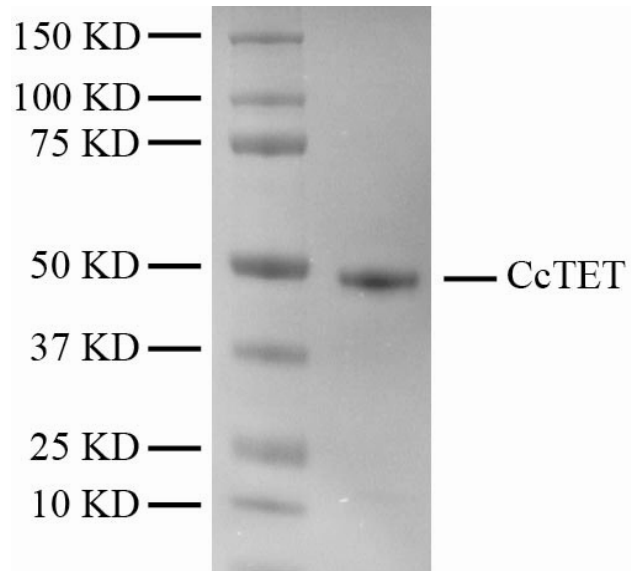
Substrate	pH	$C_{1/2V_{max}}$ ( $\mu$ M)	$V_{max}$ (M/s)
dsDNA (5mC)	7.0	1.70 ( $\pm$ 0.36)	$1.72 \times 10^{-8}$
dsDNA (5hmC)	7.0	1.72 ( $\pm$ 0.21)	$1.57 \times 10^{-8}$
dsDNA (5fC)	7.0	1.21 ( $\pm$ 0.21)	$1.02 \times 10^{-8}$





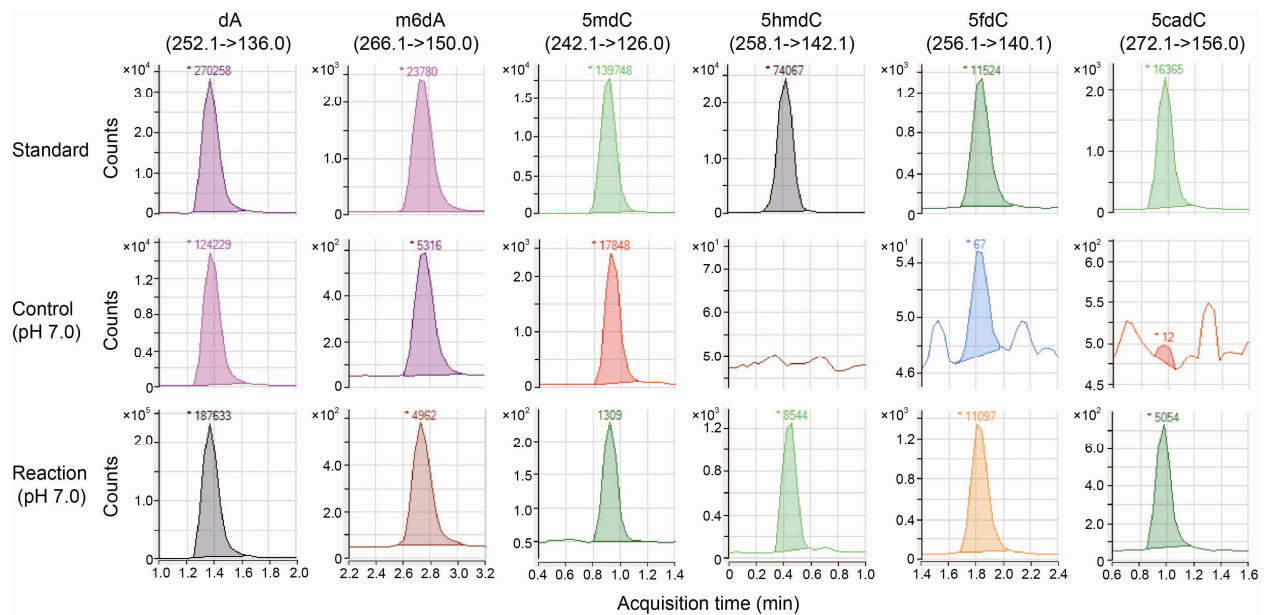
Supplementary Figure S2. 5mC oxidation assays (in the absence of ATP) at pH 5.8, 7.0 or 8.0 at 25 °C. The reaction mixture contains 1  $\mu$ M 5mC containing dsDNA substrate, 2  $\mu$ M protein (CcTET in a, b, c; CC1G\_10221 in d, e, f; CC1G\_02065 in g, h, i), 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75  $\mu$ M  $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ , 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT for 30 min at 25 °C in triplicates. The reactions were quenched by heating to 100 °C for 5 min and immediately cooled in an ice bath. The 5mC-containing dsDNA substrates were used (Sequences: 5'-GAC(5mC)GGAGT-3', 5'-GACTCCGGTCT-3').



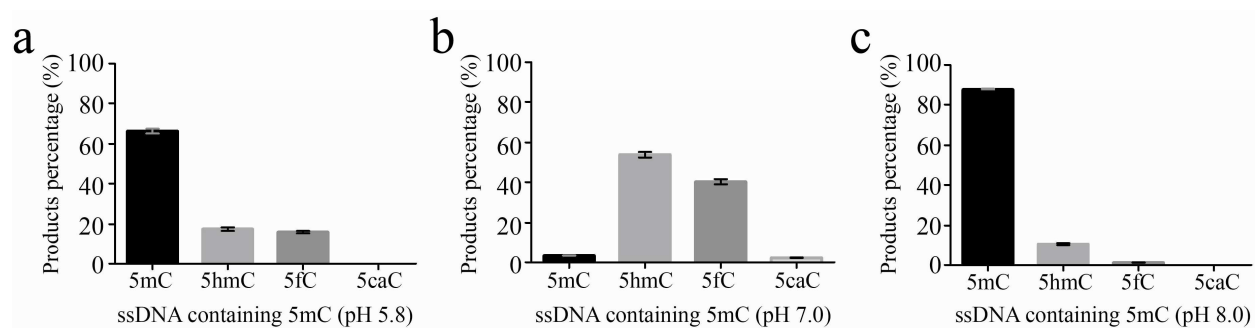


Supplementary Figure S3. The SDS page gel of purified CcTET protein.

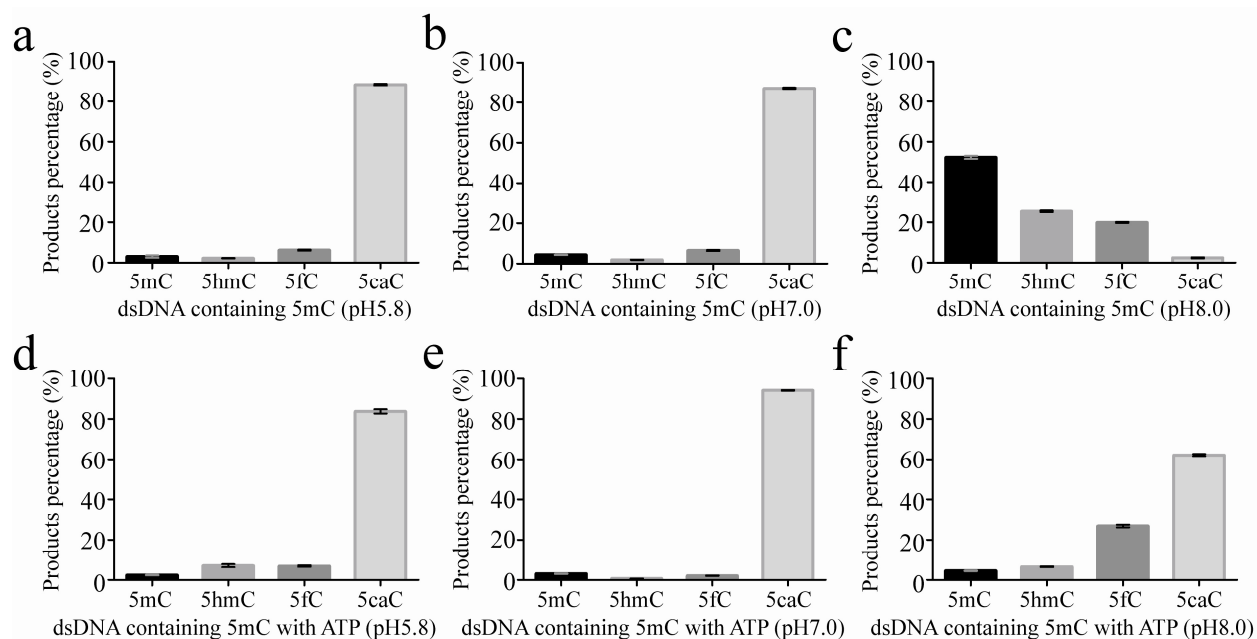




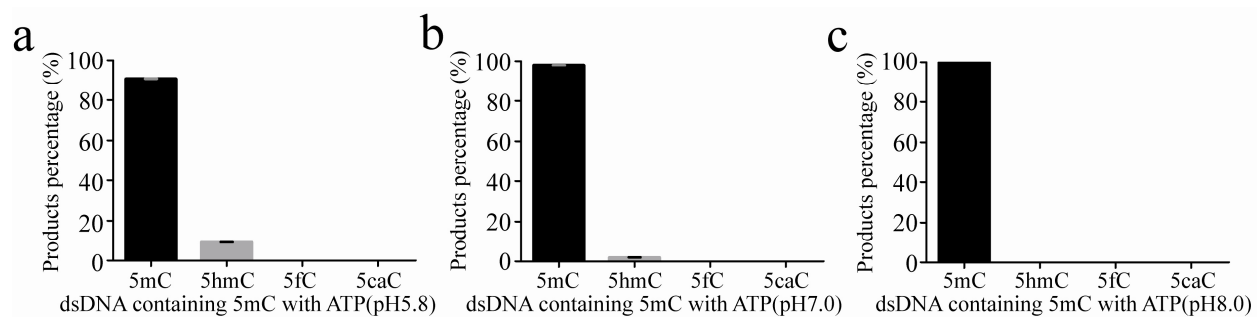
Supplementary Figure S5. Quantitative analysis of CcTET-mediated oxidation products under 37°C from 5mC-containing 9/11mer dsDNA using LC-MS/MS (Sequences: 5'-GAC(5mC)GGAGT-3', 5'-GACTCCGGTCT-3'). The nucleosides were quantified using the nucleoside to base ion mass transitions.



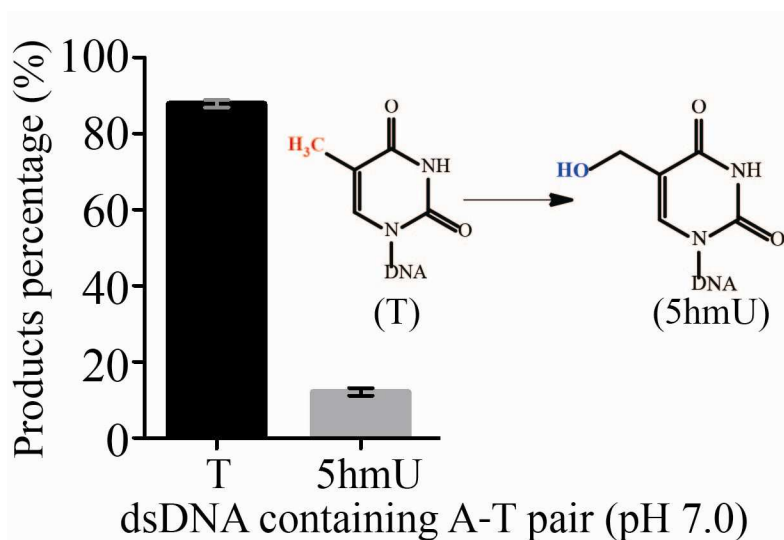
Supplementary Figure S6. CcTET oxidation assays (in the absence of ATP) with single-stranded DNA at pH 5.8, 7.0 or 8.0 at 37 °C. The reaction mixture contains 1  $\mu$ M 5mC containing ssDNA substrate, 2  $\mu$ M CcTET, 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75  $\mu$ M Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT for 30 min at 37 °C in triplicates. The reactions were quenched by being heated to 100 °C for 5 min and immediately cooled in an ice bath. The 5mC-containing ssDNA substrates were used (Sequences: 5'-GAC(5mC)GGAGT-3').



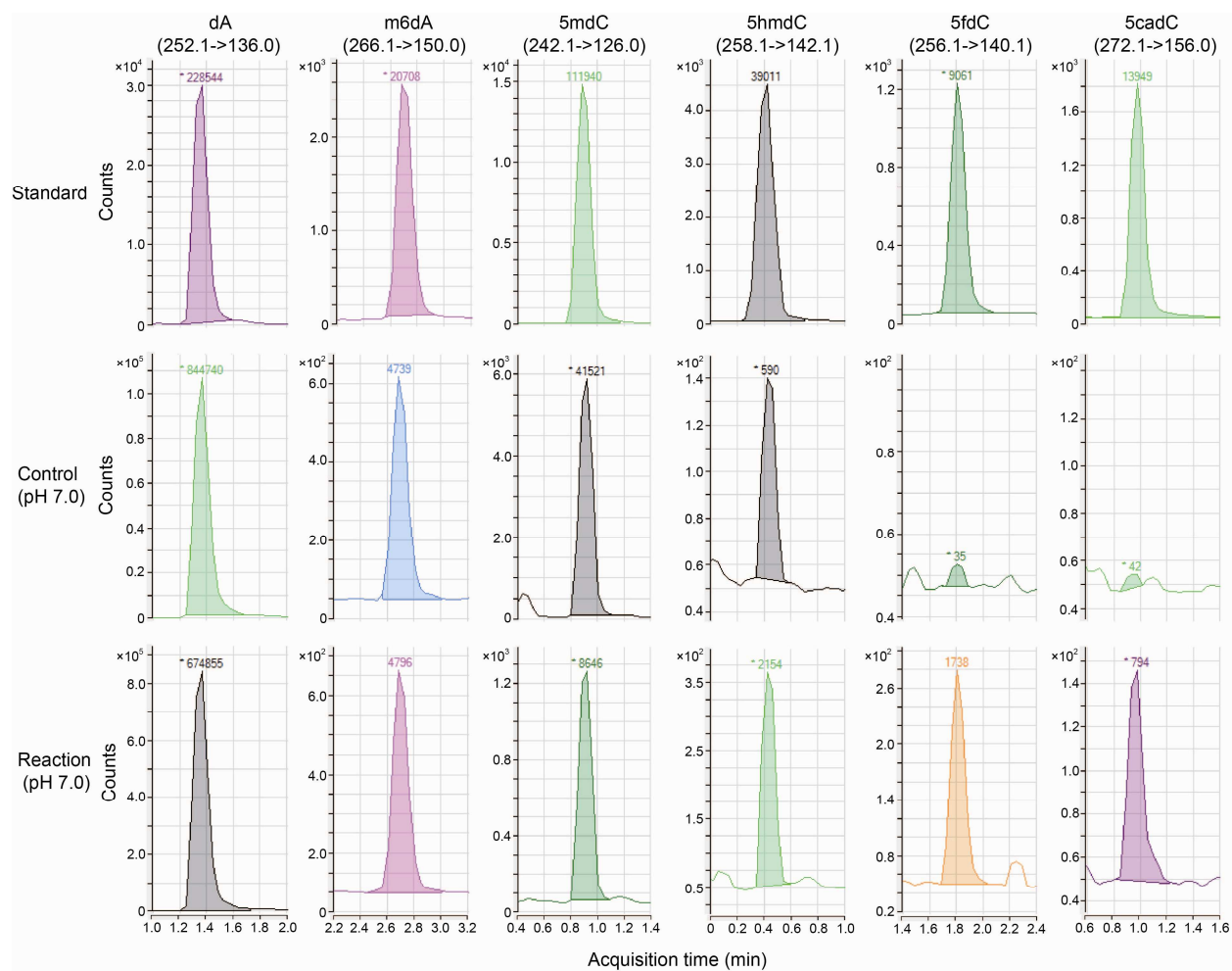
Supplementary Figure S7. Mouse TET1 oxidation assays with duplex DNA at pH 5.8, 7.0 or 8.0 at 37 °C in the absence of ATP (the reaction mixture contains 1  $\mu$ M 5mC containing dsDNA substrate, 2  $\mu$ M mTET1, 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75  $\mu$ M Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT ) (a, b, and c) or presence of ATP (the reaction mixture contains 1  $\mu$ M 5mC containing dsDNA substrate, 2  $\mu$ M mTET1, 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75  $\mu$ M Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, 100 mM NaCl, 1.2 mM ATP and 2.5 mM DTT) (d, e, and f) for 30 min at 37 °C in triplicates. The reactions were quenched by being heated to 100 °C for 5 min and immediately cooled in an ice bath. The 5mC-containing dsDNA substrates were used (Sequences: 5'-GAC(5mC)GGAGT-3', 5'-GACTCCGGTCT-3').



Supplementary Figure S8. CcTET oxidation assays (with ATP) with dsDNA at pH 5.8, 7.0 or 8.0 at 37 °C. The reaction mixture contains 1  $\mu$ M 5mC containing dsDNA substrate, 2  $\mu$ M CcTET, 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0 or pH 8.0), 75  $\mu$ M  $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ , 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, 100 mM NaCl, 1.2 mM ATP and 2.5 mM DTT for 30 min at 37 °C in triplicates. The reactions were quenched by being heated to 100 °C for 5 min and immediately cooled in an ice bath. The 5mC-containing dsDNA substrates were used (Sequences: 5'-GAC(5mC)GGAGT-3', 5'-GACTCCGGTCT-3').

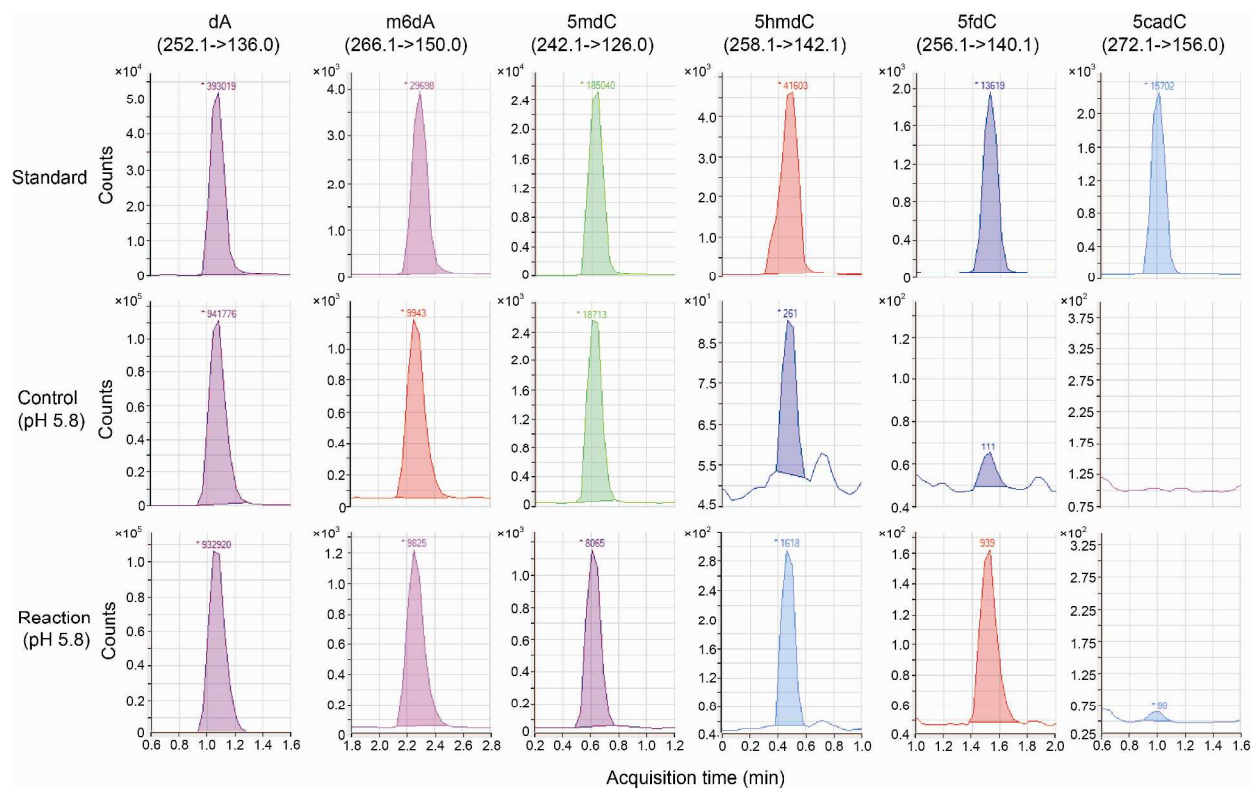


Supplementary Figure S9. CcTET oxidation assay with A·T containing duplex DNA at pH 7.0 at 25 °C in the absence of ATP (the reaction mixture contains 1 μM A·T containing dsDNA substrate, 2 μM CcTET, 50 mM HEPES, pH 7.0, 75 μM Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM α-KG, and 2.5 mM DTT ). The reactions were quenched by heating to 100 °C for 5 min and immediately cooled in an ice bath. The reaction products were digested and monitored by quantitative mass spectrometry (HPLC-MS/MS). The dsDNA substrates were used (Sequences: 5'-GACCGGAGT-3', 5'-GACTCCGGTCT-3').

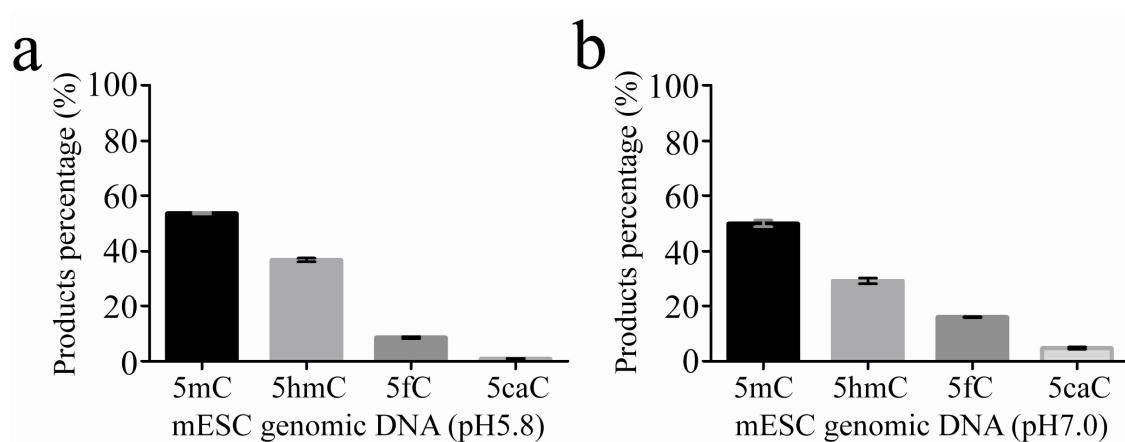


Supplementary Figure S10. Quantitative analysis of CcTET-mediated oxidation products from mESC genomic DNA at pH 7.0 using LC-MS/MS. The nucleosides were quantified using the nucleoside to base ion mass transitions.

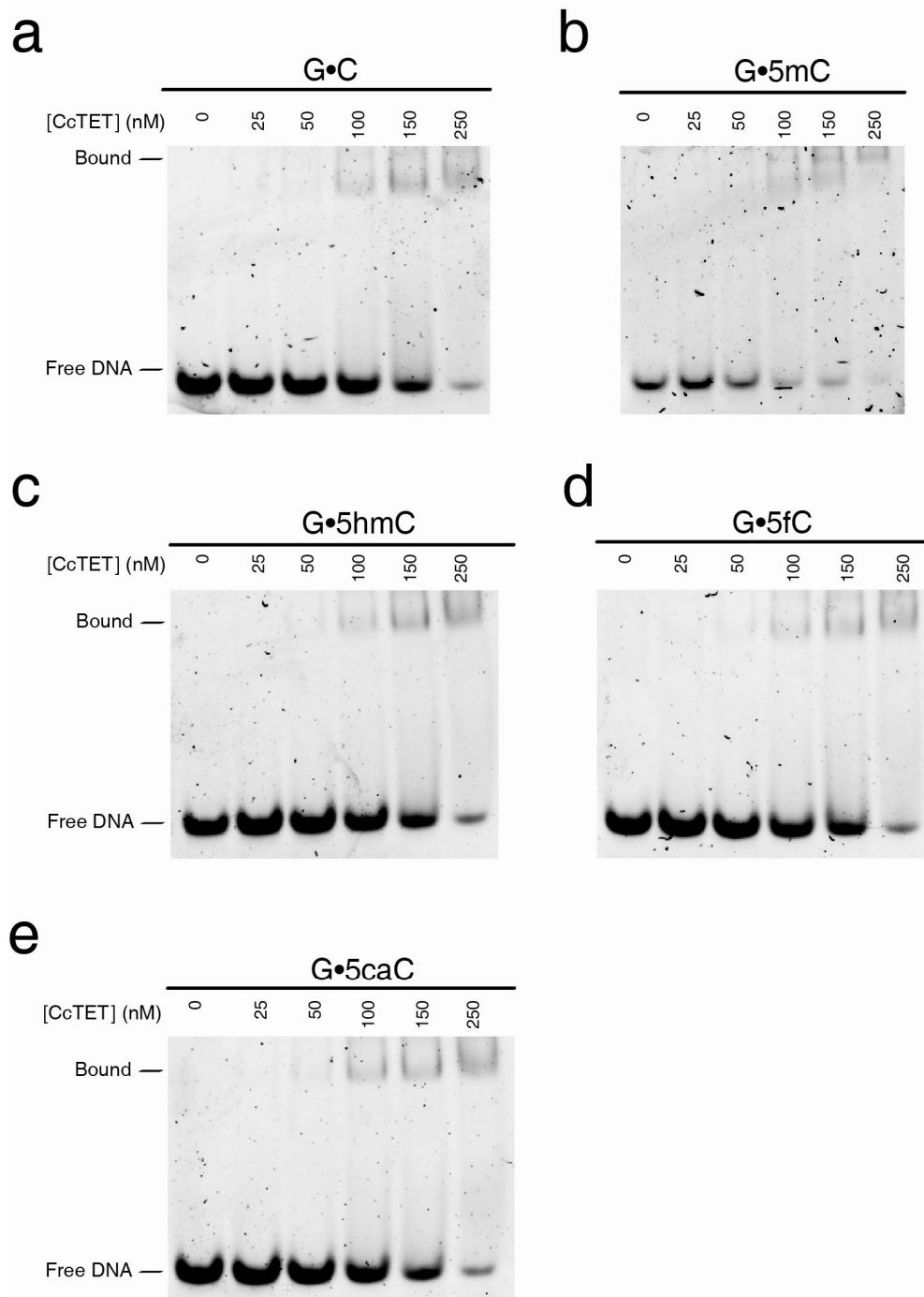




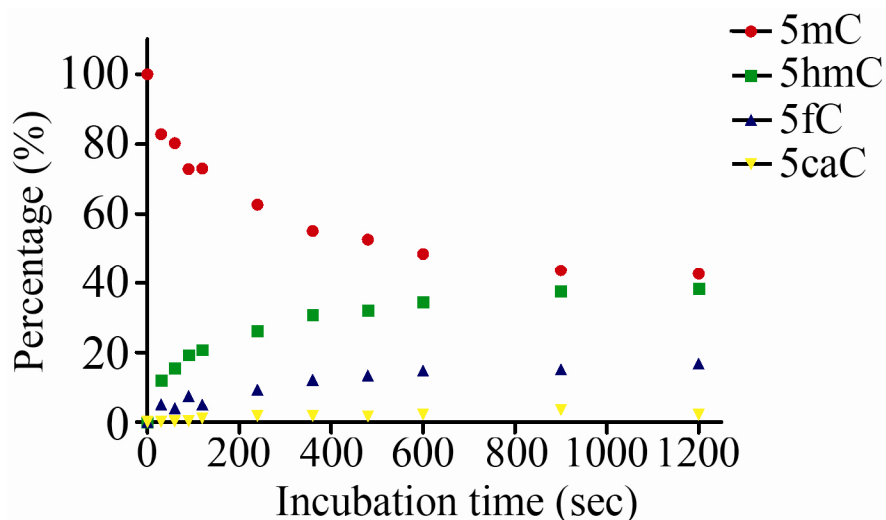
Supplementary Figure S11. Quantitative analysis of CcTET-mediated oxidation products from mESC genomic DNA at pH 5.8 using LC-MS/MS. The nucleosides were quantified using the nucleoside to base ion mass transitions.



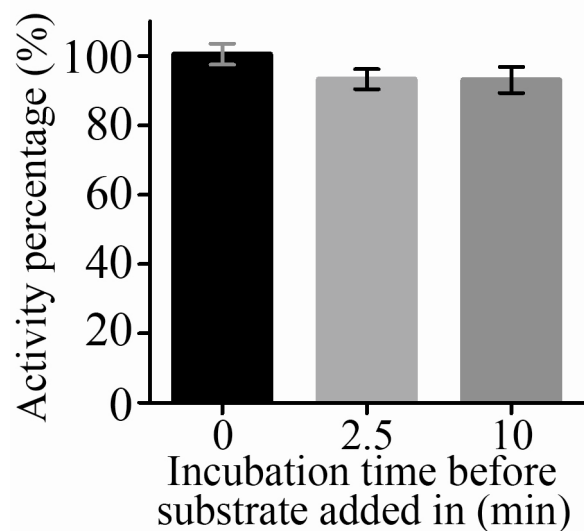
Supplementary Figure S12. The oxidation assays of CcTET (in the absence of ATP) using mESC genomic DNA at pH 5.8 or 7.0 at 37 °C. The 20  $\mu$ l reaction mixture contains 20  $\mu$ g/ $\mu$ l purified mESC genomic DNA, 20  $\mu$ M CcTET, 50 mM Bis-Tris, pH 5.8 (or 50 mM HEPES, pH 7.0), 75  $\mu$ M  $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ , 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT for 30 min at 37 °C in triplicates. The reactions were quenched by being heated to 100 °C for 5 min and immediately cooled in an ice bath.



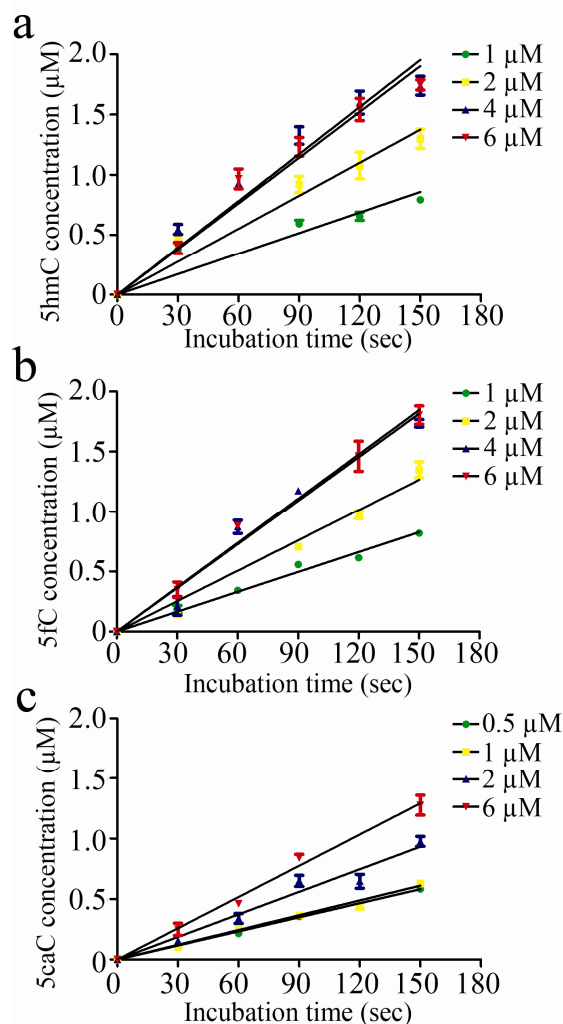
Supplementary Figure S13. The EMSA assay with 16-mer dsDNA containing G•C (a), G•5mC (b), G•5hmC (c), G•5fC (d), and G•5caC (e) base pair, respectively (Sequences: 5'-TCTGGAACGGAATTCT-3', 5'-AAGAATTCXGTTCCAG-3'. X position indicates desired cytosine derived base including C, 5mC, 5hmC, 5fC and 5caC). The experiments were performed with 50 nM dsDNA and various concentrations of CcTET as shown.



Supplementary Figure S14. Reaction progress of substrate and fraction products versus incubation time (20 min). The reactions were performed based on the CcTET oxidation activity assay (No ATP) under pH 7.0 at 37 °C. The reaction mix contains 1 μM 5mC containing dsDNA substrate and 0.2 μM CcTET, 50 mM HEPES, pH 7.0, 75 μM Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM α-KG, and 2.5 mM DTT for various incubation time (0, 30, 60, 90, 120, 240, 360, 480, 600, 900 and 1200 seconds) at 37 °C. The reactions were quenched by being heated to 100 °C for 5 min and immediately cooled in ice bath. The 5mC-containing dsDNA substrates were used (Sequences: 5'-GAC(5mC)GGAGT-3', 5'-GACTCCGGTCT-3').



Supplementary Figure S15. *CcTET inactivation assay*. 40 pmol of recombinant CcTET protein (final concentration 2  $\mu$ M) was added into 20  $\mu$ l reaction mixture containing 50 mM HEPES, pH 7.0, 75  $\mu$ M  $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ , 2 mM ascorbic acid, 1 mM  $\alpha$ -KG, and 2.5 mM DTT at 37  $^\circ\text{C}$  in triplicates for pre-incubation. A 9/11 mer duplex DNA containing one 5fCpG site on the 9mer (sequences: 5'-GAC(5fC)GGAGT-3', 5'-GACTCCGGTCT-3', final concentration 2  $\mu$ M) was employed as the substrate. 40 pmol of this dsDNA substrate (final concentration 2  $\mu$ M) was added into the reaction (containing CcTET) which has been pre-incubated for 0, 2.5 min and 10 min at 37  $^\circ\text{C}$ . The reactions were quenched after 30 min by heating to 100  $^\circ\text{C}$  for 5 min and immediately cooled in an ice bath.



Supplementary Figure S16. Reaction progress curve of fraction products versus incubation time (2.5 min). The reactions were conducted according to the CcTET oxidation activity assay (No ATP) under pH 7.0 at 37 °C. A 9/11mer dsDNA containing 5mC (a), 5hmC (b) and 5fC (c) substrates were used at various concentrations (Sequences: 5'-GACXGGAGT-3', 5'-GACTCCGGTCT-3'. X position indicates desired cytosine derived base of 5mC, 5hmC, and 5fC; the substrate concentrations were as shown). The reaction mixture contains various concentration of dsDNA substrate (1, 2, 4 and 6 µM for 5mC and 5hmC; 0.5, 1, 2 and 6 µM for 5fC), various concentration of CcTET (0.2 µM for 5mC containing substrate, 0.25 µM for 5hmC and 5fC containing substrate), 50 mM HEPES, pH 7.0, 75 µM Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>, 2 mM ascorbic acid, 1 mM α-KG, and 2.5 mM DTT for various incubation time (0, 30, 60, 90, 120, and 150 seconds) at 37 °C in triplicates. The reactions were quenched by being heated to 100 °C for 5 min and immediately cooled in an ice bath. The progress curves of 5mC to 5hmC and 5hmC to 5fC could

be slightly impacted by the subsequent oxidation step. However, the subsequent oxidation steps show slower reaction rate which may not significantly affect the kinetic analysis of the previous step. The reaction products of the subsequent step did not accumulate to a significant amount as shown below at 2.5 min time point: percentage of 5fC in 5mC to 5hmC reactions: 6.5% (substrate concentration: 1 uM), 4.7% (2 uM), 2.3% (4 uM), 1.5% (6 uM); 5caC production in 5mC to 5hmC reactions is below detection limit; percentage of 5caC in 5hmC to 5fC reactions: 9.5% (1 uM), 6.3% (2 uM), 4.5% (4 uM), 2.6% (6 uM).

## Supplementary references

- (1) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucl. Acids Res.* **1997**, *25*, 3389.
- (2) Iyer, L. M.; Tahiliani, M.; Rao, A.; Aravind, L. *Cell Cycle* **2009**, *8*, 1698.
- (3) Iyer, L. M.; Zhang, D.; Burroughs, A. M.; Aravind, L. *Nucl. Acids Res.* **2013**, *41*, 7635.
- (4) Lassmann, T.; Frings, O.; Sonnhammer, E. L. *Nucl. Acids Res.* **2009**, *37*, 858.
- (5) Cole, C.; Barber, J. D.; Barton, G. J. *Nucl. Acids Res.* **2008**, *36*, W197.
- (6) Donnelly, M. I.; Zhou, M.; Millard, C. S.; Clancy, S.; Stols, L.; Eschenfeldt, W. H.; Collart, F. R.; Joachimiak, A. *Protein Exp. Purif.* **2006**, *47*, 446.
- (7) Yu, M.; Hon, G. C.; Szulwach, K. E.; Song, C. X.; Jin, P.; Ren, B.; He, C. *Nat. Prot.* **2012**, *7*, 2159.
- (8) Yu, M.; Hon, G. C.; Szulwach, K. E.; Song, C. X.; Zhang, L.; Kim, A.; Li, X.; Dai, Q.; Shen, Y.; Park, B.; Min, J. H.; Jin, P.; Ren, B.; He, C. *Cell* **2012**, *149*, 1368.
- (9) Bennett, M. T.; Rodgers, M. T.; Hebert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 12510.
- (10) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930.
- (11) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129.
- (12) Krumova, P.; Weishaupt, J. H. *Cell. Mol. Life Sci.* **2013**, *70*, 2123.
- (13) Jia, G.; Fu, Y.; Zhao, X.; Dai, Q.; Zheng, G.; Yang, Y.; Yi, C.; Lindahl, T.; Pan, T.; Yang, Y. G.; He, C. *Nat. Chem. Biol.* **2011**, *7*, 885.
- (14) Zhang, L.; Szulwach, K. E.; Hon, G. C.; Song, C. X.; Park, B.; Yu, M.; Lu, X.; Dai, Q.; Wang, X.; Street, C. R.; Tan, H.; Min, J. H.; Ren, B.; Jin, P.; He, C. *Nat. Comm.* **2013**, *4*, 1517.