Supporting Information

An engineered split-TET2 enzyme for inducible epigenetic remodeling

Minjung Lee[§,†], Jia Li[§,†], Yi Liang[‡,†], Guolin Ma[‡,†], Jixiang Zhang[§,¶], Lian He[‡], Yuliang Liu[‡], Qian Li[§], Minyong Li[#], Deqiang Sun[§], Yubin Zhou[‡,*], Yun Huang[§,*]

§ Centre for Epigenetics and Disease Prevention, Institute of Biosciences and Technology, Department of Molecular and Cellular Medicine, College of Medicine, Texas A&M University, 2121 W Holcombe Blvd, Houston, TX 77030, USA

‡ Centre for Translational Cancer Research, Institute of Biosciences and Technology, Department of Medical Physiology, College of Medicine, Texas A&M University, 2121 W Holcombe Blvd, Houston, TX 77030, USA

¶ Department of gastroenterology, Renmin Hospital of Wuhan University, Wuhan, Hubei, 430060, China

School of Pharmacy, Shandong University, 44 Wenhua Road W., Jinan, Shandong, 250012, China

Corresponding Author

Yubin Zhou* yzhou@ibt.tamhsc.edu
Yun Huang* yun.huang@ibt.tamhsc.edu

Author Contributions

†These authors contributed equally

Table of Contents

Materials and Methods

Design and construction of CiDER. The mammalian expression vector pCMVTnT was purchased from Promega. TET2 catalytic domain (TET2CD) was amplified from TET2 cDNA by using a KOD hot start DNA polymerase (EMD Millipore, Billerica, MA, USA) and inserted between XhoI-KpnI restriction sites in the backbone of pCMVTnT. An Myc-tag(EQKLISEEDL), a nuclear localization signal from SV40 (KRPAATKKAGQAKKKK) and a linker(GGSASGGS) were added to the N-terminus of TET2CD, with subsequent insertion of mCherry tot he C-terminus between KpnI-NotI sites. Two versions of (Δ1482-1839, Δ1463-1839) truncated TET2CD constructs were made with standard PCR procedures. cDNAs encoding the chemical dimerication module FKBP-2A-FRB were purchased as a gBlocks Gene Fragments (Integrated DNA Technologies), and individually inserted into the 6 selected split sites within TET2CD (positions 1258, 1390, 1430, 1462, 1482, and 1839; primers listed below) within the AscI-XmaI restriction sites.

Primer list

L 1258	For	Cttctt GGCGCGCC CTCACCAAT CGCCGGTGTGCC
	Rev	Cttctt CCCGGG CGTGCCGTATTTCCTCAGCGTC
N 1390	For	Cttctt GGCGCGCC AAT GGCAGCACATTGGTA TGCAC
	Rev	Cttctt CCCGGG CTGCATGTTGTGCAAGTCTCTGTG
G 1430	For	Cttctt GGCGCGCC GGGAGTGTGGAAGCTCAGGAG
	Rev	Cttctt CCCGGG AAACTCATCCACGTCAGAGACTTTG
T 1463	For	Cttctt GGCGCGCC ACTTGCCGA CAAAGGAAACTAGAA G
	Rev	Cttctt CCCGGG CTTGACTGGCTCTGCTAACATCC
L 1482	For	Cttctt GGCGCGCC CTGGAGAACAGCTCAAATAAAAATG
	Reverse	Cttctt CCCGGG GGAGGAAAGCTTTTCAGCTGCAGC
6G 1839	For	Cttctt GGCGCGCC GGTGCAGAGGACAACGATGAG
	Rev	Cttctt CCCGGG AGAAGCCACACCCTGGACTAGTG

Cell culture and plasmid transfection. HeLa and human embryonic kidney HEK293T cell lines from the American Type Culture Collection (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich), supplemented with 10% heat-inactivated fetal bovine serum, 100 U/ml penicillin/ streptomycin at 37 °C with 5% CO2. Transfection was performed by using lipofectamine 3000 (Life Technologies) following the manufacturer's recommended protocol.

200-500 ng DNA was used for each well of a 24-well plate. 200 nM rapamycin or AP1903 (Sigma-Aldrich) was applied to the cells. The culture media were replaced every 24 h with fresh media containing 200nM rapamycin or AP-1903.

5hmC immunofluorescence staining and imaging. HeLa cells (4 x 10^5) plated on sterile coverslips in 24-well plates. After 24-48 h rapamycin treatment, cells were fixed with 4% paraformaldehyde in PBS for 15 min and permeabilized with 0.2% Triton X-100 in PBS for 30 min at room temperature. Next, DNAs were denatured with 3N HCl at room temprature for 15 min and neutralized with 100 mM Tris-HCl buffer (pH 8.0) for 10 min. After extensive washing with PBS, cells were blocked with 1% BSA for 30 min, and then incubated with rabbit anti-5-hmC polyclonal antibody (diluted at 1:500, Active Motif) for 2 hours at room temperature. After washing with PBS (3 times; 15 min each), FITC-conjugated anti-rabbit IgG (Sigma-Aldrich) was added to cells for 1 hour. After thoroughly washing with PBS, 250 ng/ml of 4',6-diamidino-2-phenylindole (DAPI) was added to the fixed cells and then mounted the slides for confocal imaging.

The fluorescent images were acquired by using a Nikon A1R+ confocal imaging system equipped with multiple laser sources (405/488/561/640 nm). 488 nm laser (green) was used to obtain 5hmc staining signals, 561-nm laser to excite mCherry for detecting protein expression. The NIS-Elements software was used for image analysis. The averaged FITC intensity in the nuclei of mCherry-positive cells were collected and analyzed. The data was plotted using the Prism 5 software. Images in Figure S1c were acquired using the automated high-content confocal imaging system (IN CELL Analyzer, GE Healthcare).

Flow cytometry. Cells were re-suspended in FACS buffer (PBS with 1% BSA, 2 mM EDTA) and incubated with Fc blocker for 10 min on ice. After washing with FACS buffer, cells were fixed and permeabilized using the Cell Fixation/permeabilization kit from BD Biosciences. DNA were denatured by 2N HCl and neutralized by 10 mM Tris-HCl (pH 8.0) for 20 min. Next, anti-5hmC antibody (Active Motif, 1:200) and FITC conjugated goat anti-rabbit secondary antibody (at a dilution of 1:200; Thermo Fisher Scientific) were used for 5hmC staining. Flow cytometry analysis was performed using LSRII (BD Biosciences) and data were analyzed by using the FlowJo software.

A dot-blot assay to quantify genomic 5hmC and 5mC. Dot blot assays were performed as described previously[1]. Briefly, purified genomic DNA was denatured in 0.4 M NaOH, 10 mM EDTA at 95 °C for 10 min, followed by neutralization with ice-cold 2 M ammonium acetate (pH 7.0). Two-fold serial dilutions of the denatured DNA samples were spotted on a nitrocellulose membrane in an assembled Bio-Dot apparatus (Bio-Rad) according to the manufacturer's instructions. A synthetic oligonucleotide with a known amount of 5hmC was used as standard[1,2]. The membrane was washed with 2xSSC buffer, air-dried and vacuum-baked at

80°C for 2 h. The dried membrane was blocked with 5% non-fat milk for 1 hr and incubated with an anti-5hmC antibody (1:5000, Active Motif) for 1 h at 4°C, followed by incubation with horseradish peroxidase-conjugated anti-rabbit IgG secondary antibody (1:10,000; Sigma). The membrane was visualized by West-Q Pico Dura ECL Solution (GenDEPOT). To ensure equal loading of total DNA on the membrane, the same blot was stained with 0.02% methylene blue in 0.3 M sodium acetate (pH 5.2) to visualize the total amounts of loaded DNA samples.

Western blotting. Cells were lysed with RIPA buffer (150 mM NaCl, 50 mM Tris-HCl, pH 8.0, 1% Triton X-100, 0.5% sodium deoxycholate and 0.1% SDS) supplemented with protease inhibitor cocktail (GenDEPOT), and incubated on ice for 20 min. Cell debris was removed by centrifuging at 30,000 x g for 10 min at 4°C. The protein concentration was measured by a Pierce BCA protein assay kit (Thermo Fisher Scientific). Samples were mixed with SDS sample buffer at 95°C for 10 min. Whole cell lysates were resolved on 10% or 4-12% gradient SDS-PAGE and transferred onto nitrocellulose membranes. Proteins were detected by immunoblotting in TBST (150 mM NaCl, 10 mM Tris-Cl, pH 8.0, 0.5% Tween-20) containing 5% low-fat milk followed by incubation with anti-Myc (GeneTex, 1:10,000), anti-mCherry (GeneTex , 1:3,000) or anti-GAPDH (Sigma, 1:10,000) at RT for 1 hr. Then the membrane was incubated with HRP-conjugated secondary antibodies (goat anti-mouse IgG HRP, Sigma) and proteins were detected by using the West-Q Pico Dura ECL kit (GenDEPOT).

Pull-down experiments and functional reconstitution in vitro. Myc-CiDER-mCherry encoding the split CiDER (N-terminal half tagged with Myc and C-half fused with mCherry) was transfected into HEK293T cells by using Lipofectamine 3000 (Life Technologies). 48 h post-transfection, $1X10^7$ cells were lysed in a RIPA lysis buffer (150 mM NaCl, 50 mM Tris-HCl, pH 8.0, 1% Triton X-100, 0.1% sodium deoxycholate and 0.05% SDS) supplemented with protease inhibitor cocktail (GenDEPOT) and incubated on ice for 20 min. Cell debris was removed by centrifuging at 18,000x g for 10 min at 4°C. Cell lysates were incubated with anti-Myc antibody (ab1253, Abcam) and/or a rabbit polyclonal anti-mCherry antibody (ab167453, Abcam) for 4 hrs at 4°C followed by incubation with precleared protein A/G beads (30 ul) overnight at 4°C. Protein/beads mixtures were washed with 50 mM HEPES (pH 8.0) containing 50 mM NaCl for 5 times. Then we mixed the immunoprecipitated TET2CD, N-half or both CiDER fragments with substrate (a double-stranded 5mC containing DNA oligos as we used in an earlier study [2]) in a reaction buffer (50 mM HEPES (pH 8.0), 50 mM NaCl, 1 mM α-ketoglutarate, 3.7 μM ammonium iron (II) sulfate hexahydrate, 0.1 mg ml^{-1} BSA, 1 mM ATP) at 37°C for 30 min. EDTA (11 mM) was then added to quench the reaction. DNA and protein mixtures were eluted with 50 mM HEPES (pH 8.0) containing 200 mM NaCl, 0.2% SDS with incubation at 95°C for 10 min. DNA fragments were further purified by MicroElute Cycle-Pure Kit (Omega). 5hmC level were measured by the dot-blot assay as described above.

CMS-IP sequencing library preparation. CMS-IP were performed as described previously[3,4]. Genomic DNA was isolated using a Qiagen DNeasy blood and tissue kit. Purified genomic DNA

was sonicated into ~300 bp using Covaris focused ultrasonicators. Sheared DNA was ligated with methylated adaptors using a TruSeq DNA library preparation kit (Illumina) followed by sodium bisulfite treatment (Life Technologies, Methylcode bisulfite conversion kit) to convert 5hmC to cytosine methyl sulfonate (CMS). CMS fragments were enriched using anti-CMS antibody that binds to protein A/G dynabeads. Enriched fragments were cleaned up using the phenol/ chloroform/isoamyl-alcohol method and then amplified using KAPA HiFi Uracil+ (Kapa Biosystems) polymerase with 8 PCR cycles. Amplified libraries were sequenced using Illumina NextSeq instrument (150 cycle, pair-ended).

ATAC-seq sequencing library preparation. ATAC-seq library preparation was performed as described before[5]. Briefly, nuclei were isolated in lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% IGEPAL CA-630) followed by centrifugation at 500 ×g for 10 min, 4 °C. Next, the transposition reaction was carried out by using Illumina Nextera DNA library preparation kit (37°C for 30 min). Tagmentated DNAs were purified using the Qiagen MiniElute kit. Purified DNAs were amplified with the KAPA real-time library amplification kit followed by library purification using AmpuXP beads. The quality of purified DNA libraries was checked by Agilent TapeStation and then subjected to highthroughput sequencing on an Illumina NextSeq instrument (150 cycle, pair-ended).

Bioinformatic analysis. For 5hmC peak identification, reads were mapped to hg19 assembly of the human genome using bsmap2 software with default parameters. Uniquely mapped reads were kept for the following analysis. We used a sliding window method to split the human whole genome to 300 bp size windows and count the numbers of reads mapped in each window with 100 bp spacing. By comparing with the reads from input samples (without antibody enrichment), regions with more than 2-folds enrichment are identified as a peak. For heatmap generation, we used mplot (unpublished software) to plot control (0 h) and 48 h ATAC/CMS signals on ATAC/CMS control peaks regions (normalized to 400 bp). For the ATAC-seq analysis, reads were aligned to the hg19 assembly of the human genome using bowtie2 with the '-very-sensitive' parameter. The duplicate reads were discarded and only properly paired mapped reads with mapping quality ≥30 were kept for downstream analysis. The ATAC peak calling was done by MACS2 with '-nomodel' and '-extsize 147' parameters. For the differential ATAC/CMS peaks analysis, the ATAC/CMS peaks in all groups were merged to form a merged peak set (ATAC merged peaks set: 107545 peaks; CMS merged peaks set: 231705 peaks). Next, the reads in each group falling in the peaks were counted. DESeq2 was then used to call significantly differential ATAC/CMS peaks. Significant differential peaks were identified as fold change ≥4 and q-value ≤0.05.

Plasmid sequences:

CiDER | Myc-NLS-TET2CD$_N$-linker-FRB-T2A-FKBP12-linker-TET2CD$_C$-linker-mCherry

Primary sequence |

```
          10        20        30        40        50        60        70        80        90
          |         |         |         |         |         |         |         |         |
MEQKLISEEDLKRPAATKKAGQAKKKKGGSASGGSDFPSCRCVEQIIEKDEGPFYTHLGAGPNVAAIREIMEERFGQKGKAIRIERVIYTGK
EGKSSQGCPIAKWVVRRSSSEEKLLCLVRERAGHTCEAAVIVILILVWEGIPLSLADKLYSELTETLRKYGTLTNRRCALNEERTCACQGLD
PETCGASFSFGCSWSMYYNGCKFARSKIPRKFKLLGDDPKEEEKLESHLQNLSTLMAPTYKKLAPDAYNNQIEYEHRAPECRLGLKEGRPFS
GVTACLDFCAHAHRDLHNMQNGSTLVCTLTREDNREFGGKPEDEQLHVLPLYKVSDVDEFGAPGGGGSGGGGSGGGGSILWHEMWHEGLEEA
SRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKGSGEGRGSLLTCG
DVEENPGPMGVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFDSSRDRNKPFKFMLGKQEVIRGWEEGVAQMSVGQRAKLTISPDYAYG
ATGHPGIIPPHATLVFDVELLKLEGSGSGSGSGSGSPGLGSVEAQEEKKRSGAIQVLSSFRRKVRMLAEPVKPGKKLLPGLGAEDNDEVWSD
SEQSFLDPDIGGVAVAPTHGSILIECAKRELHATTPLKNPNRNHPTRISLVFYQHKSMNEPKHGLALWEAKMAEKAREKEEECEKYGPDYVP
QKSHGKKVKREPAEPHETSEPTYLRFIKSLAERTMSVTTDSTVTTSPYAFTRVTGPYNRYIGTGSGSGSGSGSGSMVSKGEEDNMAIIKEFM
RFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGV
VTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVN
IKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK*
```

Nucleotide sequences | 2391 bp without mCh tag

```
ATGGAACAAAAACTTATTTCTGAAGAAGATCTGAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAA
AGGGAGGTTCCGCTAGCGGAGGTTCGGATTTCCCATCTTGCAGATGTGTAGAGCAAATTATTGAAAAAGATGAAGGTCC
TTTTTATACCCATCTAGGAGCAGGTCCTAATGTGGCAGCTATTAGAGAAATCATGGAAGAAAGGTTTGGACAGAAGGGT
AAAGCTATTAGGATTGAAAGAGTCATCTATACTGGTAAAGAAGGCAAAAGTTCTCAGGGATGTCCTATTGCTAAGTGGG
TGGTTCGCAGAAGCAGCAGTGAAGAGAAGCTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACACCTGTGAGGCTGCAGT
GATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGTCTCTGGCTGACAAACTCTACTCGGAGCTTACCGAGACG
CTGAGGAAATACGGCACGCTCACCAATCGCCGGTGTGCCTTGAATGAAGAGAGAACTTGCGCCTGTCAGGGGCTGGATC
CAGAAACCTGTGGTGCCTCCTTCTCTTTTGGTTGTTCATGGAGCATGTACTACAATGGATGTAAGTTTGCCAGAAGCAA
GATCCCAAGGAAGTTTAAGCTGCTTGGGGATGACCCAAAAGAGGAAGAGAAACTGGAGTCTCATTTGCAAAACCTGTCC
ACTCTTATGGCACCAACATATAAGAAACTTGCACCTGATGCATATAATAATCAGATTGAATATGAACACAGAGCACCAG
AGTGCCGTCTGGGTCTGAAGGAAGGCCGTCCATTCTCAGGGGTCACTGCATGTTTGGACTTCTGTGCTCATGCCCACAG
AGACTTGCACAACATGCAGAATGGCAGCACATTGGTATGCACTCTCACTAGAGAAGACAATCGAGAATTTGGAGGAAAA
CCTGAGGATGAGCAGCTTCACGTTCTGCCTTTATACAAAGTCTCTGACGTGGATGAGTTTGGCGCGCCgGGAGGTGGTG
GCAGCGGTGGAGGAGGTTCTGGGGGCGGTGGCTCAATTTTATGGCATGAGATGTGGCATGAGGGTTTGGAAGAGGCATC
TAGATTGTATTTCGGTGAAAGAAATGTCAAGGGAATGTTCGAAGTTTTAGAACCGTTGCACGCTATGATGGAGAGAGGT
CCACAGACTCTAAAGGAGACTTCCTTCAACCAAGCTTATGGAAGGGACCTAATGGAGGCTCAAGAATGGTGTAGAAAAT
ACATGAAAAGTGGAAATGTAAAGGACCTTACACAAGCTTGGGATCTCTACTACCATGTTTTTAGGAGAATATCTAAAGG
AAGTGGTGAGGGTAGGGGAAGTTTATTAACCTGTGGGGATGTTGAAGAAAATCCAGGTCCTATGGGCGTACAAGTTGAA
ACTATCAGCCCTGGGGACGGCAGAACCTTTCCGAAGAGGGGACAGACATGTGTTGTTCACTATACTGGAATGTTGGAAG
ATGGTAAGAAGTTCGATAGCAGCAGAGATAGGAATAAACCATTTAAATTCATGCTTGGCAAGCAAGAAGTGATTAGGGG
TTGGGAAGAAGGTGTCGCTCAAATGAGTGTAGGTCAGAGGGCTAAGTTAACAATTAGTCCTGATTATGCTTATGGCGCT
ACAGGTCATCCAGGAATCATTCCCCCACATGCTACTCTTGTTTTCGACGTTGAATTGCTCAAGCTTGAAGGATCAGGTT
CTGGATCTGGTTCAGGATCAGGCTCACCCGGGCTTGGGAGTGTGGAAGCTCAGGAGGAGAAAAAACGGAGTGGTGCCAT
TCAGGTACTGAGTTCTTTTCGGCGAAAAGTCAGGATGTTAGCAGAGCCAGTCAAGCCTGGGAAGAAGCTTCTTCCTGGG
CTTGGTGCAGAGGACAACGATGAGGTCTGGTCAGACAGCGAGCAGAGCTTTCTGGATCCTGACATTGGGGGAGTGGCCG
TGGCTCCAACTCATGGGTCAATTCTCATTGAGTGTGCAAAGCGTGAGCTGCATGCCACAACCCCTTTAAAGAATCCCAA
TAGGAATCACCCCACCAGGATCTCCCTCGTCTTTTACCAGCATAAGAGCATGAATGAGCCAAAACATGGCTTGGCTCTT
TGGGAAGCCAAAATGGCTGAAAAAGCCCGTGAGAAAGAGGAAGAGTGTGAAAAGTATGGCCCAGACTATGTGCCTCAGA
AATCCCATGGCAAAAAAGTGAAACGGGAGCCTGCTGAGCCACATGAAACTTCAGAGCCCACTTACCTGCGTTTCATCAA
GTCTCTTGCCGAAAGGACCATGTCCGTGACCACAGACTCCACAGTAACTACATCTCCATATGCCTTCACTCGGGTCACA
GGGCCTTACAACAGATATATAGGTACCGGGTCGGGTAGTGGCTCTGGTAGTGGTTCTGGTTCTATGGTGAGCAAGGGCG
AGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTT
CGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCC
CTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCC
CCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGAC
CGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGAC
GGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGG
```

GCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAA
GCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTG
GAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAG

Supplemental Figures

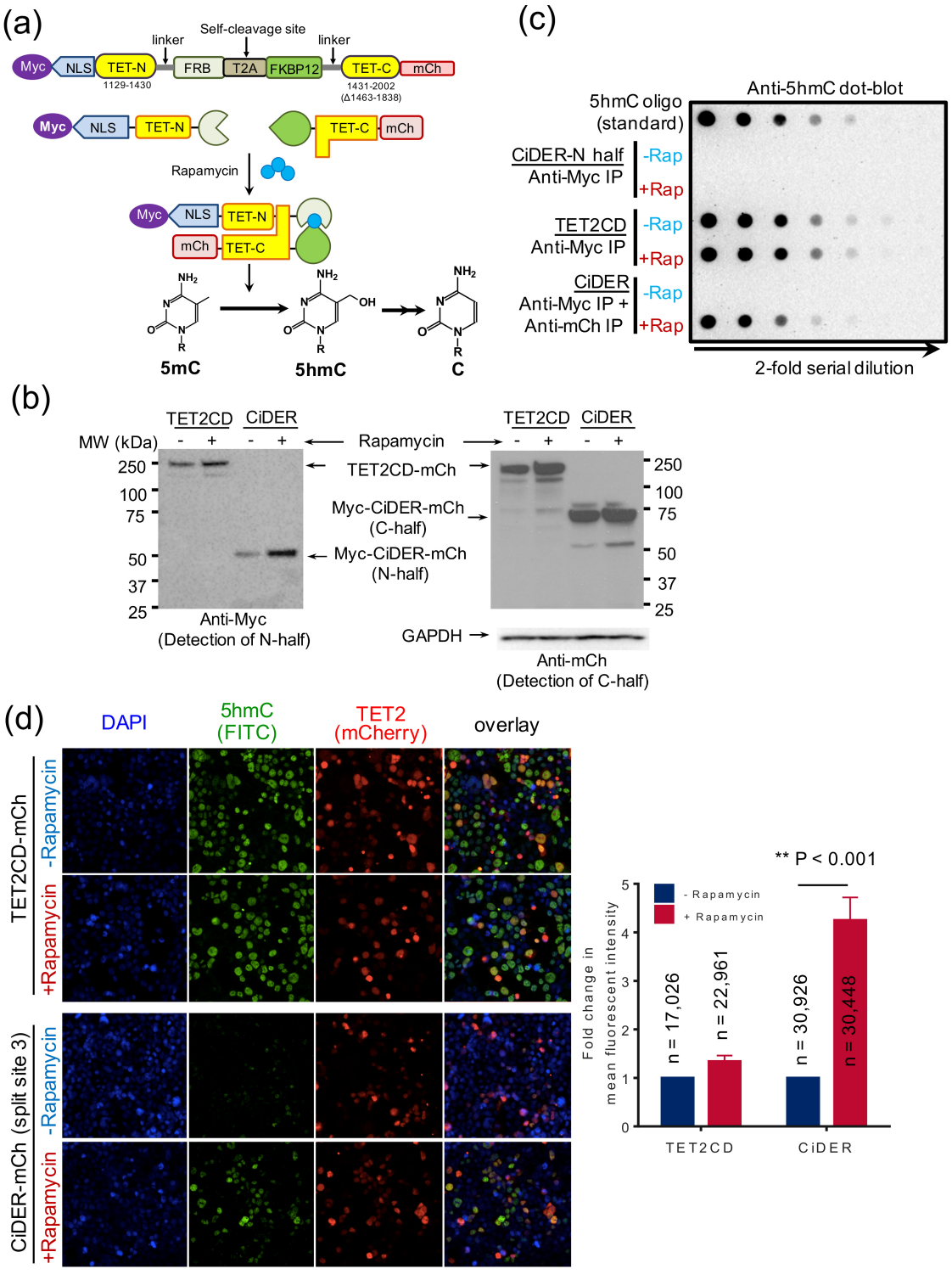

Figure S1. Design of CiDER as a split-TET2 enzyme for inducible 5mC oxidation.

(a) Diagram of CiDER. The chemical-inducible dimerization modules FRB and FKBP12, along with a self-cleaving peptide T2A, was inserted into the catalytic domain of TET2 (TET2CD with deletion of residues 1462-1839) at split site 3 (Figure 1). The N-terminal region of TET2CD (TET-N) is tagged with an Myc tag and a nuclear localization signal (NLS); whereas the C-terminal region of TET2CD (TET-C) is fused with mCherry to aid the detection of protein expression. Following the expression of CiDER in mammalian cells, the fusion protein will be cleaved into two inactive TET2CD fragments. Upon addition of rapamycin, the chemically-inducible heterodimerizaiton of FRB and FKBP12 brings the two complementary fragments of TET2CD into close proximity to restore its enzymatic function, thereby catalyzing 5mC oxidation to produce 5hmC. AP1903-inducible homodimerization of a mutant FKBP12 (F36V) can be similarly engineered to restore the catalytic activity of split-TET2CD (Figure S2).

(b) Western blotting to detect the expression of two split TET2CD fragments. The N-terminal domain was detected by the anti-Myc antibody (left; calculated MW = 51.9 kDa) whereas the C-half was immunoblotted against an anti-mCherry monoclonal antibody (right; calculated MW = 64.1 kDa).

(c) A dot-blot assay to quantify *in vitro* 5mC-to-5hmC conversion in reaction mixtures containing the following combinations: i) N-half of CiDER (negative control; enriched by an anti-Myc antibody) + 5mC dsDNA oligo (with all C replaced by 5mC); ii) Myc-TET2CD-mCherry (positive control; enriched by an anti-Myc antibody) + 5mC dsDNA oligo; or iii) both enriched fragments (enriched by antibodies against Myc and mCherry) + 5mC dsDNA oligo. The reactions were carried out in the absence (blue) or presence (red) of 1 μM rapamycin.

(d) Use of a high-content confocal imaging system to screen split-TET2 constructs and quantify 5hmC levels. HeLa cells were seeded into 96-well plates and transfected with TET2CD-mCherry (positive control) or split-TET2-mCherry constructs (CiDER as an example; red), followed by incubation with or without 200 nM rapamycin for 48 h. Cells were subsequently fixed and stained with DAPI (blue) or anti-5hmC antisera (green). The relative levels of 5hmC, plotted as fold change in the mean fluorescence intensity, were shown on the right. Numbers of cells analyzed in three independent experiments were shown above or within the bars. Data were shown as mean ± s.d.. **p<0.001.

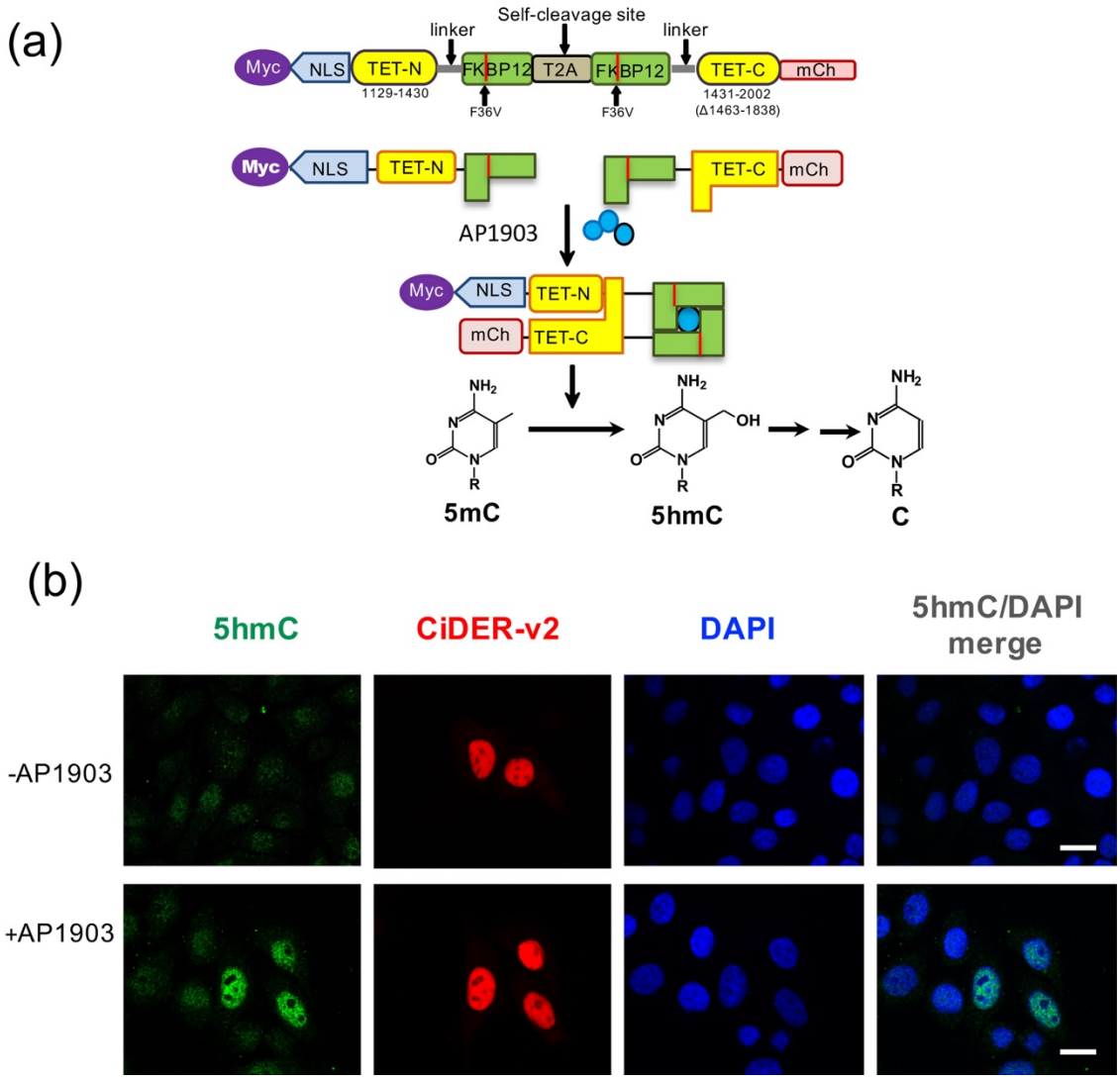

Figure S2. AP1903-inducible restoration of the enzymatic activity of a split-TET2 construct.

(a) The FRB-FKBP12 module can be replaced by an AP1903-inducible homodimerization domain FKBP12-F36V to restore the catalytic activity of the split-TET2CD protein.

(b) Representative fluorescent images of 5hmC (green), CiDER-mCh (red), and nuclear staining with DAPI (blue) in HEK293T cells before ad after 48-h incubation with AP1903 (282 ng/ml).

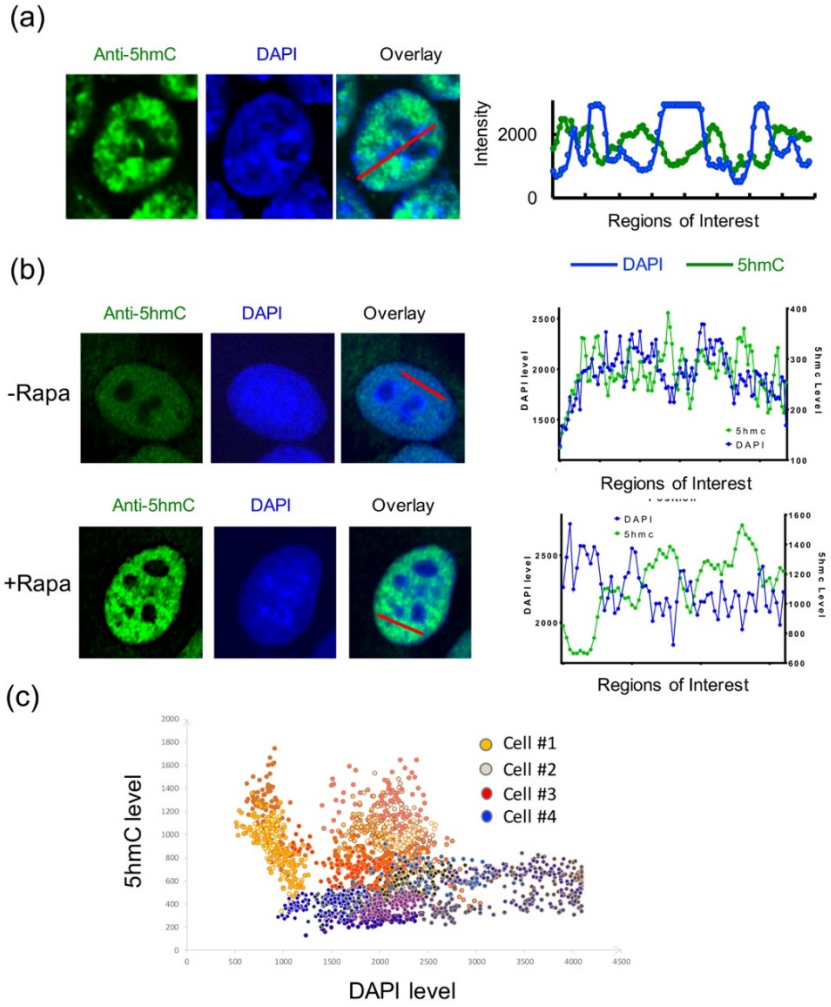

Figure S3. 5hmC and DAPI staining in the nuclei of mouse embryonic stem cells (a) and CiDER-expressing HEK293T cells (before and after rapamycin treatment; b-c).

(a) Mouse embryonic stem cells were immunostained with an anti-5hmC antibody (green). The nuclei were stained by DAPI (blue). The fluorescent signals across the red line were plotted on the right.

(b) Anti-5hmC (green) and DAPI (blue) staining in HEK293T cells expressing CiDER. Transfected HEK293T cells were cultured in the absence (0 h) or presence of (48 h) 200 nM rapamycin prior to immunostaining. The fluorescent signals across the red lines were plotted on the right.

(c) Scatter plot of 5hmC vs DAPI staining in four representative CiDER-expressing HEK293T cells after incubation with rapamycine for 48 h. In general, higher 5hmC staining was correlated with lower DAPI staining in the same cell, and *vice versa*.

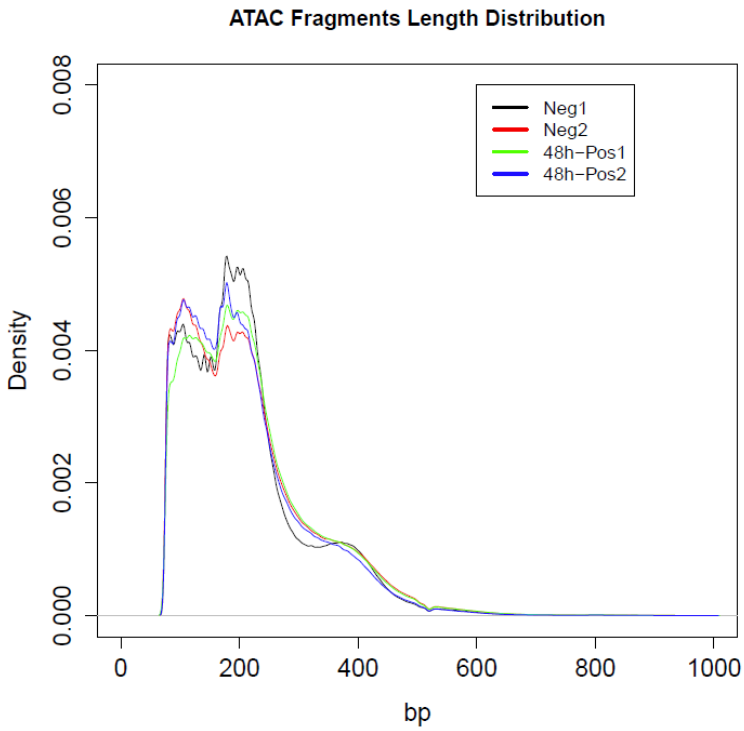

Figure S4. Fragment length distribution profiles of ATAC-seq reads. ATAC-seq fragment sizes generated from HEK293T cells expressing CiDER and treated with rapamycin for 0 or 48 hrs. The size distribution profiles confirmed the high-quality of ATAC-seq library samples.

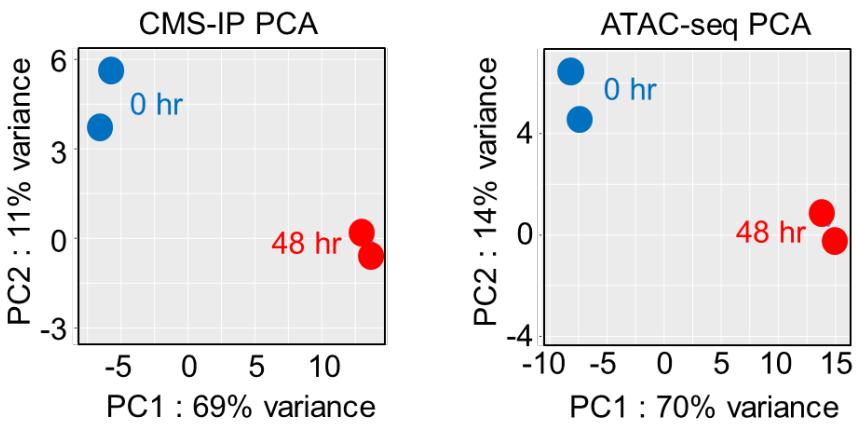

Figure S5. The principle component analysis (PCA) of CMS-IP (left) and ATAC-seq (right) results for HEK293T cells expressing CiDER (before rapamycin treatment (0 h, blue; two biological replicates) or treated for 48 h, red; two biological replicates).

Supplemental Tables

Table S1. Statistical analysis of CMS-IP-seq reads and ATAC-seq reads

CMS-IP reads

Samples	Total reads	Unique Paired Mapped Reads	Unique Paired Map Ratio	Peak number
0 hr-1	41105878	28781440	70.02%	90,477
0 hr-2	46015352	33115822	72%	108,571
48 hr-1	48383616	32455890	67.08%	103,853
48 hr-2	41646842	25820958	62%	119,091

ATAC-seq reads

Samples	Total reads	Unique Paired Mapped Reads	Unique Paired Map Ratio	Peak number
0 hr-1	96294360	34621658	35.95%	61,431
0 hr-2	49042074	18124108	36.96%	27,489
48 hr-1	97754716	36814124	37.66%	58,000
48 hr-2	81121540	32505096	40.07%	39,812

Supplemental References

1 Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., An, J., Lamperti, E. D., Koh, K. P., Ganetzky, R., Liu, X. S., Aravind, L., Agarwal, S., Maciejewski, J. P. & Rao, A. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839-843.

2 Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R. & Rao, A. (2010). The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5**, e8888.

3 Huang, Y., Pastor, W. A., Zepeda-Martinez, J. A. & Rao, A. (2012). The anti-CMS technique for genome-wide mapping of 5-hydroxymethylcytosine. *Nat Protoc* **7**, 1897-1908.

4 Pastor, W. A., Pape, U. J., Huang, Y., Henderson, H. R., Lister, R., Ko, M., McLoughlin, E. M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G. Q., Liu, X. S., Ecker, J. R., Milos, P. M., Agarwal, S. & Rao, A. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**, 394-397.

5 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218.