ENGINEERED SPLIT-TET2 ENZYME FOR INDUCIBLE EPIGENETIC REMODELING

A Thesis

by

YI LIANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Yubin Zhou
Committee Members,	Fen Wang
	Dekai Zhang
	Michael Zhu
Head of Program,	Warren Zimmer

December 2017

Major Subject: Medical Sciences

ABSTRACT

The conversion of 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC) and further oxidized species is catalyzed by the Ten-eleven translocation (TET) methylcytosine dioxygenase. Here we provided a temporal control of 5mC oxidation and subsequent remodeling of epigenetic states through a split-TET2 enzyme system in mammalian cells. This chemical-inducible epigenome remodeling tool provides a precise way to interrogate cellular systems and probe the epigenotype-phenotype relations in various biological systems.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Yubin Zhou, and my committee members, Dr. Fen Wang, Dr. Dekai Zhang, Dr. Michael Zhu, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my mother and father for their encouragement and to my wife for her patience and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis (or) dissertation committee consisting of Professor Yubin Zhou of the Department of Center for Translational Cancer Research -- IBT.

The part of the data in Section 2 was conducted by Minjung Lee from Dr. Yun Huang's Lab -- IBT.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

This work was made possible in part by National Institutes of Health grant (R01GM112003), the Welch Foundation (BE-1913), the American Cancer Society (RSG-16-215-01 TBE), the Cancer Prevention and Research Institute of Texas (RR140053, RP120348 and RP170002) and by an allocation from the Texas A&M University Health Science Center Startup Fund.

NOMENCLATURE

TAMU	Texas A&M University
IBT	Institute of Biosciences & Technology
CTCR	Center for Translational Cancer Research
TET	Ten-eleven translocation dioxygenases
5mC	5-methylcytosine
5hmC	5-hydroxymethylcytosine
DNMTs	DNA methyltransferases
5fC	5-formylmethylcytosine
5caC	5-carboxymethylcytosine
TET2CD	Catalytic domain of human TET2
DSBH	Double stranded beta-helix
FKBP12	FK506 binding protein 12
FRB	FKBP rapamycin binding domain
HEK293T	Human embryonic kidney 293T
CiDER	Chemical-inducible epigenome remodeling tool
Cas9	CRISPR associated protein 9

TABLE OF CONTENTS

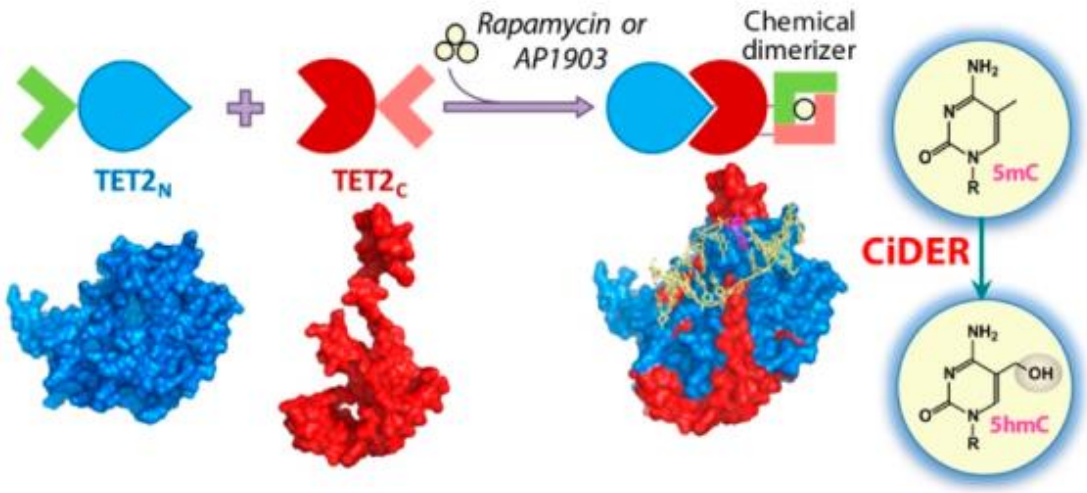
LIST OF FIGURES

Page

1. INTRODUCTION

DNA methylation, catalyzed by DNA methyltransferases (DNMTs), is a process in which a methyl group is added to the carbon 5 position of cytosine to form 5mC. 5mC has been considered as an epigenetic mark that functions in transcriptional repression, X-chromosome inactivation and transposon silencing in the mammalian genome[1]. The recovery of 5mC to cytosine is a successive process in which Tet-eleven translocation (TET) family of methylcytosine dioxygenases catalyze the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylmethylcytosine (5fC) and 5-carboxymethylcytosine (5caC)[2-4]. This process adds an additional layer of previously underappreciated epigenetic control over the mammalian genome. 5hmC catalyzed by TET proteins has been considered to serve as a rate limiting process of TET-mediated DNA demethylation[2-4], as well as a reliable epigenetic mark[5-8]. Evidence shows that DNA hydroxymethylation is highly correlated with gene expression and some function disorders[9-11]. However, it is still a challenge to establish the causal relations between epigenetic modifications on DNA and the phenotypes, owing to the lack of reliable tools to accurately control DNA methylation in genome at temporal and spatial resolution.

To overcome this challenge, we designed a chemical-inducible epigenome remodeling tool (CiDER, Scheme 1) to acquire the precise control of adding or removing epigenetic DNA modifications to further control the transcription of genes. Rather than TET1 or TET3, the catalytic domain of human TET2 (TET2CD, Figure1) has been selected as the



Scheme 1. Design of a chemical-inducible epigenome remodeling(CiDER) tool based on a split TET2 enzyme. FKBP12/FRB heterodimerization or FKBP-F36 V homodimerization modules are fused with two inactive fragments of a split TET2CD. Upon the addition of rapamycin, split TET2CD fragments reassemble into a functional methylcytosine dioxygenase to catalyze the conversion of 5mC into 5hmC and further oxidized species, thus promoting DNA demethylation to remodel the epigenetic landscapes in mammalian cells. Reprinted with permission from Engineered Split-TET2 Enzyme for Inducible Epigenetic Remodeling by Minjung Lee, Jia Li, Yi Liang, J. Am. Chem. Soc., 2017, 139 (13), pp 4659–4662, http://pubs.acs.org/doi/abs/10.1021%2Fjacs.7b01459

subject for a split epigenomic modifier for the following major reasons. First, in hematological malignancies[10], TET2 is the most frequently mutated gene. Plenty of exome sequencing data of cancer patients provide abundant information of sensitive spots that need to be avoided through the engineering process[12-13]. Second, the structure of TET2CD in complex with 5mC or 5hmC is the only one that has been determined among TET proteins[14-15], which provides more rationale to the selection of split sites. Third, there is a low complexity region (residues 1481-1843) existing in TET2CD which is possible to be replaced by a flexible GS linker without significantly compromising the catalytic activity,

and hence it offers a structural malleability of TET2CD and increases the flexibility to accommodate the insertion of foreign polypeptide sequences[15]. The replacement of this large fragment of low complexity (~1.2 kb) also helps us to minimize the size of the constructs. Then, we set out to test the idea that the two inactive fragments of split TET2CD can be restored by a chemically inducible dimerization approach.

2. RESULT

2.1 Design and optimization of Split-TET2CD constructs

To develop a split-TET2CD system, six sites have been selected on a Cys-rich region and a double-stranded beta-helix (DSBH) fold (Figure 1a,b). A FK506 binding protein 12 (FKBP12) and FKBP rapamycin binding domain(FRB)[16-17] separated by a self-cleaving T2A polypeptide sequence[18-19] was inserted between the selected split sites of TET2CD or TET2CD with a GS linker instead of the low complexity region. A mCherry-tag was added on the C-terminal of the constructs (Figure 1c).

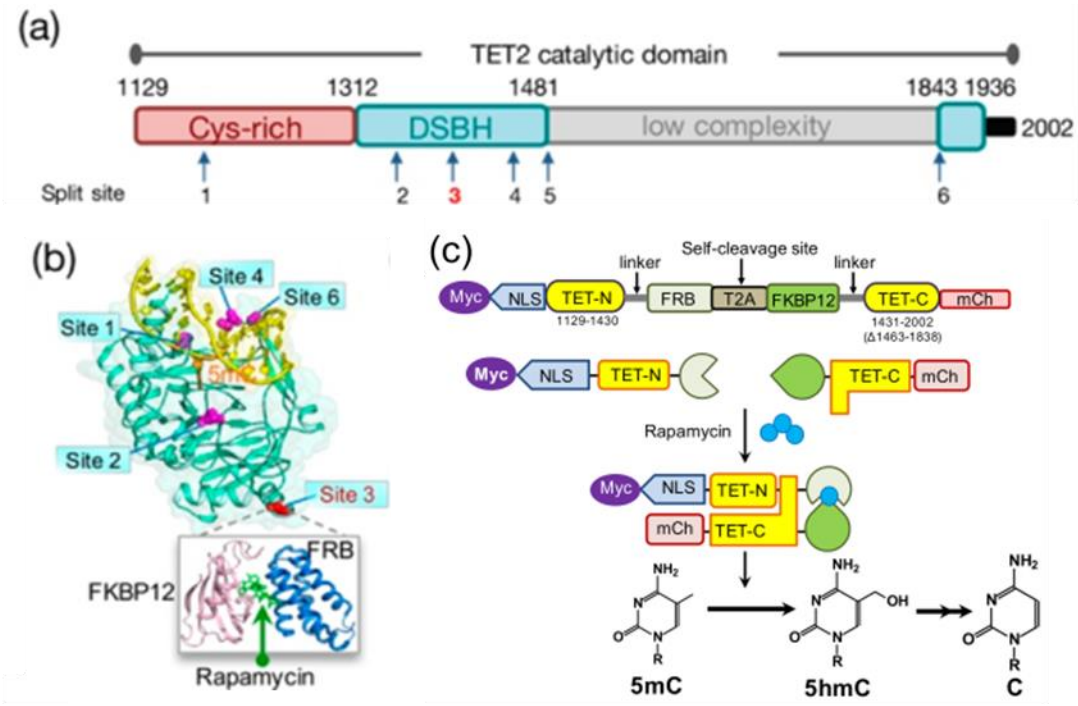
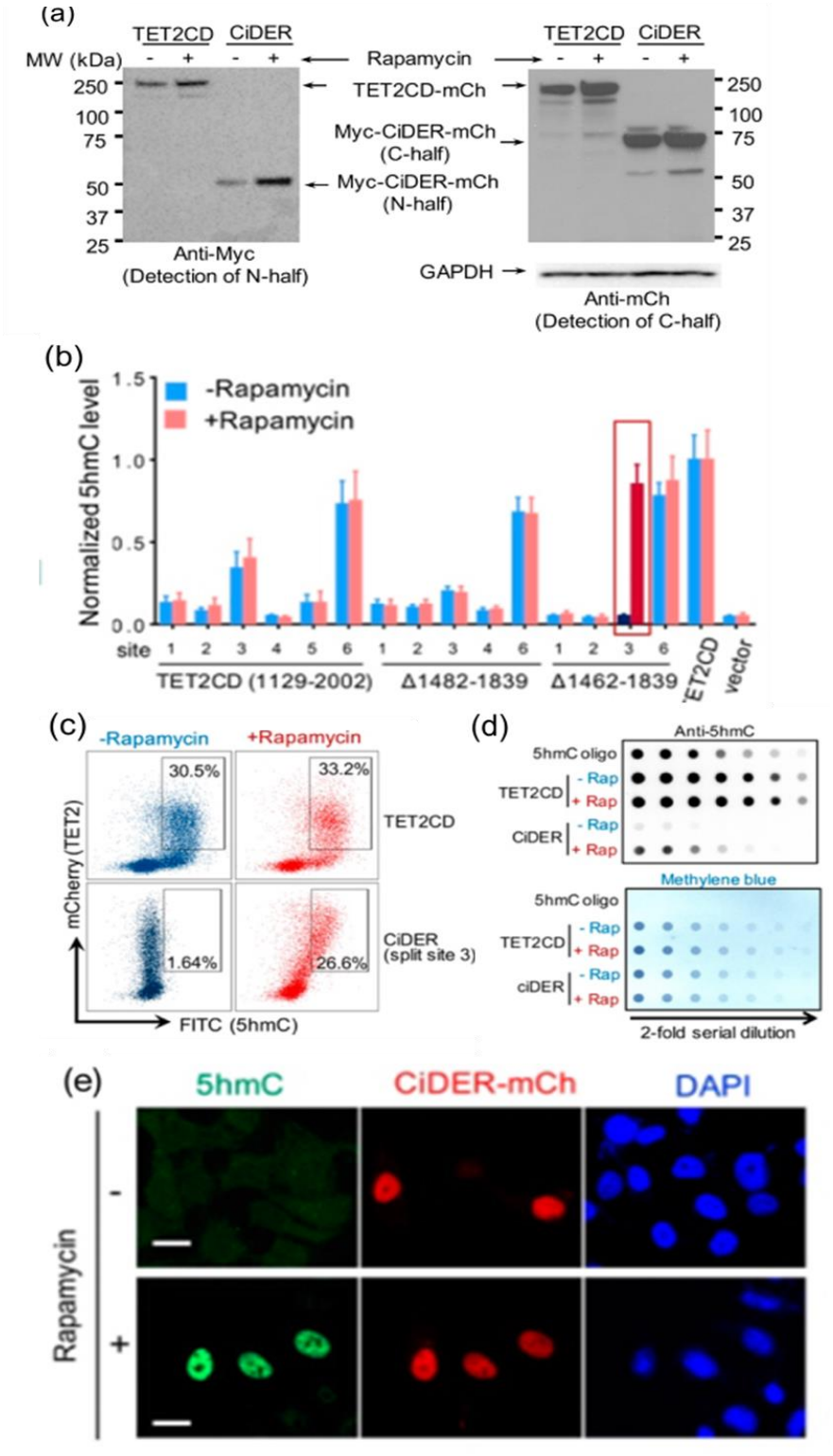


Figure 1. An engineered split-TET2 enzyme for inducible DNA hydroxymethylation in mammalian cells. (a) Domain architecture of the catalytic domain of TET2 (TET2CD; aa 1129−2002) and positions of selected split sites. DSBH, double stranded beta helix. (b) Split sites mapped to the 3D structure of TET2CD (PDB entry: 4NM6) (c) 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 (TETN) 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 heterodimerization 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. Reprinted with permission from Engineered Split-TET2 Enzyme for Inducible Epigenetic Remodeling by Minjung Lee, Jia Li, Yi Liang, J. Am. Chem. Soc., 2017, 139 (13), pp 4659–4662, http://pubs.acs.org/doi/abs/10.1021%2Fjacs.7b01459

2.2 5hmC level is increased by CiDER upon rapamycin induction in mammalian cells

Each construct was transfected into human embryonic kidney 293T (HEK293T) or Hela cells that both have extremely low basal 5hmC levels. The fusion protein was self-cleaved into two fragments once transfected into HEK293T cells (Figure 2a).

Figure 2. 5hmc level increased by CiDER. (a) 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 Chalf was immunoblotted against an anti-mCherry monoclonal antibody (right; calculated MW = 64.1 kDa). (b) Screening and optimization of split-TET2CD constructs to achieve chemical-inducible 5hmC generation in HEK293T cells. The construct with insertion of FKBP12-T2A-FRB at split site 3 and deletion of the low complexity region (Δ1462−1839) stood out as the best candidate (termed "CiDER", S1). (c) Quantification of CiDER-mediated 5hmC production by flow cytometry. HEK293T cells transfected with mCherry (mCh)-tagged CiDER or mCh-TET2CD (positive control) were mmunostained with an anti-5hmC primary antibody and an FITC-labeled secondary antibody. (d) Dot-blot assay to quantify rapamycin (200 nM)-induced changes of 5hmC levels in genomic DNA purified from HEK293T cells expressing CiDER or TET2CD. A synthetic oligonucleotide with a known amount of 5hmC was used as a positive control. The loading control was shown in the bottom panel by methylene blue staining of total amounts of input DNA. (e) Representative fluorescent images of 5hmC (green), CiDER-mCh (red), and nuclear staining with DAPI (blue) in HEK293T cells before and after rapamycin (200 nM) treatment. Reprinted with permission from Engineered Split-TET2 Enzyme for Inducible Epigenetic Remodeling by Minjung Lee, Jia Li, Yi Liang, J. Am. Chem. Soc., 2017, 139 (13), pp 4659–4662, http://pubs.acs.org/doi/abs/10.1021%2Fjacs.7b01459


(a)
TET2CD
CiDER
MW (kDa)
- + - +
Rapamycin
TET2CD-mCh
Myc-CiDER-mCh (C-half)
Myc-CiDER-mCh (N-half)
250
100
75
50
37
25
Anti-Myc (Detection of N-half)
GAPDH
Anti-mCh (Detection of C-half)
(b)
Normalized 5hmC level
-Rapamycin
+Rapamycin
site
TET2CD (1129-2002)
Δ1482-1839
Δ1462-1839
TET2CD vector
(c)
-Rapamycin
+Rapamycin
30.5%
33.2%
1.64%
26.6%
TET2CD
CiDER (split site 3)
mCherry (TET2)
FITC (5hmC)
(d)
Anti-5hmC
5hmC oligo
TET2CD
CiDER
- Rap
+ Rap
Methylene blue
ciDER
2-fold serial dilution
(e)
5hmC
CiDER-mCh
DAPI
Rapamycin

In the presence of rapamycin, the production of 5hmC was increased after incubating the enriched fragments with a synthetic 5mC-containing dsDNA oligo in vitro, indicating that the two fragments could reconstitute a functional enzyme. We further compared global 5hmC levels before and after rapamycin treatment by flow cytometry (Figure 2b,c). We identified a split-TET2CD variant (designated as CiDER for chemical-inducible epigenome remodeling tool) that shows the restoration of enzymatic activity through rapamycin induction after screening over 15 constructs. There was almost no background activity of CiDER prior to the addition of rapamycin. The total 5hmC level was restored in the cells that expressed intact integrated TET2CD proteins after rapamycin induction (Figure 2b,c). We further confirmed the generation of 5hmC by rapamycin

induction through two additional methods: immunostaining and a more quantitative dot-blot assay that measures the total amounts of 5hmC in the whole cell population. It was confirmed by both methods that rapamycin elicited robust production of 5hmC in CiDER-expressing HEK293T cells (Figure 2d,e).

2.3 Reversibility of CiDER triggered 5hmC raise

To detect the reversibility of CiDER, we monitored rapamycin-induced 5hmC level at different time points. The result showed that 5hmC level reached a maximum in about 48 hours with rapamycin treatment. The 5hmC level returned to its basal level in about 3 days after the withdrawal of rapamycin, indicating the constant dilution of 5hmC level after multiple rounds of cell division. By contrast, the 5hmC level in the group with intact TET2CD remained at high level regardless of whether rapamycin was added or removed. (Figure 3).

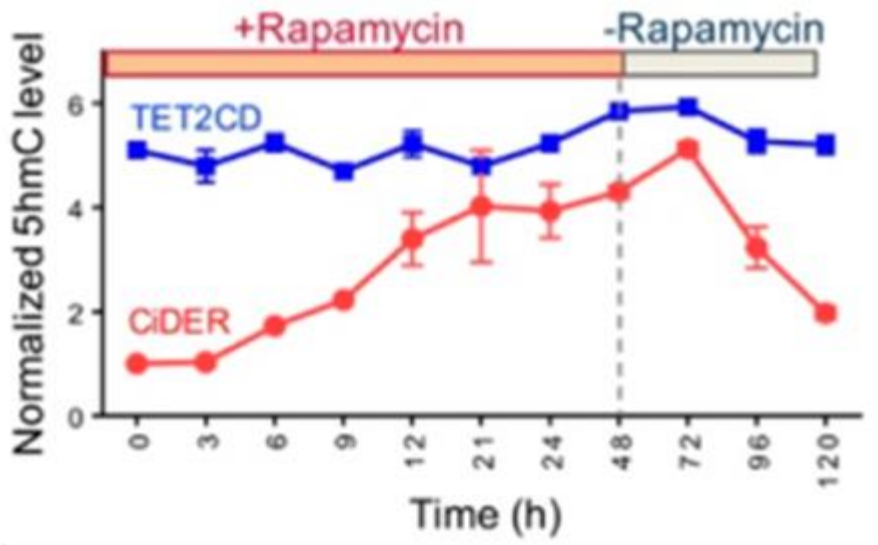


Figure 3. Time course of rapamycin (200 nM)-induced production of 5hmC in HEK293T cells expressing CiDER or TET2CD (as positive control). Rapamycin was washed away 48 h after incubation with cells. Reprinted with permission from Engineered Split-TET2 Enzyme for Inducible Epigenetic Remodeling by Minjung Lee, Jia Li, Yi Liang, J. Am. Chem. Soc., 2017, 139 (13), pp 4659–4662, http://pubs.acs.org/doi/abs/10.1021%2Fjacs.7b01459

3. DISCUSSION

To further improve CiDER system to achieve the spatial control of 5hmC level, we plan to fuse it with a catalytically inactive Cas9 or its orthologues which offer the possibility of loci-specific targeting of DNA demethylation in genome. It's quite possible that the loci-specific CiDER system could become a powerful tool for interrogating, perturbing and engineering cellular system without altering the genetic code. as well as reprogramming cell fate and disease intervention.

4. METHODS

4.1 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% CO_2. Transfection was performed by using lipofectamine 3000 (Life Technologies) following the manufacturer's recommended protocol. A total of 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.

4.2 5hmC immunofluorescence staining and imaging

HeLa cells (4×10^5) were plated on sterile coverslips in 24-well plates. After 24-48 hours of 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 temperature 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 the cells and incubated 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 was used to detect mCherry for protein localization. The NISElements 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).

4.3 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.

4.4 A dot-blot assay to quantify genomic 5hmC and 5mC

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. 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 hour and incubated with an anti-5hmC antibody (1:5000, Active Motif) for 1 hour 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.

4.5 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).

4.6 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 hours 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) 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 was measured by the dot-blot assay as described above.

5. CONCLUSION

In summary, we engineered split-TET2CD fragments based on the TET2CD/5hmC crystal structure to restrict its function temporarily, then employed the rapamycin triggered protein dimerization widget to restore the catalytic function of TET2CD to convert 5mC to 5hmC. Together, it is CiDER that offers a temporal and reversible control of 5hmC level in mammalian cells.

REFERENCES

1. Li, E.; Zhang, Y., DNA methylation in mammals. Cold Spring Harb Perspect Biol 2014, 6 (5), a019133.

2. He, Y. F.; Li, B. Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; Sun, Y.; Li, X.; Dai, Q.; Song, C. X.; Zhang, K.; He, C.; Xu, G. L., Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 2011, 333 (6047), 1303-7.

3. Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y., Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 2011, 333 (6047), 1300-3.

4. 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., Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 2009, 324 (5929), 930-5.

5. Spruijt, C. G.; Gnerlich, F.; Smits, A. H.; Pfaffeneder, T.; Jansen, P. W.; Bauer, C.; Munzel, M.; Wagner, M.; Muller, M.; Khan, F.; Eberl, H. C.; Mensinga, A.; Brinkman, A. B.; Lephikov, K.; Muller, U.; Walter, J.; Boelens, R.; van Ingen, H.; Leonhardt, H.; Carell, T.; Vermeulen, M., Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell 2013, 152 (5), 1146-59.

6. Lio, C. W.; Zhang, J.; Gonzalez-Avalos, E.; Hogan, P. G.; Chang, X.; Rao, A., Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. Elife 2016, 5.

7. Kellinger, M. W.; Song, C. X.; Chong, J.; Lu, X. Y.; He, C.; Wang, D., 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. Nat Struct Mol Biol 2012, 19 (8), 831-3.

8. Su, M.; Kirchner, A.; Stazzoni, S.; Muller, M.; Wagner, M.; Schroder, A.; Carell, T., 5-Formylcytosine Could Be a Semipermanent Base in Specific Genome Sites. Angew Chem Int Ed Engl 2016, 55 (39), 11797-800.

9. 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., Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature 2010, 468 (7325), 839-43.

10. Huang, Y.; Rao, A., Connections between TET proteins and aberrant DNA modification in cancer. Trends Genet 2014, 30 (10), 464-74.

11. Lu, X.; Zhao, B. S.; He, C., TET family proteins: oxidation activity, interacting molecules, and functions in diseases. Chem Rev 2015, 115 (6), 2225-39.

12. Delhommeau, F.; Dupont, S.; Della Valle, V.; James, C.; Trannoy, S.; Masse, A.; Kosmider, O.; Le Couedic, J. P.; Robert, F.; Alberdi, A.; Lecluse, Y.; Plo, I.; Dreyfus, F. J.; Marzac, C.; Casadevall, N.; Lacombe, C.; Romana, S. P.; Dessen, P.; Soulier, J.; Viguie, F.; Fontenay, M.; Vainchenker, W.; Bernard, O. A., Mutation in TET2 in myeloid cancers. N Engl J Med 2009, 360 (22), 2289-301.

13. Langemeijer, S. M.; Kuiper, R. P.; Berends, M.; Knops, R.; Aslanyan, M. G.; Massop, M.; Stevens-Linders, E.; van Hoogen, P.; van Kessel, A. G.; Raymakers, R. A.; Kamping, E. J.; Verhoef, G. E.; Verburgh, E.; Hagemeijer, A.; Vandenberghe, P.; de Witte, T.; van der

Reijden, B. A.; Jansen, J. H., Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat Genet 2009, 41 (7), 838-42.

14. Hashimoto, H.; Pais, J. E.; Zhang, X.; Saleh, L.; Fu, Z. Q.; Dai, N.; Correa, I. R., Jr.; Zheng, Y.; Cheng, X., Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. Nature 2014, 506 (7488), 391-5.

15. Hu, L.; Li, Z.; Cheng, J.; Rao, Q.; Gong, W.; Liu, M.; Shi, Y. G.; Zhu, J.; Wang, P.; Xu, Y., Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. Cell 2013, 155 (7), 1545-55.

16. Banaszynski, L. A.; Liu, C. W.; Wandless, T. J., Characterization of the FKBP.rapamycin.FRB ternary complex. J Am Chem Soc 2005, 127 (13), 4715-21.

17. Clackson, T.; Yang, W.; Rozamus, L. W.; Hatada, M.; Amara, J. F.; Rollins, C. T.; Stevenson, L. F.; Magari, S. R.; Wood, S. A.; Courage, N. L.; Lu, X.; Cerasoli, F., Jr.; Gilman, M.; Holt, D. A., Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. Proc Natl Acad Sci U S A 1998, 95 (18), 10437-42.

18. Ibrahimi, A.; Vande Velde, G.; Reumers, V.; Toelen, J.; Thiry, I.; Vandeputte, C.; Vets, S.; Deroose, C.; Bormans, G.; Baekelandt, V.; Debyser, Z.; Gijsbers, R., Highly efficient multicistronic lentiviral vectors with peptide 2A sequences. Hum Gene Ther 2009, 20 (8), 845-60.

19. Kim, J. H.; Lee, S. R.; Li, L. H.; Park, H. J.; Park, J. H.; Lee, K. Y.; Kim, M. K.; Shin, B. A.; Choi, S. Y., High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PLoS One 2011, 6 (4), e18556.

APPENDIX

Figure A1:

CiDER: Myc-NLS-T2S$_N$-linker-FRB-T2A-FKBP12-linker-T2S$_C$-linker-mCherry

(797 residues without mCherry tag; 1047 residues total)

MEQKLISEEDLKRPAATKKAGQAKKKKGGSASGGSDFPSCRCVEQIIEKDEGPFYTHLG
AGPNVAAIREIMEERFGQKGKAIRIERVIYTGKEGKSSQGCPIAKWVVRRSSSEEKLLC
LVRERAGHTCEAAVIVILILVWEGIPLSLADKLYSELTETLRKYGTLTNRRCALNEERT
CACQGLDPETCGASFSFGCSWSMYYNGCKFARSKIPRKFKLLGDDPKEEEKLESHLQNL
STLMAPTYKKLAPDAYNNQIEYEHRAPECRLGLKEGRPFSGVTACLDFCAHAHRDLHNM
QGAPGGGGSGGGGSGGGGSILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMER
GPQTLKETSFNQAYGRDLMEAQEWCRKYMKSGNVKDLTQAWDLYYHVFRRISKGSGEGR
GSLLTCGDVEENPGPMGVQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFDSSRDRN
KPFKFMLGKQEVIRGWEEGVAQMSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVE
LLKLEGSGSGSGSGSGSPGNGSTLVCTLTREDNREFGGKPEDEQLHVLPLYKVSDVDEF
GSVEAQEEKKRSGAIQVLSSFRRKVRMLAEPVKPGKKLLPGLGAEDNDEVWSDSEQSFL
DPDIGGVAVAPTHGSILIECAKRELHATTPLKNPNRNHPTRISLVFYQHKSMNEPKHGL
ALWEAKMAEKAREKEEECEKYGPDYVPQKSHGKKVKREPAEPHETSEPTYLRFIKSLAE
RTMSVTTDSTVTTSPYAFTRVTGPYNRYIGTGSGSGSGSGSGSMVSKGEEDNMAIIKEF
MRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSK
AYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNF

PSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQ

LPGAYNVNIKLDITSHNEDYTIVEQY

ERAEGRHSTGGMDELYK*

Annotation:

c-myc tag

NLS

FRB

T2A: self-cleaving 2A peptide

FKBP1A

mCherry

T2S: engineered minimal TET2 C+D domain composed of two components

(T2S$_N$, 1129-1430; T2S$_C$, 1431-2002/ Δ1463-1838)

ATGGAACAAAAACTTATTTCTGAAGAAGATCTGAAAAGGCCGGCGGCCACGAAAAAGGC

CGGCCAGGCAAAAAAGAAAAAGGGAGGTTCCGCTAGCGGAGGTTCGGATTTCCCATCT

TGCAGATGTGTAGAGCAAATTATTGAAAAAGATGAAGGTCCTTTTTATACCCATCTAGG

AGCAGGTCCTAATGTGGCAGCTATTAGAGAAATCATGGAAGAAAGGTTTGGACAGAAGG

GTAAAGCTATTAGGATTGAAAGAGTCATCTATACTGGTAAAGAAGGCAAAAGTTCTCAG

GGATGTCCTATTGCTAAGTGGGTGGTTCGCAGAAGCAGCAGTGAAGAGAAGCTACTGTG

TTTGGTGCGGGAGCGAGCTGGCCACACCTGTGAGGCTGCAGTGATTGTGATTCTCATCC

TGGTGTGGGAAGGAATCCCGCTGTCTCTGGCTGACAAACTCTACTCGGAGCTTACCGAG

ACGCTGAGGAAATACGGCACGCTCACCAATCGCCGGTGTGCCTTGAATGAAGAGAGAAC

TTGCGCCTGTCAGGGGCTGGATCCAGAAACCTGTGGTGCCTCCTTCTCTTTTGGTTGTT

CATGGAGCATGTACTACAATGGATGTAAGTTTGCCAGAAGCAAGATCCCAAGGAAGTTT

AAGCTGCTTGGGGATGACCCAAAAGAGGAAGAGAAACTGGAGTCTCATTTGCAAAACCT

GTCCACTCTTATGGCACCAACATATAAGAAACTTGCACCTGATGCATATAATAATCAGA

TTGAATATGAACACAGAGCACCAGAGTGCCGTCTGGGTCTGAAGGAAGGCCGTCCATTC

TCAGGGGTCACTGCATGTTTGGACTTCTGTGCTCATGCCCACAGAGACTTGCACAACAT

GCAGGGCGCGCCgGGAGGTGGTGGCAGCGGTGGAGGAGGTTCTGGGGGCGGTGGCTCAA

TTTTATGGCATGAGATGTGGCATGAGGGTTTGGAAGAGGCATCTAGATTGTATTTCGGT

GAAAGAAATGTCAAGGGAATGTTCGAAGTTTTAGAACCGTTGCACGCTATGATGGAGAG

AGGTCCACAGACTCTAAAGGAGACTTCCTTCAACCAAGCTTATGGAAGGGACCTAATGG

AGGCTCAAGAATGGTGTAGAAAATACATGAAAAGTGGAAATGTAAAGGACCTTACACAA

GCTTGGGATCTCTACTACCATGTTTTTAGGAGAATATCTAAAGGAAGTGGTGAGGGTAG

GGGAAGTTTATTAACCTGTGGGGATGTTGAAGAAAATCCAGGTCCTATGGGCGTACAAG

TTGAAACTATCAGCCCTGGGGACGGCAGAACCTTTCCGAAGAGGGGACAGACATGTGTT

GTTCACTATACTGGAATGTTGGAAGATGGTAAGAAGTTCGATAGCAGCAGAGATAGGAA

TAAACCATTTAAATTCATGCTTGGCAAGCAAGAAGTGATTAGGGGTTGGGAAGAAGGTG

TCGCTCAAATGAGTGTAGGTCAGAGGGCTAAGTTAACAATTAGTCCTGATTATGCTTAT

GGCGCTACAGGTCATCCAGGAATCATTCCCCCACATGCTACTCTTGTTTTCGACGTTGA

ATTGCTCAAGCTTGAAGGATCAGGTTCTGGATCTGGTTCAGGATCAGGCTCACCCGGGA

ATGGCAGCACATTGGTATGCACTCTCACTAGAGAAGACAATCGAGAATTTGGAGGAAAA

CCTGAGGATGAGCAGCTTCACGTTCTGCCTTTATACAAAGTCTCTGACGTGGATGAGTT

TGGGAGTGTGGAAGCTCAGGAGGAGAAAAAACGGAGTGGTGCCATTCAGGTACTGAGTT

CTTTTCGGCGAAAAGTCAGGATGTTAGCAGAGCCAGTCAAGCCCGGGAAGAAGCTTCTT

CCCGGGCTTGGTGCAGAGGACAACGATGAGGTCTGGTCAGACAGCGAGCAGAGCTTTCT

GGATCCTGACATTGGGGGAGTGGCCGTGGCTCCAACTCATGGGTCAATTCTCATTGAGT

GTGCAAAGCGTGAGCTGCATGCCACAACCCCTTTAAAGAATCCCAATAGGAATCACCCC

ACCAGGATCTCCCTCGTCTTTTACCAGCATAAGAGCATGAATGAGCCAAAACATGGCTT

GGCTCTTTGGGAAGCCAAAATGGCTGAAAAAGCCCGTGAGAAAGAGGAAGAGTGTGAAA

AGTATGGCCCAGACTATGTGCCTCAGAAATCCCATGGCAAAAAAGTGAAACGGGAGCCT

GCTGAGCCACATGAAACTTCAGAGCCCACTTACCTGCGTTTCATCAAGTCTCTTGCCGA

AAGGACCATGTCCGTGACCACAGACTCCACAGTAACTACATCTCCATATGCCTTCACTC

GGGTCACAGGGCCTTACAACAGATATATAGGTACCGGGTCGGGTAGTGGCTCTGGTAG

TGGTTCTGGTTCTATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGT

TCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAG

GGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAA

GGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCA

AGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAG

GGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCA

GGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACT

TCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAG

CGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAA

GGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGC

AGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGAC

TACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGA

CGAGCTGTACAAGTAGGCGGCCGC