Selective targeting of TET catalytic domain promotes somatic cell reprogramming

Anup Kumar Singh[a], Bo Zhao[a], Xiuhua Liu[a], Xin Wang[a], Hongzhi Li[b], Hanjun Qin[c], Xiwei Wu[c], Yuelong Ma[d], David Horne[d], and Xiaochun Yu[a,1]

[a]Department of Cancer Genetics & Epigenetics, Beckman Research Institute, City of Hope, Duarte, CA 91010; [b]Computational Therapeutic Core, City of Hope, Duarte, CA 91010; [c]Integrative Genomics Core, City of Hope, Duarte, CA 91010; and [d]Chemical Synthesis Core, City of Hope, Duarte, CA 91010

Edited by Peter A. Jones, Van Andel Institute, Grand Rapids, MI, and approved January 6, 2020 (received for review June 26, 2019)

Ten-eleven translocation (TET) family enzymes (TET1, TET2, and TET3) oxidize 5-methylcytosine (5mC) and generate 5-hydroxymethylcytosine (5hmC) marks on the genome. Each TET protein also interacts with specific binding partners and partly plays their role independent of catalytic activity. Although the basic role of TET enzymes is well established now, the molecular mechanism and specific contribution of their catalytic and noncatalytic domains remain elusive. Here, by combining in silico and biochemical screening strategy, we have identified a small molecule compound, C35, as a first-in-class TET inhibitor that specifically blocks their catalytic activities. Using this inhibitor, we explored the enzymatic function of TET proteins during somatic cell reprogramming. Interestingly, we found that C35-mediated TET inactivation increased the efficiency of somatic cell programming without affecting TET complexes. Using high-throughput mRNA sequencing, we found that by targeting 5hmC repressive marks in the promoter regions, C35-mediated TET inhibition activates the transcription of the BMP-SMAD-ID signaling pathway, which may be responsible for promoting somatic cell reprogramming. These results suggest that C35 is an important tool for inducing somatic cell reprogramming, as well as for dissecting the other biological functions of TET enzymatic activities without affecting their other nonenzymatic roles.

TET enzymes | somatic reprogramming | 5hmC | epigenetic reprogramming

DNA methylation at the 5′-position of cytosine (5-methylcytosine; 5mC) has emerged as one of the most common epigenetic modifications in genomic DNA, which plays an essential role in development, aging, and diseases (1–3). Recent studies have shown that the ten-eleven translocations (TET) family of methyl cytosine dioxygenases (TET1, TET2, and TET3) can oxidize 5mC into various derivatives (4, 5). All three members of the TET family share highly similar catalytic domains that bind DNA and convert 5mC into 5-hydroxymethylcytosine (5hmC) in the presence of cofactors including 2-oxoglutarate and Fe(II) (4). Further oxidation of 5hmC by TETs generates 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (5, 6). Both 5fC and 5caC are relatively unstable on the chromatin (5, 6) and can be recognized by thymine-DNA glycosylase (TDG) that mediates the replacement of 5fC and 5caC with unmodified cytosine for the completion of the active demethylation process (5). Hence, unlike 5fC or 5caC, 5hmC is a stable epigenetic mark in chromatin, suggesting that at least a portion of TET enzymes only catalyze the formation of 5hmC, but not 5fC or 5caC (7–9). However, the biological function of these stable 5hmC marks and their impacts on gene transcription remain elusive.

Genetic evidence suggests that TET family members are functionally redundant during embryonic development, and the loss of any member of the TET family enzyme does not induce prenatal lethality in mice. However, the combined loss of TET1 and TET2 suppresses hematopoietic stem cell differentiation (10). Knocking out all three TET proteins arrests mouse embryogenesis at or before implantation and impairs the differentiation of embryonic stem cells, suggesting that TET enzymes may play a key role in tissue-specific differentiation by controlling key genes' expression (11, 12). Therefore, TET proteins may also have a profound effect on somatic cell reprogramming, which is a process equivalent to the reversal of differentiation. Recent studies suggest that TET enzymes regulate this process of reprogramming by expressing a defined set of Yamanaka factors (Oct4, Sox2, Klf4, and Myc) in the presence of TET proteins (13–15). However, the molecular mechanism of TET enzymes in the context of somatic cell reprogramming still remains elusive. For instance, TET1 deficiency has been shown to promote somatic cell reprogramming in the presence of vitamin C. In contrast, without vitamin C, TET1 alone can replace Oct4 and facilitate somatic cell reprogramming (16, 17). Moreover, it has been reported that 5hmC may act as an independent epigenetic mark rather than just an intermediate during 5mC demethylation in the context of somatic cell reprogramming (7).

Altogether, due to the redundant nature of TET enzymes and lack of TET inhibitors, the molecular mechanism of TET enzymes in somatic cells reprogramming is understudied. In particular, without a specific inhibitor, it is difficult to distinguish the enzymatic functions of TET enzymes from their nonenzymatic roles. Here, we have identified compound 35 (C35) as a first-in-class, highly potent, and cell-permeable TET inhibitor that recognizes the catalytic core of TET enzymes and interferes with their enzymatic activities. Moreover, we found that C35 treatment

Significance

Ten-eleven translocation (TET) proteins mainly express during embryonic development and regulate tissue-specific differentiation. However, their specific role in somatic cell reprogramming is elusive. Here, we report a small molecule inhibitor, C35, that specifically targets TET catalytic domain and reduced 5-hydroxymethylcytosine (5hmC) load on the genome. Our results clearly indicate that this selective inhibition of TET catalytic activity reduced 5hmC marks genomewide and promotes somatic cell reprogramming by inducing the BMP-SMAD-ID downstream pathway. This finding will improve our current understanding about the differential role and contribution of TET catalytic and noncatalytic domains in somatic cell reprogramming. C35 could also be a key tool for the scientific community to study the selective role of TET catalytic domain on various biological processes.

Author contributions: X.Y. designed research; A.K.S., B.Z., X.L., X. Wang, H.L., and Y.M. performed research; A.K.S., H.L., H.Q., X. Wu, D.H., and X.Y. analyzed data; X.L. helped in molecular docking; X. Wang performed high-performance liquid chromatography–high-resolution mass spectrometry; and A.K.S. and X.Y. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database (accession no. GSE137282).

[1]To whom correspondence may be addressed. Email: xyu@coh.org.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1910702117/-/DCSupplemental.

First published February 5, 2020.

clearly promotes somatic cell reprogramming. Mechanistic analyses show that C35 treatment induced the expression of members of the BMP-SMAD-ID axis that are known to play key roles in somatic cell reprogramming.

Results

Identification of C35 as a TET Inhibitor In Vitro. To identify novel TET inhibitors, a virtual ligand screening pipeline (Lvspipe) was employed (18). We selected the TET2 catalytic domain (PDB: 4NM6) and performed a molecular modeling screening with the National Cancer Institute compound library that contains the structures of ~265,000 compounds. Based on the docking score for their virtual interaction and other chemical properties, we selected 40 top hits for additional in vitro biochemical screening (Fig. 1*A*). Purified catalytic domain of human TET2 was used for the in vitro screening in the presence of cofactors including 2-oxoglutarate (α-KG) and Fe(II). Using dot blotting assay with anti-5hmC antibody, we analyzed the conversion of 5mC into 5hmC in the chemically synthesized DNA oligos. Among 40 different compounds, we found C35 to be the most potent inhibitor for recombinant TET2 in vitro (Fig. 1 *B* and *C*). In addition, two other compounds, C14 and C19, were also able to suppress the enzymatic activity of TET2 during the screening; however, their inhibitory effects were not as potent as that of C35. Thus, C35 was selected for the subsequent studies. We have validated the structural identity and purity of C35 using high-performance liquid chromatography–high-resolution mass spectrometry, 13C NMR, and 1H NMR (*SI Appendix*, Figs. S1–S3), and next performed a dose course experiment to examine the efficacy of C35 using the in vitro assays, and found that half-maximal inhibitory concentration (IC_{50}) of C35 against TET2 was 1.2 μM (Fig. 1*D*). Moreover, we purified the catalytic domain of human TET1 and TET3 and examined the inhibitory effect of C35, using the same in vitro assays. We found that similar to TET2, C35 was also able to suppress the catalytic activities of TET1 and TET3 with IC_{50} 3.48 μM and 2.31 μM, respectively (*SI Appendix*, Fig. S4 *A* and *B*). Thus, these results suggest that C35 is a general TET inhibitor that targets all members of the TET family.

Next, we examined the inhibition mechanism of C35. Our molecular modeling results indicate that C35 may partially occupy the DNA substrate binding sites and block the catalytic cage of TET2 by repelling Fe(II). It also interacts with the residue close to the catalytic core of TET2 and hinders at the entrance of DNA molecules to make them inaccessible for catalytic core (Fig. 1 *E*, *Upper*

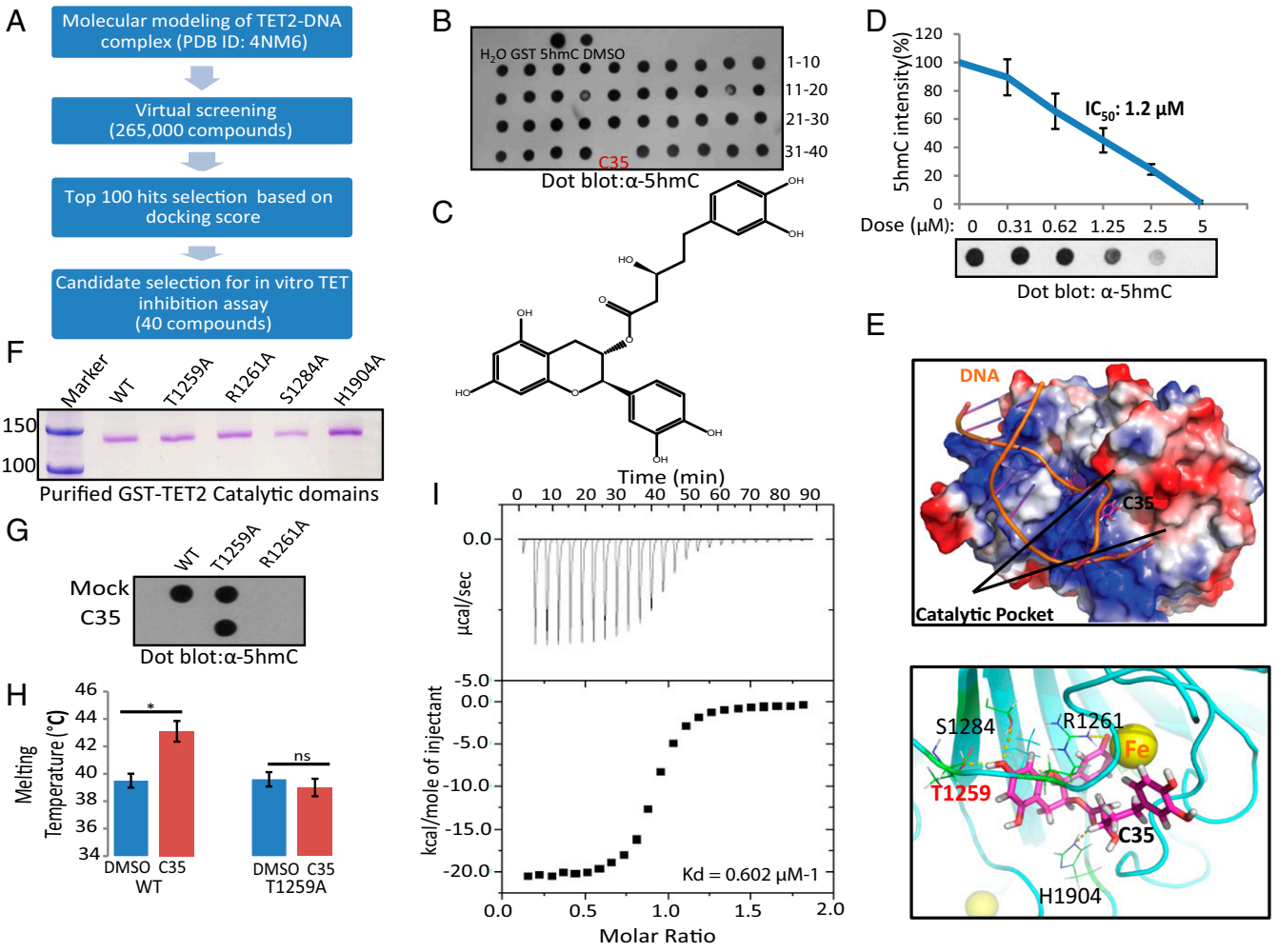

Fig. 1. Identification of C35 as a potent TET2 inhibitor. (*A*) Screening strategy. (*B*) Dot blotting by anti-5hmC antibody shows C35 as the most potent TET2 inhibitor. (*C*) The structure of C35. (*D*) In vitro dose–response curve of TET2 inhibition by C35. (*E*) In silico structural analysis of C35 binding. (*F*) CBB staining of various TET2 mutants. (*G*) C35 does not suppress the enzymatic activity of T1259A in the in vitro assays. (*H*) Ligand-protein thermal shift analysis of C35 with wild-type and T1259 mutant TET2 catalytic domain. (*I*) The binding affinity between wild-type TET2 catalytic domain and C35 was measured by isothermal titration calorimetry. *$P < 0.05$; ns, nonsignificant $P > 0.05$.

and *SI Appendix*, Fig. S4*C*). In this model, C35 forms a hydrogen bond with T1259 at a distance of 2.9 Å. It also forms other hydrogen bonds with R1261, S1284, and H1904, but with distances of more than 3 Å (Fig. 1 *E*, *Lower* and *SI Appendix*, Fig. S4*D*). To validate the binding between C35 and TET2, we replaced these residues with alanine, using site-directed mutagenesis and purified mutant proteins (Fig. 1*F*). Here we found that all these mutations abolished the enzymatic activity of TET2 except T1259A mutation, which was still able to convert 5mC into 5hmC (Fig. 1*G* and *SI Appendix*, Fig. S5). Of note, C35 was unable to suppress the enzymatic activity of T1259A mutant because this mutation is likely to abolish the binding between TET2 and C35. To further confirm this observation, we performed thermal shift assay and found that C35 was able to increase the thermal stability of TET2, but not the T1259A mutant, indicating that C35 specifically recognizes wild-type TET2, but not the T1259A mutant (Fig. 1*H*). Moreover, T1259 is well conserved in all three TET isoforms in human as well as in higher eukaryotes (*SI Appendix*, Fig. S6). To further characterize the binding between C35 and TET2, we performed isothermal titration calorimetry and found the dissociation constant (K_d) value of 0.602 μM^{-1} for binding of C35 with wild-type TET2. In contrast, the T1259A mutation drastically disrupted the interaction with C35, and the observed K_d value was more than 100 μM^{-1} (*SI Appendix*, Fig. S7).

We next analyzed whether C35 competes with α-KG for binding to TET catalytic pocket, using the in vitro TET inhibition assay with increasing concentrations of α-KG, and found that an increase in α-KG concentration was not able to rescue the inhibitory effect of C35 (*SI Appendix*, Fig. S8*A*). We further performed in silico analysis of the TET2 (PDB: 4NM6) structure docked with C35 in the presence of α-KG at multiple rotation positions, and found that the binding pockets of C35 and α-KG are significantly away from each other (*SI Appendix*, Fig. S8*B*) (19). Collectively, these results indicate that C35 is a potent TET inhibitor that specifically targets the catalytic core of TET enzymes without affecting α-KG binding and interfering with their catalytic activities.

C35 Treatment Induces Genome-Wide Loss of 5hmC in both Human and Mouse Cells. To analyze the inhibitory effect of C35 inside of cells, we have analyzed the stability of C35 against the hydrolyzing enzymes, such as esterases. We incubated the C35 in the presence of either serum-containing complete media or total whole-cell lysate for 0, 6, and 12 h and then performed liquid chromatography–mass spectrometry (LC-MS) to calculate the area of the peak corresponding to C35. Here, we found that more than half of the C35 is still stable even after 12 h of incubation (*SI*

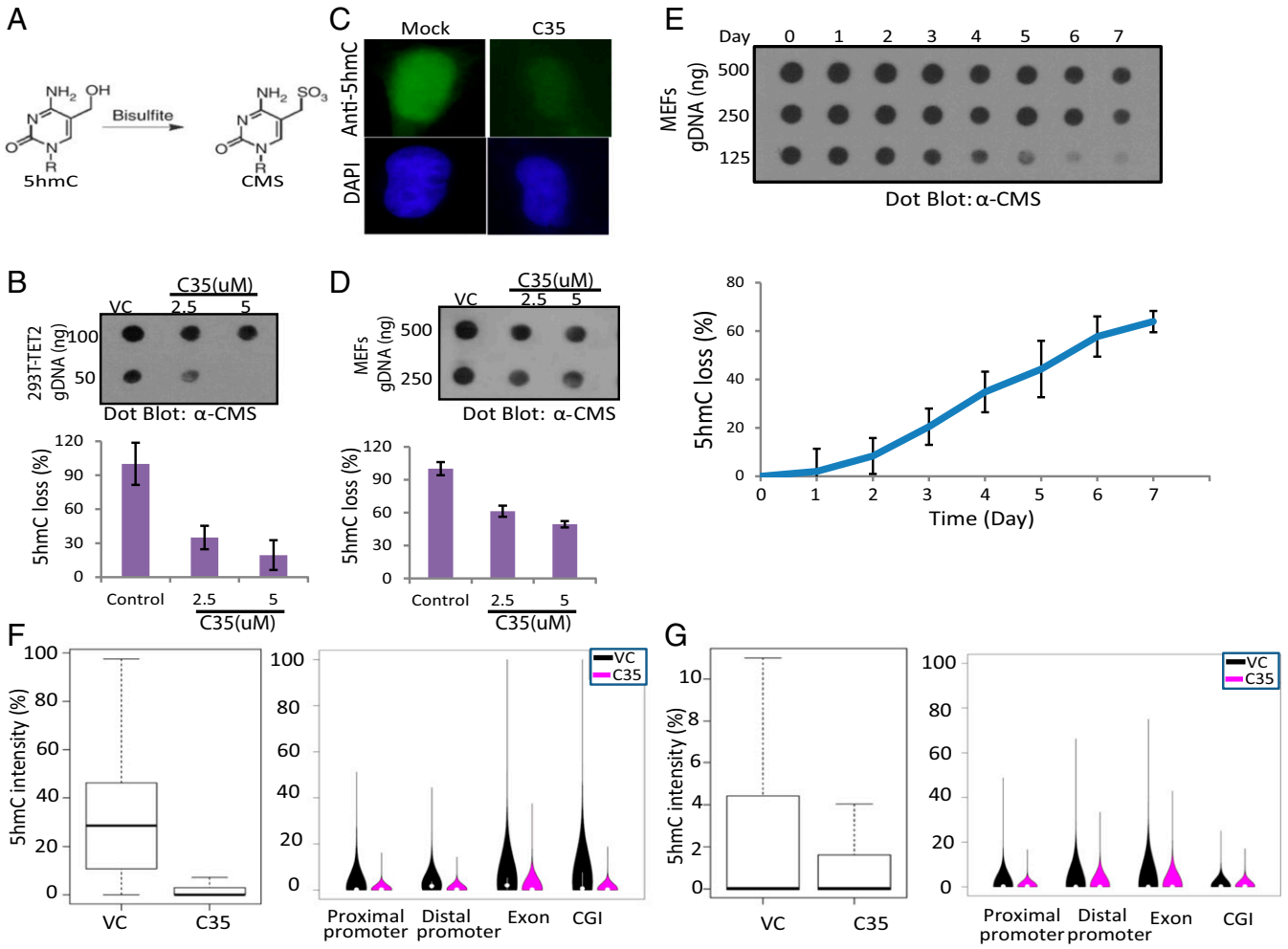

Fig. 2. C35 treatment reduces genome-wide 5hmC level in human and mouse cells. (*A*) Schematic sketch for conversion of 5hmC to CMS. (*B*) Anti-CMS dot blot shows C35 suppresses 5hmC level in HEK293T cells. (*C*) Immunofluorescence staining for 5hmC. (*D*) Dot blot indicates decrease in 5hmC level in MEF. (*E*) A time course of C35 treatment on MEFs was examined with 5 μM C35 for 7 d. (*F* and *G*) Effect of C35 on genomewide 5hmC distribution in HEK293T cells (*F*) and MEFs (*G*). CGI, CpG islands; VC, vehicle control.

Appendix, Figs. S9 and S10, respectively), indicating that it is significantly stable both in serum-containing media as well in the whole-cell lysate. We next analyzed the effect of C35 treatment on the genome-wide 5hmC level. First, we used a recently developed method to examine 5hmC on genomic DNA by converting 5hmC into cytosine-5-methylenesulfonate (CMS), using bisulfite treatment followed by dot blotting with anti-CMS antibody (Fig. 2*A*) (20). We first examined the HEK293T cells stably expressing TET2 that were established in our previous study (21). With 5 μM C35 treatment for 3 d, we harvested genomic DNA and found that 5hmC level was suppressed (Fig. 2*B*). In addition, we also performed anti-5hmC immunofluorescence staining and found a specific loss of 5hmC signal in the nuclei of 293T cells stably expressing TET2 after C35 treatment (Fig. 2*C*). Catalytic pockets of TET enzymes are well conserved between human and mouse; therefore, we also examined the efficacy of C35 in mouse embryonic fibroblasts (MEFs). Again, the basal level of 5hmC was reduced in MEFs when cells were treated with 5 μM C35 (Fig. 2*D*). We next performed a time course assay of C35 treatment on MEFs to analyze the 5hmC level. We found that the loss of 5hmC gradually occurred during the treatment (Fig. 2*E*) as a result of 5hmC being converted from 5mC by TET enzymes; 5mC level is mainly maintained by DMNT1 during DNA replication. Because C35 blocks the writers of 5hmC, loss of 5hmC is aggravated during each round of DNA replication. Thus, we observed a gradual loss of 5hmC in a time course of C35 treatment. To further validate the effect of C35 on genome-wide 5hmC distribution, we combined two different approaches usually used for high-throughput 5hmC analysis, including reduced representation bisulfite sequencing and TET-assisted bisulfite sequencing. The reduced representation bisulfite sequencing-TET-assisted bisulfite sequencing results show a global loss of 5hmC throughout the genome in response to C35 treatment, including the proximal and distal promoter regions, exons, and CpG islands in 293T-TET2 cells (Fig. 2*F*). Similar results were also observed in MEFs treated with C35 (Fig. 2*G*). Taken together, these results demonstrate that C35 is a cell membrane-permeable TET inhibitor that reduces levels of TET-mediated 5hmC in both human and mouse cells.

To further establish the selectivity of C35 on TET enzymes, we analyzed the effect of C35 on α-KG- and FE^{2+}-dependent various members of the Jumonji family of dioxygenases (including JMJD2a, JMJD2b, JMJD2c, JMJD2d, JMJD3, UTX, JARID1a, JARID1b, and JARID1c), as well as HIF1-prolyl hydroxylases (including PHD1, PHD2, and PHD3), and did not find any significant inhibitory effect on them (*SI Appendix*, Fig. S11). We next analyzed for effect of C35 treatment on common cellular off targets including G protein-coupled receptors, kinases, and Na^+/K^+ ion channels, and did not observe any significant changes in them (*SI Appendix*, Fig. S12).

C35 Treatment Promotes Somatic Cell Reprogramming. Emerging evidence suggests that TET enzymes play key roles in somatic cell reprogramming (11, 16, 17). However, the underlying molecular mechanism remains elusive. Since C35 is a potent and cell membrane permeable TET inhibitor, we used C35 to examine the molecular mechanism of TET enzymes in somatic cell reprogramming. Fresh MEFs harboring a *gfp* gene driven by endogenous *oct4* gene promoter were allowed to express canonical Yamanaka factors including Oct4, Sox2, Klf4, and Myc (OSKM) via retroviral infection. The cells were maintained in embryonic stem cell culturing media in the presence (treated) and absence (control group) of 5 μM C35. Induced pluripotent stem cell (iPSC) colonies were observed within 3 wk of ectopic expressing of OSKM. The iPSC colonies were first validated based on GFP signals that were exclusively present in the iPSC colonies (because of Oct4 expression), but not in the parental MEFs (Fig. 3*A*). Interestingly, we found that C35 treatment significantly increased the number of iPSC colonies compared

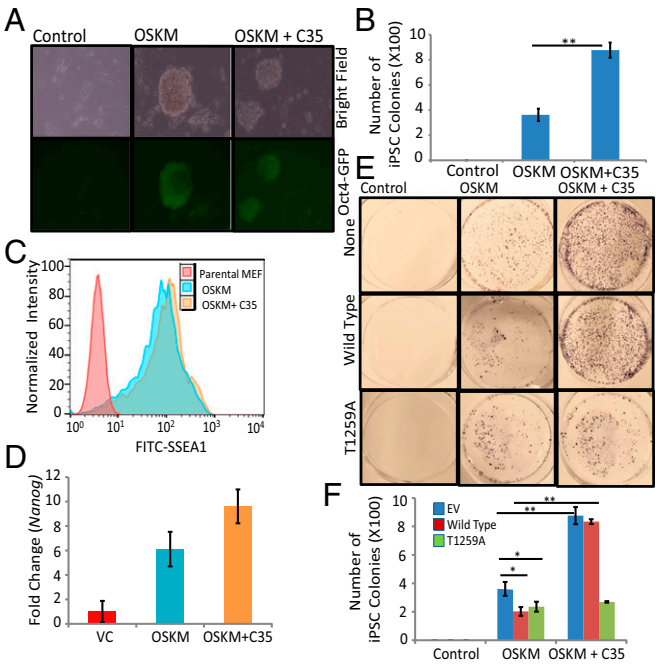

Fig. 3. C35 promotes somatic cell reprogramming. (*A*) Using viral vector, reprogramming factors OSKM were expressed in MEFs that harbor a GFP gene driven by the *oct4* promoter. After 4 wk of OSKM expression and C35 treatment, iPSCs colonies, but not the parental MEFs, showed green fluorescence. (*B*) iPSC colony number was counted. (*C* and *D*) Stem cell nature of iPSC colonies were validated by endogenous *nanog* expression (*D*) and surface expression of SSEA1 (*C*). (*E*) As shown in AP1 staining, C35 does not affect the suppression of iPSC colony by the T1259 mutation. (*F*) Quantitative analysis of the number of iPSC colonies in various conditions. VC, vehicle control. $*P < 0.05$; $**P < 0.01$.

with the OSKM expression alone (Fig. 3*B*). To further validate the cellular properties of iPSCs, we examined and found higher surface expression of stage-specific embryonic antigen 1 (SSEA1) in iPSCs than that in MEFs (Fig. 3*C*), which is a surrogate marker of iPSCs (22). Moreover, an increased *nanog* expression, which is observed during somatic cell reprogramming (22), was identified using qPCR in the iPSCs in our assays (Fig. 3*D*). We also examined alkaline phosphatase 1 (AP1) activity in the cells, which is another gold standard marker for iPSCs validation (22, 23). Again, we confirmed that C35 treatment significantly enhanced the efficiency of somatic cell reprogramming into iPSCs, as indicated by an increase in the number of AP1-positive colonies (Fig. 3*E*).

Next, we asked whether C35 treatment-induced iPSC is mediated by TET inhibition. We expressed either wild-type TET2 catalytic domain or the T1259A mutant in the MEFs (Fig. 3 *E* and *F*). In the reprogramming assays, we observed a slight decrease in the number of iPSCs colonies as compared with the empty vector. More interestingly, since C35 treatment suppresses TET2 enzymatic activity, we found that it was still able to increase the number of iPSC colonies in wild-type TET2 expressing MEFs. However, since C35 cannot suppress the enzymatic activity of the T1259A mutant (Fig. 1*F*), C35 was not able to increase in the number of iPSC colonies when cells were expressing the T1259A mutant (Fig. 3*F*). Taken together, these results suggest that C35 treatment facilitates iPCS induction by suppressing TET enzyme activities.

C35 Treatment Up-Regulates the Transcription of BMP-SMAD-ID Axis. To elucidate the mechanism by which TET inhibition facilitates somatic cell reprogramming, we examined target genes of TET enzymes by high-throughput mRNA sequencing. The mRNA was isolated from MEFs that were treated with or without C35, and was examined by Illlumina deep sequencing. A heat map based on

fold change in mRNA level in mock versus C35 treatment indicates changes of gene transcription (Fig. 4*A*). Among these genes, many of them are known to be epigenetically regulated. To dissect the effect of C35 treatment on specific molecular pathways, we performed Gene Set Enrichment Analysis and found a significant enrichment in the expression of specific members of the TGF-β family, including Bmp4/6, Smad2/3/4, Id1/2, C-myc, and so on (Fig. 4*B*). Next, we validated the results on individual genes using qPCR, with mRNAs isolated from mock and C35-treated MEFs (Fig. 4*C* and *SI Appendix*, Table S1). To exclude any off-target effect of C35 treatment, we compared the gene expression in the MEFs expressing either wild-type TET2 or the T1259A mutant. Again, the expression of wild-type TET2 did not affect the expression of genes in the BMP-SMAD-ID pathway because of the inhibitory effect of C35 treatment. However, since C35 does not suppress the enzymatic activity of the T1259A mutant, it downregulated the expression of genes in the BMP-SMAD-ID pathway (Fig. 4*E*). Taken together, these results suggest that TET enzymes regulate the BMP-SMAD-ID pathway during somatic cell programming.

Discussion

Many functions of TET enzymes are mainly a result of its catalytic activity. However, some of the TET enzyme-mediated functions are independent of enzymatic activities; rather, TET enzymes physically interact with a number of functional partners and regulate gene transcription (21, 24). Although knocking out TET enzymes has demonstrated the significance of these enzymes in epigenetic regulation, it cannot distinguish the enzymatic function of TET proteins versus their nonenzymatic functions. Thus, the biological significance and specific contribution of enzymatic and nonenzymatic activities of TET proteins remains elusive. Here, we have identified C35 as a first-in-class TET inhibitor that selectively targets TET catalytic activities, but does not abolish the TET complex (*SI Appendix*, Fig. S14).

Although we used TET2 structure for screening, C35 is able to suppress all three TET enzymes with similar IC_{50}, indicating that C35 recognizes similar sites in the catalytic domains of TET enzymes. In fact, based on the docking analysis and biochemistry validation, we found that T1259 of TET2 mediated the interaction between C35 and TET2. Moreover, this residue is conserved in all three TET family members, as well as in other species (*SI Appendix*, Fig. S6), which explains the ability of C35 to target all three TET proteins.

In the last decade, the roles of TET proteins in somatic cell reprogramming have been extensively examined, and the literature suggests that TET enzymes may regulate this process both negatively and positively during different circumstances. For instance, it has been shown that TET1 is able to promote somatic cell reprogramming by replacing Oct4, whereas in the presence of vitamin C, loss of TET1 enhances somatic cell reprogramming (16). However, such contradictions may be caused by different functional partners of TET enzymes at different stages via the nonenzymatic function of TET proteins. Using C35 treatment as a unique tool to selectively inhibit TET catalytic activity, our studies unequivocally show that the enzymatic activities of TET proteins promote somatic cell differentiation, which is in agreement

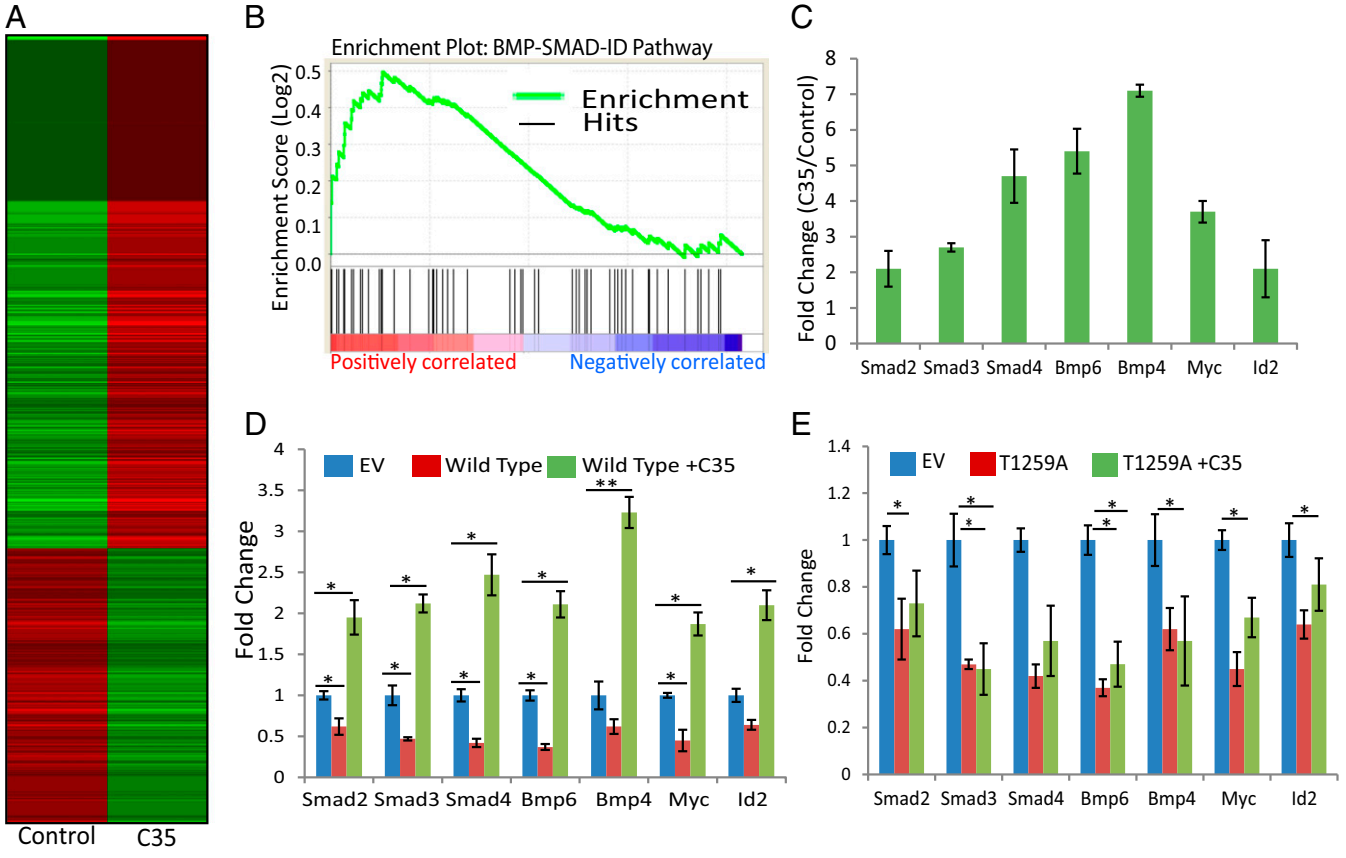

Fig. 4. C35-mediated TET inhibition induces BMP-SMAD-ID axis. (*A*) MEFs were treated with 5 μM C35, total mRNA sequencing was performed, and a comparative expression heat map was generated. (*B*) Gene Set Enrichment Analysis indicates specific enrichment of various members of the BMP-SMAD-ID pathway. (*C*) RNA sequencing results were further validated by qPCR. (*D* and *E*) As a reverse validation strategy, wild-type TET2 catalytic domain (*D*) and the T1259A mutant (*E*) were expressed in MEFs followed by C35 treatment. Expression levels of various members of the BMP-SMAD-ID pathway were analyzed by using qPCR. *$P < 0.05$; **$P < 0.01$.

with a previous report that loss of TET enzymes suppresses tissue differentiation in vivo (11, 16).

In our unbiased total mRNA sequencing study, we found that TET enzymes suppress the expression of genes in the BMP-SMAD-ID pathway, and our results indicate that C35 treatment leads to the activation of these genes. Interestingly, earlier studies from the Yamanaka group and others have shown that the BMP-SMAD-ID axis is able to accelerate the process of somatic cell reprogramming (25–27). Hence, our results align well with earlier reports and indicate that the activation of BMP-SMAD-ID axis may be the molecular basis for C35-induced somatic cell reprogramming. Furthermore, the mechanism by which TET enzymes regulate the BMD-SMAD-ID axis in the context of iPSC induction is unclear. It has been shown that TET enzyme-mediated 5hmC at the promoter regions may affect gene transcription. Moreover, our results also add a layer of understanding by revealing TET enzymes as a key missing link between the BMP-SMAD-ID signaling pathway and somatic cell reprogramming. In the mechanistic analysis, we found that suppression of TET enzymes reduces 5hmC at the promoter regions of the genes in the BMP-SMAD-ID signaling pathway, suggesting that these genes are direct targets of TET enzymes during somatic cell reprogramming.

In mammalian genomes, 5mC is a well-established and important epigenetic mark on genomic DNA, which has been shown to play important roles in imprinting, gene silencing, and chromatin remodeling. In contrast, 5hmC is considered an intermediate of oxidation of 5mC toward active DNA demethylation. However, increasing lines of evidence strongly suggest that 5hmC acts as a stable epigenetic mark and independently exhibits its own function in the genome (8, 28, 29). Since 5hmC often colocalizes with 5mC, it is likely to mediate similar function in suppression of gene transcription. Here, using C35 as an instrumental tool, we found that C35 treatment-induced expression of members in the BMP-SMAD-ID pathway is associated with a decrease in 5hmC levels, as well as MeCP2 recruitment in the respective promoter regions. Since MeCP2 is known to repress the transcription and has also emerged as an 5hmC reader, our observation strongly supports the ability of 5hmC to act as stable repressive epigenetic mark in the promoter region of various genes (8, 30). In this direction, we found that C35 treatment decreases 5hmC levels from promoters without affecting 5mC level, suggesting that 5hmC could be a stable epigenetic mark (*SI Appendix*, Fig. S13 *A* and *B*). Since we also found that loss of 5hmC associates with decreased MeCP2 and RNA polymerase II recruitment (*SI Appendix*, Fig. S13 *C* and *D*), our results support 5hmC as a transcription suppressive mark. Together, based on these preliminary results and other relevant reports, we speculate that C35-mediated 5hmC suppression may be responsible for transcriptional induction of members of BMP-SMAD-ID pathways (9, 30).

Collectively, our studies reveal C35 as a potent and cell permeable small molecule inhibitor that effectively inhibits the TET-mediated oxidation of 5mC into 5hmC without affecting the TET complex. Our further analyses indicate that C35 is able to accelerate the process of somatic cell reprogramming possibly by transcriptionally activating members of the BMP-SMAD-ID signaling pathway (*SI Appendix*, Fig. S15). To the best of our knowledge, this is a first-in-class TET inhibitor, which can be an instrumental tool to increase the efficiency of the highly inefficient process of somatic cell reprogramming. Furthermore, it can also be a very useful chemical probe for examining the role of TET enzymatic activities and the 5hmC mark on gene expression during somatic cell reprogramming and other relevant biological processes.

Materials and Methods

All the reagents, plasmids, primers, and antibodies used in this study are cataloged in the *SI Appendix*. Routine methods for virtual screening, TET activity assay, 5hmC sequencing, cell culture, transfection, protein purification, thermal shift assay, isothermal titration calorimetry, LC-MS, induction of iPSCs, induction and validation of iPSCs, high-throughput total mRNA sequencing, hMeDIP, MeDIP, ChIP, RT-PCR, flow cytometry, and immunoprecipitation are described in detail in the *SI Appendix*.

Data Availability. The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database with accession number GSE137282.

ACKNOWLEDGMENTS. We sincerely acknowledge Dr. Anjana Rao for anti-CMS sera, Dr. Rudolf Jaenisch for FUW-M2rtTA and TetO FUW OSKM constructs, and Dr. Hua Chen for providing his insight in the NMR data analysis. This work was supported in part by grants from the National Institutes of Health (CA132755 and CA130899 to X.Y.). X.Y. is a recipient of a Research Scholar Award from Tower Cancer Research Foundation.

1. K. D. Rasmussen, K. Helin, Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750 (2016).
2. L. Tan, Y. G. Shi, Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development* **139**, 1895–1902 (2012).
3. M. Ko *et al.*, Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839–843 (2010).
4. M. Tahiliani *et al.*, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
5. Y. F. He *et al.*, Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
6. S. Ito *et al.*, Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
7. C. A. Doege *et al.*, Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* **488**, 652–655 (2012).
8. M. A. Hahn *et al.*, Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep.* **3**, 291–300 (2013).
9. M. A. Hahn, P. E. Szabó, G. P. Pfeifer, 5-Hydroxymethylcytosine: A stable or transient DNA modification? *Genomics* **104**, 314–323 (2014).
10. L. Cimmino, O. Abdel-Wahab, R. L. Levine, I. Aifantis, TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell* **9**, 193–204 (2011).
11. M. M. Dawlaty *et al.*, Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell* **29**, 102–111 (2014).
12. M. M. Dawlaty *et al.*, Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* **24**, 310–323 (2013).
13. K. Takahashi, S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
14. D. Kim *et al.*, Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* **4**, 472–476 (2009).
15. K. Okita, T. Ichisaka, S. Yamanaka, Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
16. J. Chen *et al.*, Vitamin C modulates TET1 function during somatic cell reprogramming. *Nat. Genet.* **45**, 1504–1509 (2013).
17. Y. Gao *et al.*, Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
18. W. Liu *et al.*, A selective small molecule DNA2 inhibitor for sensitization of human cancer cells to chemotherapy. *EBioMedicine* **6**, 73–86 (2016).
19. M. Y. Liu *et al.*, Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine. *Nat. Chem. Biol.* **13**, 181–187 (2017).
20. Y. Huang, W. A. Pastor, J. A. Zepeda-Martínez, A. Rao, The anti-CMS technique for genome-wide mapping of 5-hydroxymethylcytosine. *Nat. Protoc.* **7**, 1897–1908 (2012).
21. Q. Chen, Y. Chen, C. Bian, R. Fujiki, X. Yu, TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* **493**, 561–564 (2013).
22. T. Brambrink *et al.*, Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2**, 151–159 (2008).
23. I.-H. Park *et al.*, Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
24. F. Zhu *et al.*, Sin3a-Tet1 interaction activates gene transcription and is required for embryonic stem cell pluripotency. *Nucleic Acids Res.* **46**, 6026–6040 (2018).
25. Y. Hayashi *et al.*, BMP-SMAD-ID promotes reprogramming to pluripotency by inhibiting p16/INK4A-dependent senescence. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13057–13062 (2016).
26. P. Samavarchi-Tehrani *et al.*, Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
27. J. Chen *et al.*, BMPs functionally replace Klf4 and support efficient reprogramming of mouse fibroblasts by Oct4 alone. *Cell Res.* **21**, 205–212 (2011).
28. W. Sun, L. Zang, Q. Shu, X. Li, From development to diseases: The role of 5hmC in brain. *Genomics* **104**, 347–351 (2014).
29. C.-X. Song, C. He, Potential functional roles of DNA demethylation intermediates. *Trends Biochem. Sci.* **38**, 480–484 (2013).
30. M. Mellén, P. Ayata, S. Dewell, S. Kriaucionis, N. Heintz, MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012).