SUMMARY

Ten-Eleven Translocation (Tet) family of dioxygenases dynamically regulates DNA methylation and has been implicated in cell lineage differentiation and oncogenesis. Yet their functions and mechanisms of action in gene regulation and embryonic development are largely unknown. Here, we report that *Xenopus* Tet3 plays an essential role in early eye and neural development by directly regulating a set of key developmental genes. Tet3 is an active 5mC hydroxylase regulating the 5mC/5hmC status at target gene promoters. Biochemical and structural studies further demonstrate that the Tet3 CXXC domain is critical for specific Tet3 targeting. Finally, we show that the enzymatic activity and CXXC domain are both crucial for Tet3’s biological function. Together, these findings define Tet3 as a transcription regulator and reveal a molecular mechanism by which the 5mC hydroxylase and DNA binding activities of Tet3 cooperate to control target gene expression and embryonic development.

INTRODUCTION

The process of vertebrate development is established through the integration of several molecular pathways controlled by key regulatory genes and complex epigenetic markings. DNA methylation at the 5-position of cytosine (5mC) is a key epigenetic mark playing crucial roles in vertebrate development (Bestor and Coxon, 1993; Bird, 1986; Reik et al., 2001). Recent studies have demonstrated that the Tet family of 5mC hydroxylases can catalyze the conversion of 5mC to 5-hydroxymethylcytosine (5hmC) (Tahiliani et al., 2009) and further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5CaC) (He et al., 2011; Ito et al., 2011). These studies also suggest that additional modification of 5mC modulated by Tet enzymes may regulate the dynamics of 5mC and its mediated gene regulation (Branco et al., 2012).

The mammalian Tet family has three members, Tet1, Tet2, and Tet3. It has been suggested that both Tet1 and Tet2 play important roles in ES cell lineage specification (Iti et al., 2010; Koh et al., 2011) and that Tet1 regulates DNA methylation and gene expression in mouse ES cells (Ficz et al., 2011; Williams et al., 2011; Wu et al., 2011; Xu et al., 2011b). Mutational inactivation of *TET2* has been reported to associate with decreased 5hmC levels in various myeloid leukemias (Delhommeau et al., 2009; Langemeijer et al., 2009), and Tet2 deficiency leads to increased hematopoietic stem cell self-renewal and myeloid transformation in mouse (Moran-Crusio et al., 2011; Quivoron et al., 2011). Recently, we and others also show that TET1 and TET2 play critical roles in other human cancers, such as melanoma and breast cancer (Hsu et al., 2012; Lian et al., 2012). In addition, Tet3 is the only Tet family member highly expressed in mouse oocytes and zygotes and is responsible for the hydroxylation of 5mC that occurs in the paternal pronucleus of advanced pronuclear-stage zygotes (Gu et al., 2011; Iqbal et al., 2011;
Wossidlo et al., 2011). Conditional knockout of Tet3 in mouse oocytes prevents resetting of DNA methylation patterns in zygotes and impairs reprogramming of transferred somatic nuclei (Gu et al., 2011). Nevertheless, Tet3−/− knockout mice are viable through development, but die on postnatal day one (Gu et al., 2011). Taken together, although the discovery of the Tet family of 5mC hydroxylases provides a potential mechanism for the dynamic regulation of DNA methylation, it remains unclear how Tet proteins are recruited to and regulate the expression of target genes, thereby providing linkage to their specific functions in early vertebrate embryonic development.

Although all Tet family members contain a conserved C-terminal catalytic domain, only Tet1 and Tet3 contain the CXXC domain, a potential DNA binding module characterized by two CXXCXXC repeats. The CXXC domains, found in other proteins such as DNMT1, MLL, and CFP1, have been shown to specifically bind to unmethylated CpG dinucleotides and participate in gene transcription regulation (Allen et al., 2006; Pradhan et al., 2008; Xu et al., 2011a). Although our previous study has suggested an important role of the CXXC domain in targeting Tet1 enzyme to specific genomic regions in ES cells (Xu et al., 2011b), the molecular mechanism and biological importance of this domain in Tet1- and Tet3-mediated transcriptional regulation of target genes remain largely unknown.

In this report, we characterize the molecular and biochemical properties and the biological function of Tet3 by using Xenopus as a model. Our study shows that Tet3 is essential for early eye and neural development in Xenopus. We also demonstrate that several master control genes essential for eye and neural development are Tet3's direct targets, mechanistically linking Tet3 function in transcriptional regulation of these key genes to the developmental phenotypes caused by Tet3 depletion. Using structural and mutational analyses and functional rescue approaches, we show that Tet3's 5mC hydroxylase and the CXXC domain-mediated DNA binding activities cooperate to regulate target gene expression during eye and neural development.

RESULTS

Identification and Characterization of Xenopus Tet3 Gene Reveal that Tet3 Is Essential for Early Eye and Neural Development

To understand the biological function of Tet proteins in early embryonic development, we investigated Tet family members in Xenopus. Database searches reveal two Tet orthologs in Xenopus tropicalis (Tet2 and Tet3), but not Tet1 (Figure S1A available online). We cloned Xenopus laevis Tet3 (xtTet3) gene and subsequently discovered two xtTet3 isoforms (HQ220207-xTet3a and HQ220208-xTet3b) that exhibit greater than 90% amino acid similarity to xtTet3 (Figure S1A). Surprisingly, despite extensive searching, we were unable to identify a Tet1 ortholog in either X. tropicalis or X. laevis. We, therefore, conclude that the Xenopus genome contains only two Tet-related genes, Tet2 and Tet3, and lacks a Tet1 gene. Sequence analysis reveals that similar to mammalian Tet3 proteins (which we cloned, validated, and deposited into GeneBank; HQ220209, human TET3; and HQ423151, mouse Tet3), Xenopus Tet3 contains a CXXC domain, a cysteine-rich domain and a double-stranded β helix (DSBH)-containing dioxygenase domain (Figure S1A).

We next examined the expression profile of xtTet3 during embryogenesis. The temporal expression pattern of xtTet3 by RT-qPCR reveals that, unlike the high level of Tet3 mRNA observed in mouse oocytes (Gu et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011), the xtTet3 mRNA level in oocytes is very low. xtTet3 mRNA is also barely detectable at stage 2 (2-cell stage) through stage 7, but increases dramatically from stage 9 (immediately after the initiation of zygotic transcription) to stage 19 and then drops at stage 25 (Figure S1B). The spatial expression pattern of xtTet3 by in situ hybridization shows that xtTet3 mRNA is barely detectable at stage 9 (Figure 1A), but clearly detected by stage 14 in the neural plate (Figure 1B). In embryo sections, we also detect xtTet3 mRNA in the neural plate and notochord (Figures S1C and S1D). Moreover, xtTet3 mRNA level remains high in the neural tube at stage 19 (Figure 1C) and in the region of the developing brain, eye, branchial arches, cement gland, and spinal cord at stage 25 (Figures 1D and 1E).

To address Tet3 function during embryogenesis, we performed loss-of-function studies by depleting endogenous xtTet3 protein by using the Morpholino antisense oligo (MO) strategy. We first confirmed that xtTet3 MOs efficiently deplete Tet3 protein (Figure 1F). Control MO or xtTet3 MOs were injected into two dorsal blastomeres of 4-cell stage embryos. At stage 35, we observe striking developmental abnormalities in 96% (194/201) of xtTet3 MOs-injected embryos, including malformation of the eye (eyeless), small head, and missing pigmentation along the lateral body, whereas control embryos develop normally (Figure 1G). The Tet3-depleted embryos die between stages 35 and 40. Importantly, these phenotypes are rescued by coinjecting the xtTet3 MOs-resistant xtTet3 mRNA (Figure 1G, right). Taken together, these data suggest that Tet3 plays an essential role in early embryogenesis, especially in early eye and neural development.

Tet3 Directly Regulates a Set of Genes Critical for Eye and Neural Development

To gain insight into the underlying molecular mechanisms linking Tet3 function to early eye and neural development, we examined the effect of Tet3 depletion on the expression of a set of key developmental genes, including pax6 (eye and neural marker), rx and six3 (eye markers), sox2 (pan-neural marker), 

sox2 (antierior neural marker), 

sox9 and snail (neural crest markers), neuroglin related 1 (ngn2) and n-tubulin (tubb2b) (primary neuron markers), and shh and ptc-1 (sonic hedgehog signaling) in stage 14 embryos. MOs were injected into one dorsal blastomere of 4-cell stage embryos so that the un.injected side could be used as an internal control. For shh and ptc-1 measurement, we injected MOs into two dorsal blastomeres of 4-cell stage embryos because the expression of shh and ptc-1 is in the midline. As shown by in situ hybridization, the expression of master eye developmental genes, pax6, rx, and six3, is greatly reduced on the xtTet3 MOs-injected side compared to the un injected side at stage 14 (Figure 2A). Reduced expression of these eye genes is also seen in the potential eye field on the Tet3-depleted side at stage 19 (Figure 2A). Depletion of Tet3 also inhibits the expression of two primary neuron markers, ngn2 and tubb2b, and two
neural crest markers, sox9 and snail, supporting a critical role of Tet3 in neural and neural crest development (Figure 2A). The expression of two major shh signaling components, shh and ptc-1, is also abolished in Tet3-depleted embryos at stage 14 (Figure 2A), whereas the expression of otx2 and sox2 shows no significant changes (Figure S2B). Importantly, the diminished expression of all affected genes is rescued by xtTet3 mRNA coinjection (Figure 2A), demonstrating the specific regulation of these genes by Tet3. Results by in situ hybridization are further independently confirmed by RT-qPCR assays (Figures 2B and S2C).

To address whether Tet3 directly regulates these key genes, we performed chromatin immunoprecipitation (ChIP) assays. We employed an epitope-tagged ChIP assay by using Flag antibody and flag-xtTet3 mRNA-injected embryos to explore Tet3 occupancy at the promoters of these genes. We detect strong Tet3 binding at the promoters of pax6, rx, six3, ptc-1, ptc-2, sox9, and ngn2 but not at the promoters of control genes, myosin light chain 2 (myl2), and cardiac actin (actc) (Figure 2C).

Finally, we tested whether ectopic expression of any one of these downstream Tet3 target genes can rescue the phenotypes caused by Tet3 depletion. We examined the rescue effects of pax6, rx, or shh overexpression. As shown in Figure S2D, none of these can rescue the phenotypes, suggesting that it may be the cumulative effect of the altered expression of a group of target genes, rather than a single gene, that causes the observed developmental defects. Nevertheless, our data strongly suggest that Tet3 is an upstream transcriptional regulator specifically and directly controlling a set of key genes important for early eye and neural development.

**Tet3 Is an Active 5mC Hydroxylase that Regulates the 5mC/5hmC Status at Target Gene Promoters**

As DNA methylation at gene promoters has regulatory effects on the expression of the associated genes, we asked whether Xenopus Tet3 modulates the 5mC/5hmC status at its target gene promoters. To address this question, we first demonstrate that Xenopus Tet3 has conserved 5mC hydroxylase activity as mammalian Tet proteins. The xtTet3 catalytic domain (CD) can convert 5mC to 5hmC in cells and in vitro (Figures 3A and 3B). Utilizing two independent methods, anti-5hmC antibody-based dot-blot and T4 Phage β-glucosyltransferase-mediated 5hmC glucosylation assays, we demonstrate that 5hmC exists in the...
genome of Xenopus embryos and is globally reduced after Tet3 depletion (Figure 3C). However, consistent with undetectable or extremely low levels of 5fC and 5CaC in ES cells and various tissues (Globisch et al., 2010; Ito et al., 2011), we failed to detect 5fC and 5CaC in Xenopus embryos (data not shown). Collectively, these data suggest that Tet3 possesses intrinsic 5mC
Figure 3. Tet3 Is an Active 5mC Hydroxylase Regulating the 5mC/5hmC Status at Target Gene Promoters
(A) xITet3 CD is sufficient to convert 5mC to 5hmC in HEK293T cells by immunofluorescence analysis. Flag-tagged xITet3 CD protein was detected using Flag antibody.
(B) xITet3a CD protein converts 5mC to 5hmC in vitro by dot-blot assay. Avidin-HRP is used to detect total biotin-labeled DNA, showing equal loading.
(C) Tet3 depletion results in globally decreased 5hmC in stage 14 embryos by dot-blot (left) and 5hmC glucosylation (right) assays. Open bar, control MO; filled bar, xITet3 MOs. Data are presented as mean ± SD (n = 3). **p < 0.01.
(D and E) hMeDIP-qPCR to detect dynamic 5hmC level changes in stage 10, 14, and 19 embryos. The targeting region for each primer set is underlined in (D). Results are shown as mean ± SD (n = 3) in (E). *p < 0.05.
(F) Site-specific 5hmC level changes by Tet3 depletion in stage 14 embryos using the EpiMark 5mC/5hmC analysis kit. Open bar, control MO; filled bar, xITet3 MOs. Red dot indicates MspI/HapII recognition site and each PCR amplified region is underlined. Arrow denotes promoter orientation. Data are shown as mean ± SD (n = 3). *p < 0.05.
(G) TAB-seq analyses of 5hmC status at the promoter of pax6 (left) and actc (right) in stage 14 embryos. The average percent at each CpG site is derived from sequencing of 30–32 clones for pax6 promoter and 24–26 clones for actc promoter.
(H) Schematic diagram of hMeDIP-MeDIP strategy.
(I) hMeDIP-MeDIP qPCR to detect 5mC level changes after Tet3 depletion in stage 14 and 19 embryos. Data are shown as mean ± SD (n = 3). *p < 0.05. The targeting region for each primer set is shown in (D).
See also Figure S3.
increasing 5hmC levels at a Tet3 target gene promoter, but not at the nontargeted actc promoter (Figures 3D and 3E). Importantly, Tet3 depletion results in a significant reduction in 5hmC at rx promoter at stages 14 and 19, thereby significantly abolishing the increase in 5hmC from stage 10 to 14 to 19 (p < 0.05), whereas no 5hmC level changes are observed at actc promoter (Figure 3E). However, we failed to detect 5hmC level changes at other Tet3 target gene promoters including pax6, six3, sox9, and ptc-2 by hMeDIP-qPCR probably due to the low 5hmC levels at those promoters and the detection limits of hMeDIP-qPCR. Thus, we applied a more sensitive CCGG site-dependent strategy, EpiMark method (Ficz et al., 2011). As exemplified by pax6 and ptc-2, depletion of Tet3 results in a significant reduction in 5hmC at CCGG sites of both promoters (p < 0.05), whereas no 5hmC changes are observed at the nontargeted myl2 and actc genes (Figure 3F). To next determine the 5hmC status at multiple Cpg sites, we employed the newly developed Tet-assisted bisulfite sequencing (TAB-Seq) approach (Yu et al., 2012), the most sensitive method to specifically detect 5hmC at base-pair resolution. We first confirm that the different amount of starting DNA (hMeDIPed DNA) will not introduce a MeDIP efficiency bias, as validated by the equal MeDIP enrichment of the human NBR2 promoter from control DNA spiked in the hMeDIPed DNA from control MO- or Tet3 MOs-injected embryos (Figure S3D). For proof-of-principle, we examined the rx gene promoter because it has the most dramatic changes in 5hmC after Tet3 depletion and can be easily detected (Figure 3E). Indeed, we clearly observe a significant 5mC level increase at rx gene promoter after Tet3 depletion by this hMeDIP-MeDIP qPCR method in stage 14 and 19 embryos (Figure 3I), whereas the 5mC level at actc promoter is not changed (Figure 3I), validating the reliability of hMeDIP-MeDIP qPCR procedure. Thus, we conclude that Tet3 regulates target gene expression, at least partially, through control of 5mC/5hmC status at the promoter of target genes.

The Tet3 CXXC Domain Possesses Unique DNA Binding Properties

In addition to the dioxygenase domain that confers 5mC hydroxylase activity, Tet3 also contains a potential DNA binding domain, the CXXC domain. In general, the selective DNA binding activity of a transcription factor serves as a key mechanism for action of the transcription factor in gene transcriptional regulation. Sequence alignment indicates that Tet CXXC domains exhibit a conserved overall structure with other CXXC domains (Figure S4A). However, Tet CXXC domains lack a short sequence motif (KFGG) (Figure S4A), which has been shown to be important for the DNA binding activity of the MLL and CFP1 CXXC domains (Allen et al., 2006; Xu et al., 2011a). To address whether the Tet3 CXXC domain may possess unique DNA binding features, we examined the DNA binding ability and specificity of the Xenopus Tet3 CXXC domain (aa 58–111) by isothermal titration calorimetry (ITC) assays. The Tet3 CXXC domain strongly binds to various C/G-rich DNA oligos but exhibits virtually no binding activity to the A/T-only DNA oligo (Figures 4A–4E and Table S1, 1–8). Significantly, these ITC results also reveal previously uncharacterized DNA binding properties of the CXXC domain. First and most importantly, the Tet3 CXXC domain strongly binds to both non-Cpg (Table S1, 5–7) and Cpg DNA oligos (Table S1, 1–4 and 9) with a slight preference for Cpg DNA oligos. Second, the Tet3 CXXC domain strongly binds to CmCGG DNA (Figure 4F and Table S1, 10). These binding properties are in stark contrast to the well-characterized DNA binding property of the CFP1 and MLL CXXC domains, which absolutely require unmethylated Cpg dinucleotides (Allen et al., 2006; Xu et al., 2011a). In contrast to the strong binding to CmCGG, CTAG and ACGT DNA oligos, the Tet3 CXXC domain does not bind to mCmCGG, mCTAG, or AmCGT DNA oligos (Figure 4G and Table S1, 11–13). These data suggest that although strict Cpg content is not required for the Tet3 CXXC domain to interact with DNA, an unmodified cytosine is
Figure 4. The Unique DNA Binding Properties of the Tet3 CXXC Domain

(A–H) Binding affinities of the Tet3 CXXC domain to various DNA oligos by ITC assays. The sequence of the central four nucleotides of each double-stranded DNA probe is shown under the corresponding panel. Detailed sequence information for all DNA oligos used in this study is listed in Table S1. NB: no binding.

(I) Representative TET3 CXXC GST pull-down sequencing results. Arrow denotes promoter orientation.

(J) Genomic distribution of the TET3 CXXC-bound loci. Promoter is defined as –2 kb to +2 kb relative to transcription start site (TSS).


(L) DNA motifs that are enriched in the TET3 CXXC-bound loci.

See also Figure S4 and Table S1.
essential for the Tet3 CXXC binding to DNA. Using the same ITC approach, we also demonstrate that the human TET3 CXXC domain has similar DNA binding properties to the Xenopus Tet3 CXXC domain (Table S1, 17–22). Thus, the CXXC domain of the Tet3 family has unique newly identified DNA binding characteristics conserved among vertebrates.

Given the similar binding properties between human and Xenopus Tet3 CXXC domains, we next determined the DNA binding specificity of the TET3 CXXC domain across the whole genome of HEK293T cells. The specific TET3 CXXC-bound genomic DNA fragments were enriched by GST pull-downs and analyzed by deep DNA sequencing, as we previously described (Xu et al., 2011b). The TET3 CXXC domain selectively binds to restricted genomic regions (Figure S4B). Importantly, consistent with Tet3 occupancy in Xenopus embryos (Figure 2C), the TET3 CXXC domain strongly and selectively binds to the promoters of PAX6, PTC1, NGN2, TUBB2B, and SHH, but not the promoters of MYL2 (Figures 4I and S4C). Furthermore, bioinformatics analyses identify 17,953 TET3 CXXC-bound peaks, more than half of which are located at gene promoters (Figure 4J). Interestingly, we also observe a significant enrichment of CpG (and a less degree of CpC) dinucleotides in TET3 CXXC-bound regions (Figure 4K). Further de novo motif discovery analyses identify three C-rich sequences among the top 15 ranking motifs within TET3 CXXC-bound regions (Figure 4L). Together with the ITC binding data, these genome-wide analyses suggest that the TET3 CXXC domain is able to bind to the unmodified C followed by A, T, C, or G with a slight preference for CpG dinucleotides. These data also indicate that the specific DNA binding activity of the TET3 CXXC domain may contribute to Tet3 targeting, thereby serving as another important mechanism for Tet3-mediated gene transcriptional regulation.

Crystal Structures of the Tet3 CXXC Domain in Complex with DNA Oligos

To gain further mechanistic insight into the unique DNA binding properties of the Tet3 CXXC domain, we next determined the crystal structures of the Xenopus Tet3 CXXC domain (aa 58–111) in complex with 5mC-containing DNA (CmCGG) and CpG-containing DNA (ACGT), respectively (Table 1). Like other CXXC domains, the Tet3 CXXC domain contains eight conserved cysteine residues coordinating two zinc ions. These two zinc ions play a structural role by holding the main unstructured CXXC domain together and forming a crescent-shaped architecture to bind DNA (Figures 5A and 5C). The DNA-binding surface is predominantly positively charged, and wedged into the major groove of DNA to extensively interact with DNA bases by means of hydrogen bonds and electrostatic interactions (Figures 5A–5D).

Surprisingly, structural comparison of the Tet3 CXXC domain in complex with CmCGG or ACGT DNA reveals that the Tet3 CXXC domain binds to the 12-mer target DNAs with one nucleotide shift. In the crystal structure of the Tet3 CXXC domain in complex with the unmethylated CpG DNA (ACGT), the Tet3 CXXC domain binds to the DNA centering on the cytosine in the CpG dinucleotide (underlined). Methylation of the cytosine in the CpG dinucleotides (CmCGG) shifts the binding to the cytosine (underlined) preceding the mCpG dinucleotides. Therefore, consistent with the ITC binding results, the Tet3 CXXC domain binds to an unmodified cytosine, which is not restricted by CpG content, distinct from the CpG-dependent binding of the MLL, CFP1, and DNMT1 CXXC domains (Allen et al., 2006; Pradhan et al., 2008; Xu et al., 2011a).

In both Tet3 CXXC-ACGT and Tet3 CXXC-CmCGG complex structures, the main chain carbonyl oxygen of residue His90 forms a hydrogen bond with the target cytosine, whereas the side chain of His90 forms another hydrogen bond with the guanine-cytosine base pair (Figures 5E and S5B). In the Tet3 CXXC-ACGT complex structure, residues Gin91 and Ser89 form hydrogen bonds with the guanine-cytosine base pair following the target cytosine-guanine base pair (Figures 5F and S5B). In contrast, in the Tet3 CXXC-CmCGG complex structure, methylation of C6 introduces steric clashes with the side chain of Gin91 and causes the side chain of Gin91 to become partially disordered. As a result, Ser89 flips peptide plane and

---

**Table 1. Data Collection and Refinement Statistics**

<table>
<thead>
<tr>
<th>PDB code</th>
<th>xTet3 CXXC-GCGCCACGGTCGCG</th>
<th>xTet3 CXXC-CmCGGTGGA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>C2</td>
<td>C2</td>
</tr>
<tr>
<td>Cell Dimensions</td>
<td>a, b, c (Å)</td>
<td>69.9, 39.5, 54.1</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>34.2-2.05</td>
<td>35.5-2.25</td>
</tr>
<tr>
<td>Number of unique HKLs</td>
<td>8829/351</td>
<td>7385/351</td>
</tr>
<tr>
<td>Number of atoms/Å&lt;sup&gt;3&lt;/sup&gt;</td>
<td>898/54.8</td>
<td>886/73.9</td>
</tr>
<tr>
<td>DNA</td>
<td>486/58.2</td>
<td>487/78.7</td>
</tr>
<tr>
<td>Protein</td>
<td>384/51.7</td>
<td>387/68.5</td>
</tr>
<tr>
<td>Zn&lt;sup&gt;2+&lt;/sup&gt;</td>
<td>2/33.2</td>
<td>2/46.8</td>
</tr>
<tr>
<td>RMSD bonds (Å/angles (°))</td>
<td>0.013/1.4</td>
<td>0.012/1.4</td>
</tr>
<tr>
<td>Ramachandran plot favored residues, no outliers (Lovell et al., 2003)</td>
<td>47 of 49</td>
<td>48 of 49</td>
</tr>
</tbody>
</table>

RMSD is an abbreviation for root mean squared deviation. Average B-factors calculated with MOLEMAN (G.J. Kleywegt, Uppsala University). The highest resolution shell is shown in parentheses.
Figure 5. The Tet3 CXXC Domain Specifically Recognizes Cytosine through a Conserved Residue His90
(A and C) Crystal structures of the Tet3 CXXC domain in complex with ACGT DNA (5'-GCCAACGT(TGGC-3')) (A) or CmCGG DNA (5'-GCCACmCGGTGGC-3') (C) in cartoon (left) and electrostatic representations (right), respectively. The double-stranded DNA sequence is shown in the middle of each corresponding panel.
loses its hydrogen bond interaction with DNA (Figures 5H and S5D). These data provide structural explanations for the binding to CmCGG DNA and the slight binding preference for CpG compared to CpH DNAs (H = A, T, C, or 5mC). Substituting the target cytosine with any other nucleotide or 5mC would introduce steric clashes with His90 (Figure S5E), consistent with the undetectable binding of the Tet3 CXXC domain to ATAT DNA (Figure 4E). To determine the role of residue His90 of Tet3, which is conserved among all Tet CXXC domains (Figure S4A), we generated the Tet3 CXXC H90A mutant and tested its DNA binding activities by ITC assay. As expected, the Tet3 CXXC H90A mutant loses its DNA binding abilities (Table S1, 14–15), supporting an essential role of His90 in DNA binding.

To further understand the novel DNA binding property of the Tet3 CXXC domain, we superimposed the complex structures of the Tet3 and CFP1 CXXC domains (Xu et al., 2011a) (Figure S6A). Although the Tet3 CXXC domain adopts a similar fold to that of the CFP1 CXXC domain, the loop region preceding His90 of Tet3 is dramatically different from that of CFP1 (Figure S6A). In the CFP1 CXXC domain, Asp189 forms three hydrogen bonds with the backbone of Lys198, Ile199, and Arg200, making the loop very rigid and only allowing CpG binding (Xu et al., 2011a). Notably, this Asp residue is highly conserved in the CXXC domains of CFP1, MLL, DNMT1, and KDM2A (Figure S4A). However, the corresponding loop between the α3 and α4 helices of the Tet3 CXXC domain is much shorter and more flexible due to lack of hydrogen bonds seen in the rigid loop of the CFP1 CXXC domain (Figure S6A). Therefore, the shorter and less rigid loop of the Tet3 CXXC domain can accommodate other nucleotides besides G, such as T, C, A, or 5mC, following the target C, which confers cytosine-specific instead of CpG-specific binding property. In addition, we also observe that the Tet3 CXXC domain binds to TCGA and mCmCGG DNA oligos (Table S1, 2 and 9) with slightly weaker binding affinities compared to other CpG DNA oligos (Table S1, 1, 3 and 4). Although the methyl group of thymine or 5mC does not cause any steric clash with the Tet3 CXXC domain, the hydrophobic methyl group points to the solvent and no residues in the Tet3 CXXC domain can recognize and stabilize it, which is not energetically favorable (Figures S6B–S6E). Taken together, the Tet3 CXXC domain binds to cytosine-containing DNA with a slight preference for a G following the target cytosine and modestly disfavoring a T or mC preceding the target cytosine.

Both 5mC Hydroxylase Activity and CXXC Domain-Mediated Specific DNA Binding Are Required for Tet3 Function in Target Gene Regulation and Embryonic Development

Having defined the regulation of 5mC/5hmC status at Tet3 target gene promoters by the 5mC hydroxylase activity and the unique DNA binding properties of the Tet3 CXXC domain, we next determined the role of these two functional domains in Tet3 function. We generated wild-type (xtTet3), CXXC-deleted (xtTet3ΔCXXC), His90-to-Ala (xtTet3H90A), and iron binding site-disrupted catalytically inactive (xtTet3ΔHD) xtTet3 expression constructs (Figure 6A). We confirm that both the wild-type and mutant xtTet3 proteins are properly expressed and localized to the nucleus (Figures S7A and S7B) and that the 5mC hydroxylase activity is retained in xtTet3ΔCXXC and xtTet3H90A but not in xtTet3ΔHD (Figure S7B). We also performed ChIP-qPCR assays to compare the occupancy of wild-type and mutant Tet3 proteins at target gene promoters. Both the CXXC domain deletion and H90A point mutation abolish the occupancy of mutant proteins at target gene promoters (Figures 6B and S7C), suggesting an essential role of the CXXC domain-mediated DNA binding in Tet3 targeting to specific genes. Importantly, although both xtTet3 and xtTet3ΔHD exhibit similar occupancy, coinjection of xtTet3, but not xtTet3ΔHD, significantly rescues the decreased 5hmC levels caused by Tet3 depletion at target gene promoters including the rx gene promoter (Figure S7D), suggesting that the enzymatic activity of Tet3 is primarily responsible for the dynamic regulation of 5hmC at its target genes.

We next used these constructs to determine which domain(s) are critical for Tet3 functions in vivo. We first employed a functional domain rescue approach to interrogate the role of Tet3 catalytic domain using pax6 gene expression as a readout. The expression of pax6 is completely rescued in 79% of xtTet3 co-injected embryos, whereas the complete rescue effect of xtTet3ΔHD coinjection only reaches 37%, significantly lower than that of xtTet3 coinjection (p < 0.01) (Figure 6C). These results are corroborated by phenotypic rescues. Only 25% of embryos are completely rescued by xtTet3ΔHD coinjection, again significantly lower than that of xtTet3 coinjection (p < 0.01) (Figure 6D). Thus, these data suggest that the 5mC hydroxylase activity of Tet3 not only controls the dynamics of 5mC/5hmC at target gene promoters but is also required for target gene regulation and biological function in early embryonic development. Next, we employed the same strategy to examine the importance of the CXXC domain/His90-mediated specific DNA binding in Tet3 function. Strikingly, xtTet3 H90A or xtTet3ΔHD coinjection shows no rescue effects on the inhibited pax6 expression or developmental phenotypes caused by Tet3 depletion (Figure 6C and 6D), highlighting an essential role of the CXXC domain-mediated DNA binding in Tet3 function in target gene regulation and embryonic development.

DISCUSSION

In this study, we report that Xenopus Tet3 is a new class of transcription regulator playing an essential role in early eye and neural development by directly regulating a set of key genes critical for these developmental processes. We also uncover that
Figure 6. Both 5mC Hydroxylase Activity and the CXXC Domain Are Important for Tet3 Function

(A) Schematic representation of xtTet3 mutants.
(B) The CXXC domain deletion disrupts Tet3 occupancy at specific gene promoters by ChIP-qPCR assay. Data are presented as mean ± SEM (n = 3). *p < 0.05.
(C) Summarized results of five independent pax6 expression rescue experiments. “Suppressed” means significantly suppressed pax6 expression in posterior and anterior neural plates; “Partially suppressed” means pax6 expression is detected but not intact in posterior and anterior neural plates; “Normal” means intact pax6 expression in posterior and anterior neural plates. **p < 0.01 compared to xtTet3 MOs, # p < 0.01 compared to xtTet3 MOs/xtTet3.
the dynamic regulation of 5mC/5hmC status at target gene promoters is an important mechanism underlying Tet3-mediated target gene regulation during embryonic development. The Tet3 CXXC domain-mediated specific DNA binding is essential for targeting Tet3 to its target genes, thus, providing another layer of regulation on the transcription of Tet3 target genes. Our findings support a molecular model of Tet3 action in target gene regulation that involves Tet3 binding to unmethylated cytosines (with a slight preference for CpG content) at target gene promoters through its CXXC domain and its intrinsic 5mC hydroxylase activity converting adjacent 5mC to 5hmC, which is an intermediate for further DNA demethylation (Figure 6E). Recruitment to a specific gene promoter and subsequent conversion of 5mC to 5hmC cooperatively activate the expression of Tet3 target genes, including those identified key developmental genes, to ensure normal and precisely regulated embryogenesis. Inactivation of either one of these functional domains therefore will have adverse impact on Tet3 function. Notably, our data also show that the catalytically inactive Tet3 mutant retains partial rescue effects in vivo, suggesting potential enzymatic activity-independent mechanisms, such as recruiting or interacting with other transcription factors, may also contribute to the full function of Tet3 in gene regulation (Figure 6E). Indeed, TET3 forms a stable complex with several critical transcription factors and histone modifiers (data not shown). It has also been shown that an enzymatic activity-independent mechanism is involved in gene regulation by Tet1 in mouse ES cells (Williams et al., 2011). Moreover, it has been suggested that 5hmC is a stable epigenetic mark that participates in regulating gene expression through unidentified mechanisms such as recruiting unidentified 5hmC specific “readers” (Branco et al., 2012; Maturase et al., 2011). Thus, this model does not exclude the possibility that Tet3 and 5hmC may participate in target gene regulation through this yet to be identified regulatory circuit.

Unlike well-characterized DNA methylation (5mC) and DNMTs, the mechanism by which 5hmC and the enzymatic activity of TET proteins contribute to gene regulation has been elusive. Our findings here offer a biological model demonstrating that the dynamic regulation of 5mC/5hmC by the Tet family of 5mC hydroxylases has an important role in gene regulation during early eye and neural development. We show that Tet3 is an active 5mC hydroxylase highly expressed in the region of the developing brain, eye, and spinal cord in Xenopus embryos. As exemplified by rx and pax6, our data clearly show that Tet3 target genes have significant alterations in 5hmC status at their promoters after Tet3 depletion. Importantly, 5hmC level alterations and dysregulated gene expression can be completely rescued by wild-type Tet3. Yet the rescue effect by the catalytically inactive Tet3 mutant is significantly impaired, highlighting the importance of 5mC hydroxylase activity in Tet3 function. Furthermore, consistent with the eye developmental phenotypes in Xenopus, we find that homozygous deletion of the Tet3 catalytic domain in pure B6 genetic background mice (Gu et al., 2011) results in eye-related phenotypes including the eyelid open at birth (EOB) phenotype (data not shown), suggesting an evolutionally conserved function in eye development from vertebrate to mammals and the critical role of the 5mC hydroxylase activity in Tet3 function. Taken together, our study clearly indicates that precise regulation of the dynamics of DNA modification status at specific gene loci by Tet3-mediated conversion of 5mC to 5hmC is an important and conserved epigenetic mechanism for target gene regulation.

How epigenetic enzymes are targeted and/or confined to their functional sites is a fundamental question in understanding the mechanism underlying epigenetic transcription regulation. In this study, our biochemical characterizations reveal that the Tet3 CXXC domain binds to DNA in a cytosine-dependent manner with a slight preference for CpG dinucleotides, distinct from the CpG-dependent binding of other well-characterized CXXC domains. We also notice that the DNA binding property of the Tet3 CXXC domain is different from that of the TET1 CXXC domain that binds to unmodified C or 5mC- or 5hmC-modified CpG-rich DNA (Xu et al., 2011b), suggesting that Tet1 and Tet3 may have different functions in gene regulation besides their distinct expression patterns during early embryonic development (Tan and Shi, 2012). Moreover, our genome-wide mapping and de novo motif analyses of the TET3 CXXC domain binding sites strongly suggest that the TET3 CXXC domain selectively binds to several consensus sequences, therefore offering a mechanism for targeting Tet3 to its target genes. Finally, the crystal structure analysis of the Tet3 CXXC domain in complex with DNA provides the following structural mechanisms underlying its novel DNA binding ability and specificity. (1) It reveals that the major binding force of the Tet3 CXXC domain to CmCGG DNA is through the specific interaction between residue His90 and the unmethylated cytosine, which explains why the Tet3 CXXC domain can still bind to partially methylated DNA, as long as there is an unmethylated cytosine available in the target sequence; (2) it explains the slight binding preference of the Tet3 CXXC domain for CpG compared to CpH (H = A, T, C, or 5mC) and the modest disfavor for a T or 5mC preceding the target C, suggesting that sequences flanking the target C contribute to the overall binding affinity and specificity of the Tet3 CXXC domain.

The importance of this CXXC domain-mediated specific DNA binding activity of Tet3 is further demonstrated in our functional studies using the Tet3 mutants with the CXXC domain deletion (∆CXXC) or H90A point mutation. These mutants lose the specific association with Tet3 target gene promoters and functional rescue abilities in vivo. Taken together, our study reports a novel DNA binding property and functionality of the CXXC domain. Please refer to the related text for more details. See also Figure S7.
domain that is essential for Tet3 function both in gene regulation and embryonic development. It not only provides a key mechanistic layer of Tet3-mediated target gene regulation but also significantly advances our current understanding of the molecular and biological function of the Tet family CXXC domains. Noteworthy, even though the CXXC domain is essential for targeting Tet3 to specific genomic regions, we do not exclude the involvement of other potential cellular mechanisms for targeting or recruiting Tet3 to its functional sites. We favor the hypothesis that Tet3 is likely in complex with many other cellular factors including sequence-specific DNA binding transcription factors and cofactors, to execute its molecular, cellular, and biological functions. Therefore, the associated transcription factors may in part coordinate with the Tet3 CXXC domain and enzymatic activity and contribute to the overall mechanism of action of Tet3 in gene regulation and embryonic development.

EXPERIMENTAL PROCEDURES

Embryo Manipulation and Microinjection

X. laevis eggs were artificially fertilized with testis homogenate and cultivated in 0.1× MMR as previously described (Sakano et al., 2010). Capped synthetic mRNAs were generated by in vitro transcription with SP6 polymerase. Embryos were transferred to 3% Ficol 400 in 0.1× MMR and injected embryos were cultured in 0.1× MMR until the desired stage. For phenotype experiments, 80 ng of xTet3 MOs (40ng of each xTet3 MO) or control MO were injected into two dorsal blastomeres of 4-cell stage embryos. For phenotype rescue experiments, 1 ng of mRNA was coinjected with 80 ng of xTet3 MOs into two dorsal blastomeres of 4-stage cell embryos. For pax6, rx, six3, sox2, sox9, snail, tubb2b, and ngn2 in situ hybridization experiments, 160 ng of ngn2 MOs or control MO were injected into one dorsal blastomere of each 4-cell stage embryo. Because the expression of shh and ptc-1 is on the midline, for shh and ptc-1 in situ hybridization experiments, two dorsal blastomeres of each 4-cell stage embryo were injected with 160 ng of xTet3 MOs or control MO. In all injection studies, 200 pg of nuc-β-gal RNA were coinjected as the injection tracer.

Cryocrystallization

Each pair of single-stranded DNAs was mixed with a molar ratio of 1:1 and annealed to form double-stranded DNAs. Before cryocrystallization, purified xTet3 CXXC protein was mixed with different DNAs in a molar ratio of 1:1.2. Crystals of xTet3 CXXC in complex with aCGT DNA (GCAACGTTGGC) were obtained via sitting drop vapor diffusion; 1.0 μl of complex was mixed with 1.0 μl of well solution containing 0.1 M HEPES (pH 7.5), 0.2 M NaCl, 30% PEG 1500, against 800 μl of reservoir buffer at 18°C. Crystals grow to a mountable size in 3 days. Crystals of xTet3 CXXC in complex with CmCGG DNA (GCAACGTTGGC) were obtained in a similar way in the buffer containing 0.1 M HEPES (pH 7.5), 0.2 M NaCl, 30% PEG 1500, 5% MPD. Both crystals were flash-frozen in liquid nitrogen directly without cryoprotectant.

ACKNOWLEDGMENTS

We thank R. Harland for helpful discussions, A Khorasan for critical reading of the manuscript and Q. Xu and L. Jiang for statistical analyses. We also thank R. Moon, W. Harris, J. Yang, X. Wang, M. Vetter and J.-P. Saint-Jeannet for valuable discussions and technical help. This work was supported by NIH grants GM078458 to Y.G.S., GM067641 to Y.K., and GM074241 to X.H., Brigham and Women’s Hospital Biomedical Research Institute fund to sustain research excellence to Y.G.S., and Chinese Ministry of Education Project 985 to Biomedical Core Facility, Fudan University. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. J.G.A. was funded partially by CNPq (Brazil). Y.G.S. is a Pew Scholar. X.H. acknowledges support by Boston Children’s Hospital Intellectual and Developmental Disabilities Research Center (P30 HD-18655). The Structural Genomics Consortium is a registered charity (No. 1097737) that receives funds from Canadian Institutes for Health Research, Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Eli Lilly, Pfizer, Novartis Research Foundation, Life Technologies, Ontario Innovation Trust, Ontario Ministry for Research and Innovation and Wellcome Trust (JM). Argonne National Laboratory is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

REFERENCES


