Biochemical and Biophysical Research Communications 439 (2013) 373–377

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

# Characterization of the nuclear localization signal of the mouse TET3 protein

Peng Xiao, Xiao-long Zhou, Hong-xiao Zhang, Kai Xiong, Yun Teng, Xian-ju Huang, Rui Cao, Yi Wang, Hong-lin Liu *

*College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, People's Republic of China*

ARTICLE INFO

*Article history:*
Received 14 August 2013
Available online 30 August 2013

*Keywords:*
TET3
NLS
NIH 3T3

ABSTRACT

DNA demethylation is associated with gene activation and is mediated by a family of ten-eleven translocation (TET) dioxygenase. The TET3 protein is a 1668-amino-acid DNA demethylase that is predicted to possess five nuclear localization signals (NLSs). In this paper, we used a series of green fluorescent protein-tagged and mutation constructs to identify a conserved NLS (KKRK) embedded between amino acid 1615 and 1618 of mouse TET3. The KKRK sequence facilitates the cytoplasmic protein's translocation into the nucleus. Additionally TET3 may be imported into the nucleus by importin-α and importin-β.

## 1. Introduction

Nuclear proteins, imported into the nucleus through nuclear pore complexes (NPCs), have important roles in biological events, including transcription and DNA demethylation [1,2]. Small molecules can be imported into NPCs by passive diffusion, such as ions and proteins of up to 20–40 kDa, whereas large proteins are imported by active translocation, a selective process using proteins belonging to the importin family [3]. This selective and active nuclear protein transport is mediated by a nuclear localization signal (NLS) [4].

NLSs are classified into canonical and non-canonical types. The first canonical type consists of a single cluster of basic amino acids such as SV40 large T antigen [5]. The second canonical type contains bipartite sequences that carry two clusters of positively charged amino acid residues separated by a spacer region, such as the NLS of *Xenopus laevis* nucleoplasmin [6]. Canonical types, import of which is mediated by importin-α and importin-β, can be predicted computationally [7–9]. Import of a portion of the non-canonical type is directly mediated by importin-β2, and is termed PY-NLS [10]. Other non-canonical types are still unknown.

Recent studies have demonstrated that the TET family of 5mC hydroxylases can convert 5mC to 5-hydroxymethylcytosine (5hmC) [11], and further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5CaC) [12,13]. These studies also suggest that additional modification of 5mC converted by TET enzymes may regulate the dynamics of 5mC and its mediated gene regulation [14]. Analysis of the TET3 amino acid sequence has revealed that TET3 contains several conserved domains, including a CXXC domain and a catalytic domain that is typical of Fe(II)- and 2-oxoglutarate (2OG)-dependent dioxygenase [15]. TET3 also acts as a transcription factor located in the nucleus [16]. Although TET3 is known to be targeted to the nucleus, the NLS responsible for TET3 has not yet been identified. In this report, we characterize a traditional NLS embedded in TET3.

## 2. Materials and methods

### 2.1. Plasmid construction

Fragments encoding the KKRK peptide and various mouse TET3 (GenBank: NM_183138) fragments were generated by polymerase chain reaction(PCR) from total cDNA from adult ovaries. Fragment primer pairs contained *Eco*RI and *Bgl*II restriction sites. The mouse glutathione S-transferase (GST, GenBank: X65021.1) and pig growth hormone (GH) were PCR-amplified from total cDNA from adult ovaries and plasmid pTRE-Tight-BI-GH-rtTA [17], respectively. Fragment primer pairs contained *Bsp*EI and *Bgl*II restriction sites. The fragment encoding KKRK or the TET3 fragment excised from the PCR product by *Eco*RI and *Bgl*II was ligated into the pEGFP vector (Clontech, USA) that had been previously digested with *Eco*RI and *Bgl*II. The fragment encoding GST or GH excised from the PCR product by digestion with *Bsp*EI and *Bgl*II was ligated into pEGFP, pEGFP-TET3 1310–1668 or pEGFP-KKRK, previously

---

*Abbreviations:* TET, ten-eleven translocation dioxygenase; TET1, ten-eleven translocation dioxygenase 1; TET3, ten-eleven translocation dioxygenase 3; PCR, polymerase chain reaction; GH, growth hormone; GST, glutathione S-transferase; NLS, nuclear localization signal; EGFP, enhanced green fluorescent protein.

* Corresponding author. Fax: +86 13813983156.
*E-mail addresses:* liuhonglin@263.net, 2011105020@njau.edu.cn (H.-l. Liu).

0006-291X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.08.075

digested with *Bsp*EI and *Bgl*II. Site-directed mutagenesis were carried out by CloneMutation (Vazyme, China). All constructs were confirmed by DNA sequencing. Primers used for PCR are shown in Table S1 in the supplementary material.

### 2.2. Cell culture and transfection

NIH 3T3 cells were cultured in growth medium containing GB11995 (Gibco, USA), 10% fetal bovine serum (FBS; Gibco, USA) and penicillin/streptomycin (Gibco, USA). Transfection was carried out using the Lipofectamine 2000 reagent (Invitrogen, USA) according to the manufacturer's procedure.

### 2.3. Microscopy

NIH 3T3 cells were seeded onto glass over slides, grown at 37 °C, and transiently transfected with each construct. Cells were fixed with 4% paraformaldehyde for 20 min at room temperature, washed twice with phosphate-buffered saline (PBS) and permeabilized with methanol for 10 min at −20 °C. Cells were then washed with PBS and stained with 4,6-diamidino-2-phenylindole (DAPI) for 15 min at room temperature. After washing with PBS, cells were visualized by confocal laser scanning microscopy.

## 3. Results

### 3.1. The C-terminus is required for TET3 nuclear localization

TET3 catalyzes 5mC into 5hmC in the nucleus. Numerous reports show that TET3 is localized in the nucleus [1,2,16,18]. However, TET3 (~180 kDa) cannot diffuse into the nucleus. One or more functional NLSs are thought to be embedded in TET3.

Structurally, TET3 contains an entire catalytic domain from amino acid 697 to 1668 [2,15]. Therefore, we fused a fragment from amino acid 664 to 1668 in the C-terminus of TET3 to the C-terminus of enhanced green fluorescent protein (EGFP) and examined the subcellular distribution of this fusion protein in NIH 3T3 cells. As shown in Fig. 1C, the C-terminal fragment (pEGFP-C1-TET3 664–1668) of TET3 was able to direct EGFP to the nucleus. We also fused a fragment from amino acid 1 to 339 and amino acid 333 to 672 in the N-terminus of TET3 to the C-terminus of EGFP and examined the subcellular distribution of these two fusion proteins in NIH 3T3 cells. As shown in Fig. 1C, the N-terminal fragments (pEGFP-C1-TET3 1–339 and pEGFP-C1-TET3 333–672) of TET3 did not direct EGFP to the nucleus. To determine the precise position of the NLS, we constructed the pEGFP-C1-TET3 995–1668 and pEGFP-C1-TET3 1310–1668. As shown in Fig. 1C, these fragments directed EGFP to the nucleus. To confirm that the NLS is embedded within the fragment between amino acid 1310 and 1668, we also constructed pEGFP-C1-TET3 664–1323. As shown in Fig. 1C, the fragment (pEGFP-C1-TET3 664–1323) did not direct EGFP to the nucleus. Therefore, the NLS located in TET3 is embedded within the fragment from amino acid 1310 to 1668. All fusion proteins were too large for nuclear pore complex (NPC) transport of small molecules through passive diffusion.

### 3.2. Prediction and mutation of the NLS in TET3

We used the web-based computer software, PSORT (available from the PSORT web site), to search for the NLS within TET3. There were five putative NLS regions containing multiple positively charged amino acid residues located at amino acids 439–445, 442–445, 443–446, 1615–1618 and 1617–1633, designated NLS1, NLS2, NLS3, NLS4 and NLS5, respectively (Fig. 2A). Because NLS1, NLS2 and NLS3 lie outside the region mapped by the truncation experiment, we focused on two NLSs within the C-terminus, NLS4 and NLS5.

To determine whether these two NLSs were functional, site-directed mutagenesis was performed to generate three mutant constructs. As shown in Fig. 2C, the mutant K1630 N/E1631 N/K1632 N/K1633 N did not impair the nuclear localization of pEGFP-C1-TET3 1310–1668. However, the mutant K1615 N/K1616 N/R1617 N/K1618 N disrupted the nuclear localization of pEGFP-C1-TET3 1310–1668. The double mutation of KEKK and KKRK showed the same distribution pattern as the KKRK mutation. Thus, KKRK is crucial in mediating nuclear localization of pEGFP-C1-TET3

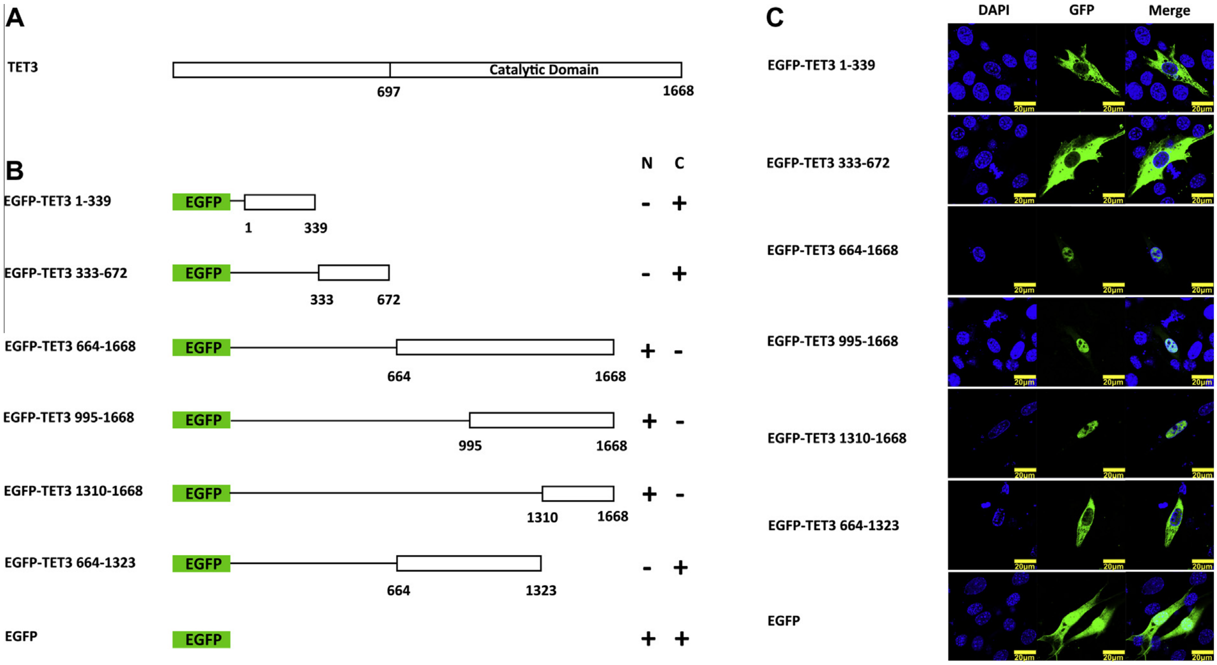


**Fig. 1.** The C-terminus of TET3 is sufficient for nuclear localization of EGFP. (A) Structure of TET3 protein. (B) Schematic diagram for the EGFP-TET3 fragment fusion proteins and their respective subcellular localizations. (C) EGFP and six deletion constructs were generated and transiently transfected into NIH 3T3 cells. Forty-eight hours after transfection, the subcellular localizations of EGFP and EGFP-tagged proteins were visualized by fluorescence microscopy.

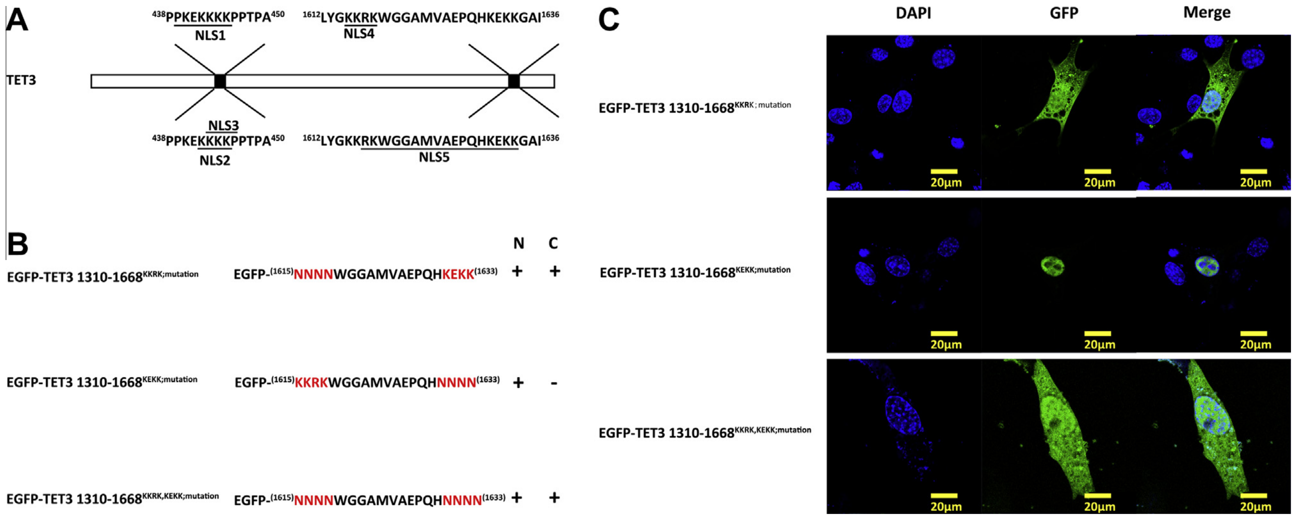


**Fig. 2.** Mapping of functional NLSs in mouse TET3 by site-directed mutagenesis. (A) Schematic representations of TET3 and five NLSs (underlined) as predicted by the PSORTII software. (B) Determination of the functional NLS by mutational analysis. (C) NIH 3T3 cells were transiently transfected with each of the three mutant constructs derived from pEGFP-TET3 1310–1668 by site-directed mutagenesis and examined by fluorescence microscopy.

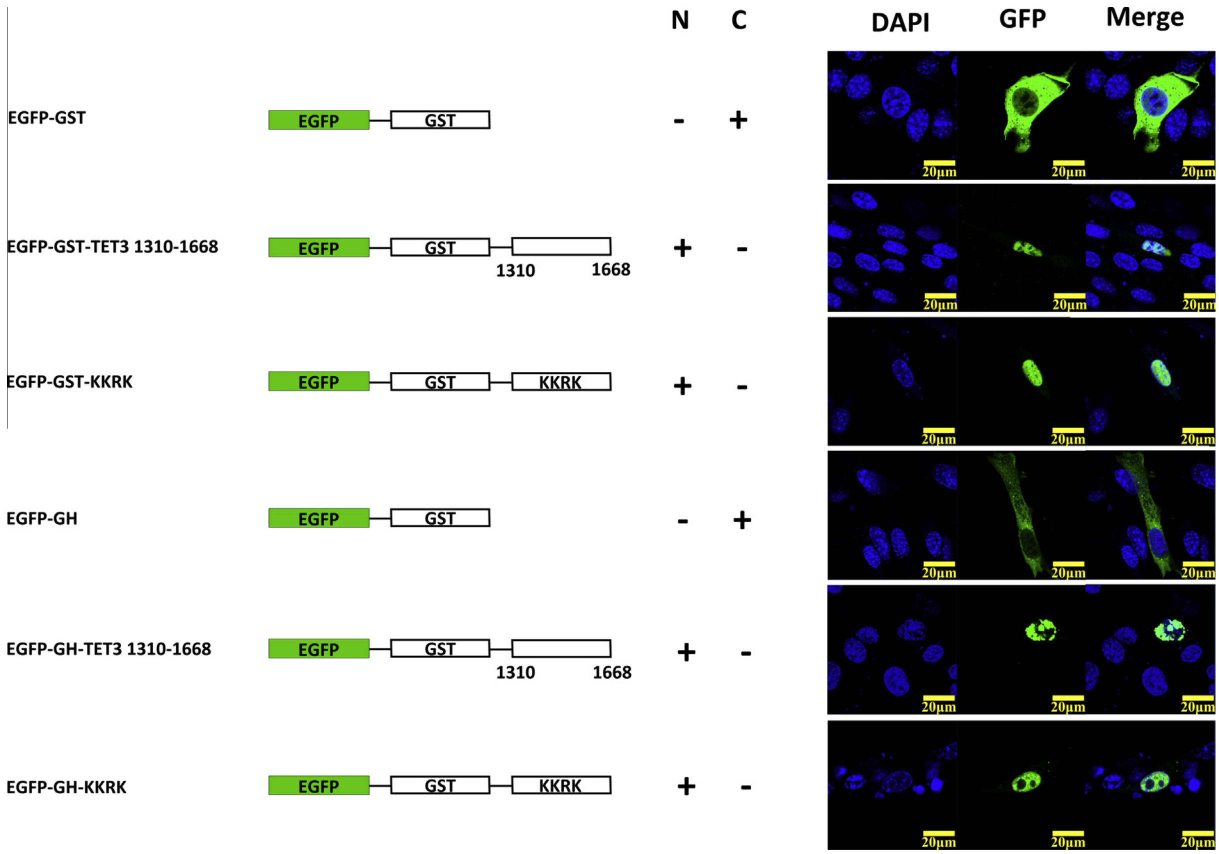
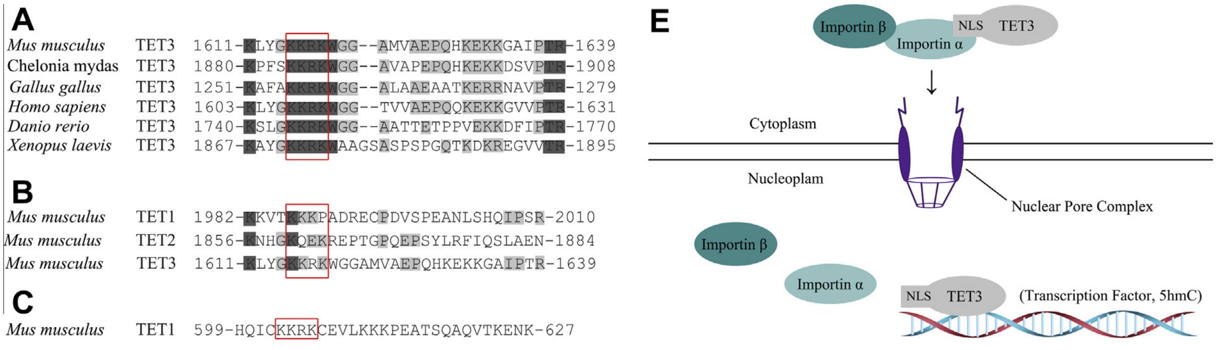


**Fig. 3.** The NLS in TET3 is sufficient for nuclear localization of EGFP-GST and the EGFP-GH fusion protein. (A) Schematic diagram of the fusion proteins and their respective subcellular localizations. (B) Six fusion proteins were generated and transiently transfected into NIH 3T3 cells. Forty-eight hours after transfection, the subcellular localizations of EGFP-tagged proteins were visualized using fluorescence microscopy.

1310–1668, and corresponds to the classic consensus sequence K(K/R)X(K/R) of a monopartite NLS.

### *3.3. KKRK can target cytoplasmic proteins to the nucleus*

To determine whether the peptide KKRK could efficiently target cytoplasmic protein to the nucleus, we constructed fusion proteins containing the KKRK tag. Since GFP is a small protein, a small GFP-KKRK polypeptide may be localized to the nucleus due to passive diffusion. To distinguish passive diffusion from active import, we used GFP-GST and GFP-GH fusion proteins as tags. The GFP-GST fusion protein was not localized in the nucleus because the fusion protein has a higher molecular mass, and neither EGFP nor GST contain an active NLS. In contrast, insertion of the C-terminal 359 residues of TET3 (pEGFP-C1-TET3 1310–1668) or KKRK conferred nuclear localization on the otherwise cytoplasmic GFP-GST,



D

| Protein | Mutation or Trucation | Receptor |
|---|---|---|
| FMIP | $^{7}$KKRK$^{10}$/KTGK | |
| $\delta_B$-CaM Kinase | $^{328}$KKRK$^{331}$/NKRK or $^{328}$KKRK$^{331}$/KNRK | |
| CUL4B | $^{37}$KKRK$^{40}$/KARK or $^{37}$KKRK$^{40}$/KKAK | importin α1, importin α3, importin α5, importin β1 |
| DNA helicase Q1 | $^{640}$GAKKRKIDDA$^{649}$ | importin α2 (Rch1), importin α3 (Qip1), importin α5 (NPI-1) |

**Fig. 4.** Alignment of TET3 NLSs with other proteins. (A) The NLS from residue 1615 to 1618 is highly conserved. Sequence alignment of mouse TET3 with orthologous proteins in six species was carried out using the ClustalW alignment software. Missing amino acids are denoted by hyphens. Identical and similar residues are highlighted in black and gray, respectively. The NLS of TET3 is boxed, and residue numbers corresponding to each species' sequence are also shown. (B) Sequence alignment of mouse TET1, TET2 and TET3 was carried out using the ClustalW alignment software. (C) NLS of mouse TET3 is found in mouse TET1 from residue 603 to 606. (D) NLS of mouse TET3 is also reported as the only NLS in other proteins using mutation or truncation methods. (E) Function and potential mechanism of TET3 import into the nucleus.

demonstrating that KKRK could target the protein to the nucleus (Fig. 3). We also tested KKRK using EGFP-GH as a tag. The EGFP-GH fusion protein was localized in the cytoplasm because neither GFP nor GH contain an active NLS. Fusion of EGFP-GH to amino acids 1310–1668 of TET3 or KKRK resulted in nuclear localization (Fig. 3).

### *3.4. KKRK is conserved in TET3 orthologs and is also responsible for other proteins targeting to the nucleus*

Alignment of the KKRK region with TET3 orthologous proteins from other species showed that $^{1615}$KKRK$^{1618}$ is highly conserved among the species examined (Fig. 4A). However, alignment of the KKRK region with the other two members of the TET family showed that KKRK is not conserved (Fig. 4B). Surprisingly, we found that KKRK is located in another region of the mouse TET1 protein (Fig. 4C).

We were curious whether KKRK was responsible for nuclear localization of other proteins reported in earlier studies. As shown in Fig. 4D, previous reports have suggested that KKRK is a functional NLS in other proteins using mutation or truncation experiments [19–22].

## 4. Discussion

Nuclear proteins, including transcription factors and nuclear enzymes, are synthesized in the cytoplasm and must be transported into the nucleus to be functional. Nuclear transport usually requires the presence of a NLS in the cargo and importins, which recognize and bind to the NLS. Generally, classic NLSs are composed of single stretches of amino acids (monopartite) or two short stretches separated by a spacer region (bipartite). Mouse TET3 possesses five potential classical NLSs, but only one mediates the nuclear localization of TET3. In this study, we determined that the NLS of mouse TET3 is located in the C-terminus using deletion, mutants and fusion proteins analysis. Importantly, this functional NLS, KKRK, is highly conserved in the TET3 protein of other species. These results demonstrate that KKRK may target TET3 in other species for nuclear localization.

The KKRK region is not conserved in mouse TET1 and TET2, which are other members of the TET family. This demonstrates that the KKRK region may not be responsible for nuclear localization of TET1 and TET2. Surprisingly, we found that KKRK is located from amino acid 603 to 606, which suggests that this region may be responsible for nuclear localization of TET1.

In this study, we added a small KKRK peptide into the C-terminus of the cytoplasmic fusion protein EGFP-GH and EGFP-GST. Fusion proteins containing the KKRK peptide were targeted to the nucleus. These results show that the small peptide KKRK is a useful NLS, targeting cytoplasmic proteins to the nucleus.

From previous reports, we also know that KKRK is responsible for nuclear localization of CUL4B, FMIP, $\delta_B$-CaM kinase, and DNA helicase Q1 using deletion or mutation methods [19–22]. Previous studies have determined that transport of these proteins is mediated by importin-α1, -α2, -α3, -α5, and -β1 (Fig. 4D); this indicates that TET3 may also be mediated by importin-α and importin-β.

TET3 act as a transcription factor and dioxygenase catalyzing 5mC into 5hmC in the nucleus [2,16]. TET3 is synthesized in the cytoplasm and weighs as much as 180 kDa, which disallows it through the NPC passively. Based on earlier reports and our findings of this NLS, we propose that KKRK of TET3 recognizes and binds to importin-α and -β, and targets TET3 to the nucleus. In the nucleus, TET3 depolymerizes from importin-α and -β, and fulfills its function as a transcription factor or dioxygenase (Fig. 4E).

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (31072028) and the key project of the Chinese National Programs for Fundamental Research and Development (973 Program 2014CB138502).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2013.08.075.

## References

[1] S. Ito, A.C. D'Alessio, O.V. Taranova, K. Hong, L.C. Sowers, Y. Zhang, Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification, Nature 466 (2010) 1129–1133.

[2] T.-P. Gu, F. Guo, H. Yang, H.-P. Wu, G.-F. Xu, W. Liu, Z.-G. Xie, L. Shi, X. He, S.-g Jin, The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes, Nature 477 (2011) 606–610.

[3] D. Gorlich, I.W. Mattaj, Nucleocytoplasmic transport, Science 271 (1996) 1513–1518.

[4] F. Melchior, L. Gerace, Mechanisms of nuclear protein import, Curr. Opin. Cell Biol. 7 (1995) 310–318.

[5] D. Kalderon, W.D. Richardson, A.F. Markham, A.E. Smith, Sequence requirements for nuclear location of simian virus 40 large-T antigen, Nature 311 (1984) 33–38.

[6] J. Robbins, S.M. Dilwortht, R.A. Laskey, C. Dingwall, Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence, Cell 64 (1991) 615–623.

[7] D.S. Goldfarb, A.H. Corbett, D.A. Mason, M.T. Harreman, S.A. Adam, Importin α: a multipurpose nuclear-transport receptor, Trends Cell Biol. 14 (2004) 505–514.

[8] A. Lange, R.E. Mills, C.J. Lange, M. Stewart, S.E. Devine, A.H. Corbett, Classical nuclear localization signals: definition, function, and interaction with importin α, J. Biol. Chem. 282 (2007) 5101–5105.

[9] Y. Chook, G. Blobel, Karyopherins and nuclear import, Curr. Opin. Struct. Biol. 11 (2001) 703–715.

[10] B.J. Lee, A.E. Cansizoglu, K.E. Süel, T.H. Louis, Z. Zhang, Y.M. Chook, Rules for nuclear localization sequence recognition by karyopherinβ2, Cell 126 (2006) 543–558.

[11] M. Tahiliani, K.P. Koh, Y. Shen, W.A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L.M. Iyer, D.R. Liu, L. Aravind, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, Science 324 (2009) 930–935.

[12] S. Ito, L. Shen, Q. Dai, S.C. Wu, L.B. Collins, J.A. Swenberg, C. He, Y. Zhang, Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine, Science 333 (2011) 1300–1303.

[13] Y.-F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, Science 333 (2011) 1303–1307.

[14] M.R. Branco, G. Ficz, W. Reik, Uncovering the role of 5-hydroxymethylcytosine in the epigenome, Nat. Rev. Genet. 13 (2011) 7–13.

[15] H. Wu, Y. Zhang, Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation, Genes Dev. 25 (2011) 2436–2452.

[16] Y. Xu, C. Xu, A. Kato, W. Tempel, J.G. Abreu, C. Bian, Y. Hu, D. Hu, B. Zhao, T. Cerovina, Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *xenopus* eye and neural development, Cell 151 (2012) 1200–1213.

[17] B. Jiang, D. Yu, L. Wang, F. Dong, H. Kaleri, X. Wang, N. Ally, J. Li, H. Liu, Doxycycline-regulated growth hormone gene expression system for swine, Genet Mol. Res. 11 (2012) 2946–2957.

[18] T. Nakamura, Y.-J. Liu, H. Nakashima, H. Umehara, K. Inoue, S. Matoba, M. Tachibana, A. Ogura, Y. Shinkai, T. Nakano, PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos, Nature 486 (2012) 415–419.

[19] Y. Zou, J. Mi, J. Cui, D. Lu, X. Zhang, C. Guo, G. Gao, Q. Liu, B. Chen, C. Shao, Characterization of nuclear localization signal in the N terminus of CUL4B and its essential role in cyclin E degradation and cell cycle progression, J. Biol. Chem. 284 (2009) 33320–33332.

[20] M. Srinivasan, C.F. Edman, H. Schulman, Alternative splicing introduces a nuclear localization signal that targets multifunctional CaM kinase to the nucleus, J. cell biol. 126 (1994) 839–852.

[21] A. Mancini, A. Koch, A.D. Whetton, T. Tamura, The M-CSF receptor substrate and interacting protein FMIP is governed in its subcellular localization by protein kinase C-mediated phosphorylation, and thereby potentiates M-CSF-mediated differentiation, Oncogene 23 (2004) 6581–6589.

[22] Y. Miyamoto, N. Imamoto, T. Sekimoto, T. Tachibana, T. Seki, S. Tada, T. Enomoto, Y. Yoneda, Differential modes of nuclear localization signal (NLS) recognition by three distinct classes of NLS receptors, J. Biol. Chem. 272 (1997) 26375–26381.