

领域前沿 · 中国



徐彦辉, 复旦大学生物医学研究院研究员、博士生导师; 复旦大学附属肿瘤医院兼职教授; 中国生物物理学会理事; 上海生物物理学会理事。1999年获清华大学生物科学与技术系学士学位, 2004年获清华大学生物科学与技术系博士学位, 2004~2007年在普林斯顿大学分子生物学系做博士后。2008年回国, 受聘于复旦大学生物医学研究院。先后任职副研究员、研究员、附属肿瘤医院兼职教授。长期从事表观遗传调控的结构生物学研究, 阐明其发挥功能的分子机理, 并为相关疾病的治疗提供理论基础。主持与承担国家“973”计划(课题组长)、国家自然科学基金重点项目、面上项目、“重大新药创制”专项、上海市科委项目等十几项课题。迄今为止, 已在国内外学术刊物上发表论文近40篇。自2008年组建课题组以来, 已在国内外学术刊物上以通讯作者发表多篇SCI论文, 包括Cell、Genes & Dev、Mol Cell、Proc Natl Acad Sci、Cell Res、J Biol Chem等。徐彦辉研究员先后入选中组部“青年拔尖人才”、教育部“新世纪人才”、上海市科委“浦江人才计划”、上海市教委“曙光学者”。获得了“明治生命科学奖(杰出奖)”、中国生物物理学会“贝时璋青年生物物理学家奖”、上海市生物物理学会“上海市生物物理学科青年科技英才”等奖项。

哺乳动物TET2蛋白氧化5-甲基胞嘧啶的 结构生物学研究

程净东 徐彦辉*

(复旦大学生物医学研究院, 复旦大学附属肿瘤医院, 上海 200032)

DNA甲基化是表观遗传学最重要的修饰之一, 哺乳动物的DNA甲基化修饰主要发生在胞嘧啶第5位碳原子上, 称为5-甲基胞嘧啶(5-methylcytosine, 5mC)。DNA甲基化参与了诸多的生物学过程, 包括基因印迹、X染色体的失活、基因组稳定性、转座子及逆转录转座子的沉默及组织特异性基因的沉默等^[1]。DNA甲基化模式的紊乱可以引起多种神经退行性疾病、免疫系统疾病、实体肿瘤及白血病。1948年, Hotchkiss^[2]在从牛胸腺组织提取的DNA中首次观察到了DNA的甲基化现象。在之后几十年的研究中, 科学家们发现DNA甲基化修饰不仅存在于高等哺乳动物, 还广泛分布于古菌、细菌和其他真核生物中。

与组蛋白的甲基化类似, DNA甲基化的建立、维持和去除存在动态变化。DNA甲基化模式可以通过复制来维持, 也可以通过有性生殖传递给下一代^[3]。DNA甲基化主要由DNA甲基化转移酶(DNA methyltransferase, DNMT)催化完成, 主要发生在CpG位点。甲基化模式的建立主要由DNMT3a、DNMT3b和DNMT3I蛋白负责, 在DNA复制过程中, DNA甲基化的维持由DNMT1蛋白负责。在哺乳动物中, DNA的甲基化去除主要发生在两次重编程过程中, 一次是在配子发生过程中, 一次是从受精到胚胎发育的早期^[4]。这两个过程中都发生了显著的DNA主动去甲基化。DNA上发生的主动去甲基化过程对于配子发生和胚胎的发育至关重要, 但催化该反应的酶一直没有被鉴定出来。

2009年,《Science》杂志的一篇文章揭开了DNA去甲基化研究的新篇章。研究者发现, TET1蛋

*通讯作者。Tel: 021-54237294, E-mail: xuyh@fudan.edu.cn

*Corresponding author. Tel: +86-21-54237294, E-mail: xuyh@fudan.edu.cn

网络出版时间: 2014-02-13 17:09

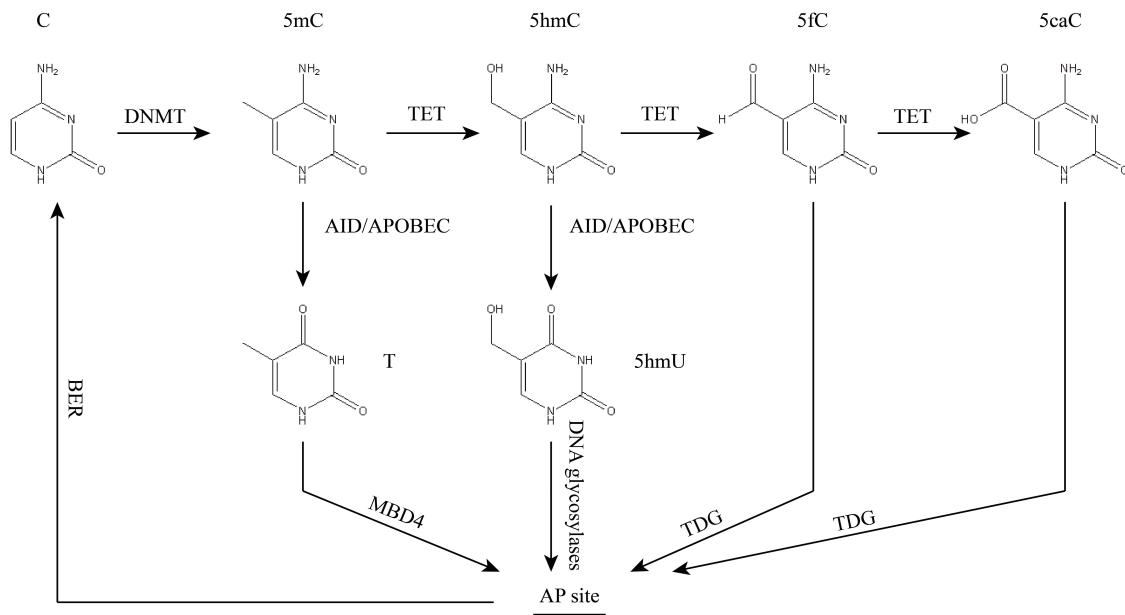
URL: <http://www.cnki.net/kcms/doi/10.11844/cjcb.2014.03.9001.html>

白可以将5-甲基胞嘧啶(5mC)氧化成5-羟甲基胞嘧啶(5-hydroxymethylcytosine, 5hmC)^[5]。后续的研究还发现, TET家族的蛋白还能进一步将5-羟甲基胞嘧啶(5hmC)氧化成5-甲酰胞嘧啶(5-formylcytosine, 5fC)和5-羧甲基胞嘧啶(5-carboxylcytosine, 5caC)^[6](图1)。5-羟甲基胞嘧啶(5hmC)是5-甲基胞嘧啶(5mC)的羟基化形式, 最早于1953年在噬菌体中发现^[7], 后来在脊椎动物的脑组织中也发现了这种修饰^[8]。5-羟甲基胞嘧啶(5hmC)是重要的表观遗传调控修饰, 与肿瘤的发生有密切关联。5-甲酰胞嘧啶(5fC)和5-羧基胞嘧啶(5caC)被人体内的一种糖苷水解酶TDG识别, 启动细胞内碱基错配切除修复, 将其替换成胞嘧啶, 最终实现DNA去甲基化^[9]。TET蛋白介导的DNA去甲基化成为近年来表观遗传研究领域的热点。

TET蛋白最早发现于2002年。研究人员发现, 在一种急性髓系白血病(acute myelocytic leukemia, AML)中, 第10号和11号染色体发生易位, 导致*MLL*基因与一个新基因发生融合, 该基因由此被命名为

ten eleven translocation(TET)。哺乳动物TET蛋白家族包括TET1、TET2和TET3。其结构都包括一个保守的C端的催化区和一个N端的调节区。C端的催化区属于Fe²⁺和 α 酮戊二酸(α -KG)依赖的双加氧酶家族, 是氧化5mC的催化结构域。与其他双加氧酶不同的是, TET蛋白催化结构域还包含一个TET蛋白特有的半胱氨酸富集结构域(Cysteine-rich), 该结构域对于保持酶活性是必需的^[10]。TET1与TET3的N端具有CXXC结构域, 有研究表明, 该结构域对于TET3准确的染色体定位是至关重要的^[11]。虽然TET2没有CXXC结构域, 但是后来的研究发现, TET2与CXXC4蛋白的相互作用可以辅助TET2蛋白的基因定位^[12]。

最近的研究表明, 血液系统肿瘤患者的基因中TET2蛋白的突变率非常高, 这些突变影响了TET2蛋白的活性, 对于白血病的发生起着非常重要的作用^[13]。因此, 研究TET蛋白的三维结构, 揭示其识别底物和催化反应的分子机制, 将对开发靶向TET蛋白的特异性药物有重要意义。我们课题组在国际上首次成功地解析了TET2-DNA复合物的晶体结构



DNA甲基化主要是由DNA甲基转移酶(DNMT)建立和维持。5-甲基胞嘧啶(5mC)可以被TET蛋白氧化生成5-羟甲基胞嘧啶(5hmC)、5-甲酰胞嘧啶(5fC)和5-羧甲基胞嘧啶(5caC)。5-甲基胞嘧啶(5mC)和5-羟甲基胞嘧啶(5hmC)可以被脱氨酶AID(activation-induced cytidine deaminase)识别脱氨生成胸腺嘧啶(T)和5-羟甲基尿嘧啶(5hmU)。5-甲酰胞嘧啶(5fC)、5-羧甲基胞嘧啶(5caC)、5-羟甲基尿嘧啶(5hmU)以及5-甲基胞嘧啶(5mC)脱氨形成的G/T错配可以被糖苷水解酶识别水解, 形成AP位点(apurinic/aprimidinic site), 从而启动碱基错配修复途径, 完成去甲基化。

DNA methylation (5mC) is established and maintained by DNA methyltransferases(DNMT). 5mC can be oxidated by the TET family of dioxygenases to generate 5hmC, 5fC and 5caC. Alternatively, 5mC and 5hmC may be further deaminated to become T or 5hmU by AID/APOBEC deaminases. 5hmU, 5fC, 5caC and G/T mismatch (generated by 5mC deamination) can be excised from DNA by DNA glycosylases such as TDG and MBD4. Through base excision repair pathway, 5mC can be replaced by cytosine (C).

图1 TET蛋白介导的DNA去甲基化反应示意图

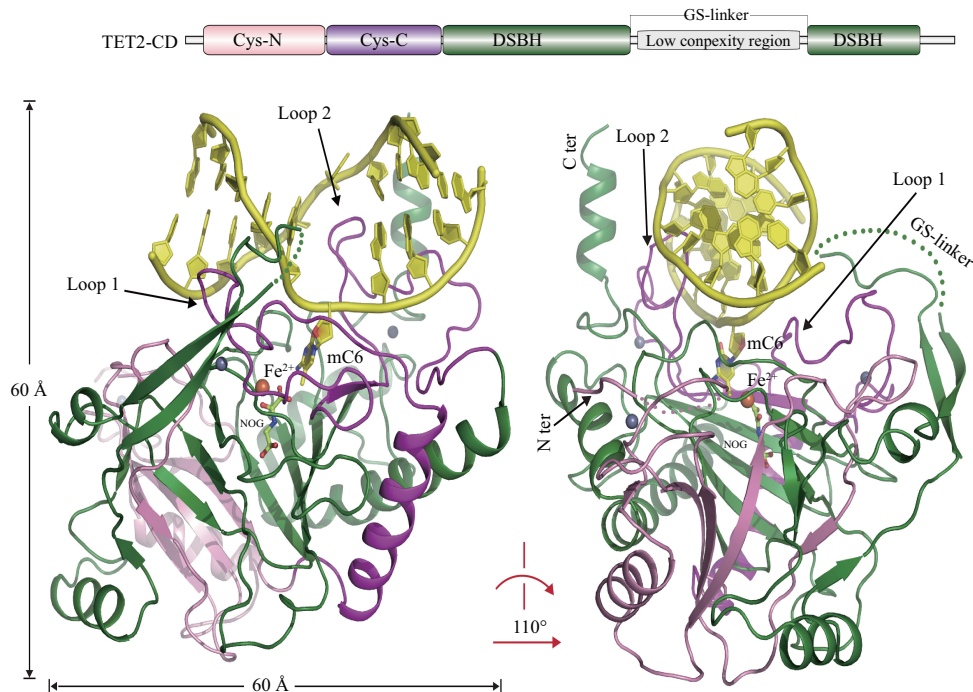
Fig.1 Proposed models of TET-initiated DNA demethylation pathways

(分辨率2.02 Å)^[10](图2)。

我们的研究表明, TET2蛋白具有保守的双加氧酶的催化中心, 形成double-stranded beta helix(DSBH)折叠, Fe²⁺和NOG(α -KG类似物)位于DSBH结构域形成的催化中心。Cysteine-rich结构域被分割成了Cystine-rich氨基端(Cys-N)和羧基端(Cys-C)两个部分。Cysteine-rich结构域将DSBH结构域包裹在中心。二者通过多个结合面的作用, 在三个锌指结构域的螯合作用下, 构成了一个完整的催化结构域。Cysteine-rich结构域既维持催化结构域的稳定性, 又负责与DNA底物的结合。含有甲基化CpG的DNA底物位于催化中心的正上方, 被DSBH和Cysteine-rich结构域牢牢地锚定。Cysteine-rich结构域中的loop1和loop2两个片段参与了DNA的识别和结合。其中, loop2的疏水氨基酸还参与了5mCpG附近DNA的解旋。在DSBH和Cysteine-rich结构域的共同作用下, 5-甲基胞嘧啶(5mC)被翘出DNA双螺旋, 通过碱基翻转机制插入TET蛋白催化中心, 将甲基指向催化反应中心, 在氧气和 α -KG的参与下完成氧化反应。

我们的研究发现, TET2通过TET家族中高度保守的氨基酸, 特异性地识别CpG二核苷酸。由于哺乳动物中5-甲基胞嘧啶主要存在于CpG二核苷酸中, 结构的发现很好地解释了TET对DNA底物的选择性。TET对CpG二核苷酸周围的DNA序列没有明显的偏好性, 这也提示我们TET2蛋白在染色体的定位可能会受到其相互作用蛋白的辅助。插入到催化中心5-甲基胞嘧啶被几个特定的氨基酸稳定到一个固定的构象, 使其甲基朝向具有催化活性的Fe²⁺和 α -KG。有趣的发现是甲基的周围并没有疏水相互作用的存在, 这一特点也解释了该催化中心能够将亲水性的5hmC和5fC继续催化。

通过进一步的结构和生化实验分析, 我们发现白血病病人体内TET2的突变大部分都发生在对TET结构或功能至关重要的氨基酸上。病人来源的氨基酸突变参与DNA的结合、5mC的翻转以及催化中心的组成。这些突变有的降低TET2蛋白稳定性, 有的影响底物DNA的识别, 有的削弱TET2蛋白酶活力。这一研究也很好地将TET2的功能与白血病的发生机制建立了联系。



使用丝带模式表示的TET2催化结构域和底物DNA三维结构(两种不同的视角)。图中底物DNA被标记成黄色, Cysteine-rich结构域被标记成粉色和紫色, DSBH核心标记成绿色。甲基化胞嘧啶(mC6)、铁离子和N-oxalylglycine(NOg)在图中标记出。识别DNA的loop1和loop2以及催化结构域的N端和C端也在图中标出。

Ribbon representation of TET2-DNA structure in two different views. The DNA is colored in yellow, cysteine-rich domain is colored in pink and violet, DSBH core is colored in green. mC6, Fe ion and NOG are marked. DNA-interacting loops (loop1 and loop2), and N and C termini are indicated.

图2 TET2-DNA复合物三维结构示意图(根据参考文献[10]修改)

Fig.2 Overall structure of TET2-DNA complex (modified from reference [10])

上述研究揭示了TET蛋白催化结构域的空间构成,阐明了TET2发生催化反应,特异性识别5-甲基胞嘧啶的分子机制,为TET蛋白家族的结构与功能研究打开了一扇大门。高分辨率TET2-DNA复合物的晶体结构也为进一步设计靶向TET蛋白的药物奠定了重要的结构基础。

TET蛋白酶活性的调节机制将是后续研究中的一个重要问题。首先,TET蛋白的N端调节区和DBSH结构域里的无序区域的功能都还未知。但是这两段区域在TET蛋白家族内都存在,说明他们对于TET蛋白的功能非常重要。这两段区域的不保守性,恰好可以解释三种TET蛋白功能的非冗余性。其次,结合蛋白对TET的功能也很重要。已经有研究证明,TET相互作用蛋白可以调节其染色体定位和酶活。新的相互作用蛋白的鉴定和功能研究还有待于我们去完成。再次,体内代谢小分子对TET的酶活也有调节作用,从而影响体内表观遗传学的改变,引发疾病。总之,对TET蛋白的研究才刚刚开始,回答上述问题将有助于我们对TET功能深入理解,并为治疗因TET失活而引起的疾病奠定理论基础。

另外,是否存在DNA甲基化脱羧酶仍是个谜。迄今为止,我们知道的DNA主动去甲基化过程主要有以下几个途径:一是通过脱氨途径。5-甲基胞嘧啶(5mC)和5-羟甲基胞嘧啶(5hmC)分别在脱氨酶(activation-induced cytidine deaminase, AID)的作用下被转化成胸腺嘧啶(T)和5-羟甲基尿嘧啶,然后通过碱基错配修复途径完成去甲基化。二是通过TET蛋白的氧化作用。最终生成的5-甲酰胞嘧啶(5fC)和5-羧基胞嘧啶(5caC),可以被体内的糖苷水解酶识别,然后通过碱基切除修复途径完成去甲基化^[14]。这一类途径与植物中的DNA去甲基化非常相似^[15],但是体内很有可能存在一类脱羧酶,它们可以识别5-羧基胞嘧啶(5caC),催化脱羧最终生成胞嘧啶。这需要科学家们继续去发现。

参考文献 (References)

- 1 Robertson KD. DNA methylation and human disease. *Nat Rev Genet* 2005; 6(8): 597-610.
- 2 Hotchkiss RD. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem* 1948; 175(1): 315-32.
- 3 Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 2005; 74: 481-514.
- 4 Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science* 2001; 293(5532): 1089-93.
- 5 Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; 324(5929): 930-5.
- 6 Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 2011; 333(6047): 1300-3.
- 7 Wyatt GR, Cohen SS. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochem J* 1953; 55(5): 774-82.
- 8 Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J* 1972; 126(4): 781-790.
- 9 He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 2011; 333(6047): 1303-7.
- 10 Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, *et al.* Crystal Structure of TET2-DNA complex: Insight into TET-mediated 5mC oxidation. *Cell* 2013; 155(7): 1545-55.
- 11 Xu Y, Xu C, Kato A, Tempel W, Abreu JG, Bian C, *et al.* Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for *Xenopus* eye and neural development. *Cell* 2012; 151(6): 1200-13.
- 12 Ko M, An J, Bandukwala HS, Chavez L, Aijö T, Pastor WA, *et al.* Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* 2013; 497(7447): 122-6.
- 13 Tan L, Shi YG. Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development* 2012; 139(11): 1895-902.
- 14 Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev* 2011; 25(23): 2436-52.
- 15 Zhu JK. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genetics* 2009; 43: 143-66.