

DOI: 10.13376/j.cblls/2017125

文章编号: 1004-0374(2017)10-0934-05



徐彦辉, 复旦大学附属肿瘤医院研究员, 复旦大学生物医学研究院兼职教授。国家杰出青年科学基金获得者, 教育部“长江学者”特聘教授。主要从事表观遗传调控蛋白的结构与功能研究, 系统地阐明了DNA甲基化和组蛋白甲基化关键酶的催化、特异性底物识别和酶活性调节的分子机制, 并为靶向药物设计奠定了基础。以通讯作者在 *Cell*、*Nature*、*Mol Cell*、*Cell Res* 等杂志发表多篇论文。

哺乳动物DNA甲基化修饰的动态调控机制

徐彦辉

(复旦大学生物医学研究院, 上海 200032)

摘要: DNA甲基化是生命体最主要的表观遗传修饰之一。哺乳动物DNA甲基化主要发生在胞嘧啶第五位碳原子上, 称为5-甲基胞嘧啶(5-methylcytosine, 5mC)。哺乳动物DNA甲基化由从头DNA甲基转移酶DNMT3A/3B在胚胎发育早期建立, 甲基化模式的维持由DNA甲基转移酶DNMT1实现。TET家族蛋白氧化5-甲基胞嘧啶起始DNA的去甲基化过程。这些DNA甲基化修饰酶精确调节DNA甲基化的动态过程, 在整个生命发育过程中发挥重要作用, 其失调也与多种疾病发生密切相关。现结合国内外同行研究进展, 介绍课题组近年来对DNA甲基化修饰酶的结构与功能研究。

关键词: DNA甲基化; DNA甲基转移酶; TET

中图分类号: Q523 **文献标志码:** A

Dynamic regulation of mammalian DNA methylation

XU Yan-Hui

(Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China)

Abstract: DNA methylation is one of the key epigenetic modifications and mainly occurs at the carbon 5 position of cytosine (5-methylcytosine, 5mC). Mammalian DNA methylation is established by *de novo* DNA methyltransferases DNMT3A/3B in early embryonic development. The patterns of methylation are propagated to daughter cells by maintenance DNA methyltransferase DNMT1 during replication. TET proteins oxidize 5mC into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), and initiate DNA demethylation. These proteins play essential roles in various biological processes during development and their dysregulation is involved in many pathological processes. Here, we summarized structural and functional studies of these enzymes.

Key words: DNA methylation; DNA methyltransferase; TET

收稿日期: 2017-09-04

基金项目: 国家自然科学基金项目(914190301)

通信作者: E-mail: xuyh@fudan.edu.cn

表观遗传学 (epigenetics) 研究的是 DNA 序列不发生改变的可遗传的基因组功能改变, 通过对细胞内外信号的整合, 与遗传信息共同作用于生命体, 从而决定细胞和个体的表型^[1]。DNA 甲基化是发生在 DNA 上的共价修饰, 是最重要的表观遗传修饰之一。DNA 的甲基化修饰广泛存在于古细菌、细菌和真核生物等物种中, 而在如酵母等生物中缺失。在哺乳动物中, DNA 甲基化主要发生在胞嘧啶-鸟嘌呤二核苷酸 (CpG) 位点上。此外, 少量非 CpG 位点的甲基化修饰也存在于胚胎干细胞 (embryonic stem cells, ESCs) 中。据统计, 在人类基因组中, 有 70%~80% 的 CpG 位点发生了甲基化修饰。未甲基化的 CpG 二核苷酸主要富集于基因启动子区域, 通常成簇存在, 被称为 CpG 岛 (CpG islands, CGIs)。

哺乳动物 DNA 甲基化修饰是可逆并且动态变化的。DNA 甲基化修饰的建立与维持主要由两类 DNA 甲基转移酶 (DNA methyltransferases, DNMTs) DNMT1 和 DNMT3A/3B 完成^[2-4]。DNA 甲基化的去除机制是近年来的研究热点, 目前认为是 TET (ten-eleven translocation, TET) 家族蛋白氧化 5-甲基胞嘧啶 (5-methylcytosine, 5mC) 起始这一关键过程。

在哺乳动物胚胎发育过程中, 有两次主要的 DNA 甲基化去除后重建的过程。第一次是在受精之后, 雄原核甲基化水平快速降低, 雌原核甲基化水平随 DNA 复制发生被动去甲基化。从囊胚期开始, 基因组的甲基化修饰被广泛建立起来, 主要由 DNMT3 家族的酶 DNMT3A 和 DNMT3B 催化^[5-6]。第二次主要过程是在原生殖细胞 (primordial germ cells, PGCs) 发育过程中发生的 DNA 去甲基化, 而后再次通过 DNMT3A 和 DNMT3L (DNMT3-like protein) 在印记基因和转座元件上重新建立甲基化^[5-6]。上述两个过程的 DNA 甲基化动态变化对个体发育和生殖都是至关重要的。

DNA 的甲基化修饰是可逆的。哺乳动物有两种主要 DNA 去甲基化机制: 主动去甲基化和 DNA 复制依赖的被动去甲基化。在受精卵雌原核基因组的去甲基化过程中, DNA 双链数目剧增, 原有甲基化修饰被逐渐稀释, 呈现为被动的去甲基化的过程。TET 家族蛋白在 DNA 主动去甲基化过程中起关键作用。

在本文中, 我们将依次叙述哺乳动物 DNA 甲基化模式的建立者 DNMT3 家族的结构与功能, 以及 TET 蛋白的结构及其底物偏好性的机制研究。

1 DNMT3A的活性调控机制研究

哺乳动物 DNMT3 家族的成员, 包括 DNMT3A、DNMT3B 和 DNMT3L。其中, DNMT3A/3B 具有高度同源性, 在体内和体外均能催化 DNA 的甲基化^[7]; 而 DNMT3L 不仅同源性相对较低, 而且由于催化区域内重要结构域的缺失, 不具备转移甲基基团的活性, 但它能够直接结合 DNMT3A/3B 并提高两者的酶活, 并且对配子早期发育过程中的甲基化模式的建立有着重要作用^[8]。

DNMT3A 与 DNMT3B 具有相似的结构域, 包括 N 端无规则序列、PWWP 结构域 (Pro-Trp-Trp-Pro)、ADD 结构域 (ATRX-DNMT3-DNMT3L) 和 C 端保守的催化结构域 (catalytic domain, CD)。其中, PWWP 结构域能识别 36 位赖氨酸三甲甲基化的组蛋白 H3 (H3K36me3) 的修饰并指导 DNMT3 蛋白的亚细胞定位^[9]。ADD 结构域由一个 C2-C2 锌指结构域和一个植物同源结构域 (plant homeodomain, PHD) 组成, 能识别并结合组蛋白 H3 的 N 端, 同时 4 位赖氨酸 (K4) 甲基化可破坏该结合。因此, ADD 被认为特异性识别 K4 未甲基化的组蛋白 H3 (H3K4me0)^[10-11]。CD 结构域与 DNA 结合, 以 S-腺苷甲硫氨酸 (SAM) 为甲基供体, 催化生成 5 位甲基化的胞嘧啶。

前期研究表明, DNMT3A 和 DNMT3L 的催化结构域相互结合形成四聚体^[12], 其中两个 DNMT3A 的 CD 结构域形成催化中心相邻的二聚体; 两个 DNMT3L 的 CD-like 结构域位于 DNMT3A 的两侧, 可能对 DNMT3A 的催化中心区域起到稳定作用。DNMT3L 的 ADD 结构域能够与组蛋白 H3 结合, 这一相互作用会被 H3K4 位点的甲基化修饰所破坏^[10,13]。此外, 组蛋白 H3 也通过结合 Dnmt3a 的 ADD 结构域来调控 Dnmt3a 的活性: 体外合成的 H3K4me0 的组蛋白多肽可以激活 Dnmt3a 的活性, 而 H3K4me3 则削弱它的激活效果, 当突变掉 ADD 结构域中参与识别 H3K4 位点的关键氨基酸后, 这种对 H3K4 甲基化修饰状态的敏感性消失了^[14-15]。

我们受到上述研究的启发, 利用生化和结构生物学方法研究组蛋白修饰与 DNA 甲基化修饰的相互调控^[16]。我们首先发现, DNMT3A 可以被组蛋白 N 端的多肽 (H3K4me0) 所激活, 但 H3K4me3 无法激活该蛋白。同时, 在核小体水平上, 未修饰的核小体也可以激活 DNMT3A, 但含有 H3K4me3 的核小体无法激活。更为有趣的是, CD 结构域无法被 H3K4me0 多肽激活, 但其本身的活性与激活后

DNMT3A 的活性基本相当, 在 ADD 结构域存在的情况下, CD 的活性受到抑制, 表明 DNMT3A 蛋白存在活性的自抑制, 加入 H3K4me0 多肽, 可以解除自抑制并激活 DNMT3A。

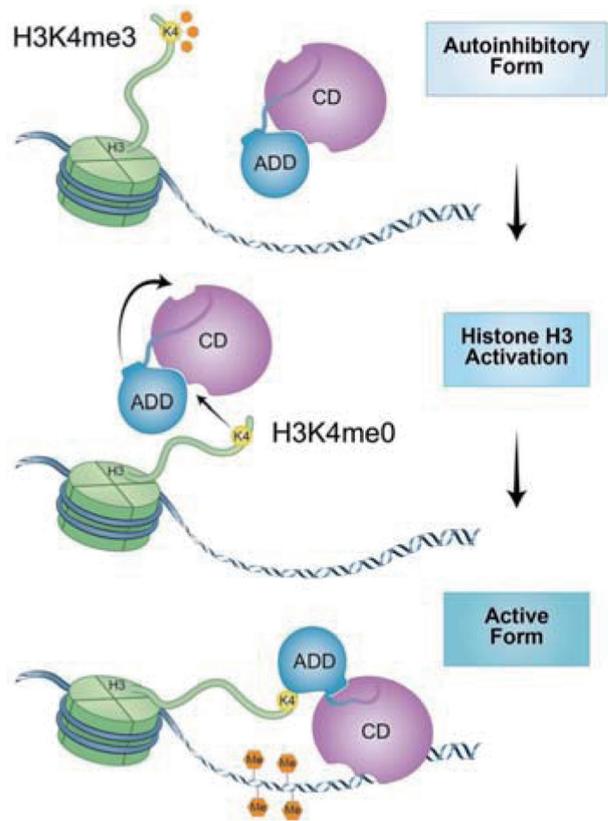
我们进一步解析了激活状态的 DNMT3A 结构, 发现组蛋白 H4me0 多肽结合在 ADD 结构域, 其结合方式与已知的 ADD-H3 多肽结构中发现的一致, 最关键的结合是通过第四位赖氨酸与 ADD 上的氨基酸 D529 和 D531 形成氢键。将 K4 甲基化或者突变 D529/D531 都会破坏两者的结合。在此激活状态下的结构分析发现, DNMT3A 的构象与抑制态有很大不同, ADD 结合在 CD 的另外一个表面, 提示组蛋白 H3K4me0 的结合可能会诱导 DNMT3A 发生构象变化, 使 ADD 结构域从抑制状态的结合位点解离, 变为激活状态的结合方式。在激活状态下, ADD 不阻挡 CD 与 DNA 的结合, 从而释放了 DNMT3A 的自抑制。这一构象变化进一步被电镜与磁共振实验所验证 (图 1)。

我们的这一研究工作很好地说明了组蛋白 H3 通过调节 DNMT3A 的构象激活其酶活性, 并指出核小体的组蛋白招募和激活 DNMT3A, DNMT3A 具有较高酶活性的区域需要组蛋白 H3 的 N 端, 因此, 组蛋白就对 DNMT3A 的工作范围做了限制。这样的多重调控机制有助于使 DNMT3A 在特定的区域发挥功能, 避免了不必要的基因组甲基化发生。同时, 也从结构生物学的角度支持了基因组分析中发现的 DNA 甲基化与 H3K4me3 负相关的现象。

2 TET蛋白的结构研究及其底物偏好性机制

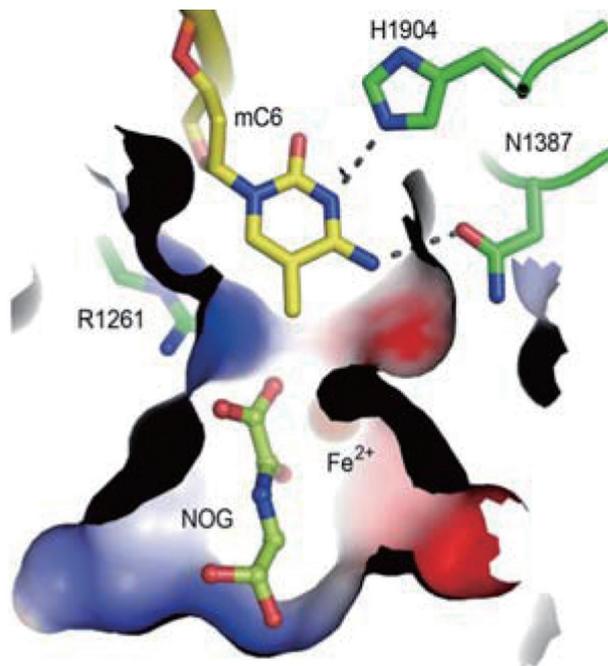
TET 蛋白可将 5mC 连续氧化生成 5-羟甲基胞嘧啶 (5-hydroxymethylcytosine, 5hmC)、5-醛基胞嘧啶 (5-formylcytosine, 5fC) 和 5-羧基胞嘧啶 (5-carboxycytosine, 5caC)^[17-19]。哺乳动物 TET 家族蛋白共有 TET1/TET2/TET3 三个成员, 属于 Fe(II) 和 α -酮戊二酸 (α -ketoglutarate, α -KG) 依赖的双加氧酶超家族^[19-20]。它们都包含有 1 个 C 端催化结构域, 该结构域又包含有 1 个特征性的半胱氨酸富集结构域 (Cys-rich) 和 1 个双链 β 螺旋结构域 (double-stranded β helix, DBSH)。

我们对 TET2 和 5mC-DNA 复合物的晶体学研究阐述了 TET 蛋白识别底物 DNA 并催化反应的机制^[21] (图 2)。我们获得了 TET2-5mC-DNA 复合物高分辨率 (2.0 Å) 晶体结构, 发现 Cys-Rich 结构域缠绕着 DSBH 催化核心, 并被 DSBH 天然地分成



DNMT3A 蛋白本身处于自抑制状态。H3K4me3 存在的情况下, DNMT3A 不结合核小体。当存在 H3K4me0 的情况下, H3 结合并且诱导 DNMT3A 发生构象改变, 成为激活态, 使其可以结合并催化 DNA 的甲基化。

图1 DNMT3A自抑制及其被组蛋白H3激活模式图^[16]



催化中心中, 5mC 被 TET 特异性识别。

图2 TET2-5mC-DNA复合物结构^[21]

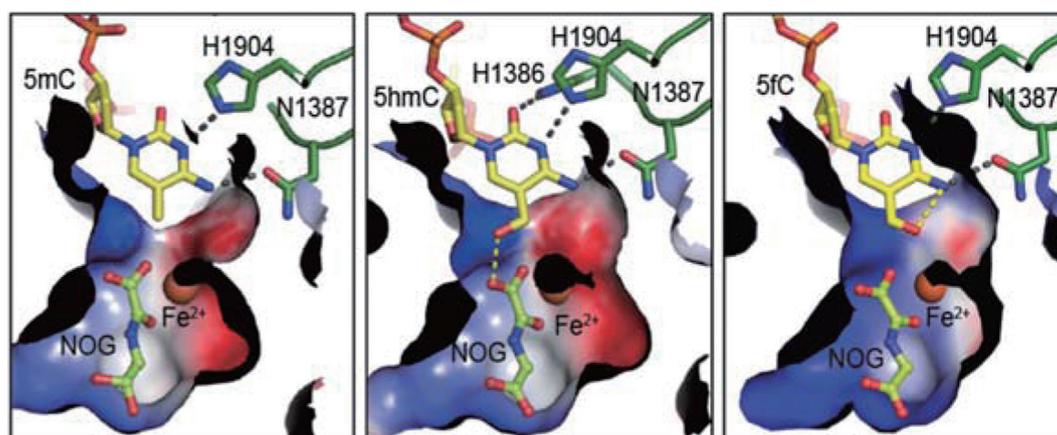
了 Cys-N 和 Cys-C 两部分, 3 个锌指结构使 Cys-Rich 和 DSBH 结构域的柔性区域相互连接, 进而形成一个紧凑的催化结构域。底物 DNA 位于催化结构域核心的上方, Cys-Rich 结构域的一段 loop 伸入 DNA 双螺旋中, 通过碱基翻转机制将其“撬开”, 使 5mC 碱基翻出 DNA 双螺旋并插入到 DSBH 催化核心。催化核心通过氢键和碱基堆积力特异性识别 5mC, 5mC 上的甲基朝向 Fe(II) 和 NOG (α -KG 类似物), 分子氧被 Fe(II) 活化后进攻甲基的 C-H 键, 最后将 5mC 氧化成 5hmC。在白血病中常见的 TET2 突变多数都发生在 DNA 识别、锌以及铁螯合作用区域。

我们进一步的生化研究发现, TET 蛋白对 5mC-DNA 的活性远远高于 5hmC/5fC-DNA, 也就是说 TET 氧化 5mC 的 3 个反应步骤中, 第一步的反应迅速, 而第二、第三步效率要低很多^[22](图 3)。细胞内 3 种反应产物的浓度也支持我们的观测结果。大量的测序研究发现, 细胞中 5hmC 的含量远远高于 (10~100 倍) 5fC 和 5caC 的含量, 由此, 研究人员认为 5hmC 很可能是一类新型的表观遗传修饰。我们由此推测, 很可能是 TET 本身的酶活偏好性导致了这一现象。我们通过大量生化分析发现, 人源的 TET1 和 TET2 都具有对 5mC 的底物偏好性, 该性质与酶和底物 DNA 结合能力等因素无关。我们前期已经获得了 TET2 与 5mC-DNA 的高分辨 (2.0 Å) 晶体结构, 我们进一步解析了 TET2 与 5hmC-DNA (1.80 Å) 及 TET2 与 5fC-DNA (1.97 Å) 的晶体结构。3 个高分辨晶体结构, 使我们可以很准确地建立了 5mC/5hmC/5fC 在催化中心的结构模型。结

构分析发现, 3 个结构中, TET2 蛋白和 DNA 的整体构象基本没有变化。在催化中心, 3 种核苷酸的胞嘧啶取向也比较一致, 关键的不同点在于甲基的取代基, 即甲基、羟甲基、醛基的取向和化学环境不同。在催化中心, 5mC 的甲基与周围的氨基酸没有特定的相互作用, 这与其疏水性是一致的。5hmC 中的羟基与 NOG (α -KG 类似物) 形成氢键, 而 5fC 中的醛基与胞嘧啶中的氮原子形成分子内氢键。

TET 属于 Fe(II) 和 α -KG 依赖的双加氧酶, 我们进一步通过生化研究, 证明了消氢作用是其催化过程的决速步骤, TET 对 3 种不同底物的催化反应中, 消氢作用的效率从高到低依次为 5mC>5hmC>5fC, 与酶活反应效率一致。通过结构分析与分子动力学模拟催化过程, 我们提出 TET 底物偏好性的可能机制是: 在催化过程中, 5mC 的甲基与胞嘧啶相连的 C-C 键可以自由旋转, 由此被消除的氢原子可以处于优势取向, 使消氢过程很容易发生。而在 5hmC 和 5fC 中, 由于羟基和醛基与周围环境的作用, 限制了碳相连的氢原子, 使其远离高价态的铁氧中间体, 导致消氢作用效率低, 酶活性减弱。其他 Fe(II) 和 α -KG 依赖的双加氧酶催化不同底物时, 也可能有类似的机制。

这一研究很好地证明了细胞内 5hmC 的水平远远高于 5fC 和 5caC, 很可能是由于 TET 本身的底物偏好性导致的, 说明 5hmC、5fC 相对稳定, 一旦产生就不容易发生进一步的氧化反应。如果要跨越这个障碍, 就需要 TET 在特定的基因组区域达到较高的浓度或者被调控蛋白或小分子所激活。



三个 TET-DNA 复合物的结构比对。结合生化实验及分子动力学分析, 发现 TET 对 5mC 催化活力高是因为其消氢作用效率高于另外两种反应。

图3 TET底物偏好性机制

3 展望

DNA 甲基化的精确动态调控对细胞的命运决定和个体发育至关重要, 其异常与多种肿瘤的发生发展密切相关, 我们对这一复杂过程的理解还有很多工作要做, 如 DNMT3A/3B 如何识别底物 DNA, 底物识别偏好性是怎样形成的。5mC 氧化产物 5hmC/5fC/5caC 是否具有表观遗传特性, 还是仅仅作为中间产物。尤其是 5hmC 在神经系统中的含量非常高, 而在某些肿瘤中含量又很低, 是否具有特定的功能。5hmC 作为潜在的表观标记是否具有特异性的识别或者排斥蛋白。这些问题的回答将使我们更为深入地理解 DNA 甲基化与生命发育过程的关系。

[参 考 文 献]

- [1] Allis CD, Jenuwein T, Reinberg D, et al. Epigenetics [M]. New York: Cold Spring Harbor Laboratory Press, 2007
- [2] Kim JK, Samaranyake M, Pradhan S. Epigenetic mechanisms in mammals. *Cell Mol Life Sci*, 2009, 66: 596-612
- [3] Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*, 2005, 74: 481-514
- [4] Cheng X, Blumenthal RM. Mammalian DNA methyltransferases: a structural perspective. *Structure*, 2008, 16: 341-50
- [5] Kafri T, Ariel M, Brandeis M, et al. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev*, 1992, 6: 705-14
- [6] Monk M, Boubelik M, Lehnert S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*, 1987, 99: 371-82
- [7] Okano M, Bell D, Haber DA, et al. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell*, 1999, 99: 247-57
- [8] Bourc'his D, Xu GL, Lin CS, et al. Dnmt3L and the establishment of maternal genomic imprints. *Science*, 2001, 294: 2536-9
- [9] Dhayalan A, Rajavelu A, Rathert P, et al. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem*, 2010, 285: 26114-20
- [10] Ooi SK, Qiu C, Bernstein E, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature*, 2007, 448: 714-7
- [11] Otani J, Nankumo T, Arita K, et al. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep*, 2009, 10: 1235-41
- [12] Jia D, Jurkowska RZ, Zhang X, et al. Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. *Nature*, 2007, 449: 248-51
- [13] Zhou BO, Zhang RR, Zhou JQ, et al. The N-terminus of histone H3 is required for *de novo* DNA methylation in chromatin. *Proc Natl Acad Sci USA*, 2009, 106: 22187-94
- [14] Li BZ, Huang Z, Cui QY, et al. Histone tails regulate DNA methylation by allosterically activating *de novo* methyltransferase. *Cell Res*, 2011, 21: 1172-81
- [15] Zhang Y, Lurkowska R, Soeroes S, et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res*, 2010, 38: 4246-53
- [16] Guo X, Wang L, Li J, et al. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature*, 2015, 517: 640-4
- [17] Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, 2011, 333: 1300-3
- [18] He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 2011, 333: 1303-7
- [19] Tahilian M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 2009, 324: 930-5
- [20] Iyer LM, Tahiliani M, Rao A, et al. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle*, 2009, 8: 1698-71
- [21] Hu L, Li Z, Cheng J, et al. Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell*, 2013, 155: 1545-55
- [22] Hu L, Lu Z, Cheng J, et al. Structural insight into substrate preference for TET-mediated oxidation. *Nature*, 2015, 527: 118-22