领域前沿・中国



罗成博士,中国科学院上海药物研究所研究员,博士生导师。1994年本科毕 业于福州大学化学系, 2001年毕业于复旦大学化学系, 获得物理化学硕士学 位(主修金属表面催化与量子化学)。2004年毕业于中国科学院上海药物研 究所, 获有机化学博士学位(主修药物设计与药物化学); 2005年起赴美国宾 夕法尼亚大学Wistar研究所著名表观遗传结构生物学家Ronen Marmorstein 教授实验室从事表观遗传修饰酶的化学生物学研究。2008年3月至今,工 作于中国科学院上海药物研究所。罗成研究员基于理论模拟策略,结合药 物化学、结构生物学和化学生物学等实验验证手段,针对重大疾病或通路 (特别是表观遗传修饰酶和蛋白质-蛋白质相互作用)开展创新药物靶标及 其先导化合物的发现、确证与重要功能蛋白质动态调控机理研究,揭示了 多个重要靶标的调控机制;发现一批具有开发前景的先导化合物(五类以上 为该酶世界首次发现:包括铜伴侣蛋白抑制剂DC-AC50、RNA去甲基化酶 FTO抑制剂、靶向核抗原EBNA1/DNA作用界面抑制剂等); 两个候选化合 物进入了系统的临床前评价。共发表SCI论文120余篇,其中近5年以通讯或 第一作者身份(含共同通讯或第一)在Nature、Nature Chem、Cell Res、Adv Mater、 J Am Chem Soc、 Proc Natl Acad Sci USA、 Nucleic Acid Res、 Cancer Res和J Med Chem等著名杂志发表SCI论文近60篇。已发表的研究成果曾两 次作为中国科学院研究重大进展入选《中国科学院院刊》封面报道、申请 新化学实体专利16项(美国专利一项、PCT两项),获得授权专利四项,部分 专利已经实现转让。

阐明TET家族蛋白质底物偏好性机制

卢俊彦¹ 胡璐璐² 程净东² 王 晨¹ 徐彦辉^{2*} 罗 成^{1*} (中国科学院上海药物研究所, 新药研究国家重点实验室, 上海 201203;²复旦大学生物医学研究院, 上海 200032)

摘要 基因组中胞嘧啶的甲基化修饰是重要的表观遗传标记,其动态变化参与了多种重要生物学过程。TET(ten-eleven translocation)家族蛋白质介导了5-甲基胞嘧啶(5-methylcytosine, 5mC)的连续氧化,相继生成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-carboxylcytosine, 5caC)三种产物。生化实验结果表明,虽然TET2可以连续催化5mC、5hmC和5fC的氧化,但其对不同底物的催化效率具有明显差异,针对5mC的催化效率最高,而针对5fC的最低。这一特性可能对维持基因组甲基化状态稳定具有重要意义。然而,生化与结构生物学实验均显示,TET2对不同底物结合与识别能力无明显差异。分子模拟与QM/MM计算结果表明,整个反应循环中的第三步(氢抽提)为限速步骤,且能垒趋势与实验观测反应效率一致,并预测氢抽提反应的能垒差异主要来源于不同底物在反应中间体时取向不同。进一步的同位素动力学效应实验确证了氢抽提步骤为整个反应的限速步骤,并且停留光谱实验证

^{*}通讯作者。Tel: 021-54237880, E-mail: xuyh@fudan.edu.cn; Tel: 021-50271399, E-mail: cluo@simm.ac.cn

^{*}Corresponding authors. Tel: +86-21-54237880, E-mail: xuyh@fudan.edu.cn; Tel: +86-21-50271399, E-mail: cluo@simm.ac.cn 网络出版时间: 2016-01-11 10:46:15 URL: http://www.cnki.net/kcms/detail/31.2035.Q.20160111.1046.006.html

TET2底物偏好性源于底物碱基5-位取代基自身的性质,并且证实5hmC修饰由于不易于被TET2继续氧化而可在基因组中保持稳定。这对深入理解基因组去甲基化修饰的分子机制及对TET2及其家族蛋白小分子调控剂的研发具有重要意义。

关键词 表观遗传; DNA甲基化; TET家族蛋白质; QM/MM

Elucidation of the Mechanism for the Substrate Preference of TET Proteins

Lu Junyan¹, Hu Lulu², Cheng Jingdong², Wang Chen¹, Xu Yanhui^{2*}, Luo Cheng^{1*}

(¹Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; ²Shanghai Medical College of Fudan University, Shanghai 200032, China)

Methylation of cytosine in DNA is considered as an important epigenetic marker, which takes Abstract part in various biological processes. Ten-eleven translocation (TET) proteins are key players involving in DNA demethylation through mediating the processive oxidation of 5-methylcytosine (5mC). It can convert 5mC to 5-hydroxymethylcytosine (5hmC), then to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). According to previous reports and our biochemical experiments, although TET2 could catalyze the oxidation of 5mC, 5hmC and 5fC, the catalytic efficiencies are different. The efficiency for 5mC is the highest while the efficiency for 5fC is the lowest. Biochemical experiments and crystal structures all showed TET2 could recognize and bind to its different substrates with similar binding affinities. Structure modelling and QM/MM calculations suggested different orientations of the substrates bases on the hydrogen abstraction reaction step may result in substrate preference. Further biochemical experiments such as Kinetic isotope effect (KIE) experiments and Stopped-flow spectroscopy experiments validated the hypothesis and indicated the difference in catalytic efficiencies indeed resulted from the differences in hydrogen abstraction rate. Our studies for the first time demonstrate that the substrate preference of TET2 results from the intrinsic value of its substrates at their 5mC derivative groups and suggest that 5hmC is relatively stable and less prone to further oxidation by TET proteins. In addition, it will light up the road for the development of small molecule regulators against TET proteins.

Keywords epigenetics; DNA methylation; ten-eleven translocation (TET) proteins; QM/MM

表观遗传(epigenetics)通常指在DNA序列不 发生改变的情况下,基因表达发生了可遗传的变 化。表观遗传的主要分子机制包括:DNA修饰(DNA modification)、RNA干扰(RNA interference)、组 蛋白变体(histone variant)以及组蛋白修饰(histone modification)。其中,DNA甲基化修饰是最基本也 是目前研究最为深入的一种表观遗传修饰。DNA 甲基化修饰参与生物体多种重要的生理过程,如胚 胎发育、X染色体沉默以及基因组印记等^[1-2]。哺 乳动物体内的DNA甲基化通常发生于CpG二核苷 酸中胞嘧啶的五位碳原子上,形成5-甲基胞嘧啶 (5-methylcytosine, 5mC)。CpG二核苷酸在基因组 中呈不对称分布,富含CpG的区域被称为CpG岛。 在人类基因组中,大约60%的基因启动子区存在于 CpG岛。基因启动子区CpG岛的甲基化,可以使得 染色质结构变得更为紧密,从而直接导致基因转录 活性降低。此外,还存在一些包含甲基化CpG结合 结构域(methylation CpG binding domain, MBD)的蛋 白质,可特异识别基因组中甲基化的DNA,并招募 其他表观遗传修饰酶或染色质重塑因子(chromatin remodeling factors),间接导致转录活性的降低^[3]。在 真核细胞中,特别是高等生物体内,甲基化和非甲基 化基因的转录活性可相差达10⁶倍。高度甲基化的 基因,如女性两条X染色体其中一条X染色体上的基 因,常处于失活状态;而一直处于活性转录状态的管 家基因则始终保持低水平的甲基化^[4-6]。DNA甲基 化的异常也与多种疾病(尤其是肿瘤)有着密切的联 系。抑癌基因的过度甲基化或原癌基因的甲基化不 足,可导致抑癌基因的沉默与原癌基因的激活,从而 促进肿瘤的发生、发展^[7]。

DNA甲基化修饰主要由两大类DNA甲基转移 酶(DNA methyltransferases)负责:维持性DNA甲基 转移酶——DMNT1(DNA methyltransferases 1)和从 头DNA甲基转移酶——DMNT3a、DMNT3b。作为 DNA甲基化修饰的逆过程(DNA的去甲基化修饰) 在相当长的时间内被认为是一种被动的过程,即在 DNA多轮复制的过程中甲基化修饰会逐渐丢失。直 到最近几年,一批在DNA主动去甲基化过程中扮演 关键作用的酶被相继发现,人们才开始认识到DNA 主动去甲基化过程在表观遗传调控中发挥的重要作 用。这些酶主要包括:活化诱导胞嘧啶核苷脱氨酶 (activation induced cytidine deaminase, AID), TET 家族蛋白质以及胸腺DNA糖苷酶(thymine-DNA glycosylase, TDG)^[8-11]。这其中, TET家族蛋白质是 最为年轻的一个成员,在2009年才被发现[12]。但由 于TET家族蛋白质在胚胎发育和肿瘤生物学中扮演 的重要角色,已经成为表观遗传领域的研究热点。

TET家族蛋白质均属于铁-酮戊二酸依赖的/ 双加氧酶超家族,其主要生理功能是催化5-甲基 胞嘧啶(5mC)中甲基的氧化,生成5-羟甲基胞嘧啶 (5-hydroxymethylcytosine, 5hmC)^[13-15]。有报道表 明,小鼠胚胎干细胞(mouse embryonic stem cells, mESCs)以及诱导多功能干细胞(induced pluripotent stem cells, iPSCs)中广泛存在的5hmC标记主要来源 就是TET(ten-eleven translocation)蛋白质催化5mC 的氧化[13-15]。与DNA的甲基化修饰不同, DNA的羟 甲基化修饰主要参与了基因表达激活。TET蛋白 质可以进一步以5hmC为底物,将其连续氧化生成5-醛基胞嘧啶(5-formylcytosine, 5fC)和5-羧基胞嘧啶 (5-carboxylcytosine, 5caC)^[15]。而这两种5hmC的氧 化产物可以被TDG特异性识别并进行剪切。在碱基 切除修复(base excision repair)的机制作用下,最终完 成DNA主动去甲基化的过程^[8]。

尽管TET蛋白质在DNA主动去甲基化过程中的作用已被广泛认可,然而越来越多的证据表明, TET蛋白质催化产生的5hmC不仅仅只是一个DNA 去甲基化过程的中间产物。近年来,多种特异性 的5hmC结合蛋白相继被发现^[16-17];同时,研究发现, 5hmC标记同5mC一样,可以在长期稳定的基因组中 存在^[18]。因而, 5hmC可能作为一种新的、稳定的表

观遗传标记,直接参与基因转录调控,而无需依赖 DNA去甲基化过程。由此也可见, TET蛋白质在转 录调控中扮演的重要角色。而TET蛋白质是如何在 DNA去甲基化过程以及保持5hmC稳定之间保持平 衡的呢?根据之前的文献报道以及我们的生化实 验结果发现,虽然TET蛋白质可以催化5mC连续氧 化生成5hmC、5fC以及5caC,其每一步的催化效率 并不相同。人类TET1/2蛋白质、小鼠Tet2蛋白质以 及线虫中的NgTet更容易催化5mC生成5hmC,而催 化5hmC生成5fC以及进一步催化5fC生成5caC的能 力则较弱^[19-21]。这一特性可能决定了TET蛋白质在 通常情况下倾向于仅将5mC催化产生5hmC,只有当 TET蛋白质在基因组某一区域明显富集时,才能显 著地将5mC连续氧化为5fC和5caC,介导DNA的去 甲基化过程。由于TET蛋白质对于底物的偏好性可 能与其生理功能有着密切的联系,我们对TET蛋白 质底物偏好性的机制产生了兴趣。

通过多种生物化学实验手段,如荧光偏振 (fluorescence polarization, FP)以及表面等离子共振 (Surface plasmon resonance, SPR)等, 我们发现, TET 蛋白质与包含不同碱基(5mC、5hmC和5fC)的底物 DNA片段的结合能力并没有显著差异,说明TET蛋 白质的底物偏好性不大可能来自于其对底物结合 能力的差异。同时,我们还获得了TET蛋白家族的 一个重要成员TET2及其与不同底物结合的晶体结 构(图1A)。复合物晶体结构也显示, TET2与5mC、 5hmC和5fC的结合模式基本相同,说明TET2催化 口袋也可适应不同底物碱基的结合。但是,我们 发现,在TET2的催化中心,参与反应的碱基上5-位 修饰基团取向在不同的复合物结构中略有不同(图 1A~图1C)。因此,我们猜测,TET2对于底物碱基催 化效率的差异是否来源于催化过程中底物5-位修 饰基团取向不同所导致的催化能垒的差异? 由于 TET蛋白质属于铁-酮戊二酸依赖的双加氧酶超家 族,因而TET蛋白质与其他铁--酮戊二酸依赖的双 加氧酶家族成员,如负责烷基化碱基修复的AlkB蛋 白、胸腺嘧啶双加氧酶(thymine dioxygenase)以及组 蛋白去甲基化酶(jumonji domain-containing histone demethylases, JMJD)家族等, 可能具有相似的催化机 制^[22]。有研究表明, TET涉及的催化机制较为复杂, 至少涉及脱羧反应(decarboxylation)、氧桥异裂(O-O heterolysis)、氢抽提(hydrogen abstraction)以及羟基



A~C: TET2与5mC、5hmC和5fC的结合模式示意图。TET2活性口袋区域与不同碱基的结合模式基本相似,但碱基的5-位取代基朝向有所不同。 D: TET蛋白质催化循环示意图。

A~C: the binding mode between TET2 and 5mC, 5hmC and 5fC. The catalytic pocket of TET2 can accommodate different bases, but the orientations of the 5-substitution group are different. D: the catalytic cycle of TET proteins.

图1 TET2与不同碱基结合的共晶结构以及催化循环示意图(根据参考文献[21]修改)

Fig.1 Structure of TET2 binds to different substrates and an illustration of catalytic cycle of TET2 (modified from reference [21])

再结合(hydroxyl rebind)这四步连续反应,还涉及众 多的反应中间体(图1D)。因而,仅采用实验方法难 以直接对每步反应的效率进行测量。

于是,我们从理论模拟出发,采用了量子力学/ 分子力学(QM/MM)混合模拟的手段^[23],分别计算 了TET2催化5mC、5hmC和5fC每一步反应的能垒。 根据计算结果显示,整个反应循环中的第三步(氢抽 提反应)所需要跨越的能垒最高,因此认为,氢抽提 反应是TET2催化底物氧化的限速步骤。同时我们 发现,TET2催化不同底物氧化的前两步反应的能垒 均较低且没有明显的差异,而催化氢抽提反应的能 垒则是5mC>5hmC>5fC(图2A)。根据过渡态理论 (transition state theory),能垒越高,反应速率越慢^[24], 因而TET2催化不同底物氢抽提反应的能垒大小与 实验测得的反应速率相一致。同时,通过对模拟得 到的中间体结构进行分析,我们发现,氢抽提反应的 能垒差异可能主要来源于不同底物在反应中间体时 取向不同,即为5mC甲基上氢的朝向更适合反应,而 5fC和5hmC上的氢更加远离反应中心,不适合氢抽 提反应的进行。

进一步,我们采用生化实验手段对理论模拟结 果进行了验证。根据氢抽提反应的原理,在反应进 行前,需要产生一个活性的高价铁-氧中间体—— Fe(IV)=O。这一中间体具有很高的反应活性,可夺 取底物上氢原子,使底物的C-H键断裂。根据反应 机理,我们首先将5mC上的氢原子(H¹)采用其同位 素氘(H²)代替,进行酶活性测试。根据同位素动力 学效应(kinetic isotope effect),用H²代替H¹会使氢抽 提反应的速率显著降低,若氢抽提为反应的限速步 骤,则5mC的催化效率也会显著降低^[25]。与预测相 吻合的是,TET2催化被H²置换的5mC的效率显著降 低,说明氢抽提这一步骤的确是整个TET2催化循环



A: QM/MM计算发现, TET2催化5mC发生氢抽提反应的能全最低而催化5fC的能全最高; B: 同位素动力学效应实验显示, 氢抽提反应为TET2催化循环限速步骤; C: 停流光谱实验通过检测Fe(IV)=O在318 nm处的信号, 证实5mC氢抽提反应速率最快而5fC速率最慢。 A: QM/MM calculations indicated the barrier for the hydrogen abstraction reaction of 5mC was the lowest while the barrier for 5fC was the highest; B: kinetic isotope effect suggested the hydrogen abstraction was the rate-limiting step in the catalytic cycle of TET2; C: through monitoring the signaling of absorption at 318 nm, the stopped-flow spectroscopy experiments validated that the hydrogen abstraction rate for 5fC was the lowest.



的限速步骤(图2B)。为验证TET2催化不同底物氢 抽提反应的效率是否有差异,我们接下来采用停流 光谱法(stopped-flow spectroscopy)检测整个反应过 程中Fe(IV)=O基团存在的时间。由于Fe(IV)=O基 团在318 nm处有特异性吸收信号,若该信号持续的 时间长,则说明Fe(IV)=O消耗速率慢,氢抽提反应效 率低。同样,与预测结果相吻合的是,5fC在318 nm 处吸收信号最强且持续的时间最长,5hmC次之,而 5mC的吸收信号最弱、持续时间最短,说明氢抽提 反应的效率的确是5mC最高而5fC最低。

这一研究首次揭示了TET蛋白质对其三种底物 的不同催化活性的分子机制,阐明了TET蛋白质底 物偏好性原理,为基因组中5-羟甲基胞嘧啶稳定存 在提供了分子水平的解释。此发现解决了困扰表观 遗传学领域的一个难题,也为研究其他蛋白质逐步 催化反应的分子机制提供了新思路和新方法。该 研究论文("Structural insight into substrate preference for TET-mediated oxidation",晶体结构揭示TET蛋白 介导的氧化反应底物偏好性机制)于2015年10月29 日在线发表于国际顶级学术期刊《自然》^[21]。此外, 由于目前还未见TET家族蛋白质选择性小分子调控 剂的报道,针对TET蛋白质结构和催化机制的研究, 也有助于基于结构和基于机理的TET小分子调控 剂的发现。目前,通过结合计算机辅助药物设计和 酶活性测试,我们已经获得了一批极具潜力的针对 TET蛋白质的小分子抑制剂。对这些小分子化合物 的进一步确证、优化和改造将有希望为TET家族蛋 自质的化学生物学研究和靶向TET家族蛋白质的药 物设计研究提供优良的工具分子和先导化合物。

参考文献 (References)

- 1 Bird AP. CpG-rich islands and the function of DNA methylation. Nature 1986; 321(6067): 209-13.
- 2 Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 1993; 366(6453): 362-5.
- 3 Wade PA. Methyl CpG-binding proteins and transcriptional repression. BioEssays 2001; 23(12): 1131-7.
- 4 Cedar H. DNA methylation and gene activity. Cell 1988; 53(1): 3-4.
- 5 Imamura M, Miura K, Iwabuchi K, Ichisaka T, Nakagawa M, Lee J, et al. Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells. BMC Dev Biol 2006; 6: 34.
- 6 Kass SU, Pruss D, Wolffe AP. How does DNA methylation

repress transcription? Trends Genet 1997; 13(11): 444-9.

- 7 Baylin SB. DNA methylation and gene silencing in cancer. Nat Clin Pract Oncol 2005; 2 Suppl 1: S4-11.
- 8 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.
- 9 Mohr F, Döhner K, Buske C, Rawat VP. TET Genes: New players in DNA demethylation and important determinants for stemness. Exp Hematol 2011; 39(3): 272-81.
- 10 Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell 2011; 8(2): 200-13.
- 11 Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 2010; 463(7284): 1042-7.
- 12 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.
- 13 Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 2011; 473(7347): 398-402.
- 14 Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 2009; 324(5929): 929-30.
- 15 Xu CR, Cole PA, Meyers DJ, Kormish J, Dent S, Zaret KS. Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas. Science 2011; 332(6032): 963-6.
- 16 Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N, MeCP2

binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 2012; 151(7): 1417-30.

- 17 Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I, et al. Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. PLoS One 2011; 6(6): e21306.
- 18 Bachman M, Uribe-Lewis S, Yang XP, Williams M, Murrell A, Balasubramanian S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat Chem 2014; 6(12): 1049-55.
- 19 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.
- 20 Hashimoto H, Pais JE, Zhang X, Saleh L, Fu ZQ, Dai N, et al. Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. Nature 2014; 506(7488): 391-5.
- 21 Hu L, Lu J, Cheng J, Rao Q, Li Z, Hou H, et al. Structural insight into substrate preference for TET-mediated oxidation. Nature 2015; 527(7576): 118-22.
- 22 Hausinger RP. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. Crit Rev Biochem Mol Biol 2004; 39: 21-68.
- 23 Friesner RA, Guallar V. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. Annu Rev Phys Chem 2005; 56: 389-427.
- 24 Lasaga AC. Transition state theory. Rev Mineral (United States) 1981; 8.
- 25 Price JC, Barr EW, Hoffart LM, Krebs C, Bollinger JM. Kinetic dissection of the catalytic mechanism of taurine: α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. Biochemistry 2005; 44(22): 8138-47.