

组蛋白 H3K36 位点甲基转移酶识别组蛋白底物的分子机制研究进展*

曹英华 王占新[†]

(细胞增殖及调控生物学教育部重点实验室, 北京师范大学生命科学学院, 100875, 北京)

摘要 组蛋白 H3 第 36 位赖氨酸的甲基化修饰在染色质上含量丰富, 与活跃转录以及 DNA 损伤修复等重要生理过程相关. H3K36 位点可以被一甲基化、二甲基化和三甲基化 3 种形式修饰, 目前已知的负责组蛋白 H3K36 三甲基化修饰的人源蛋白是 SETD2, 负责组蛋白 H3K36 二甲基化修饰的酶包含 NSD1、NSD2 和 NSD3 和 ASH1L 共 4 名成员. 这些 H3K36 甲基转移酶都具有非常特异的 H3K36 位点选择性, 因此, 对调控体内 H3K36 甲基化修饰的水平和分布十分重要. 此外, 它们的表达异常与人类的多种疾病相关. 因此, 解析组蛋白 H3K36 甲基转移酶识别并修饰组蛋白底物的分子机制, 对揭示这些酶参与的表观遗传调控机制及其在体内的生理功能都具有十分重要的意义. 早期的研究使得人们对组蛋白 H3K36 甲基转移酶催化底物的机制有了较深入的认识, 但是由于解析的修饰酶与底物复合物的结构较少, 对这些酶特异识别组蛋白底物分子机制的认识尚有很多不足. 近年来, 随着冷冻电镜技术的应用, H3K36 甲基转移酶与核小体底物的复合物结构相继取得了突破, 极大地推进了人们对这些酶识别并催化组蛋白底物分子机制的认识. 本文以这几个组蛋白 H3K36 甲基转移酶为主要目标, 对其分子机制的最新进展进行介绍总结.

关键词 H3K36 甲基转移酶; SETD2; NSD 家族蛋白; ASH1L; 核小体底物

中图分类号 Q71

DOI: 10.12202/j.0476-0301.2021323

0 引言

真核生物的基因组 DNA 以染色质的形式存储遗传信息. 染色质的最小结构单元是核小体. 核小体的核心颗粒(NCP)是由 145~146 碱基对的 DNA 缠绕着由 H2A、H2B、H3 和 H4 组成的组蛋白八聚体形成的一个圆盘状结构^[1]. 核小体核心颗粒之间由连接 DNA 相连. 核小体之间既可以像珠串一样松散地连接形成易于转录的常染色质, 也可以被进一步压缩成具有高级结构的对转录起抑制作用的异染色质. 组蛋白修饰可以调控染色质的结构和基因表达, 是表观遗传调控的重要方式^[2]. 组蛋白修饰有很多种, 例如赖氨酸可以被乙酰化、甲基化和泛素化修饰. 组蛋白八聚体中被核小体 DNA 包裹在内结构性较强的部分被称为组蛋白核心, 而暴露在核小体 DNA 包裹的圆盘外的组蛋白的 N 端和 C 端由于柔性较强, 被称为组蛋白尾巴. 组蛋白修饰主要发生在组蛋白尾巴上. H3K36 在核小体上的位置非常特殊, 它虽位于组蛋白 H3 的尾巴上, 但是与核小体 DNA 离得很近, 因此修饰或识

别 H3K36 位点的蛋白通常通过与核小体核心颗粒相互作用发挥功能. 此外, H3K36 还位于靠近连接 DNA 的地方, 这也预示着 H3K36 位点修饰及与其相互作用的蛋白很有可能参与核小体连接 DNA 的动态调控. 关于组蛋白修饰相关的研究和综述很多, 本文主要总结组蛋白 H3K36 位点修饰酶特异性识别组蛋白底物的最新研究进展.

组蛋白 H3K36 位点主要通过甲基化修饰的方式参与表观遗传调控. H3K36 位点在其特异的组蛋白赖氨酸甲基转移酶的催化下, 可以形成一甲基化 (H3K36me1)、二甲基化 (H3K36me2) 和三甲基化 (H3K36me3) 3 种修饰状态, 其中甲基基团由蛋白质赖氨酸甲基转移酶的通用辅酶 S-腺苷甲硫氨酸 (SAM) 提供. SAM 失去活跃的甲基基团后的产物是 S-腺苷高半胱氨酸 (SAH). 目前已知的 H3K36 位点特异的组蛋白赖氨酸甲基转移酶分为 2 类^[3]. 一类是负责三甲基化修饰的酶, 人源蛋白只有 SETD2 一个成员. 在体外, SETD2 能够催化 H3K36 位点的一甲基、二甲基和三甲基化修饰, 但是其在胞内的主要功能是

* 国家自然科学基金资助项目(31870725, 32071204)

[†] 通信作者: 王占新(1977—), 男, 博士, 教授. 研究方向: 表观遗传调控的分子机制. E-mail: wangz@bnu.edu.cn

收稿日期: 2021-12-28

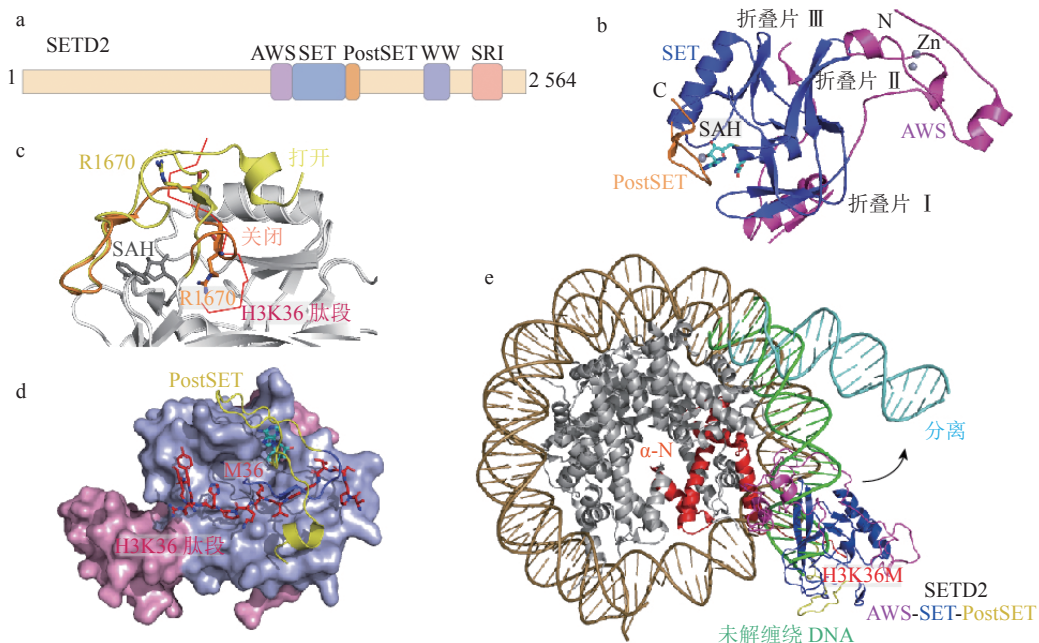
调控染色质的 H3K36me3 水平^[4]。另一类是负责二甲基化修饰的酶, 人源蛋白包括 ASH1L 和 NSD 家族的 NSD1、NSD2 和 NSD3 共 4 个成员。这些二甲基转移酶在体外实验中也具有一甲基化修饰的活性, 但是在胞内的功能主要是调控染色质的 H3K36me2 修饰水平。目前还未发现专门负责 H3K36 位点一甲基化修饰的酶, 所以 H3K36me1 修饰通常被认为是二甲基化和三甲基化修饰的中间产物, 或者是二甲基化修饰被组蛋白去甲基酶作用的产物。SMYD2^[5] 和 SETMAR^[6] 也被报道可以修饰组蛋白 H3K36 的甲基化, 但是 SMYD2 对组蛋白底物的选择性并不高, 而对非组蛋白如 p53 有较高的选择性^[7]; SETMAR 在 DNA 的损伤修复中发挥作用, 但是有质谱实验表明它并不能甲基化 H3K36^[8]。此外, 由于 SMYD2 和 SETMAR 都不能以核小体为底物进行修饰, 因此 SMYD2 和 SETMAR 不再被认为是 H3K36 特异的组蛋白赖氨酸甲基转移酶^[3]。

1 三甲基化修饰酶——SETD2 的底物识别机制

酵母蛋白 Set2 最先被发现具有组蛋白 H3K36 三甲基化修饰酶的活性^[9], 随后其人源同源蛋白 SETD2 被发现具有同样的活性^[10]。到目前为止, SETD2 是已知哺乳动物细胞中唯一可以催化组蛋白 H3K36 三甲基

化修饰的甲基转移酶。H3K36me3 修饰在小鼠的多种组织细胞中占组蛋白 H3 的 5% 左右^[11], 是活跃转录延伸的标志修饰。SETD2 与转录延伸过程中的 RNA 聚合酶 II (RNAPII) 相互作用^[12], 在活跃转录的基因上留下 H3K36me3 修饰, 这种基因上的 H3K36me3 修饰可以防止基因内部的错误起始, 保障转录的准确性^[13]。此外, SETD2 还可以甲基化修饰 α -tubulin 的 K40 位点^[14] 和 STAT1 的 K525 位点^[15], 这些非组蛋白修饰分别在维持微管的稳定和 INF α 依赖的抗病毒免疫应答中发挥重要功能。

人源 SETD2 含有 2564 个氨基酸残基, 包含 1 个由 AWS-SET-PostSET 这 3 个紧密联系在一起的结构域组成的催化片段, 1 个 WW 结构域和 1 个 SRI 结构域 (图 1-a)。WW 结构域主要负责介导蛋白质之间的相互作用^[16], SETD2 的 WW 结构域可以与 Huntingtin 蛋白相互作用, 因此与“亨廷顿病”相关^[17]; SETD2 还可以通过其 WW 结构域与 p53 相互作用调控基因的表达^[18]。SETD2 的 SRI 结构域命名源于酵母 Set2 的同源结构域, 而且 SRI 结构域的功能从酵母到人类都十分保守, 都能够与活跃转录的磷酸化的 RNA 聚合酶 II 的 C 末端尾巴相互作用^[19-20], 这种相互作用把 SETD2 的功能与转录延伸过程偶联起来。SETD2 的核心催化结构域由 AWS-SET-PostSET 区段组成, 其中 SET



a. 人源 SETD2 的结构组成示意。b. SETD2 的 AWS-SET-PostSET 催化结构域与配体 SAH 的复合物晶体结构示意图 (PDB 号 4H12)。其中 AWS、SET 和 PostSET 结构域分别标为粉色、蓝色和橙色。SAH 显示为棒状模型, 锌离子显示为灰色圆球。c. 自抑制状态与激活状态的 SETD2 催化结构域结构对比。橙色 (关闭) 为自抑制构象 (PDB 号 4H12), 黄色 (打开) 为结合肽段后的激活构象 (PDB 号 5JJY)。含有 H3K36 位点的小肽用红色线条显示, 其余部分用灰色标记。d. SETD2 与含有 H3K36 的肽段复合物的表面结构示意图 (PDB 号 5JJY)。e. SETD2 与核小体复合物的结构示意图。红色标记 H3.3 的 N 端 α 螺旋。解缠绕和未解缠绕的核小体出/入口处的 DNA 分别用青色和绿色标记 (PDB 号 7EA8, 5AVC)。

图 1 SETD2 识别组蛋白底物的分子机制

结构域是蛋白质赖氨酸甲基转移酶的典型结构域,而 AWS 结构域和 PostSET 结构域则分别位于 SET 结构域之前和之后,对 SET 结构域发挥活性至关重要,因此 SETD2 主要通过 AWS-SET-PostSET 区段识别并催化底物。

SETD2 催化结构域与配体 SAH 的复合物结构以及含有致癌突变 H3.3K36M/H3K36I 多肽底物和 SAH 的复合物结构的相继解析^[21-22],为揭示 SETD2 催化底物的分子机制提供了重要线索。SETD2 的 SET 结构域主要由 3 个三角形排布的 β 折叠片组成的核心加 1 个外周分布的长 α 螺旋组成(图 1-b),这种结构在组蛋白甲基转移酶中十分保守。而分别位于 SET 结构域 N 端和 C 端的 AWS 和 PostSET 结构域与中心的 SET 结构域紧密交错排布形成一个结构整体,小分子 SAH 配体位于 SET 和 PostSET 结构域之间的一个保守的口袋中(图 1-b)。不同于其他赖氨酸甲基转移酶的是,SETD2 催化结构域的活性中心被其 PostSET 区的一段无规卷曲(称为自抑制无规卷曲)所结合而呈现一种自抑制构象(图 1-c、d),这种构象下该自抑制无规卷曲结合在底物肽段结合的位置上,其上的精氨酸 Arg1670 侧链插到底物 H3K36 的侧链插入的口袋中(图 1-c),因此 SETD2 的催化活性受到抑制。结合底物后,SETD2 的自抑制无规卷曲向外侧移动,其 Arg1670 侧链外翻暴露出 H3K36 侧链的结合口袋,同时伴随着 PostSET 的 C 末端片段的位置调整,SETD2 的催化结构域由自抑制状态转变为“激活态”(图 1-c、d)。

虽然 SETD2 与小肽底物的复合物结构可以解释其催化底物的机制,但是生化研究表明 SETD2 更倾向于催化核小体底物^[3],而基于肽段的研究无法解释其选择核小体底物的机制。随后真菌 Set2^[23]以及人源 SETD2^[24]与核小体复合物的冷冻电镜结构相继被解析,揭开了 SETD2 蛋白催化核小体底物的分子机制。真菌 Set2 与人源 SETD2 识别核小体底物的分子机制非常类似。以人源 SETD2 为例,SETD2 主要结合在核小体的侧面靠近连接 DNA 的地方(图 1-e)。SETD2 的结合使得核小体出/入口处的 DNA 部分解缠绕,然后 SETD2 的催化结构域结合到解缠绕的 DNA 下暴露出来的组蛋白表面上,从而激活并发挥催化功能(图 1-e)。这种核小体 DNA 的解缠绕是 SETD2 能够以核小体为底物发挥催化功能所必需的,因为 SETD2 的催化结构域结合的组蛋白片段有一部分被核小体出/入口处的 DNA 挡住了,只有核小体出/入口处的 DNA 部分解缠绕,才能使足够长的组蛋白 H3 暴露出来被 SETD2 识别(图 1-e)。相反,如果加入组蛋白 H1,

组蛋白 H1 与连接 DNA 的相互作用可以稳定核小体的紧密构象,抑制核小体出/入口处 DNA 的解缠绕,阻碍核小体与 SETD2 的结合,从而抑制 SETD2 介导的 H3K36 三甲基化修饰^[24]。在与核小体的相互作用中,SETD2 不仅识别晶体结构中观测到的组蛋白尾巴片段,其 AWS 结构域还对组蛋白 H3 的 N 端 α 螺旋进行识别,同时 SETD2 的催化结构域还对内外 2 圈核小体 DNA 进行识别(图 1-e)。这些额外的相互作用,使得 SETD2 对核小体底物具有更高的亲和力和催化活性。值得注意的是,H2B 的 K120 位点(酵母中是 H2B 的 K123 位)的泛素化修饰会增强 Set2 对核小体的活性^[23, 25],结构表明 H2B 的 K120 位上修饰的泛素在空间上与 Set2 的催化结构域有相互作用,因此可能稳定了 Set2 在核小体上的结合。

此外,酵母 Set2 的 SRI 结构域和人源 SETD2 的 SRI 结构域都能够增强各自蛋白对核小体的催化活性^[24],但是由于解析的结构中并不能看到 SRI 结构域,因此其促进酶活机制仍有待研究。除此之外,对 SETD2 的底物序列特异性进行的探索发现,相比于野生型 H3,一些氨基酸突变的底物可以显著提高酶活^[26],但是由于这些序列并不存在于组蛋白中,暗示 SETD2 在选择非组蛋白底物时可能会识别不同的蛋白质序列或者三级结构。

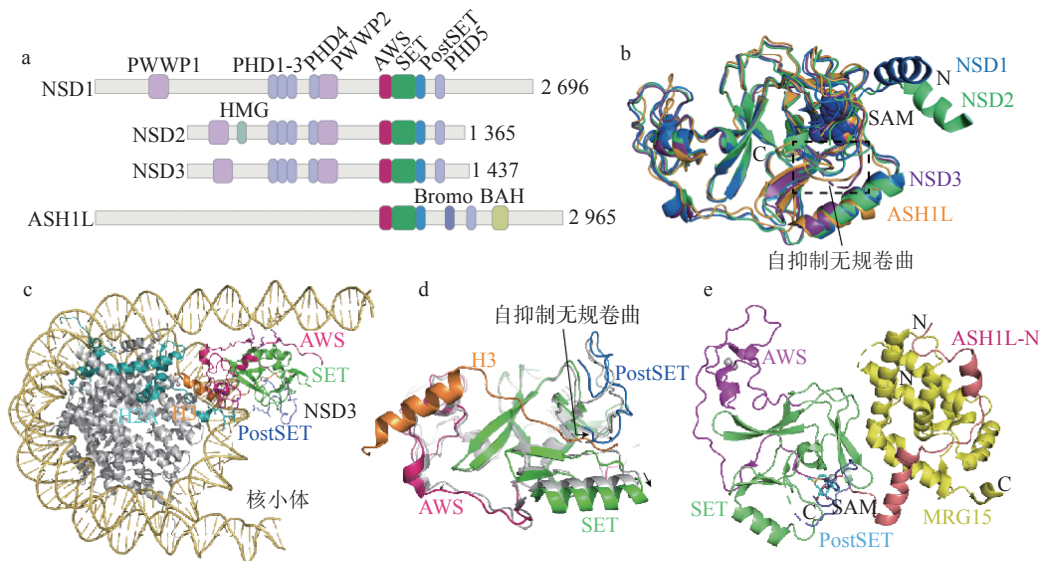
2 二甲基化修饰酶——NSD 家族蛋白的底物识别机制

人源 NSD 家族蛋白由 NSD1、NSD2(WHSC1、MMSET)和 NSD3(WHSC1L1)3 个成员组成,都是组蛋白 H3K36 的二甲基转移酶,参与多种重要的生理功能并且与许多人类疾病有关^[13]。例如:染色体易位产生的融合蛋白 NUP98-NSD1 和 NUP98-NSD3 可导致急性髓系白血病;而 NSD2 融合蛋白可导致多发性骨髓瘤的发生^[27-29],NSD2 的高表达会导致整个基因组 H3K36me2 修饰水平的异常增高,从而导致与之拮抗的 H3K27 甲基化修饰水平的下调,引起多种基因的异常表达以致癌症的发生^[30]。

NSD 家族蛋白具有保守的结构域组成,它们都含有 2 个 PWWP 结构域、5 个 PHD 结构域和 1 个 AWS-SET-PostSET 催化结构区域(图 2-a)。NSD 家族蛋白的 PWWP 结构域能够识别含有 H3K36m2/3 修饰的组蛋白^[31-32],NSD2 的 PHD1-2 结构域与其在染色质上的招募及催化活性相关^[33],但是由于相关的结构数据缺失,其调控机制未知。NSD2 和 NSD3 的 PHD5 可以识别未修饰的 H3K4,而 NSD3 的 PHD5 对 H3K9me3 修饰具有一定的亲和力^[34],这种对组蛋白识别的差异可

能有助于 NSD 家族蛋白对底物的选择性. 与 SETD2 类似, NSD 家族蛋白的催化活性也是由 AWS-SET-PostSET 这一催化结构域来完成. NSD 家族蛋白的 AWS-SET-PostSET 催化结构域晶体结构表明^[35-37], 该家族蛋白的催化核心、SAM 分子结合的方式和结合位点与 SETD2 类似, 主要差别是 NSD 家族蛋白与 SETD2 催化结构域的 N 端 α 螺旋方向略有不同(图 2-b). NSD 家族蛋白催化结构域在未结合底物时也呈现一种自抑制构象(图 2-b), 这种自抑制构象下, SET 和 PostSET 结构域连接处的无规卷曲通过氢键作用结合在底物结合的通道上, 阻碍了组蛋白底物的结合. 自抑制构象对 NSD 家族蛋白底物催化活性有着较大的影响, 因为以合成的组蛋白小肽为底物时, NSD 家族蛋白的催化活性非常低而且选择性也很差^[37-38], 而 SETD2 对小肽底物仍有较高的活性. 只有以核小体为底物时, NSD 家族蛋白才能被激活并发挥最大的催化活性. 这种核小体选择性的催化机制在 NSD2 和 NSD3 与核小体复合物的冷冻电镜结构解析后得以揭示^[39-40]. 在以合成的小肽为底物时, 可能由于亲和力不够, 小肽无法挤开 PostSET 的自抑制无规卷曲, 因此 NSD 家族蛋白难以激活. 在以核小体为底物时, NSD 家族蛋白通过与核小体 DNA 和组蛋白的多重相互作用而激活. 同 SETD2 与核小体复合物的结构类似, NSD2 或 NSD3 与核小体的结合也会导致核小体出入口处 DNA 的解缠绕, 这样 NSD 家族蛋白的催化结构域夹在内外 2 圈核小体 DNA 之间, 其 AWS 结构域

和 SET 结构域分别与外圈和内圈 DNA 的小沟结合(图 2-c), 增强了 NSD 家族蛋白对核小体的亲和力. 其次, NSD 家族蛋白的催化结构域与 SETD2 一样, 也和组蛋白 H3 的 N 端 α 螺旋有紧密的相互作用(图 2-c), 这部分 H3 结合位点是由于核小体出入口处的 DNA 解缠绕才得以暴露. 最后, 组蛋白 H3 的 N 末端尾巴从核小体的 2 圈 DNA 中伸出之后, 正好结合在 NSD 家族蛋白催化结构域的底物结合通道中, 挤开了自抑制无规卷曲, 因此完成了 NSD 蛋白的激活(图 2-c、d). 与自抑制的 NSD 催化结构域相比较, 结合核小体的 NSD 催化结构域由于核小体的稳定作用, 整体向外以比较伸展的方式结合在组蛋白底物上(图 2-d), 这些信息对 NSD 家族蛋白被核小体底物激活的原理作出了合理的解释. 此外, NSD 蛋白还与 H2A 的 C 末端尾巴相互作用, 因此 H2A 的 K119 位点泛素化修饰可能引起空间位阻而抑制 NSD 蛋白的活性(图 2-c), 解释了之前观测到的生化现象^[41]. 与 Set2 不同的是, H2B 的 K120 泛素化修饰不但不能激活 NSD 蛋白, 还对 NSD 的活性起抑制作用, 可能原因是 H2B 的 K120 泛素化修饰会影响非激活态下 NSD 蛋白与核小体的非特异结合^[39]. 此外值得注意的是, NSD 家族蛋白的 PostSET 结构域后面的一段碱性区被报道有助于该家族蛋白结合核小体以及催化底物^[42], 但是解析的结构中都观测不到该区段的结构, 表明这段碱性区可能通过非特异性的增强与核小体的相互作用来增强 NSD 蛋白的催化活性.



a. NSD 家族蛋白和 ASH1L 的结构示意. b. NSD1 (蓝色)、NSD2 (绿色)、NSD3 (紫色) 和 ASH1L (橘色) 的催化结构域叠合 (PDB 号 3OOI、5LSU、5UPD 和 4YNM). SAM 分子用球状模型显示. c. NSD3 与核小体复合物的结构示意 (PDB 号 7CRP). 组蛋白 H3 和 H2A 分别显示为橙色和青色. d. NSD3 催化结构域的自抑制构象 (PDB 号 5UPD) 与激活构象 (PDB 号 7CRP) 的叠合对比图. e. ASH1L 与 MRG15 的复合物结构 (PDB 号 6INE). AWS、SET 和 PostSET 结构域分别用浅紫色、淡绿色和浅蓝色显示, 橘红色显示的是 ASH1L 催化结构域 N 端之前的保守序列, 与之结合的 MRG15 用黄色显示.

图 2 NSD 家族蛋白和 ASH1L 识别组蛋白底物的分子机制

3 二甲基化修饰酶——ASH1L 的底物识别机制

果蝇 Ash1 及哺乳动物同源物 Ash1-like (ASH1L) 都具有 H3K36 二甲基化修饰活性^[43], 并能拮抗 Polycomb 复合物的功能, 即 H3K36 的甲基化修饰抑制 H3K27 位点的甲基化^[44]. 此外, ASH1L 对 MLL1 染色体易位引起的急性白血病发病起着重要的作用^[45].

ASH1L 包含 1 个 Bromo(bromodomain)、1 个 PHD 和 1 个 BAH 结构域, 以及 1 个 AWS-SET-PostSET 催化结构域(图 2-a). ASH1L 催化结构域的晶体结构^[46]表明其具有与 NSD 家族蛋白催化结构域相似的特征, 而且未结合底物的 ASH1L 的催化结构域也呈现自抑制的构象(图 2-b). 对其自抑制无规卷曲的动态变化研究表明, 该无规卷曲的构象变化对于解除自抑制很重要^[47]. 与 NSD 家族蛋白类似, 由于自抑制构象的存在, ASH1L 对合成的小肽底物有着很低的活性, 而与核小体底物的相互作用会促进其自抑制的解除并发挥较高的催化活性^[47].

此外, 近年研究发现, 果蝇 Ash1、Mrg15 与 Caf1/Nurf55 形成 3 元复合物^[48-49], 体外实验证实 Mrg15/MRG15 能够提高 Ash1/ASH1L 的甲基转移酶活性^[48]. 在 ASH1L 的催化结构域与 MRG15 的复合物结构^[50-51]中发现, MRG15 结合在 ASH1L 催化结构域 N 端保守的序列上后会使其自抑制无规卷曲变得不稳定而易于从底物结合口袋中解离出来(图 2-e), 通过变构的方式激活 ASH1L 的催化活性. 不过由于没有获得与底物的复合物结构, MRG15 激活 ASH1L 的活性尤其是其对核小体底物的催化机制仍不清楚.

4 总结与展望

人源组蛋白 H3K36 赖氨酸甲基转移酶主要包含三甲基修饰酶 SETD2、ASH1L 以及 NSD1-3 等 4 个二甲基修饰酶, 这些蛋白虽然生理功能、相对分子质量大小和结构域组成有很大差别, 但是它们在底物识别以及催化的分子机制上有很多相同的地方.

1) 这些蛋白的催化核心都是由 AWS-SET-PostSET 这 3 个紧密结合在一起的结构域组成, 而其他结构域虽然不直接参与催化, 但可以通过结合染色质或者招募其他蛋白的方式影响催化功能.

2) 这些蛋白的催化活性中心在未结合底物时都呈现一种自抑制的构象, 这种自抑制构象在蛋白修饰酶中经常出现, 可以防止酶对非特异底物的修饰.

3) 这些 H3K36 甲基转移酶都以核小体为底物时

才能被激活, 并且发挥最大的活性.

4) 除了 ASH1L 与核小体的复合物结构未知外, 其他的 H3K36 甲基转移酶催化核小体底物的时候都会导致核小体出/入口处 DNA 的解缠绕, 这样可以使 H3K36 上游的一段组蛋白序列暴露出来, 便于修饰酶催化结构域的识别.

H3K36 甲基转移酶的催化结构域识别 H3 的 N 端 α 螺旋、H3K36 附近的序列以及核小体 DNA 来增强对核小体的结合, 发挥催化功能. H3K36 甲基转移酶的这种核小体结合特性与大多数蛋白通过核小体 H2A/H2B 的酸性区结合核小体的方式有很大差别. 尽管对这些酶分子机制的研究取得了很大进展, 但是尚有很多问题有待研究.

1) H3K36 位点甲基转移酶甲基化底物的级数调控机制, 比如 NSD 家族蛋白负责催化作用的活性氨基酸与 SETD2 相应位点的氨基酸类似, 但是 NSD 家族蛋白只能修饰到二甲基化而不能三甲基化 H3K36 的原因未知.

2) SETD2 可以甲基化非组蛋白底物, 同样地, 对 SETD2 和 NSD1 底物序列的筛选表明在肽段水平可以找到活性优于组蛋白的底物, 那么这些组蛋白修饰酶的非组蛋白底物选择性、其催化机制及生理意义仍有许多未知.

3) 这些 H3K36 组蛋白赖氨酸甲基转移酶都是多结构域多功能的大分子相对质量蛋白, 除去催化结构域, 其他结构域的结构和功能尚有许多未知. 因此, 对 H3K36 组蛋白赖氨酸甲基转移酶修饰非组蛋白底物的机制研究, 以及对其非催化结构域的结构与功能研究是该领域有待深入研究的重要方向. 此外, 这些酶的异常与人类的多种癌症相关, 例如对 NSD 家族蛋白的研究表明, 催化活性增强的突变是该家族蛋白致癌的重要原因, 因此靶向这些酶的催化结构域或者靶向其他结构域但是能够影响酶的活性或稳定性的药物研发仍将是该领域研究的热点.

4) H3K36 组蛋白赖氨酸甲基转移酶在体内受到多种因子的调控, 对这些酶与其调节因子相互作用的分子机制研究, 以及对调节因子调控机制的研究将会极大地促进我们深入理解组蛋白 H3K36 甲基转移酶的分子机制及生理功能, 并为靶向这些蛋白相关的疾病提供更多的解决方案.

5 参考文献

- [1] LUGER K, MÄDER A W, RICHMOND R K, et al. Crystal structure of the nucleosome core particle at 2.8 Å resolution[J]. *Nature*, 1997, 389(6648): 251

- [2] JENUWEIN T, ALLIS C D. Translating the histone code[J]. *Science*, 2001, 293(5532): 1074
- [3] HUSMANN D, GOZANI O. Histone lysine methyltransferases in biology and disease[J]. *Nature Structural & Molecular Biology*, 2019, 26(10): 880
- [4] MCDANIEL S L, STRAHL B D. Shaping the cellular landscape with Set2/SETD2 methylation[J]. *Cellular and Molecular Life Sciences*, 2017, 74(18): 3317
- [5] YI X, JIANG X J, FANG Z M. Histone methyltransferase SMYD2: ubiquitous regulator of disease[J]. *Clinical Epigenetics*, 2019, 11(1): 112
- [6] SHAHEEN M, WILLIAMSON E, NICKOLOFF J, et al. Metnase/SETMAR: a domesticated primate transposase that enhances DNA repair, replication, and decatenation[J]. *Genetica*, 2010, 138(5): 559
- [7] HUANG J, PEREZ-BURGOS L, PLACEK B J, et al. Repression of p53 activity by Smyd2-mediated methylation[J]. *Nature*, 2006, 444(7119): 629
- [8] CARLSON S M, MOORE K E, SANKARAN S M, et al. A proteomic strategy identifies lysine methylation of splicing factor snRNP70 by the SETMAR enzyme[J]. *Journal of Biological Chemistry*, 2015, 290(19): 12040
- [9] STRAHL B D, GRANT P A, BRIGGS S D, et al. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression[J]. *Molecular and Cellular Biology*, 2002, 22(5): 1298
- [10] EDMUNDS J W, MAHADEVAN L C, CLAYTON A L. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation[J]. *The EMBO Journal*, 2008, 27(2): 406
- [11] LI J, AHN J H, WANG G G. Understanding histone H3 lysine 36 methylation and its deregulation in disease[J]. *Cellular and Molecular Life Sciences*, 2019, 76(15): 2899
- [12] SUN X J, WEI J, WU X Y, et al. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase[J]. *Journal of Biological Chemistry*, 2005, 280(42): 35261
- [13] WAGNER E J, CARPENTER P B. Understanding the language of Lys36 methylation at histone H3[J]. *Nature Reviews Molecular Cell Biology*, 2012, 13(2): 115
- [14] PARK I Y, POWELL R T, TRIPATHI D N, et al. Dual chromatin and cytoskeletal remodeling by SETD2[J]. *Cell*, 2016, 166(4): 950
- [15] CHEN K, LIU J, LIU S, et al. Methyltransferase SETD2-mediated methylation of STAT1 is critical for interferon antiviral activity[J]. *Cell*, 2017, 170(3): 492
- [16] SUDOL M, CHEN H I, BOUGERET C, et al. Characterization of a novel protein-binding module: the WW domain[J]. *FEBS Letters*, 1995, 369(1): 67
- [17] FABER P W, BARNES G T, SRINIDHI J, et al. Huntingtin interacts with a family of WW domain proteins[J]. *Human Molecular Genetics*, 1998, 7(9): 1463
- [18] XIE P, TIAN C Y, AN L G, et al. Histone methyltransferase protein SETD2 interacts with p53 and selectively regulates its downstream genes[J]. *Cellular Signalling*, 2008, 20(9): 1671
- [19] KIZER K O, PHATNANI H P, SHIBATA Y, et al. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation[J]. *Molecular and Cellular Biology*, 2005, 25(8): 3305
- [20] LI J X, MOAZED D, GYGI S P. Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation[J]. *The Journal of Biological Chemistry*, 2002, 277(51): 49383
- [21] ZHENG W H, IBÁÑEZ G, WU H, et al. Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2[J]. *Journal of the American Chemical Society*, 2012, 134(43): 18004
- [22] YANG S, ZHENG X D, LU C, et al. Molecular basis for oncohistone H3 recognition by SETD2 methyltransferase[J]. *Genes & Development*, 2016, 30(14): 1611
- [23] BILOKAPIC S, HALIC M. Nucleosome and ubiquitin position Set2 to methylate H3K36[J]. *Nature Communications*, 2019, 10: 3795
- [24] LIU Y, ZHANG Y, XUE H, et al. Cryo-EM structure of SETD2/Set2 methyltransferase bound to a nucleosome containing oncohistone mutations[J]. *Cell Discovery*, 2021, 7: 32
- [25] VLAMING H, MOLENAAR T M, VAN WELSEEM T, et al. Direct screening for chromatin status on DNA barcodes in yeast delineates the regulome of H3K79 methylation by Dot1[J]. *eLife*, 2016, 5: e18919
- [26] SCHUHMACHER M K, BELDAR S, KHELLA M S, et al. Sequence specificity analysis of the SETD2 protein lysine methyltransferase and discovery of a SETD2 super-substrate[J]. *Communications Biology*, 2020, 3: 511
- [27] WANG G G, CAI L, PASILLAS M P, et al. NUP98-NSD1 links H3K36 methylation to *Hox-A* gene activation and leukaemogenesis[J]. *Nature Cell Biology*, 2007, 9(7): 804
- [28] JAJU R J, FIDLER C, HAAS O A, et al. A novel gene, *NSD1*, is fused to *NUP98* in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia[J]. *Blood*, 2001, 98(4): 1264
- [29] ROSATI R, LA STARZA R, VERONESE A, et al. *NUP98* is fused to the *NSD3* gene in acute myeloid leukemia

- associated with t(8;11)(p11.2;p15) [J]. *Blood*, 2002, 99(10): 3857
- [30] BRITO J L R, WALKER B, JENNER M, et al. *MMSET* deregulation affects cell cycle progression and adhesion regulons in t(4;14) myeloma plasma cells[J]. *Haematologica*, 2009, 94(1): 78
- [31] ZHANG M M, YANG Y X, ZHOU M Q, et al. Histone and DNA binding ability studies of the NSD subfamily of PWWP domains[J]. *Biochemical and Biophysical Research Communications*, 2021, 569: 199
- [32] WU H, ZENG H, LAM R, et al. Structural and histone binding ability characterizations of human PWWP domains[J]. *PLoS One*, 2011, 6(6): e18919
- [33] HUANG Z, WU H P, CHUAI S, et al. NSD2 is recruited through its PHD domain to oncogenic gene loci to drive multiple myeloma[J]. *Cancer Research*, 2013, 73(20): 6277
- [34] HE C, LI F D, ZHANG J H, et al. The methyltransferase NSD3 has chromatin-binding motifs, PHD5-C5HCH, that are distinct from other NSD (nuclear receptor SET domain) family members in their histone H3 recognition[J]. *The Journal of Biological Chemistry*, 2013, 288(7): 4692
- [35] QIAO Q, LI Y, CHEN Z, et al. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation[J]. *Journal of Biological Chemistry*, 2011, 286(10): 8361
- [36] TISI D, CHIARPARIN E, TAMANINI E, et al. Structure of the epigenetic oncogene *MMSET* and inhibition by *N*-alkyl sinefungin derivatives[J]. *ACS Chemical Biology*, 2016, 11(11): 3093
- [37] MORRISON M J, BORIACK-SJODIN P A, SWINGER K K, et al. Identification of a peptide inhibitor for the histone methyltransferase WHSC1[J]. *PLoS One*, 2018, 13(5): e0197082
- [38] LI Y, TROJER P, XU C F, et al. The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate[J]. *Journal of Biological Chemistry*, 2009, 284(49): 34283
- [39] LI W, TIAN W, YUAN G, et al. Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases[J]. *Nature*, 2021, 590(7846): 498
- [40] SATO K, KUMAR A, HAMADA K, et al. Structural basis of the regulation of the normal and oncogenic methylation of nucleosomal histone H3 Lys36 by NSD2[J]. *Nature Communications*, 2021, 12: 6605
- [41] YUAN G, MA B, YUAN W, et al. Histone H2A ubiquitination inhibits the enzymatic activity of H3 lysine 36 methyltransferases[J]. *Journal of Biological Chemistry*, 2013, 288(43): 30832
- [42] ALLALI-HASSANI A, KUZNETSOVA E, HAJIAN T, et al. A basic post-SET extension of NSDs is essential for nucleosome binding in vitro[J]. *Journal of Biomolecular Screening*, 2014, 19(6): 928
- [43] DORIGHI K M, TAMKUN J W. The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in *Drosophila*[J]. *Development (Cambridge, England)*, 2013, 140(20): 4182
- [44] SCHMITGES F W, PRUSTY A B, FATY M, et al. Histone methylation by PRC2 is inhibited by active chromatin marks[J]. *Molecular Cell*, 2011, 42(3): 330
- [45] ZHU L, LI Q, WONG S H K, et al. ASH1L links histone H3 lysine 36 dimethylation to MLL leukemia[J]. *Cancer Discovery*, 2016, 6(7): 770
- [46] AN S, YEO K J, JEON Y H, et al. Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implications for the regulatory mechanism[J]. *The Journal of Biological Chemistry*, 2011, 286(10): 8369
- [47] ROGAWSKI D S, NDOJ J, CHO H J, et al. Two loops undergoing concerted dynamics regulate the activity of the ASH1L histone methyltransferase[J]. *Biochemistry*, 2015, 54(35): 5401
- [48] HUANG C, YANG F, ZHANG Z, et al. Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 Trithorax group protein function in *Drosophila*[J]. *Nature Communications*, 2017, 8: 1649
- [49] SCHMÄHLING S, MEILER A, LEE Y, et al. Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC[J]. *Development (Cambridge, England)*, 2018, 145(7): dev163808
- [50] LEE Y, YOON E, CHO S, et al. Structural basis of MRG15-mediated activation of the ASH1L histone methyltransferase by releasing an autoinhibitory loop[J]. *Structure*, 2019, 27(5): 846
- [51] HOU P N, HUANG C, LIU C P, et al. Structural insights into stimulation of Ash1L's H3K36 methyltransferase activity through Mrg15 binding[J]. *Structure*, 2019, 27(5): 837

Histone substrate recognition by histone H3K36 lysine methyltransferases

CAO Yinghua WANG Zhanxin

(Key Laboratory of Cell Proliferation and Regulation Biology of Ministry of Education, College of Life Sciences, Beijing Normal University, 100875, Beijing, China)

Abstract Histone H3 methylation at Lys36 is widespread on chromatin and is closely related to active transcription, DNA damage and repair. H3K36 can be methylated into mono-, di- and tri-methylation states. Currently, SETD2 is the only known human methyltransferase responsible for H3K36 tri-methylation, while NSD1, NSD2, NSD3 and ASH1L are responsible for H3K36 di-methylation. H3K36 methyltransferases are very specific for substrate histones, and thus are important for the distribution and methylation status of H3K36. In addition, their aberrant expressions are closely related to human diseases. Therefore, deciphering substrate recognition and catalysis mechanisms of these H3K36 methyltransferases are important to understand the epigenetic regulation mechanisms and the physiological roles of these enzymes *in vivo*. Previous studies on H3K36 methyltransferases greatly facilitated the understanding of their catalytic mechanisms. However, due to limited structural data on complexes of these enzymes with their bound substrates, questions related to their histone substrate recognition remain unanswered. With recent breakthroughs in cryo-EM technology, most complex structures of these H3K36 methyltransferases with their bound nucleosomal substrates are now solved, greatly advancing the understanding of their substrate recognition and catalysis mechanisms. This review article will focus on human H3K36 methyltransferases, with a particular emphasis on the molecular mechanisms of these enzymes.

Keywords H3K36 methyltransferase; SETD2; NSD family proteins; ASH1L; nucleosomal substrate

【责任编辑: 武 佳】