



天然无序区域在Ash1/Ash1L组蛋白甲基转移酶 活化过程中的调控作用*

杨静熊俊文 黄永棋** (湖北工业大学生物工程系,武汉 430068)

摘要 组蛋白甲基化修饰与基因的转录调控密切相关,其中果蝇的Ash1以及哺乳动物的同源蛋白Ash1L是催化组蛋白H3上36位赖氨酸进行单甲基化和双甲基化的甲基转移酶.由于存在一个自抑制环区阻挡了底物的结合位点,Ash1/Ash1L本身的组蛋白甲基转移酶活性是很低的.但与Mrg15结合后,Ash1/Ash1L从原来的自抑制态转变为活化态.最近,Ash1L与Mrg15结合后复合物的晶体结构被成功解析出来,使之可以揭示Ash1L与Mrg15之间的特异相互作用以及Mrg15激活Ash1L的机理.我们通过比较Ash1L/Mrg15复合物的两个晶体结构来讨论Ash1L分子中的天然无序区域在其活化过程中所起的重要调控作用.

关键词 组蛋白甲基转移酶, Ash1, Mrg15, 天然无序蛋白质, 构象转化 中图分类号 Q518.2, Q518.4 **DOI**: 10.16476/j.pibb.2019.0109

真核生物基因的表达受到多种机制的调控,其 中一种机制是通过对核小体组蛋白的翻译后修饰进 行的. 核小体是染色体的基本组成单元,由146个 碱基对的双链 DNA 片段缠绕在组蛋白八聚体上而 形成. 在核小体的组蛋白八聚体中分别含有两个 H2A、H2B、H3和H4. 这些组蛋白的N端区域并没 有被 DNA 缠绕, 而是伸展在核小体外面, 因此表 现出高度的结构柔性. 组蛋白的这些N端区域还可 以进行多种翻译后修饰,如甲基化、乙酰化、磷酸 化和泛素化等. 在这些翻译后修饰中, 组蛋白H3 的甲基化修饰与基因的表达调控密切相关,研究表 明,阻断组蛋白H3的9号位赖氨酸和27号位赖氨 酸的甲基化使得着丝粒功能异常,导致染色体在有 丝分裂的过程中无法进行正常分离[1-2]. 另外,组 蛋白H3的4号位赖氨酸和36号位赖氨酸的甲基化 修饰在转录激活中起着重要作用[3].

组蛋白H3的甲基化修饰可以由多种组蛋白甲基转移酶来催化完成,其中果蝇Ash1及其哺乳动物同源蛋白Ash1L是催化组蛋白H3中第36号位赖氨酸修饰的甲基转移酶^[48]. Ash1/Ash1L是典型的含有多个有序结构域和无序区域的蛋白质. 尽管早

期生物化学分析和截短实验表明 Ash1/Ash1L的催化单元是有序的 SET 结构域,但是单独的 Ash1/Ash1L或者 SET结构域的甲基转移酶活性却是非常低 [4]. 最近,Lee 等 [9] 和 Hou 等 [10] 两个研究团队分别解析了 Ash1L 与其调控因子 Mrg15 复合物的晶体结构. 他们的结果清晰地展示了 Mrg15 与 Ash1L的结合作用,以及 Ash1L 从抑制态转变为活化态过程中发生的局部"有序-无序"的结构变化.

在本文中,我们将以天然无序蛋白质为主要背景,在对Lee等^[9]和Hou等^[10]研究成果进行分析评述的基础上,阐述天然无序结构区域在调控Ashl/AshlL组蛋白甲基转移酶活性的过程中所起的作用。

1 天然无序蛋白质

通常,人们认为蛋白质只有在折叠成天然三维 结构后才能发挥其生理功能.这就是蛋白质的"序

Tel:15207114174, E-mail: yqhuang@hbut.edu.cn 收稿日期: 2019-07-08, 接受日期: 2019-08-02

^{*} 国家自然科学基金资助项目(21603121).

^{**} 诵讯联系人.

列-结构-功能"范式.这个范式对蛋白质研究起着重要的指导作用,尤其是在蛋白质工程、药物设计和酶学研究等方面.然而,随着对蛋白质结构表征数据的积累,人们开始发现有些蛋白质并不具有一个确定的三维结构:如在X射线晶体衍射图中,蛋白质的有些区域缺失电子密度;在远紫外圆二色对蛋白质的二级结构进行表征时有些蛋白质含有大量的无规卷曲;在核磁共振谱图中,蛋白质的有些区域所对应的化学位移严重重叠在一起.现在,人们把这一类与传统认识不同的新蛋白质类型称为天然无序蛋白质(intrinsically disordered proteins)[11-12].

1.1 天然无序蛋白质的序列与结构特点

对经过实验表征确认的天然无序蛋白质进行氨基酸序列分析,可以发现天然无序蛋白质在氨基酸组成上具有显著的分布特征.相比于传统有序蛋白质,天然无序蛋白质疏水残基(尤其是芳香性氨基酸残基)的含量较低,而带电残基的含量较高^[13].因此,天然无序蛋白质具有较弱的折叠驱动力,无法形成或者维持一个稳定的疏水结构.在疏水性-净电荷图上,天然无序蛋白质分布于疏水性较低、电荷较高的区域.除了疏水性和电性的特点外,天然无序蛋白质还含有更多的倾向于形成无序结构的残基^[13].基于天然无序蛋白质氨基酸序列上的特征,人们已经开发出了数十种预测方法,可以对不同物种中、不同功能中天然无序蛋白质的含量进行系统的分析^[14].

虽然天然无序蛋白质在整体上不具有一个确定的三维结构,但它们却可以形成一些局部的二级结构,如α螺旋.这些局部的二级结构在天然无序蛋白质的分子识别过程中可能起着重要的作用,如启动结合过程的发生,它们往往被称为分子识别基元 (molecular recognition feature) [15].

尽管单独存在时是无序的,但很多天然无序蛋白质与靶标结合后会折叠成一个有序的结构.这个过程一般被称为折叠与结合的耦合(coupled folding and binding)^[16].如转录因子环腺苷酸应答因子结合蛋白CREB中的激酶诱导结构域KID与靶标KIX结构域结合后,折叠成两段α螺旋.尽管天然无序蛋白质与靶标结合后变得更加有序,但这并不意味着整个天然无序蛋白质都发生了折叠.在很多时候,除了参与结合的局部区域外,天然无序蛋白质上的其他区域仍然会保存无序的状态.

1.2 天然无序蛋白质的功能

利用天然无序蛋白质预测方法对基因组进行生

物信息学分析发现,各种生物体内都普遍存在天然无序蛋白质,如天然无序蛋白质在古细菌、细菌和真核生物中的含量分别为2.0%、4.2%和33.0% [17]. 进一步对这些天然无序蛋白质进行功能分析,发现天然无序蛋白质广泛参与了基因的转录调控、蛋白质的翻译后修饰、细胞信号转导以及细胞中无膜结构的组装等重要生理过程 [18]. 由于天然无序蛋白质广泛参与了重要的生物学过程,因此天然无序蛋白质功能的异常也往往与众多疾病,包括癌症、神经退行性疾病和心血管疾病等的发生具有密切的关联 [18-22].

2 天然无序区域对Ash1/Ash1L组蛋白甲基 转移酶活性的调控作用

2.1 Ash1/Ash1L组蛋白甲基转移酶

组蛋白甲基转移酶负责催化组蛋白的甲基化修 饰,一般它们都以SET结构域作为催化活性结构 域.有些组蛋白甲基转移酶上还含有其他一些调控 区域,对调控组蛋白甲基转移酶催化活性具有重要 作用. 高等真核生物中专门对组蛋白H3的第36号 位赖氨酸进行甲基化修饰的甲基转移酶有哺乳动物 的 Ash1L、NSD1/2/3 和 SETD2 以及果蝇的同源蛋 白 Ash1、Nsd 和 Set2 [23]. Ash1/Ash1L 是一个三空 腔结构蛋白质 (trithorax-group protein), 它通过对 组蛋白H3中第36号位赖氨酸进行甲基化修饰来调 控染色体的转录活性,从而在发育过程起着抵抗 Hox 基因多梳沉默 (polycomb silencing) 的作 用^[24]. 因此, 突变的 Ash1 基因可以导致果蝇在发 育过程中发生Hox基因表达的缺陷以及同源异形转 化. Ash1/Ash1L蛋白分子中含有多个有序结构域和 无序区域.生物化学分析和截短实验表明,有序的 SET 结构域是 Ash1/Ash1L 执行组蛋白甲基转移酶 功能的主要活性单元,而相邻的AWS结构域和PS 结构域则对 SET 的催化活性起着调控作用 [4]. 但 是,在单独存在的情况下,Ash1/Ash1L的甲基转 移酶活性是非常低的. 这是因为在SET结构域的C 端存在一个postSET片段,其上面有一个环区(AI 环)可以与SET结构域中的底物结合位点进行结 合,从而阻挡了底物组蛋白质H3与Ash1/Ash1L的 结合[4].

2.2 天然无序区域对Ash1/Ash1L活性的调控

由于单独存在时 Ash1/Ash1L 的甲基转移酶活性是非常低的,因此 Ash1/Ash1L 甲基转移酶活性的发挥需要外来因素的调控.确实,人们已经发现

Ash1/Ash1L可以和多个其他蛋白质发生相互作用,其中与Mrg15结合可以使得Ash1/Ash1L从非活性状态转变为活性状态^[25-26].进一步的研究表明,Ash1/Ash1L与Mrg15之间的相互作用,主要是通过一段位于Ash1/Ash1L蛋白AWS结构域N端的含有FxLP保守序列的天然无序区域和Mrg15蛋白C

端的 MRG 结构域介导实现的(图 1a)[25-26]. 但是,Mrg15激活 Ash1/Ash1L 的分子机理却一直不清楚. 最近, Ash1L/Mrg15 复合物晶体结构的成功解析 [9-10],使得人们可以在原子水平上认识 Ash1L 与Mrg15之间的特异识别作用,为最终阐明 Mrg15激活 Ash1L 的分子机理提供了可能.

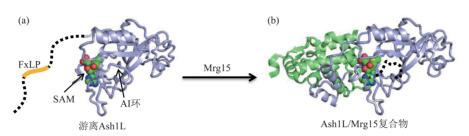


Fig. 1 Structural disorder-order transition of Ash1L upon Mrg15 binding 图1 Ash1L与Mrg15结合后发生的结构上无序-有序的转变

Ash1L蛋白以浅蓝色显示; Mrg15蛋白以浅绿色显示; Ash1L上无序的区域用虚线表示. 蛋白质结构数据库编号分别为: (a) Ash1L 3OPE, (b) Ash1L/Mrg15复合物6INE.

对 Ash1L/Mrg15 复合物的结构进行分析可以发现, AshL 主要是通过 FxLP 保守序列及其两端的邻

近区域(下文中简称为MDR)与Mrg15进行结合(图1,图2),与Mrg15的疏水凹槽形成了广泛的

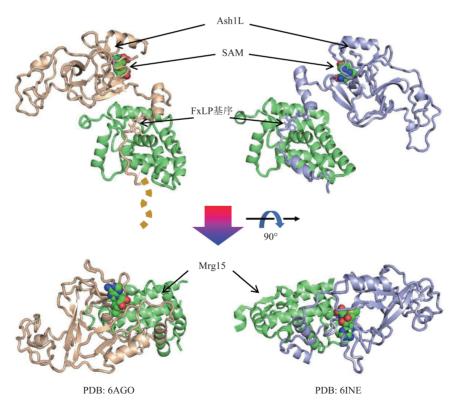


Fig. 2 Comparison of the two Ash1L/Mrg15 complex structures 图2 Ash1L/Mrg15复合物两种结构的比较

通过对Mrg15进行比对将这两个结构进行了叠合. Mrg15分子用浅绿色显示; PDB 6AGO中的Ash1L分子用浅橙色显示; 其中, Ash1L N端的 无序区域用虚线显示; PDB 6INE中的Ash1L分子用浅蓝色显示.

疏水相互作用. 但是我们注意到, MDR在两个复合 物中所形成的结构是不完全相同的 (图2): MDR 的C端在6AGO中形成了一个3圈的α螺旋,在 6INE中则形成了一个4圈的α螺旋; MDR的N端 在6AGO中形成了无规卷曲以及部分的无序结构, 在6INE中则形成了一个α螺旋(图2). 因为MDR 在两个复合物中的结构存在较大的差异并存在部分 无序的结构,我们推测 MDR 的 N 端在结合前很可 能是无序的. 因此, Ash1L蛋白上的MDR区域可能 是通过进行"折叠-结合耦合"的方式来实现 Ash1L与调控因子Mrg15之间的特异性结合.但是, 由于Mrg15的MRG结构域并没有与Ash1L的SET 结构域发生明显的结合作用, 因此FxLP保守序列 与 Mrg15 的结合并不是引起 Ash1L 活化的直接原 因. 通过对 Ash1L SET 结构域与 Mrg15 结合前后的 结构进行比较分析,可以发现 SET 结构域整体上并 没有发生明显的结构变化(图1).唯一一点主要 的变化在AI环区. AshIL单独存在时, AI环区的结 构是有序的,占据了底物的结合口袋(图1a).但 是,与Mrg15结合后,AI环区则变成了无序的, 底物的结合口袋释放出来(图1b).因此,通过这 种局部"有序-无序"的结构变化, SET结构域从 抑制状态转变为活性状态.

虽然 AI 环区的无序化可以解释 Ash1L 抑制状态的解除,但 AI 环为什么在 Mrg15 结合后变得无序了呢?目前对这个问题还没有明确的答案.因为

Mrg15主要是与 Ash1L 上含有 FxLP 保守序列的无 序区域结合的, 其与SET结构域尤其是AI环区之 间并没有直接的接触作用. 那么这种变构效应是如 何传递的呢? Lee 等 [9] 认为, Mrg15与Ash1L结合 后使得H2193的取向发生变化以及Y2207变得无 序,削弱了AI环与SET结构域的作用(图3). Hou 等 [10] 则认为 Trp2152 和 Gln2266 的构象在结合 后发生的变化是引起AI环无序的主要原因(图 3). 比较 SAM 在 Ash1L 和在 AshL/Mrg15 复合物中 的结构还可以发现, SAM 在这些结构中的具体位 置及其与周边残基的相互作用也存在细微的差别 (图3). 因此也有可能是Mrg15结合后引起SAM结 合位点结构发生变化,从而对AI环的相互作用产 生扰动, 使其无序. 另外, 通过比较 Ash1L/Mrg15 复合物的两个晶体结构, 我们发现, SET结构域相 对于Mrg15分子的取向在复合物的两个结构中是完 全相反的, 二者刚好发生了180°的扭转(图2). 目前我们并不能确定这两种构象产生的原因.有可 能这两种构象都是Ash1L/Mrg15复合物的活性构 象,但也不能排除是在结晶过程中晶体堆积使得 SET结构域的取向发生了扭转.不管是哪一种情况, 复合物的这两个结构都表明,在Ash1L与Mrg15结 合后,SET结构域相对于Mrg15来说是具有一定的 结构活动自由的,只有这样才能形成这两种不同的 构象.

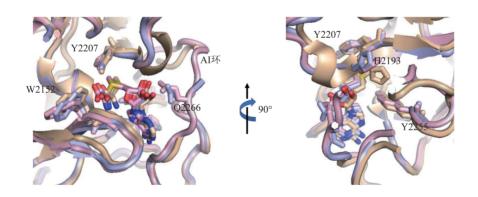


Fig. 3 Comparison of SAM binding pocket in the Ash1L and Ash1L/Mrg15 structures
图3 Ash1L和AshL/Mrg15复合物中SAM结合口袋的比较

SAM分子用不同颜色显示了其元素组成. PDB 6AGO中的Ash1L分子用浅橙色显示; PDB 6INE中的Ash1L分子用浅蓝色显示; PDB 3OPE中的Ash1L分子用浅紫红色显示.

3 研究展望

综上所述,通过解析Ash1L/Mrg15复合物的晶 体结构, Lee 等 [9] 和 Hou 等 [10] 的研究工作为阐明 Mrg15激活 Ash1/Ash1L 的分子机理提供了非常重 要的信息.结构分析表明Ash1/Ash1L分子中的天然 无序区域在调控其组蛋白甲基转移酶活性的过程中 起着关键作用,包括: a. 含有FxLP保守序列的区 域通过"折叠-结合耦合"的方式实现了与激活因 子Mrg15的特异性识别; b. AI环区局部的"有序-无序"结构转变释放出了SET结构域中的底物结 合口袋; c. SET结构域与Mrg15分子间的动态相互 作用可能在控制 Ash1L 的活性中起着调控作用. 在 未来的研究中,通过单分子荧光共振能量转移 (single-molecule Förster resonance energy transfer) 和顺磁驰豫增强(paramagnetic relaxation enhancement)等技术将可以进一步阐明 Ash1L/ Mrg15之间的动态相互作用.

参考文献

- [1] Mellone B G, Ball L, Suka N, *et al.* Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. Curr Biol, 2003, **13**(20): 1748-1757
- [2] Pengelly A R, Copur O, Jackle H, *et al.* A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. Science, 2013, **339**(6120): 698-699
- [3] Kingston R E, Tamkun J W. Transcriptional regulation by trithorax-group proteins. Cold Spring Harb Perspect Biol, 2014, 6(10): a019349
- [4] 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 Biol Chem, 2011, 286(10): 8369-8374
- [5] Dorighi K M, Tamkun J W. The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila. Development, 2013, 140(20): 4182-4192
- [6] Miyazaki H, Higashimoto K, Yada Y, et al. Ash11 methylates Lys36 of histone H3 independently of transcriptional elongation to counteract polycomb silencing. PLoS Genet, 2013, 9(11): e1003897
- [7] Tanaka Y, Katagiri Z, Kawahashi K, et al. Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene, 2007, 397(1-2): 161-168
- [8] Yuan W, Xu M, Huang C, et al. H3K36 methylation antagonizes

- PRC2-mediated H3K27 methylation. J Biol Chem, 2011, **286**(10): 7983-7989
- [9] Lee Y, Yoon E, Cho S, *et al.* Structural basis of MRG15-mediated activation of the ASH1L histone methyltransferase by releasing an autoinhibitory loop. Structure, 2019, **27**(5): 846-852
- [10] Hou P, Huang C, Liu C P, et al. Structural insights into stimulation of Ash1L's H3K36 methyltransferase activity through Mrg15 binding. Structure, 2019, 27(5): 837-845
- [11] Habchi J, Tompa P, Longhi S, et al. Introducing protein intrinsic disorder. Chem Rev, 2014, 114(13): 6561-6588
- [12] Uversky V N. Natively unfolded proteins: a point where biology waits for physics. Protein Sci, 2002, 11(4): 739-756
- [13] Tompa P. Intrinsically unstructured proteins. Trends Biochem Sci, 2002, **27**(10): 527-533
- [14] Nielsen J T, Mulder F A A. Quality and bias of protein disorder predictors. Sci Rep, 2019, 9(1): 5137
- [15] Mohan A, Oldfield C J, Radivojac P, et al. Analysis of molecular recognition features (MoRFs). J Mol Biol, 2006, 362(5): 1043-1059
- [16] Wright P E, Dyson H J. Linking folding and binding. Curr Opin Struct Biol, 2009, 19(1): 31-38
- [17] Dunker A K, Lawson J D, Brown C J, et al. Intrinsically disordered protein. J Mol Graphics Modell, 2001, 19(1): 26-59
- [18] Wright P E, Dyson H J. Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Biol, 2015, 16(1): 18-29
- [19] Uversky V N, Oldfield C J, Dunker A K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu Rev Biophys, 2008, 37(1): 215-246
- [20] Uyar B, Weatheritt R J, Dinkel H, *et al.* Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer? Mol Biosyst, 2014, **10**(10): 2626-2642
- [21] Babu M M. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. Biochem Soc Trans, 2016, 44(5): 1185-1200
- [22] Kulkarni P, Uversky V N. Intrinsically disordered proteins in chronic diseases. Biomolecules, 2019, 9(4): 147
- [23] Wagner E J, Carpenter P B. Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol, 2012, 13(2): 115-126
- [24] Papp B, Muller J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev, 2006, 20(15): 2041-2054
- [25] Huang C, Yang F, Zhang Z, et al. Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 Trithorax group protein function in Drosophila. Nat Commun, 2017, 8(1): 1649
- [26] Schmahling S, Meiler A, Lee Y, et al. Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. Development, 2018, 145(7): 2041-2054