泛素化核小体复合物的结构机制研究进展

徐子昱, 艾华松, 孙茂铤, 田家昆, 刘磊*

清华大学化学系, 清华-北大生命科学联合中心, 生命有机磷化学与化学生物学教育部重点实验室, 北京 100084

*通讯作者, E-mail: liu@mails.tsinghua.edu.cn

收稿日期: 2019-12-29; 接受日期: 2020-02-19; 网络版发表日期: 2020-03-09

国家重点研发计划(编号: 2017YFA0505200)和国家自然科学基金(编号: 91753205, 21761142008, 21532004)资助项目

摘要 核小体组蛋白(H2A、H2B、H3、H4)的泛素化修饰介导了一系列效应蛋白与核小体发生特异性的相互作用, 并调控细胞的表现遗传。近年来得益于一些新的化学方法与生物方法的发展, 如蛋白质化学修饰等样品获取技术以及冷冻电子显微镜等结构观测技术, 一些泛素化核小体与效应蛋白的复合物结构得以被解析, 揭示了泛素调控染色质表现遗传过程的分子机制。本文对近期报道的六种泛素化核小体与效应蛋白的复合物进行了归纳, 对比分析了泛素化核小体对效应蛋白的特异性识别与活化机制, 并总结了组蛋白上的关键相互作用位点。

关键词 翻译后修饰, 泛素, 核小体, 结构机制, 染色质

1 引言

真核生物细胞中遗传物质以DNA形式集中在细胞核的染色质中。染色质是一种高度压缩的DNA和蛋白质的复合物, 以单核小体(nucleosome core particle, NCP)作为基本结构单元。核小体之间通过DNA双链和连接组蛋白进行连接和有序排列^[1]。核小体相关的分子机制研究从20世纪70年代开始。伴随观测技术手段的不断进步, 核小体相关各层级结构逐步得到了更加精确的解析。1997年, Luger等^[2]通过晶体衍射技术获得了分辨率为2.8 Å的单核小体晶体结构, 如图1所示。

单核小体中心是组蛋白八聚体, 由两个H2A/H2B组蛋白二聚体和一个(H3-H4)₂组蛋白四聚体构成, 单核小体之间则由组蛋白H1连接。多聚体间的组蛋白凭借H4和H2B之间的 α -螺旋相互作用(图1(b))。在组蛋白

盘的侧面, 缠绕着约1.75圈、总长度约145 bp的DNA(图1(a, c))。单核小体拥有一个二重旋转轴, 是标记核小体DNA碱基对位置的基准点。取二重旋转轴指向的DNA位置作为SHL0, 此后取每10个碱基对为1个SHL单位, 将缠绕DNA的位置用SHL-7~7表示(图1(d))。

核小体在染色质相关的各级调控过程中发挥着关键作用, 以DNA缠绕方式调节染色质一级结构, 还能作为体内信号传导的中枢来招募信号分子, 对遗传信息的复制和转录过程起到重要的调控作用^[3,4]。此外, 核小体可以自组装, 排列形成染色质二级结构。真核细胞的染色质相关行为, 如转录等均涉及染色质二级结构变化^[5,6]。

核小体调控基因表达等生理过程可以通过组蛋白翻译后修饰来实现。组蛋白具有多个赖氨酸残基, 这些赖氨酸可以被亲电试剂进行基团修饰, 这些修饰包括

引用格式: Xu Z, Ai H, Sun M, Tian J, Liu L. Advances on structural mechanisms of ubiquitinated nucleosome complexes. *Sci Sin Chim*, 2020, 50: 415–436, doi: 10.1360/SSC-2019-0170

图1 单核小体结构(PDB: 1kx5). (a) 从核小体盘面的垂线向下观察; (b) H2B与H4组蛋白之间通过四重 α -螺旋相互作用; (c) 沿核小体双重轴观察; (d) 核小体DNA的SHL位置. 灰色: DNA; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4 (网络版彩图)

Figure 1 Structure of NCP. (a) Viewed down the DNA superhelical axis; (b) four helix-bundle formed by histone H2B and H4; (c) viewed from the dyad axis; (d) SHL location representation of nucleosomal DNA. Gray: DNA; orange: H2A; paleyellow: H2B; green: H3; blue: H4 (color online).

甲基化、乙酰化和泛素化等. 其中泛素化修饰大约占据总修饰的16%, 是一种在组蛋白上广泛存在且与染色质的各种活动均紧密相关的修饰^[7]. 泛素分子是一种高度保守的由76个氨基酸组成的蛋白, 分子量大约为8.5 kD, 其结构如图2(a)所示. 由于泛素分子的大体积, 相比于大部分的小基团修饰, 泛素分子可通过多个相互作用界面与其他蛋白相互作用. 例如, 它可以用两个疏水界面(分别以Ile36和Ile44为中心)招募一些蛋白, 其碳端也常为去泛素化酶(deubiquitinating enzymes, DUB)识别的位点(图2(b))^[8,9].

在真核生物中, 泛素的主要功能是通过E1/E2/E3三类泛素化酶的催化, 以共价键连接到蛋白质上^[10]. 泛素标记的蛋白质被运送到蛋白酶体, 导致蛋白质的分解. 除此之外, 泛素还能够作为胞吞的信号分子对胞吞进行控制, 或者通过改变构象的方式控制泛素连接蛋白的活性^[11]. 而在核小体组蛋白上的泛素化修饰主要为单泛素修饰, 不均匀地分布在组蛋白上, 其中H2A和H2B组蛋白上的泛素化修饰占绝大部分. 在脊

图2 泛素(PDB:1ubq)的整体结构与主要结合位点. 绿色: 碳端; 青色: 144疏水口袋; 黄色: 136疏水口袋(网络版彩图)

Figure 2 Overall structure and binding hotspots on ubiquitin. Green: C-terminal; cyan: 144-centered hydrophobic patch; yellow: 136-centered hydrophobic patch (color online).

椎动物中, 约有11%的H2A和1%~1.5%的H2B组蛋白被泛素化修饰, 而酵母中则有约10%的H2B组蛋白被泛素化修饰. 相比之下, 在HeLa细胞中仅有0.3%的H3组蛋白和不到0.1%的H4组蛋白被泛素化修饰^[12-14].

这些组蛋白上不同位点的泛素化修饰具有不同的生理功能. 丰度最高的两个修饰位点分别是H2AK119和H2BK120, 其中H2AK119的泛素化修饰导致转录抑制, 而H2BK120泛素化修饰倾向于促进转录等过程^[15,16]. 除此之外, 其他位点上的泛素化修饰也具有不同的生理作用. 例如, H2AK13/15泛素化修饰参与DNA断裂修复过程, H3K23泛素化修饰是DNA复制中维持DNA甲基化的必要条件, H4K91泛素化修饰则在DNA双链断裂修复过程中起到保护作用^[17-19]. 核小体泛素化修饰也可以调控核小体上其他位点的修饰活性, 例如H2BK120ub能够通过促进H3K4和H3K79的甲基化来调节基因激活或沉默, 而H2AK15ub和H4K20双甲基化修饰共同促进DNA双链断裂修复^[20-22]. 在核小体上, 一种修饰参与调控核小体另一修饰的过程称为修饰间的串扰(cross-talk)^[23].

不仅如此, 泛素还可在核小体动态结构水平上影响染色质, 如H2BK34泛素化能够通过空间因素促进核小体的解离^[24]. 核小体的泛素化修饰也能影响相邻与核小体之间的相互作用, 从而调节染色质的高级结构^[25,26]. 大量生化实验表明组蛋白修饰与相关效应蛋白(writers、erasers、readers)之间存在特异性的识别与激活机制. 解析修饰核小体与效应蛋白复合物的原子或近原子分辨率结构, 能够揭示这些重要生命化学过程的分子机制, 加深理解生命的化学本质, 并对发展化学活性干预分子用于疾病诊疗提供知识基础.

伴随体外化学合成泛素化组蛋白类似物和组装核

小体新方法的发展以及冷冻电子显微镜(Cryo-EM)等结构观测技术, 近期研究者对一系列重要的泛素化核小体与效应蛋白复合物开展了研究. 这些复合物包括: SAGA复合物中负责去泛素化修饰的DUB结构域与H2BK120ub核小体之间的复合物^[27]; DNA双链断裂修复过程中53BP1和H2AK15ub、H4K20me2双修饰核小体形成的复合物^[28]; 转录延长过程中染色质重塑酶Chd1与H2BK120ub核小体形成的复合物^[29]; 甲基转移酶Dot1L和COMPASS/MLL分别与H2BK120ub核小体之间形成的复合物; 甲基转移酶Set2结合H2BK120ub核小体形成的复合物^[30-38]. 本文对体外获取泛素化修饰组蛋白的方法和基于泛素化核小体的复合物结构进行了归纳, 总结分析了泛素化核小体对效应蛋白的异性识别与活化机制.

2 化学合成泛素化核小体

研究泛素化核小体复合物的结构及其活化机制时, 需要在体外重组泛素化单核小体, 再令其与招募蛋白结合形成复合物, 最后利用电子显微镜等技术观察复合物的结构. 理论上, 可以利用纯化分离真核细胞裂解液的方法得到泛素化的核小体, 但是通过这种方法提取的核小体经常同时含有甲基化等其他修饰. 细胞裂解液中的去泛素酶也会对核小体去泛素化, 现有的分离手段下得到的泛素化核小体不纯. 使用体外重组泛素化酶进行对组蛋白的酶促泛素化反应时, 由于反应的区域选择性不强, 往往产率较低^[39,40]. 为克服以上方法的局限性, 需要开发更有效率的化学合成方法进行体外环境下泛素化组蛋白的制备.

化学合成修饰组蛋白的方法包含三类: 半胱氨酸参与的蛋白修饰过程、非天然氨基酸替换、蛋白质交联^[41]. 化学合成蛋白主要分为全合成与半合成两种方法^[42]. 蛋白全合成需要将目标蛋白分为若干个比较短的多肽链, 再将固相合成的这些多肽链化学连接成一条蛋白^[43-48]. Brik等^[49]成功完成了H2BK34ub组蛋白的化学全合成(图3). 首先通过固相合成的方法合成5个H2B组蛋白片段, 并且将K34修饰位点的赖氨酸替换成一种非天然氨基酸(δ -巯基赖氨酸), 这种非天然氨基酸可以以较高的选择性便捷地与泛素分子形成肽键^[50]. 组蛋白片段之间交联后最后通过一个酰肼中间体得到完整组蛋白^[43], 在形成完整组蛋白骨架后进行

泛素化, 最后脱硫, 即可得到天然键合方式的H2BK34ub组蛋白. 此外, 诸如H2AK13/K15ub组蛋白、H2AK119ub组蛋白、H3K4me3组蛋白等也可用全合成方法获取^[51,52].

近年来, 对于单修饰(如泛素化修饰)的组蛋白的化学合成主要采用半胱氨酸和蛋白质交联共同参与的半合成(semi-synthesis)方法^[42]. 1998年开始兴起的表达蛋白连接(expressed protein ligation, EPL)的方法^[53], 是建立在自然化学连接(native chemical ligation, NCL)和酶的重组表达之上的一种半合成方法. 它通过交联化学合成的较短的多肽和重组表达的蛋白质, 能够简便地合成含有非天然氨基酸或者翻译后修饰的分子量较大的蛋白^[54]. 2008年, Muir等^[55]利用EPL方法制备了具有天然键合方式的H2BK120ub组蛋白. 制备过程主要分为三步: 将重组表达的泛素(1~75)交联到固相化学合成的短肽H2B(117-124)K120-Ub76上; 脱保护后用重组表达的H2B(1-116)与上述蛋白交联; 最后脱去巯基(图4). 该方法存在产率低、反应速率慢、交联辅基制备繁琐等缺点^[56]. 随后Muir等^[56,57]对方法进行了改进, 摒弃了繁杂的辅基制备, 简化了交联步骤, 优化了产量和效率, 得到了ubG76A替换的H2BK120ub组蛋白. 由于甘氨酸和丙氨酸的结构相似, 该方法得到的泛素化核小体复合物具有和原复合物相同的活性^[57]. 随后, Muir等^[58]利用类似方法成功制备了另一个广泛存在的泛素修饰H2AK119ub(ubG76(A))组蛋白. 除了位于组蛋白末端的泛素化修饰H2AK119ub和H2BK120ub外, 如H2BK34ub等位于组蛋白中端的泛素化修饰的获取也可通过半制备合成法获取^[59]. 近年来, 许多利用非天然交联键合方式的泛素化组蛋白交联方法已经被提出^[60,61].

由于泛素化核小体的体外重构过程并不严格要求体外重组的泛素化核小体是天然的, 因此可采用质粒表达氨基酸替换的组蛋白和泛素, 再用交联剂(如二氯丙酮等)对二者进行交联来制备带修饰的组蛋白. 目前在文献中常用的几种交联方法包括二氯丙酮交联和二硫键交联^[60,61]. 2009年, Muir等^[60]注意到二硫键容易形成和被还原的特点, 选择利用二硫键交联的方法, 成功重组了能够活化hDot1L甲基转移酶的H2BK120ub泛素化核小体的类似物, 二硫键的键长比天然异肽键长约2.4 Å (图5(b)). 他们注意到泛素分子中没有半胱氨酸, 组蛋白八聚体也仅有一个在H3上的半胱氨酸, 因

图 3 化学全合成H2BK34ub组蛋白示意图^[49] (网络版彩图)

Figure 3 Total synthesis of histone H2BK34ub [49] (color online).

图 4 表达蛋白连接合成H2BK120ub组蛋白的示意图^[55] (网络版彩图)

Figure 4 Expressed protein ligation scheme of native H2BK120ub synthesis [55] (color online).

此若将需要进行泛素化的赖氨酸位点替换为亲核性更强的半胱氨酸, 这样泛素可以选择性地结合半胱氨酸. 由于二硫键的不稳定性, 这种方法的适用范围有限. 2014年, Long等^[61]选择二氯丙酮作为交联剂, 将泛素的76位甘氨酸和组蛋白泛素化位点的赖氨酸均替换为半胱氨酸, 通过二氯丙酮交联(图5(c)). 这种方法形成的交联虽然比天然异肽交联多一根键的距离, 并且多一个羧基, 可能会导致蛋白的结合发生变化, 但是交联的化学性质稳定、不易水解, 也不对去泛素酶敏感, 是一种适合泛素化核小体结构观察的交联方法.

3 SAGA-DUB-H2BK120ub-NCP复合物

人源核小体H2B的120位赖氨酸泛素化H2BK120ub (对应酵母源H2BK123ub)是一个重要的翻译后修饰, 在RNA聚合酶II介导的转录激活过程的调控、转录延长、DNA复制和损伤修复过程中均发挥重要作用^[10,62,63]. 人源H2BK120ub的生成是由RNF20/40泛素连接酶介导的(而在酵母中对应的Bre1泛素连接酶生成H2BK123ub). 该位点泛素化修饰可以被SAGA (Spt-Ada-Gcn5 acetyltransferase), 一种含有19

个蛋白质亚基的复合物进行去泛素化^[64,65]. 该位点上的泛素化-去泛素化过程可进一步对组蛋白H3上的4位和79位赖氨酸的甲基化(H3K4me/H3K79me)过程进行调控, 或调节高级染色质的松散程度, 从而调控基因转录启动^[26,64]. 对于SAGA-核小体复合物分子结构机制的研究可以帮助揭示基因转录调控的化学原理.

SAGA是一种多功能的复合蛋白, 其催化去泛素化反应的核心模块被称为DUB区域, 包含Ubp8、Sgf11、Sus1和Sgf73四个亚基. Sgf73不具有催化活性, 但它连接了DUB结构域与SAGA其他部件, 并且调节了复合物的空间结构^[66,67]. Morgan等^[27,62]利用二氯丙酮对泛素(UbG76C)和组蛋白H2BK120C交联得到泛素化修饰的核小体. 在DUB中将Ubp8活性位点146位半胱氨酸替换为丙氨酸, 加强DUB对核小体的结合. 经过体外重组构筑了结合比为2:1的SAGA-DUB-H2BK120ub-NCP复合物并获得了晶体. 使用X射线衍射法解析了复合物的晶体结构(图6)^[27], 分辨率为3.9 Å^[27]. SAGA分别位于核小体盘面的两侧, 二重轴上远离DNA出入口的位置, 形成轴为二重轴的对称结构. 每个DUB模块与核小体盘的相互作用面积达1000 Å², 而核小体上的泛素与DUB的活性中心的接

图5 交联H2B组蛋白与泛素的方法. (a) 体内的交联过程; (b) 二氯丙酮作为交联剂的交联过程^[60]; (c) 二硫键交联过程^[61](网络版彩图)

Figure 5 Schemes of crosslinking between histone H2B and ubiquitin. (a) Native crosslinking; (b) dichloroacetone-directed crosslinking ^[60]; (c) disulfide-directed crosslinking ^[61] (color online).

触面积达 $1650 \text{ \AA}^2$, 说明这些接触面可能存在蛋白之间的相互作用^[27]. 值得一提的是, 在2020年, Wang等^[68]成功重组了完整SAGA复合物与泛素化核小体的复合物结构, 结构表明, 完整SAGA与泛素化核小体相互作用界面的结构特征与下文中观察到的SAGA-DUB与泛素化核小体的结构特征基本相同.

相比于独立的DUB结构域和泛素化核小体, 复合物中没有显示出明显的结构差别^[67]. 但是通过观察Durand等^[69]获得的未经过修饰的核小体与完整SAGA复合物SAGA-NCP的电子显微镜结构, 发现该复合物与SAGA-DUB-H2BK120ub-NCP的结构有许多不同. 在SAGA-NCP结构中, 除了Sgf11亚基, 没有其他结构域与核小体具有相互作用, 且核小体取向具有波动^[69]. 然而, 在SAGA-DUB-H2BK120ub-NCP复合物中, DUB和泛素化核小体具有多重相互作用, 不仅包括Ubp8的泛素特异性结合, 还有Ubp8的数个残基以及Sgf11的锌指结构域分别与H2B碳端螺旋和H2A/H2B酸性口袋进行的结合^[70].

Ubp8是SAGA中特异性结合泛素的亚基. 核小体H2BK120位的赖氨酸, 以及连接在这里的泛素同时落入Ubp8亚基的凹陷中, 结合面积为 $376 \text{ \AA}^2$ (图7)^[27]. 根据电子密度建模, Ubp8的233位酪氨酸指向H2B的碳端螺旋, 与H2B的H106和S109形成相互作用. Ubp8的374位精氨酸则与核小体DNA产生静电相互作用^[27]. 生化实验证明, Y233、R374的替换均降低了去泛素化

活性, 但不影响Sgf11与非特异性底物K48连接的双泛素的作用^[27]. 因此, Ubp8在功能上与Sgf11有相似之处, 均通过与核小体特定定位点作用促进DUB的结合, 从而介导H2BK120ub去泛素化.

Sgf11通过其锌指(zinc finger)结构与核小体H2A/H2B酸性口袋产生静电相互作用, 面积约为 $610 \text{ \AA}^2$ (图8)^[27]. 这一发现否定了之前的假说, 即该锌指(包含78、84、91位精氨酸)与核小体DNA相互作用^[71]. 为了验证Sgf11锌指对SAGA识别泛素化核小体的重要性, Morgan等^[27]将Sgf11中6个精氨酸R78、84、91、95、98、99分别替换成丙氨酸, 发现R78、84、91替换后SAGA无法将核小体去泛素化, 而R95、98、99的

图7 Ubp8与泛素化核小体相互作用. (a) Ubp8的Y233与H2A相互作用; (b) Ubp8的R374与核小体DNA相互作用. 灰色: DNA; 浅黄色: H2B; 砖红色: Ubp8 (网络版彩图)

Figure 7 Interactions between Ubp8 and ubiquitinated nucleosome. (a) Y233 of Ubp8 interacts with residues on H2A; (b) R374 of Ubp8 interacts with DNA. Gray: DNA; paleyellow: H2B; brickred: Ubp8 (color online).

图6 SAGA-DUB-H2BK120ub-NCP复合物(PDB: 4zux). 灰色: DNA; 紫色: 泛素; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4; 砖红色: Ubp8; 青色: Sgf11; 粉红色: Sgf73; 草绿色: SUS1 (网络版彩图)

Figure 6 Structure of SAGA-DUB binding to H2BK120ub-modified nucleosome core particle. Gray: DNA; purple: ubiquitin; orange: histone H2A; paleyellow: H2B; forest: H3; blue: H4; brickred: Ubp8; deepteal: Sgf11; pink: Sgf73; grass: SUS1 (color online).

图8 Sgf11亚基结合H2A/H2B酸性口袋. 橙色: H2A; 浅黄色: H2B; 深青色: Sgf11; 砖红色: Ubp8 (网络版彩图)

Figure 8 Sgf11 binds H2A/H2B acidic patch. Orange: H2A; paleyellow: H2B; deepteal: Sgf11; brickred: Ubp8 (color online).

替换不产生显著影响^[27]。同时, 将酸性口袋中的带负电的E46替换成精氨酸, 会导致去泛素化活性降低降低为原来的百分之一^[27]。上述替换均不影响SAGA对K48双泛素的去泛素化能力, 因此上述实验证明SAGA也通过Sgf11的锌指结构域与泛素化核小体的酸性口袋特异性结合促进去泛素化^[27]。

综上所述, SAGA-DUB-H2BK120ub-NCP复合物揭示了SAGA被泛素化核小体招募的结构机制: 泛素通过促进DUB的Ubp8与核小体结合, 进而在Sgf11与核小体酸性口袋的相互作用帮助下, 招募SAGA对核小体进行去泛素化。Morgan等^[27]证明SAGA-DUB不仅可以对核小体, 还能对H2A/H2B组蛋白的二聚体去泛素化。这个结论确认了SAGA可以在转录过程中, 对不同存在方式的组蛋白进行去泛素化。

4 53BP1-ubme-NCP复合物

DNA双链断裂(double strand break, DSB)修复是一个重要的生理过程。该过程广泛受到核小体翻译后修饰如甲基化、乙酰化和泛素化等调控^[72]。在触发DNA双链断裂修复时, E3泛素连接酶RNF168对核小体组蛋白H2A上13和15位赖氨酸进行泛素化。随后, 核小体H2AK15ub修饰协同H4K20me2修饰, 共同招募53BP1蛋白(一种重要的DNA双链断裂调控分子), 促进53BP1在DSB区域聚集进而招募RNF169或Rad51等修复因子的活性, 从而激活DNA双链断裂修复^[62,73,74]。

53BP1利用其TTD (tandem Tudor domain)与UDR (ubiquitination-dependent recruitment motif)结构域分别识别H4K20me2与H2AK15ub修饰^[75,76]。2016年, Wilson等^[28]重构了含H2AK15ub和H4K20me2修饰以及601序列DNA的单核小体(记为ubme-NCP)^[77]。通过半胱氨酸S烷基化合成了H4K20me2的类似物H4K20Cme2组蛋白。利用RNF168酶促反应合成了H2AK15ub组蛋白, 并对H2A的13和36赖氨酸进行替换, 防止这些位点被泛素化^[28,78,79]。随后, 选取了53BP1蛋白最小的截短体, 即由TTD和UDR结构域组成的1484~1631片段, 形成2:1结合比的53BP1-ubme-NCP复合物(图9)。

Wilson等^[28]利用冷冻电子显微镜解析了53BP1-ubme-NCP复合物的结构, 分辨率为4.5 Å。复合物中核小体的结构与单核小体基本没有区别^[2], 核小体盘的

上下两边各结合一分子53BP1形成2:1结合比的复合物。53BP1跨越核小体的表面, 穿过H4、H2B组成的裂口以及泛素和核小体构成的疏水口袋, 在N端的UDR结合域与H2A/H2B酸性口袋结合(图10)^[28]。

尽管核小体部分的结构未发生明显变化, 但是与未招募53BP1的NCP-ubme核小体(分辨率为7.7 Å)相比, 53BP1-ubme-NCP复合物上的泛素位置更加固定^[28]。这一现象的原因是泛素与H2B的碳端 α -螺旋和53BP1的UDR结构域产生了相互作用, 进而稳定了泛素的构象。在53BP1-ubme-NCP复合物中, 泛素的R42、R72、L8等与H2B组蛋白 α -螺旋上的S123和K116产生相互作用(图11)。此外, 53BP1的招募同时也受到其他组蛋白修饰的影响。对H2BK119位苏氨酸进行了谷氨酸替换, 会大幅减弱NCP-ubme和53BP1之间的复合, 表明有丝分裂等过程可以通过磷酸化H2B的119位苏氨酸来阻止53BP1对NCP-ubme的识别^[28,80]。此外, 值得一提的是, 不同于组蛋白H2AK13/15ub的修饰, CLR4^{WDR70}介导的H2BK116ub修饰也被发现可以促进DNA损伤修复过程, 机制上, H2Bub招募同源重组通路的关键蛋白Exo1, 同时拮抗53BP1的招募进而促进DNA损伤修复。但是, 体外的H2BK120ub pull down实验证明53BP1能够与H2BK120ub结合, 这可能是由于泛素化位点H2Bub与H2AK15ub在核小体空间

图9 53BP1-ubme-NCP复合物(PDB: 5kgf)。灰色: 核小体DNA; 紫色: 泛素; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4; 青色: 53BP1(网络版彩图)

Figure 9 Structure of 53BP1 binding to H2AK15ub and H4K20me2-modified nucleosome. Gray: DNA; purple: ubiquitin; orange: H2A; paleyellow: H2B; green: H3; blue: H4; cyan: 53BP1 (color online).

图 10 53BP1在核小体表面的位置。灰色: 核小体DNA; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4; 紫色: 泛素; 青色: 53BP1 (网络版彩图)

Figure 10 Location of 53BP1 on the ubiquitinated nucleosome. Gray: DNA; orange: H2A; paleyellow: H2B; forest: H3; blue: purple: ubiquitin; H4; cyan: 53BP1 (color online).

图 11 H2B与泛素之间相互作用。(a) 泛素R42和H2B的R72与S123之间相互作用;(b) 泛素L8与H2B的K116之间相互作用。浅黄色: H2B; 青色: 53BP1; 紫色: 泛素(网络版彩图)

Figure 11 Interactions between ubiquitin and histones. (a) H2B interacts with R42 of ubiquitin; (b) H2B interacts with L8 of ubiquitin. paleyellow: H2B; cyan: 53BP1; purple: ubiquitin (color online).

定位上距离邻近造成的假阳性结果。体内H2Bub可能结合了同源重组相关蛋白, 阻碍了53BP1的招募, 详尽的阻碍机制有待进一步深入研究^[28,81,82]。

泛素在复合物中特异性结合UDR的尾端环结构。UDR中I1617、L1619和泛素L8、I44和V70的疏水口袋产生疏水相互作用(图12), 而UDR的D1620和D1616分别与泛素K6和H68产生氢键作用。向UDR的碳端方向看去, K1628、R1627、R1630等正电性残基与核小体酸性口袋中H2A的E61、E64、D90、E92等产生氢键相互作用。这些残基的替换都会导致ubme-NCP与53BP1的结合受到阻碍^[28], 尤其是R1627替换成丙氨酸还会进一步影响53BP1在DNA双链断裂位点上的招募^[28]。

组蛋白H2AK13和K15均能进行泛素化修饰, 而53BP1特异性识别K15而非K13泛素化修饰的结构基

图 12 UDR与NCP-ubme作用。(a) UDR与泛素疏水作用;(b) UDR与泛素静电作用;(c) UDR与组蛋白酸性口袋静电作用。橙色: H2A; 浅黄色: H2B; 青色: 53BP1; 紫色: 泛素(网络版彩图)

Figure 12 Interactions between UDR and NCP-ubme. (a) Hydrophobic interaction between UDR domain and NCP-ubme; (b) electrostatic interaction between UDR domain and NCP-ubme; (c) electronic interaction between UDR domain and H2A/H2B acidic patch. Orange: H2A; paleyellow: H2B; cyan: 53BP1; purple: ubiquitin (color online).

础来自于H2A组蛋白形成的“正电钳”结构。组蛋白H2A的R11和R17横跨核小体DNA的SHL4.5方位, 与DNA通过静电相互作用结合。将R11和R17的“正电钳”做一个“位移”, 即将K9和K15替换为R, 替换后K13ub同样可以招募53BP1, 说明53BP1特异性结合K15泛素化修饰的结构基础是H2A组蛋白氮末端形成的R/KxxxKxR结构域^[28]。这片结构域通过与核小体DNA作用, 使得泛素分子与53BP1的作用更加稳定。

53BP1-ubme-NCP的电子显微镜结构揭示了双修饰核小体对53BP1的多价结合, 代表了一种独特的通过多价结合的方式识别修饰核小体的模型。53BP1的UDR结构域与泛素和组蛋白酸性口袋结合, 并且H2A的DNA钳结构固定了二者的结合; 与此同时, 53BP1与H4氮末端结合, 识别H4K20me2。这一研究初步阐明了53BP1在DNA双链断裂修复中产生功能的分子机制, 但是诸如53BP1识别后RNF169取代招募的分子机制, 仍需进一步研究来阐明。

5 Chd1-H2BK120ub-NCP复合物

2018年, 首个关于受泛素化调控的染色质重塑酶的核小体复合物结构得到了解析^[29]。真核细胞染色质

中遗传物质的复制、修复和转录过程需要调控染色质中核小体的空间位置, 调控此类活动的蛋白被称为染色质重塑蛋白. 一种依赖于ATP调控染色质中蛋白质与DNA的相互作用的Snf2蛋白, 是染色质重塑蛋白中的一个家族^[83]. 酵母Chd1蛋白是该族成员, 可以与转录因子互相作用, 控制染色质在基因组水平上的空间形态, 从而改变该染色质区域相关生理活动^[84-86]. 核小体翻译后修饰是调控染色质空间结构的因素之一, H2BK120泛素化已被证明能够帮助招募组蛋白伴侣因子FACT激活转录, 也是H3K4和H3K79甲基化修饰的必要条件^[6,87]. 因此组蛋白修饰可能与Chd1被核小体招募相互串扰^[88], 前人已证明H2BK120泛素化可以促进Chd1对核小体的滑位作用^[89].

Chd1包括与连接DNA结合的DNABD结构域、ATP酶活性中心的两片叶结构域以及染色域共四个亚基(图13)^[90,91]. Farnung等^[92]报道了Chd1和未修饰核小体复合物的结构(记为Chd1-NCP), 而Sundaramoorthy等^[29]报道了Chd1与泛素化修饰核小体的复合物结构(记为Chd1-H2BK120ub-NCP). 工作中采用了半胱氨酸S烷基化的方式制备甲基化赖氨酸的类似物^[93], 使用二氯丙酮交联的方式引入H2BK120泛素, 并使用ADP-BeF模拟过渡态. 为引入DNABD结合所需的连接DNA, 将核小体DNA不对称延长14 bp作为连接DNA, 最后用重新设计的非洲爪蟾的组蛋白与601序列DNA

进行核小体的重构^[29,77].

Chd1-H2BK120ub-NCP复合物结合比为1:1, 分辨率为4.5 Å^[29]. 复合物中Chd1结构并未发生明显变化, 复合物结构相比没有泛素化的核小体也没有明显的区别^[92,94]. 然而连接DNA所在的端部发生了一定程度的偏离核小体方向的解离. Chd1的两片ATP酶叶和染色域在核小体DNA的SHL1~2方位的侧面结合, 而DNABD与核小体外部的连接DNA作用. 令人惊讶的是, Chd1与H2A、H2B之间并没有明显的接触, 仅与H3和H4末端有相互作用, 与泛素也存在一定的距离. 这与人研究的SAGA-DUB和53BP1结合泛素化核小体的结构存在很大的差异^[27,28].

Chd1染色域与核小体以及Chd1其他结构域具有多价相互作用. 在没有结合核小体时, Chd1染色域通过与ATP酶叶的相互作用, 抑制ATP酶活性. 而在复合物中, 核小体与染色域相互作用, 解除ATP酶叶的抑制构象, 促进转录发生^[94]. 在复合物中, 染色域的K240、R241、R277和K247与核小体DNA的SHL1部分作用(图14), 可能对核小体DNA的部分解离具有促进作用. 以染色域K329为中心的面与以DNABD结构域中D1200和P1201为中心的面互相作用; K345与DNABD上的一个富含天冬氨酸和谷氨酸等酸性残基的小螺旋结构进行静电作用; R341与ATP酶叶I上的E468和Y469相互作用.

图 13 Chd1-核小体复合物. (a) Chd1序列^[92]; (b) 1:1结合比的Chd1-NCP复合物电子显微镜结构(PDB: 5o9g); (c) 1:1结合比的Chd1-H2BK120ub-NCP复合物电子显微镜结构(PDB: 6ftx). 灰色: DNA; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4; 紫色: 泛素; 深紫色: DNABD结合域; 粉色: 染色域; 青色: ATP酶叶I; 棕色: ATP酶叶II (网络版彩图)

Figure 13 Chd1 sequence and structure of Chd1 binding to H2BK120ub nucleosome. (a) Sequence of Chd1 ATPase chromatin remodelling enzyme^[92]; (b) structure of Chd1 binding to nucleosome; (c) structure of Chd1 binding to H2BK120ub nucleosome. Gray: DNA; orange: H2A; pale yellow: H2B; green: H3; blue: H4; purple: ubiquitin; deep purple: DNA binding domain; pink: chromodomain; deeptea: ATPase lobe I; sand: ATPase lobe II (color online).

图 14 染色域与其他结构域的相互作用。(a) 染色域与DNA相互作用; (b) 染色域与DNABD相互作用; (c) 染色域与DNABD的酸性螺旋作用; (d) 染色域与ATP酶叶I相互作用。灰色: DNA; 深紫色: DNABD结合域; 粉色: 染色域; 青色: ATP酶叶I (网络版彩图)

Figure 14 Interactions of chromodomain with other subunits. (a) Chromodomain interacts with DNA; (b) chromodomain interacts with DNA binding domain; (c) chromatin interacts with acidic helix of DNA binding domain; (d) chromodomain interacts with ATPase lobe I. Gray: DNA; deep purple: DNA binding domain; pink: chromodomain; deeptea: ATPase lobe I (color online).

结合核小体之后, ATP酶叶结构域的构象发生了从抑制到活化的变化。首先, ATP酶叶II靠近叶I, 导致保守的Walker盒基序的K407、R804和R807组成的ATP结合结构域由抑制构象转变为活性构象, 使得ADP-BeF能够结合到上述残基构成的口袋中(图15)^[29,94,95]。在处于抑制构象时, ATP结合口袋外ATP酶叶II的一个富含正电残基的面与染色域的一条长螺旋中的酸性残基之间互相作用。结合底物后, ATP酶叶II旋转了40°, 其酸性表面转向染色域的酸性螺旋^[29]。ATP酶叶II的构象变化也使其与核小体DNA(图16)、组蛋白H4的尾端和H3的 α 1-螺旋相互接触, 即K632和K642与H3的 α 1-螺旋的D81和E73相互作用。生化实验证明K632~K646段肽环的缺失使得Chd1丧失核小体滑动的活性^[29]。此外, ATP酶叶II的E669、D725和D729与H4尾端作用高度保守, 证实了前人假说, 即H4尾端能够调控Chd1活性, 以及H4尾端切除对于核小体滑动和ATP酶活性均有抑制作用的实验结果^[96]。

复合物中组蛋白与Chd1基本没有互相作用, 但是Chd1的存在可能对H3组蛋白的构象产生影响: 在两个H3中, 靠近解离DNA的H3分辨率更高, 可能是由于H3的氮端延长之后与连接DNA互相作用。此外, 每个核小体上连有2个泛素, 其分辨率低于复合物的其他组分。在Chd1结合侧的泛素信号更强, 且位置偏离核小体表面, 向SHL6方位的DNA移动, 利用其K48、

图 15 结合染色体后ATP结合域的活化(黄点处为ATP结合域)。(a) 在没有结合核小体时的Chd1 (缺少DNABD)的惰性结构(PDB: 3mwy); (b) 结合核小体时Chd1的活性结构。深紫色: DNABD结合域; 粉色: 染色域; 深青色: ATP酶叶I; 棕色: ATP酶叶II (网络版彩图)

Figure 15 Activation of ATP binding domain (yellow dots). (a) Chd1 in loose conformation without ubiquitinated nucleosome binding; (b) Chd1 in activated conformation with ubiquitinated nucleosome binding. Deep purple: DNA binding domain; pink: chromodomain; deeptea: ATPase lobe I; sand: ATPase lobe II (color online).

图 16 ATP酶叶II与核小体互相作用。(a) 叶II与DNA相互作用; (b) 叶II与H3相互作用; (c) 叶II与H4的氮端作用。灰色: DNA; 森林绿: H3; 蓝色: H4; 粉色: 染色域; 深青色: ATP酶叶I; 棕色: ATP酶叶II (网络版彩图)

Figure 16 ATPase lobe II interacts with nucleosome. (a) Interaction between lobe II and nucleosomal DNA; (b) interaction between lobe II and H3; (c) interaction between lobe II and H4 N-terminal. Gray: DNA; forest: H3; blue: H4; pink: chromodomain; deeptea: ATPase lobe I; sand: ATPase lobe II (color online).

R54和N60与DNA相互作用。这些相互作用可能参与了DNA松弛的过程, 促进Chd1和核小体的结合, 解释了H2BK120ub修饰增强Chd1活性的实验现象^[89]。

ATP的结合是导致整个构象转变以及染色体重新定位的驱动因素^[29]。因此我们可以提出Chd1重塑核小体的分子机制, 即ATP结合到ATP酶叶产生构象变化, 诱导染色域构象变化, 最后导致DNA滑位。虽然Chd1

与H2A和H2B均无相互作用, 但失去H2A/H2B的二聚体会导致核小体部分解构与DNA解离, 影响Chd1中DNABD结合连接DNA. 反之, Chd1结合核小体DNA也可以促进核小体组蛋白六聚体招募单独的H2A/H2B二聚体, 形成正常的八聚体组蛋白^[89]. 这种特性揭示了Chd1具有在不同阶段调节核小体组装和核小体排布的能力.

6 Dot1L-H2BK120ub-NCP复合物

核小体H3K79甲基化修饰在人体内具有重要的作用, 是调控转录延长^[97]、DNA损伤修复^[98,99]和结合Sir蛋白导致端粒沉默^[100]的关键因素^[22,101]. 该修饰受到H2BK120位点(在酵母中为H2BK123)泛素化修饰的调控, 该调控过程是组蛋白翻译后修饰之间的串扰的一个重要案例^[21,55,97,102]. H3K79甲基化修饰是在Dot1L甲基转移酶的作用下, 赖氨酸侧链氨基氮原子亲核进攻甲基供体S-腺苷甲硫氨酸(SAM)而发生的. Dot1L最早作为一种阻止酵母端链沉默的蛋白被发现^[103], 总共含有1527个氨基酸, 其核心催化结构则由偏向氮端的高度保守甲基转移酶结构域和富含赖氨酸可以特异性定位至基因组位点的两个主要功能结构域组成(图

17)^[104~106]. Dot1L在人体发育过程中的作用不可替代, 其异常定位能导致H3K79的超甲基化和异位基因表达, 是小儿白血病的致病因素之一, 可用于研发靶向药物^[107~109].

体外替换和敲除实验证明了H4的氮端碱性区域可以激活Dot1L甲基化酶的活性, 同时泛素的L71和L73两个疏水残基在招募Dot1L起到关键作用^[110,111]. 2019年, 多个研究组得到了Dot1L结合H2BK120ub修饰核小体的电子显微镜结构^[30~34], 揭示了Dot1L被泛素化核小体招募的分子机制. Worden等^[30]解析了三种不同状态的Dot1L和H2BK120ub核小体复合物结构, 分别为Dot1L与核小体结合比为1:1和2:1的两种“活化状态”的复合物(分别记为aDot1L-H2BK120ub-NCP和2aDot1L-H2BK120ub-NCP)和结合比为1:1的Dot1L的催化结构远离H3K79组蛋白的“准备阶段”复合物(记为pDot1L-H2BK120ub-NCP). 他们采用非洲爪蟾的组蛋白和人源的Dot1L蛋白片段(2~416片段, 包含甲基转移催化结构域和特异性结合DNA的结构域, 代替完整Dot1L蛋白). H2BK120C与泛素UbG76C采用二氯丙酮交联, 再与其他组蛋白和601序列DNA进行重组^[61,77]. 在活化状态复合物中, 为增加组蛋白对Dot1L的亲

图 17 Dot1L和H2BK120ub核小体复合物的电子显微镜结构. (a) Dot1L的部分序列结构^[106]; (b) 2aDot1L-H2BK120ub-NCP结构(2:1, 核小体下部的结合未显示, PDB: 6nj9); (c) pDot1L-H2BK120ub-NCP结构(1:1, PDB: 6nog). 灰色: DNA; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4; 青色: Dot1L; 紫色: 泛素(网络版彩图)

Figure 17 Structure of Dot1L binding to H2BK120ub nucleosome. (a) Sequence information of Dot1L; (b) structure of Dot1L binding to nucleosome in active state; (c) structure of Dot1L binding to ubiquitinated nucleosome in poised state; Gray: DNA; orange: H2A; paleyellow: H2B; forest: H3; blue: H4; cyan: Dot1L; purple: ubiquitin (color online).

性, H3K79赖氨酸被替换为正亮氨酸^[30]. 其他团队复合物构建方法与此类似, 均使用二氯丙酮交联进行组蛋白体外修饰^[27,31~34,61].

比较准备阶段和活化阶段的复合物结构, 发现准备状态的Dot1L蛋白对组蛋白H4尾端部分缺少分辨率, 表明Dot1L蛋白在活性条件下才能结合H4的尾端. 在活化阶段中, H3组蛋白的构象发生了变化, 拉近了H3K79与Dot1L的距离, 使得Dot1L构象更加紧密, 靠近核小体. 在复合物中, 泛素通过与Dot1L的直接作用来完成对其的招募. 泛素与Dot1L的L320~330残基组成的螺旋疏水结构具有接触面积为637 Å²的相互作用(图18)^[30]. Dot1L的L322与F326, 以及邻近的L284和I290疏水残基与泛素的I36、L71和L73组成一组平行的疏水作用, F326的替换对结合活性影响最为显著^[30,31].

Dot1L与核小体盘面主要依靠精氨酸锚与组蛋白酸性口袋产生相互作用. Dot1L的R282接触H2A/H2B的酸性口袋H2AE56和H2BE113两个负电残基, 而R278残基与H2AE61和E64相互作用. 虽然组蛋白H2AE61和64作为蛋白招募作用位点已有类似报道^[27,28], 但对于R282, 不同于已知的与酸性口袋结合, 它结合了酸性口袋中并非典型结合位点的H2AE56和H2BE113, 是一种新型的结合范式. 除此以外, 在活性状态复合物中还出现了H4与Dot1L的相互作用. Dot1L组成氮端环的残基22~32向外位移了约4.6 Å的距离, 形成了一个凹槽固定住H4尾端^[30]. 这方面的作用力主要为17~19残基的静电作用, 其中H18和R17与Dot1L的

图 18 2aDot1L-H2BK120ub-NCP复合物中Dot1L与泛素和酸性口袋的相互作用. (a) Dot1L与泛素相互作用; (b) Dot1L与H2A/H2B酸性口袋相互作用. 橙色: H2A; 浅黄色: H2B; 青色: Dot1L; 紫色: 泛素(网络版彩图)

Figure 18 Dot1L interacts with ubiquitin and histone acidic patch. (a) Interaction between Dot1L and ubiquitin; (b) interaction between Dot1L and histone acidic patch. Orange: H2A; paleyellow: H2B; cyan: Dot1L; purple: ubiquitin (color online).

N126、S304、E138和Y115作用, 而R19与H3的T80、Q76和K79Nle作用. 体外敲除RHR三个残基的核小体将丧失对Dot1L的结合能力^[111,112], 将17~19RHR或20~22KVL全部突变为丙氨酸, 或者将Dot1L蛋白E138替换均可全部消除Dot1L甲基转移的活性^[31,34] 电子显微镜结构为上述生化现象提供了理论基础, 证明了H4末端对Dot1L结合核小体活性的重要作用.

Dot1L的激活是通过自身组蛋白进行变构而完成的. 在pDot1L-H2BK120ub-NCP结构中, H3K79的侧链结构平伏在核小体盘面上, 对甲基转移酶具有化学惰性(图19)^[113]. 然而, 2aDot1L-H2BK120ub-NCP的H3K79构象发生了有趣的变化, 即H3发生了构象旋转, 导致H3K79赖氨酸侧链方向改变为垂直核小体盘面, 直接指向SAM结合位点. 这一构象变化受到多种因素推动, 例如H4的尾端R19可以通过静电作用, Dot1L的F131和W305也可以通过范德华力稳定H3K79的激活构象. 并且F131通过推动K79Nle的构象旋转, 自身也发生构象变化, 占据了K79Nle原来的位置. 此外, H3K79环附近的Dot1L富含疏水残基, 形成了一个疏水通道, 连接K79和甲基给体SAM. 这一结构也解释了Dot1L激活H3K79甲基化的机制, 即由于被疏水基团包裹的

图 19 在准备状态(浅绿)和活化状态(森林绿)的H3K79构象改变. (a) H3K79Nle侧链由惰性构象活化为活性构象; (b) H3K79Nle的活性构象使得氨基靠近SAM; (c) H3K79Nle与SAM的反应中心被疏水基团(浅蓝)所包裹. 森林绿色: H3; 蓝色: H4; 青色: Dot1L; 红色: SAM; 白色: Dot1L的疏水残基(网络版彩图)

Figure 19 Conformational change in poised state (light green) and active state (forest). (a) Activation of H3K79Nle; (b) active state brings H3K79Nle to SAM; (c) catalytic center is wrapped by a hydrophobic pocket. Forest: H3; blue: H4; cyan: Dot1L; red: SAM; white: hydrophobic residues (color online).

赖氨酸具有更低的 pK_a , 因此其无质子化的中性形态更稳定, 导致亲核性更强^[114]. 将参与组成疏水通道的F131和W305分别替换为丙氨酸, 发现能够全部消除Dot1L的活性^[30,34]. 以上生化实验证明了疏水通道的形成对于Dot1L激活的必要性^[30].

综上所述, 核小体上的泛素、H2A/H2B酸性口袋和H4尾端的多重作用使得Dot1L在核小体上稳定结合, 同时H4和Dot1L的一系列疏水残基维持活性状态中H3K79指向甲基化试剂SAM的直立构象稳定, 并形成疏水通道增加赖氨酸的亲核性, 促进K79的甲基化. 这些关于Dot1L结合泛素化核小体的研究提出了泛素化核小体通过自身的组蛋白上残基侧链的构象变化来符合复合物的活性构象完成招募的新机制.

7 MLL1/compass-H2BK120ub-NCP复合物

H3K4甲基化修饰(H3K4me1/2/3)是一种高度保守的、通过招募转录因子来调控细胞转录启动等生理活动的翻译后修饰^[21,88,101,115,116]. 该修饰在人体中受到6种不同的SET1/MLL家族甲基转移酶蛋白调控, 包括SET1A、B和MLL1~4, 其中MLL1和MLL2在启动子位点聚集, 承担H3K4三甲基化修饰的任务^[117]; 而MLL3和MLL4主要聚集在增强子区域对H3K4进行单甲基化^[118,119]. 相比于MLL家族蛋白, 催化能力更强的SET1类蛋白则在体内作为广泛的甲基化修饰蛋白存在^[120]. 研究证明此类蛋白的调控对胚胎干细胞的发育、造血和神经发生都具有不可替代的作用^[121~124], 而人类的精神分裂症、自闭症和混合谱系白血病等疾病均与这类蛋白的突变有密切联系^[124~126].

与H3K79甲基化相似, SET1/MLL家族蛋白介导的H3K4甲基化过程受到H2BK120ub调控^[20,22,40,122]. SET1/MLL复合物结构呈Y型, 催化位点SET位于其中心(图20), 依靠SET甲基转移活性中心对组蛋白进行甲基化^[127]. 人源MLL活性中心包括WDR5(Cps30)、RbBP5(Cps50)、ASH2L(Cps60)、DPY30(Cps25)等亚基, 对应酵母甲基转移酶COMPASS的活性中心则还有Cps40^[128,129]. 泛素化修饰对COMPASS与核小体的亲和力没有显著影响, 只影响其甲基化活性^[128].

2019年报道了多个有关MLL族蛋白结合泛素化核小体的结构及其激活机制(图21). Xue等^[35]获取了人源MLL1甲基转移酶结合泛素化核小体的电子显微镜

图 20 COMPASS复合物的结构及其活性中心(PDB: 6chg). 青色: Cps50; 粉色: Cps30; 淡蓝色: Set1; 红色: Cps60; 绿色: Cps25_a; 棕色: Cps25_b (网络版彩图)

Figure 20 Structure of COMPASS complex and its active domain. Cyan: Cps50; pink: Cps30; lightblue: Set1; red: Cps60; green: Cps25_a; sand: Cps25_b (color online).

图 21 MLL1蛋白结合泛素化核小体的复合物. 灰色: DNA; 紫色: 泛素; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4 (网络版彩图)

Figure 21 Structure of MLL1 complex binding to ubiquitinated nucleosome. Gray: DNA; purple: ubiquitin; orange: H2A; paleyellow: H2B; forest: H3; blue: H4 (color online).

结构(记为MLL1-H2BK120ub-NCP), 同时Hsu等^[36]和Worden等^[37]报道了COMPASS复合物结合泛素化核小体的电子显微镜结构(记为COMPASS-H2BK120ub-NCP). 在核小体重组中, Xue等^[35]采用二氯丙酮交联UbG76C和非洲爪蟾的组蛋白H2BK120C制备泛素化

组蛋白, 再与601序列DNA重组核小体, 与MLL1各亚基组装得到复合物, 最终结构的分辨率为3.2 Å^[77]. Hsu等^[36]采用了乳酸克鲁维斯酵母表达的COMPASS蛋白的甲基转移活性中心和601序列DNA核小体, 使用形成二硫键的交联方式重组核小体, 分辨率为3.5 Å^[77]. Worden等^[37]选取酿酒酵母菌(*Saccharomyces cerevisiae*)的COMPASS复合物, 将组蛋白H3K4进行替换成正亮氨酸Nle, 并用二氯丙酮交联泛素, 再与非洲爪蟾DNA组装核小体复合物, 最后得到分辨率为3.95 Å、结合比为2:1的复合物电子显微镜结构. 由于高度同源性, MLL1和COMPASS-H2BK120ub-NCP的电子显微镜结构在许多方面具有相似性. 以下关于蛋白亚基与核小体之间作用的描述统一采用MLL1-H2BK120ub-NCP的复合物结构为准^[35].

在复合物中, 泛素化核小体主要通过DNA、泛素分子和组蛋白招募MLL1复合物. 首先是核小体DNA通过不同的位点与MLL1复合物相互作用, 包括位于SHL4方位上与RbBP5的四个环上精氨酸的相互作用, 以及位于SHL7方位上与MLL1的R3916、K3924相互作用. 虽然没有直接观察到ASH2L与DNA的相互作用, 但是通过局部电子显微镜解析, ASH2L的前-SPRY结构域^[35]共在三个位点上固定MLL1复合物, 其中一个位点由YRYIL序列构成, 其位于一个β-链结构中, 而残基的替换分析确认了这个β-链结构对核小体招募的必要性^[35]. 相比之下, COMPASS-H2BK120ub-NCP由于存在多余的Cps40, 其上的正电聚集的富集赖氨酸的环可以与DNA作用, 导致其与MLL-H2BK120ub-NCP的结构有所不同^[36,37].

MLL1与COMPASS复合物与泛素分子的作用存在差异. 在MLL1-H2BK120ub-NCP中, 泛素的作用是识别RbBP5 (图22). RbBP5被核小体盘面与泛素分子夹心形成双重结合, 泛素的L44疏水面与P239疏水的五元环形成疏水相互作用, 同时疏水面周边也有一系列的静电作用. 而在COMPASS-H2BK120ub-NCP中, 虽然泛素与Cps50 (RbBP5)也具有疏水作用, 但泛素还与Set1和Cps60之间存在相互作用, 形成面积高达930 Å²的相互作用面^[37]. 在结合过程中, Cps50发生构象伸展, 通过泛素分子的I44中心的疏水口袋来相互作用. Set1蛋白则与Dot1L-H2BK120ub-NCP复合物类似, 与泛素的I36疏水口袋结合^[36].

H2A和H2B组蛋白也参与了MLL1与COMPASS的

图22 MLL1复合物中RbBP5与泛素作用. (a) 泛素与RbBP5的静电作用; (b) 泛素与RbBP5的疏水作用. 紫色: 泛素; 青色: RbBP5 (网络版彩图)

Figure 22 RbBP5 subunit interacts with ubiquitin on histone. (a) Electronic interaction between RbBP5 and ubiquitin; (b) hydrophobic interaction between RbBP5 and ubiquitin. Purple: ubiquitin; cyan: RbBP5 (color online).

招募. 在MLL-H2BK120ub-NCP中, H2B与H4形成的裂缝容纳了RbBP5蛋白亚基的部分进入. RbBP5通过双环结构与其他组蛋白相互作用(图23). 其中一个环与H2B的V118产生疏水作用, 另一个环则通过H3K79有小部分接触. 与H2B的疏水作用在酵母COMPASS中是保留的^[36], 但在COMPASS复合物中的Cps50与H3K79形成的氢键在MLL复合物中却没有被观测到^[36]. H2A与MLL1亚基形成多价的互相作用, 如

图23 核小体组蛋白与MLL1复合物作用. (a) RbBP5与H3、H2B之间作用; (b, c) MLL1与H2A之间作用. 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 青色: RbBP5; 淡蓝色: MLL1 (网络版彩图)

Figure 23 Interaction between histones and MLL1 complex. (a) RbBP5 interacts with H3 and H2B; (b, c) MLL1 subunit interacts with H2A. Orange: H2A; paleyellow: H2B; forest: H3; cyan: RbBP5; lightblue: MLL1 (color online).

R3821与H2A的D72形成静电作用。值得注意的是,在COMPASS复合物中,与MLL1同源的Set1结构域的结合对COMPASS调控H3K4甲基化起到重要的作用。

在COMPASS复合物中,Set1富含精氨酸的RxR结构域在激活甲基转移活性时起到关键作用^[127]。在单独COMPASS结构中RxR结构域是无序的,无泛素化核小体不能使得COMPASS显示甲基化活性,但是敲除掉RxR结构域的COMPASS却对无泛素化核小体有甲基化活性。这说明完整的COMPASS的酶活性在没有泛素介导的作用下,被RxR结构域自抑制^[36,128]。而与泛素化核小体结合后,RxR结构域位于泛素和核小体盘面中间,并变构成 α -螺旋与H2A/H2B酸性口袋接触(图24),且泛素参与稳定了 α -螺旋周边部分的构象^[36,37]。因此,RxR与核小体的作用诱导了这一段肽链发生了构象改变,从而加强了甲基转移活性。这一性质在MLL复合物中没有观察到,说明了SET1/MLL此类酶在不同物种同源物间的不同激活机制。

为探究MLL1复合物在识别泛素化核小体后对H3K4甲基化修饰的分子机制,Xue等^[35]制备了MLL1结合未修饰核小体的复合物并解析了电子显微镜结构(PDB: 6kix、6kiz)。观察复合物时发现了两种结合方

图 24 COMPASS复合物中RxR结构域与核小体盘面作用。(a, b) RxR结构域与H2A、H2B作用;(c) Set1亚基与泛素作用。紫色: 泛素; 橙色: H2A; 浅黄色: H2B; 蓝色: Set1 (网络版彩图)

Figure 24 COMPASS RxR domain interacts with histone surface. (a, b) RxR domain interacts with H2A (orange) /H2B (paleyellow); (c) Set1 subunit interacts with ubiquitin. Purple: ubiquitin; orange: H2A; paleyellow: H2B; lightblue: Set1 (color online).

式,其中一种与MLL-H2BK120ub-NCP区别不大,而另一种结合方式发生了26°的旋转(图25)^[35]。实验显示这种新结合方式和原有方式相比,与核小体的结合面积减小了约2/3^[35]。考虑到泛素化核小体招募MLL1甲基化的速率大于无修饰核小体^[35],推测H2BK120ub和H3K4me可能串扰机制的如下:在无修饰核小体上MLL1采用两种方式结合,而在泛素修饰核小体中泛素通过与MLL1相互作用,提升了第一种结合方式的稳定性,使得MLL1催化位点接近H3K4的构象占比显著增大,催化修饰速率加快^[35]。

8 Set2-H2BK120ub-NCP复合物

H3上36位赖氨酸的甲基化修饰(H3K36me)是一种高度保守性的翻译后修饰,在基因转录区域的下游大量出现^[21,130]。该修饰被Set2甲基转移酶所调控^[131],其受到RNA聚合酶II的招募,在转录活化的区域对核小体H3K36进行甲基化修饰^[132,133]。与Dot1L相同,Set2需要甲基化给体SAM,而H2BK120ub与Set2催化H3K36me之间可能具有的串扰机制,即泛素化核小体对Set2招募和活性产生促进^[134]。研究表明,Set2酶和H3K36me修饰参与了DNA损伤修复^[135,136]、组蛋白互作^[137]、转录抑制^[131,138]等生理过程,而Set2的突变将会导致急性白血病^[139]、肿瘤等疾病发生^[140,141]。

图 25 MLL1的第一种(不透明)和第二种(半透明)结合方式对比。青色: RbBP5; 粉色: WDR5; 淡蓝色: MLL1; 红色: ASH2L (网络版彩图)

Figure 25 Comparison between the two binding modes for MLL1 to recognize nucleosome. Cyan: RbBP5; pink: WDR5; lightblue: MLL1; red: ASH2L (color online).

近年来Set2催化中心的结构得到了表征。在Set与SAH或H3多肽结合中,发现Set2的催化中心由AWS和SET两个结构域构成(图26)^[142]。但是,对于H2BK120ub和Set2激活H3K36me修饰之间的串扰机制仍不清楚。2019年, Bilokapic等^[38]制备了H2B泛素化核小体与Set2组成的复合物(Set2-H2BK120ub-NCP), 阐明了Set2被泛素化核小体招募的结构机制。Bilokapic等^[38]将H3K36替换为甲硫氨酸, 使用二氯丙酮将泛素UbG76C与非洲爪蟾H2BK120C交联, 之后采用八聚体组蛋白和601序列DNA组建重组核小体, 与Set2形成复合物^[38,77], 得到了分辨率为3.8 Å、结合比1:1的复合物结构^[38](图27)。

通过3D重构,发现Set2蛋白依然仅有AWS和SET结构域具有较高分辨率的信号, 这些信号分布在核小体的DNA出入核小体的位点附近^[38]。相比于单独的核小体^[2], 在Set2-H2BK120ub-NCP电子显微镜结构中核小体DNA在出入口处呈现出更松弛的部分解绑状态。H3K36的位置正好位于核小体DNA的出口处, 因此Set2的结合需要DNA进行部分解离, 以减小结合的空间位阻。

Set2与泛素化核小体结合由Set2与DNA、组蛋白之间的多组作用共同促进。首先Set2蛋白与核小体DNA产生静电作用, 稳定DNA松弛构象(图28)。其中Set2酶的高度保守R117和核小体DNA的松弛区域, 即SHL5.5方位作用, 同时Set2的K241、K322也与核小体DNA的SHL1方位作用。此外, H3和H2A在招募Set2并

图 26 Set2是催化甲基化修饰K36的甲基转移酶。(a) Set2蛋白的催化结构序列; (b) Set2蛋白结合SAH的结构。棕色: Set2; 红色: SAH^[142] (网络版彩图)

Figure 26 Set2 is a methyl transferase catalyzing methyl modification on histone H3 lysine 36 residue. (a) Sequence of Set2 catalytic domain; (b) Set2 structure binding SAH, only AWS and SET domains are visible under observation. Sand: Set2; red: SAH ^[142] (color online).

图 27 Set2-H2BK120ub-NCP复合物结构(PDB: 6px3). (a) 泛素结合在核小体盘面正上方; (b) Set2在其中一个DNA出口处、核小体盘面边缘处结合。灰色: 核小体DNA; 橙色: H2A; 浅黄色: H2B; 森林绿色: H3; 蓝色: H4; 紫色: 泛素; 棕色: Set2蛋白(网络版彩图)

Figure 27 Structure of Set2 binding to ubiquitinated nucleosome. (a) Ubiquitin is located above the histone surface; (b) Set2 binds to the DNA at SHL1 position. Gray: DNA; orange: H2A; paleyellow: H2B; forest: H3; blue: H4; purple: ubiquitin; sand: Set2 (color online).

图 28 Set2的碱性残基与DNA的两个区域相互作用。灰色: 核小体DNA; 森林绿色: H3; 紫色: 泛素; 棕色: Set2 (网络版彩图)

Figure 28 Set2 utilizes its basic residues to bind nucleosomal DNA. Gray: DNA; forest: H3; purple: ubiquitin; sand: Set2 (color online).

且促进其H3K36甲基化的过程中起到了重要作用。H3通过多重作用结合Set2, 其K56、R52与Set2的D125、N151相互作用(图29), H3 α -N螺旋的I51、L48疏水面与Set2的I150、A153的疏水面也产生作用。H3构象因为与Set2结合而发生变化, 其氮端37~41残基的位置被DNA和Set2的作用固定^[38]。相比于Set2与H3多肽结合的晶体结构^[142], H3P38之后的残基在复合物结构中发生构象变化, 参与了与Set2和核小体的互相作用。由于前人证实H3P38参与K36甲基化调控^[143], 因此Set2可能通过促进H3氮端尾部的变构来进行H3K36甲基化调控。

与其他泛素化核小体复合物不同, Set2蛋白与H2A/H2B酸性口袋没有相互作用。实际上H2A组蛋白末端的K119通过静电力与Set2的D146结合。此外,

图 29 Set2与组蛋白作用. (a) Set2与H3静电作用; (b) Set2与H3疏水作用; (c) Set2与H2A之间静电作用; (d) Set2与H2A疏水作用. 橙色: H2A; 森林绿色: H3; 棕色: Set2 (网络版彩图)

Figure 29 Interaction between Set2 and histones. (a) Interaction between Set2 and H3; (b) hydrophobic interaction between Set2 and H3; (c) static interaction between Set2 and H2A; (d) the hydrophobic patch comprises Set2 and H2A. Orange: H2A; forest: H3; sand: Set2 (color online).

H2A的I111、L116能够与Set2的I150、A153形成疏水作用. 核小体H2A的碳端和H3的 α -N螺旋形成了一个由4个残基组成的疏水口袋. 生化实验证明口袋中疏水残基的替换可以消除Set2甲基化活性^[38]. H2AK119ub这一利用大体积修饰破坏疏水口袋结构的修饰, 也能破坏Set2甲基化的活性^[144]. 此疏水口袋具有潜在的与H2A/H2B酸性口袋所类似的广泛招募的能力, 可能是潜在的一种核小体招募目标蛋白的通用范式.

泛素在Set2-H2BK120ub-NCP电子显微镜结构中分辨率较低, 说明了其位置的灵活性. 泛素的63~67位残基与Set2的AWS结构域距离最近, 但多肽骨架之间的距离仍然高达约15 Å (图30), 说明泛素和Set2之间没有直接相互作用. 生化实验表明, 泛素修饰确实能够促进Set2的H3K36甲基化, 而且不会显著增加Set2对核小体的结合^[38]. 这一现象的结构基础亟待后续研究来阐明.

9 总结

伴随冷冻电子显微镜等分析技术和蛋白质化学合成方法的不断发展^[60,61,145~148], 泛素化核小体招募蛋白

图 30 泛素骨架与Set2骨架距离. 紫色: 泛素; 棕色: Set2 (网络版彩图)

Figure 30 Distance from peptide skeleton of ubiquitin to peptide skeleton of Set2. Purple: ubiquitin; sand: Set2 (color online).

的结构机制近期得到了重要的进展. 本文归纳了近年来解析的6类识别泛素化核小体的复合物的结构机制, 归纳出了一些泛素化核小体招募的共同范式. 泛素在部分复合物中采用它的两个泛素疏水口袋进行招募; 核小体H2A/H2B的酸性口袋作为组蛋白上最经典和广泛的招募结构域, 参与了包括Dot1L和MLL1在内的多种复合物的结合. DNA也通过其带有负电的特征, 稳定了复合物的构象.

上述六类泛素化核小体复合物通过不同的机制对招募蛋白进行活化. 其中一部分复合物是通过使招募蛋白发生构象变化来对其进行活化的, 典型的例子有MLL1复合物的竞争结合模式和COMPASS复合物的自抑制解除模式; 而部分复合物则能够通过改变自身组蛋白的构象, 活化被招募蛋白所修饰的残基, 如Dot1L能够通过改变H3K79残基的空间取向而催化H3K79的甲基化过程, 这一种新的活化方式也改进了人们对组蛋白结构和翻译后修饰的认识, 对于人们探究其他“惰性”位点的修饰时提供参考. 除了这两种激活方式以外, 在Chd1和Set2的泛素化核小体复合物中, 发现了DNA的部分解绑, 这也说明了核小体可以通过调整DNA缠绕程度来招募一些目标蛋白. 值得注意的是, 这两种复合物中招募蛋白与泛素分子的相互作用相比其他复合物更少, 在这些复合物中, 泛素化核小体促进目标蛋白招募的具体机制仍需要进一步的研究.

在未来, 随着越来越多受到泛素化核小体蛋白招

募和活化的蛋白出现, 会有更多染色质相关生理过程的调控机制得到阐明。但是目前的研究依然面临许多挑战。首先, 目前发现的泛素化核小体复合物结构中, 只包含H2A和H2B泛素化的复合物, H3和H4泛素化的核小体复合物还没有报道。除了H3和H4泛素化位点和丰度均较少, 形成复合物数量不多, 这还可能是因为H3和H4泛素化组蛋白的体外合成相对不成熟, 导致无

法重组蛋白。其次, 目前对泛素化核小体结构和蛋白调节机制的观测仍停留在单核小体层面。在多核小体层面, 招募蛋白和泛素化核小体的行为是否会出现不同或受到限制, 现有技术还不足以解答。以上这些问题的解答, 将使得我们能够更加深入地认知泛素修饰组蛋白在染色质复杂过程中产生的生物功能与采用的分子机制。

参考文献

- Andrews AJ, Luger K. *Annu Rev Biophys*, 2011, 40: 99–117
- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. *Nature*, 1997, 389: 251–260
- McGinty RK, Tan S. *Chem Rev*, 2015, 115: 2255–2273
- Jenuwein T, Allis CD. *Science*, 2001, 293: 1074–1080
- Orphanides G, Reinberg D. *Nature*, 2000, 407: 471–476
- Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D. *Science*, 2003, 301: 1090–1093
- Tweedie-Cullen RY, Reck JM, Mansuy IM. *J Proteome Res*, 2009, 8: 4966–4982
- Winget JM, Mayor T. *Mol Cell*, 2010, 38: 627–635
- Vijay-kumar S, Bugg CE, Cook WJ. *J Mol Biol*, 1987, 194: 531–544
- Hershko A, Ciechanover A. *Annu Rev Biochem*, 1998, 67: 425–479
- Mukhopadhyay D, Riezman H. *Science*, 2007, 315: 201–205
- West MHP, Bonner WM. *Nucl Acids Res*, 1980, 8: 4671–4680
- Robzyk K, Recht J, Osley MA. *Science*, 2000, 287: 501–504
- Wang H, Zhai L, Xu J, Joo HY, Jackson S, Erdjument-Bromage H, Tempst P, Xiong Y, Zhang Y. *Mol Cell*, 2006, 22: 383–394
- Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. *Nature*, 2004, 431: 873–878
- Hwang WW, Venkatasubrahmanyam S, Ianculescu AG, Tong A, Boone C, Madhani HD. *Mol Cell*, 2003, 11: 261–266
- Mattioli F, Vissers JHA, van Dijk WJ, Ikpa P, Citterio E, Vermeulen W, Martijn JA, Sixma TK. *Cell*, 2012, 150: 1182–1195
- Nishiyama A, Yamaguchi L, Sharif J, Johmura Y, Kawamura T, Nakanishi K, Shimamura S, Arita K, Kodama T, Ishikawa F, Koseki H, Nakanishi M. *Nature*, 2013, 502: 249–253
- Yan Q, Dutt S, Xu R, Graves K, Juszczynski P, Manis JP, Shipp MA. *Mol Cell*, 2009, 36: 110–120
- Sun ZW, Allis CD. *Nature*, 2002, 418: 104–108
- Briggs SD, Xiao T, Sun ZW, Caldwell JA, Shabanowitz J, Hunt DF, Allis CD, Strahl BD. *Nature*, 2002, 418: 498
- Hyun K, Jeon J, Park K, Kim J. *Exp Mol Med*, 2017, 49: e324
- Suganuma T, Workman JL. *Cell*, 2008, 135: 604–607
- Krajewski WA, Li J, Dou Y. *Nucleic Acids Res*, 2018, 46: 7631–7642
- Debelouchina GT, Gerecht K, Muir TW. *Nat Chem Biol*, 2017, 13: 105–110
- Fierz B, Chatterjee C, McGinty RK, Bar-Dagan M, Raleigh DP, Muir TW. *Nat Chem Biol*, 2011, 7: 113–119
- Morgan MT, Haj-Yahya M, Ringel AE, Bandi P, Brik A, Wolberger C. *Science*, 2016, 351: 725–728
- Wilson MD, Benlekbi S, Fradet-Turcotte A, Sherker A, Julien JP, McEwan A, Noordermeer SM, Sicheri F, Rubinstein JL, Durocher D. *Nature*, 2016, 536: 100–103
- Sundaramoorthy R, Hughes AL, El-Mkami H, Norman DG, Ferreira H, Owen-Hughes T. *eLife*, 2018, 7: e35720
- Worden EJ, Hoffmann NA, Hicks CW, Wolberger C. *Cell*, 2019, 176: 1490–1501.e12
- Anderson CJ, Baird MR, Hsu A, Barbour EH, Koyama Y, Borgnia MJ, McGinty RK. *Cell Rep*, 2019, 26: 1681–1690.e5
- Jang S, Kang C, Yang HS, Jung T, Hebert H, Chung KY, Kim SJ, Hohng S, Song JJ. *Genes Dev*, 2019, 33: 620–625
- Valencia-Sánchez MI, De Ioannes P, Wang M, Vasilyev N, Chen R, Nudler E, Armache JP, Armache KJ. *Mol Cell*, 2019, 74: 1010–1019.e6

- 34 Yao T, Jing W, Hu Z, Tan M, Cao M, Wang Q, Li Y, Yuan G, Lei M, Huang J. *Cell Res*, 2019, 29: 330–333
- 35 Xue H, Yao T, Cao M, Zhu G, Li Y, Yuan G, Chen Y, Lei M, Huang J. *Nature*, 2019, 573: 445–449
- 36 Hsu PL, Shi H, Leonen C, Kang J, Chatterjee C, Zheng N. *Mol Cell*, 2019, 76: 712–723.e4
- 37 Worden EJ, Zhang X, Wolberger C. *eLife*, 2020, 9: e53199
- 38 Bilokapic S, Halic M. *Nat Commun*, 2019, 10: 3795
- 39 Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D. *Mol Cell*, 2005, 20: 601–611
- 40 Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, Shilatifard A, Muir TW, Roeder RG. *Cell*, 2009, 137: 459–471
- 41 Qi YK, Ai HS, Li YM, Yan B. *Front Chem*, 2018, 6: 19
- 42 Maity SK, Jbara M, Brik A. *J Pept Sci*, 2016, 22: 252–259
- 43 Fang GM, Li YM, Shen F, Huang YC, Li JB, Lin Y, Cui HK, Liu L. *Angew Chem Int Ed*, 2011, 50: 7645–7649
- 44 Fang GM, Wang JX, Liu L. *Angew Chem Int Ed*, 2012, 51: 10347–10350
- 45 Tang S, Liang LJ, Si YY, Gao S, Wang JX, Liang J, Mei Z, Zheng JS, Liu L. *Angew Chem Int Ed*, 2017, 56: 13333–13337
- 46 Wang JX, Fang GM, He Y, Qu DL, Yu M, Hong ZY, Liu L. *Angew Chem Int Ed*, 2015, 54: 2194–2198
- 47 Zheng JS, Tang S, Qi YK, Wang ZP, Liu L. *Nat Protoc*, 2013, 8: 2483–2495
- 48 He Q, Li J, Qi Y, Wang Z, Huang Y, Liu L. *Sci China Chem*, 2017, 60: 621–627
- 49 Siman P, Karthikeyan SV, Nikolov M, Fischle W, Brik A. *Angew Chem Int Ed*, 2013, 52: 8059–8063
- 50 Zhang B, Deng Q, Zuo C, Yan B, Zuo C, Cao XX, Zhu TF, Zheng JS, Liu L. *Angew Chem Int Ed*, 2019, 58: 12231–12237
- 51 Li J, Li Y, He Q, Li Y, Li H, Liu L. *Org Biomol Chem*, 2014, 12: 5435–5441
- 52 Ai H, Guo Y, Sun D, Liu S, Qi Y, Guo J, Qu Q, Gong Q, Zhao S, Li J, Liu L. *ChemBioChem*, 2019, 20: 221–229
- 53 Muir TW, Sondhi D, Cole PA. *Proc Natl Acad Sci USA*, 1998, 95: 6705–6710
- 54 Si Y, Liang L, Tang S, Qi Y, Huang Y, Liu L. *Sci China Chem*, 2018, 61: 412–417
- 55 McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW. *Nature*, 2008, 453: 812–816
- 56 McGinty RK, Köhn M, Chatterjee C, Chiang KP, Pratt MR, Muir TW. *ACS Chem Biol*, 2009, 4: 958–968
- 57 Holt M, Muir T. *Annu Rev Biochem*, 2015, 84: 265–290
- 58 Fierz B, Kilic S, Hieb AR, Luger K, Muir TW. *J Am Chem Soc*, 2012, 134: 19548–19551
- 59 Chu GC, Pan M, Li J, Liu S, Zuo C, Tong ZB, Bai JS, Gong Q, Ai H, Fan J, Meng X, Huang YC, Shi J, Deng H, Tian C, Li YM, Liu L. *J Am Chem Soc*, 2019, 141: 3654–3663
- 60 Chatterjee C, McGinty RK, Fierz B, Muir TW. *Nat Chem Biol*, 2010, 6: 267–269
- 61 Long L, Furgason M, Yao T. *Methods*, 2014, 70: 134–138
- 62 Morgan MT, Wolberger C. *Curr Opin Struct Biol*, 2017, 42: 75–82
- 63 Bonnet J, Devys D, Tora L. *Drug Discovery Today-Technol*, 2014, 12: e19–e27
- 64 Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NCT, Kao CF, Pillus L, Shilatifard A, Osley MA, Berger SL. *Genes Dev*, 2003, 17: 2648–2663
- 65 Papai G, Frechard A, Kolesnikova O, Crucifix C, Schultz P, Ben-Shem A. *Nature*, 2020, 577: 711–716
- 66 Köhler A, Schneider M, Cabal GG, Nehrbass U, Hurt E. *Nat Cell Biol*, 2008, 10: 707–715
- 67 Samara NL, Datta AB, Berndsen CE, Zhang X, Yao T, Cohen RE, Wolberger C. *Science*, 2010, 328: 1025–1029
- 68 Wang H, Dienemann C, Stützer A, Urlaub H, Cheung ACM, Cramer P. *Nature*, 2020, 577: 717–720
- 69 Durand A, Bonnet J, Fournier M, Chavant V, Schultz P. *Structure*, 2014, 22: 1553–1559
- 70 Barbera AJ, Chodaparambil JV, Kelley-Clarke B, Joukov V, Walter JC, Luger K, Kaye KM. *Science*, 2006, 311: 856–861
- 71 Koehler C, Bonnet J, Stierle M, Romier C, Devys D, Kieffer B. *J Biol Chem*, 2014, 289: 8989–8999
- 72 Lukas J, Lukas C, Bartek J. *Nat Cell Biol*, 2011, 13: 1161–1169
- 73 Panier S, Boulton SJ. *Nat Rev Mol Cell Biol*, 2014, 15: 7–18
- 74 Hu Q, Botuyan MV, Cui G, Zhao D, Mer G. *Mol Cell*, 2017, 66: 473–487.e9
- 75 Fradet-Turcotte A, Canny MD, Escribano-Díaz C, Orthwein A, Leung CCY, Huang H, Landry MC, Kiteviski-LeBlanc J, Noordermeer SM, Sicheri F, Durocher D. *Nature*, 2013, 499: 50–54
- 76 Botuyan MV, Lee J, Ward IM, Kim JE, Thompson JR, Chen J, Mer G. *Cell*, 2006, 127: 1361–1373
- 77 Lowary PT, Widom J. *J Mol Biol*, 1998, 276: 19–42

- 78 Simon MD, Chu F, Racki LR, de la Cruz CC, Burlingame AL, Panning B, Narlikar GJ, Shokat KM. *Cell*, 2007, 128: 1003–1012
- 79 Mattioli F, Uckelmann M, Sahtoe DD, van Dijk WJ, Sixma TK. *Nat Commun*, 2014, 5: 3291
- 80 Baarends WM, Wassenaar E, Hoogerbrugge JW, Schoenmakers S, Sun ZW, Grootegoed JA. *J Cell Sci*, 2007, 120: 1841–1851
- 81 Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D. *Cell*, 2006, 125: 703–717
- 82 Zeng M, Ren L, Mizuno KI, Nestoras K, Wang H, Tang Z, Guo L, Kong D, Hu Q, He Q, Du L, Carr AM, Liu C. *Nat Commun*, 2016, 7: 11364
- 83 Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. *Cell*, 2013, 154: 490–503
- 84 Simic R, Lindstrom DL, Tran HG, Roinick KL, Costa PJ, Johnson AD, Hartzog GA, Arndt KM. *EMBO J*, 2003, 22: 1846–1856
- 85 Pointner J, Persson J, Prasad P, Norman-Axelsson U, Strålfors A, Khorosjutina O, Krietenstein N, Peter Svensson J, Ekwall K, Korber P. *EMBO J*, 2012, 31: 4388–4403
- 86 Gkikopoulos T, Schofield P, Singh V, Pinskaya M, Mellor J, Smolle M, Workman JL, Barton GJ, Owen-Hughes T. *Science*, 2011, 333: 1758–1760
- 87 Fleming AB, Kao CF, Hillyer C, Pikaart M, Osley MA. *Mol Cell*, 2008, 31: 57–66
- 88 Strahl BD, Allis CD. *Nature*, 2000, 403: 41–45
- 89 Levendosky RF, Sabantsev A, Deindl S, Bowman GD. *eLife*, 2016, 5: e21356
- 90 Sharma A, Jenkins KR, Héroux A, Bowman GD. *J Biol Chem*, 2011, 286: 42099–42104
- 91 Ryan DP, Sundaramoorthy R, Martin D, Singh V, Owen-Hughes T. *EMBO J*, 2011, 30: 2596–2609
- 92 Farnung L, Vos SM, Wigge C, Cramer P. *Nature*, 2017, 550: 539–542
- 93 Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. *Mol Cell Biol*, 2005, 25: 3305–3316
- 94 Hauk G, McKnight JN, Nodelman IM, Bowman GD. *Mol Cell*, 2010, 39: 711–723
- 95 Gorbalenya AE, Koonin EV. *Curr Opin Struct Biol*, 1993, 3: 419–429
- 96 Ferreira H, Flaus A, Owen-Hughes T. *J Mol Biol*, 2007, 374: 563–579
- 97 Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, Vakoc AL, Kim JE, Chen J, Lazar MA, Blobel GA, Vakoc CR. *Mol Cell Biol*, 2008, 28: 2825–2839
- 98 Giannattasio M, Lazzaro F, Plevani P, Muzi-Falconi M. *J Biol Chem*, 2005, 280: 9879–9886
- 99 Huyen Y, Zgheib O, DiTullio Richard A. J, Gorgoulis VG, Zacharatos P, Petty TJ, Sheston EA, Mellert HS, Stavridi ES, Halazonetis TD. *Nature*, 2004, 432: 406–411
- 100 Ng HH, Feng Q, Wang H, Erdjument-Bromage H, Tempst P, Zhang Y, Struhl K. *Genes Dev*, 2002, 16: 1518–1527
- 101 Worden EJ, Wolberger C. *Curr Opin Struct Biol*, 2019, 59: 98–106
- 102 Ng HH, Xu RM, Zhang Y, Struhl K. *J Biol Chem*, 2002, 277: 34655–34657
- 103 Singer MS, Kahana A, Wolf AJ, Meisinger LL, Peterson SE, Goggin C, Mahowald M, Gottschling DE. *Genetics*, 1998, 150: 613–632
- 104 Sawada K, Yang Z, Horton JR, Collins RE, Zhang X, Cheng X. *J Biol Chem*, 2004, 279: 43296–43306
- 105 Kuntimaddi A, Achille NJ, Thorpe J, Lokken AA, Singh R, Hemenway CS, Adli M, Zeleznik-Le NJ, Bushweller JH. *Cell Rep*, 2015, 11: 808–820
- 106 Min J, Feng Q, Li Z, Zhang Y, Xu RM. *Cell*, 2003, 112: 711–723
- 107 Bernt KM, Zhu N, Sinha AU, Vempati S, Faber J, Krivtsov AV, Feng Z, Punt N, Daigle A, Bullinger L, Pollock RM, Richon VM, Kung AL, Armstrong SA. *Cancer Cell*, 2011, 20: 66–78
- 108 Chen CW, Armstrong SA. *Exp Hematol*, 2015, 43: 673–684
- 109 McLean CM, Karemaker ID, van Leeuwen F. *Leukemia*, 2014, 28: 2131–2138
- 110 Holt MT, David Y, Pollock S, Tang Z, Jeon J, Kim J, Roeder RG, Muir TW. *Proc Natl Acad Sci USA*, 2015, 112: 10365–10370
- 111 Altaf M, Utley RT, Lacoste N, Tan S, Briggs SD, Côté J. *Mol Cell*, 2007, 28: 1002–1014
- 112 Fingerman IM, Li HC, Briggs SD. *Genes Dev*, 2007, 21: 2018–2029
- 113 Isom DG, Castaneda CA, Cannon BR, Garcia-Moreno E. B. *Proc Natl Acad Sci USA*, 2011, 108: 5260–5265
- 114 Sinha KK, Gross JD, Narlikar GJ. *Science*, 2017, 355: eaaa3761
- 115 Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A. *J Biol Chem*, 2002, 277: 28368–28371
- 116 Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. *Nat Cell Biol*, 2004, 6: 73–77
- 117 Wang P, Lin C, Smith ER, Guo H, Sanderson BW, Wu M, Gogol M, Alexander T, Seidel C, Wiedemann LM, Ge K, Krumlauf R, Shilatifard A.

- Mol Cell Biol*, 2009, 29: 6074–6085
- 118 Hu D, Gao X, Morgan MA, Herz HM, Smith ER, Shilatifard A. *Mol Cell Biol*, 2013, 33: 4745–4754
 - 119 Eissenberg JC, Shilatifard A. *Dev Biol*, 2010, 339: 240–249
 - 120 Wu M, Wang PF, Lee JS, Martin-Brown S, Florens L, Washburn M, Shilatifard A. *Mol Cell Biol*, 2008, 28: 7337–7344
 - 121 Yang W, Ernst P. *Int J Hematol*, 2017, 105: 7–16
 - 122 Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. *Cell*, 2007, 129: 823–837
 - 123 Sze CC, Cao K, Collings CK, Marshall SA, Rendleman EJ, Ozark PA, Chen FX, Morgan MA, Wang L, Shilatifard A. *Genes Dev*, 2017, 31: 1732–1737
 - 124 Shilatifard A. *Annu Rev Biochem*, 2012, 81: 65–95
 - 125 Krivtsov AV, Armstrong SA. *Nat Rev Cancer*, 2007, 7: 823–833
 - 126 Sze CC, Shilatifard A. *Cold Spring Harb Perspect Med*, 2016, 6: <https://doi.org/10.1101/cshperspect.a026427>–833
 - 127 Kim J, Kim JA, McGinty RK, Nguyen UTT, Muir TW, Allis CD, Roeder RG. *Mol Cell*, 2013, 49: 1121–1133
 - 128 Qu Q, Takahashi YH, Yang Y, Hu H, Zhang Y, Brunzelle JS, Couture JF, Shilatifard A, Skiniotis G. *Cell*, 2018, 174: 1117–1126.e12
 - 129 Hsu PL, Li H, Lau HT, Leonen C, Dhall A, Ong SE, Chatterjee C, Zheng N. *Cell*, 2018, 174: 1106–1116.e9
 - 130 Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. *J Biol Chem*, 2005, 280: 17732–17736
 - 131 Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, Allis CD. *Mol Cell Biol*, 2002, 22: 1298–1306
 - 132 Li B, Howe LA, Anderson S, Yates Iii JR, Workman JL. *J Biol Chem*, 2003, 278: 8897–8903
 - 133 Li J, Moazed D, Gygi SP. *J Biol Chem*, 2002, 277: 49383–49388
 - 134 Vlaming H, van Welssem T, de Graaf EL, Ontoso D, Altelaar AFM, San-Segundo PA, Heck AJR, van Leeuwen F. *EMBO Rep*, 2014, 15: 1077–1084
 - 135 Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li GM. *Cell*, 2013, 153: 590–600
 - 136 Pai CC, Deegan RS, Subramanian L, Gal C, Sarkar S, Blaikley EJ, Walker C, Hulme L, Bernhard E, Codlin S, Bähler J, Allshire R, Whitehall S, Humphrey TC. *Nat Commun*, 2014, 5: 4091
 - 137 Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL. *Nature*, 2012, 489: 452–455
 - 138 Landry J, Sutton A, Hesman T, Min J, Xu RM, Johnston M, Sternglanz R. *Mol Cell Biol*, 2003, 23: 5972–5978
 - 139 Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, Davies H, Edkins S, Hardy C, Latimer C, Teague J, Andrews J, Barthorpe S, Beare D, Buck G, Campbell PJ, Forbes S, Jia M, Jones D, Knott H, Kok CY, Lau KW, Leroy C, Lin ML, McBride DJ, Maddison M, Maguire S, McLay K, Menzies A, Mironenko T, Mulderrig L, Mudie L, O'Meara S, Pleasance E, Rajasingham A, Shepherd R, Smith R, Stebbings L, Stephens P, Tang G, Tarpey PS, Turrell K, Dykema KJ, Khoo SK, Petillo D, Wondergem B, Anema J, Kahnoski RJ, Teh BT, Stratton MR, Futreal PA. *Nature*, 2010, 463: 360–363
 - 140 Lu C, Jain SU, Hoelper D, Bechet D, Molden RC, Ran L, Murphy D, Venneti S, Hameed M, Pawel BR, Wunder JS, Dickson BC, Lundgren SM, Jani KS, De Jay N, Papillon-Cavanagh S, Andrulis IL, Sawyer SL, Grynspan D, Turcotte RE, Nadaf J, Fahiminiyah S, Muir TW, Majewski J, Thompson CB, Chi P, Garcia BA, Allis CD, Jabado N, Lewis PW. *Science*, 2016, 352: 844–849
 - 141 Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, Wedge DC, Cooke SL, Gundem G, Davies H, Nik-Zainal S, Martin S, McLaren S, Goodie V, Robinson B, Butler A, Teague JW, Halai D, Khatri B, Myklebost O, Baumhoer D, Jundt G, Hamoudi R, Tirabosco R, Amary MF, Futreal PA, Stratton MR, Campbell PJ, Flanagan AM. *Nat Genet*, 2013, 45: 1479–1482
 - 142 Zhang Y, Shan CM, Wang J, Bao K, Tong L, Jia S. *Sci Rep*, 2017, 7: 43906
 - 143 Nelson CJ, Santos-Rosa H, Kouzarides T. *Cell*, 2006, 126: 905–916
 - 144 Yuan G, Ma B, Yuan W, Zhang Z, Chen P, Ding X, Feng L, Shen X, Chen S, Li G, Zhu B. *J Biol Chem*, 2013, 288: 30832–30842
 - 145 Murata K, Wolf M. *Biochim Biophys Acta*, 2018, 1862: 324–334
 - 146 Yang J, Zhao J. *Sci China Chem*, 2018, 61: 97–112
 - 147 Bi X, Pasunooti KK, Liu CF. *Sci China Chem*, 2018, 61: 251–265
 - 148 Qi YK, Si YY, Du SS, Liang J, Wang KW, Zheng JS. *Sci China Chem*, 2019, 62: 299–312

Advances on structural mechanisms of ubiquitinated nucleosome complexes

Ziyu Xu, Huasong Ai, Maoshen Sun, Jiakun Tian, Lei Liu^{*}

Tsinghua-Peking Center for Life Sciences, Ministry of Education Key Laboratory of Bioorganic Phosphorus Chemistry and Chemical Biology, Department of Chemistry, Tsinghua University, Beijing 100084, China

**Corresponding author (email: lliu@mail.tsinghua.edu.cn)*

Abstract: The ubiquitination of nucleosome histones (H2A, H2B, H3, H4) mediates a series of selective interactions between effector proteins and nucleosomes, and regulates the epigenetics of cells. In recent years, thanks to the development of new chemical and biological methods, sample acquisition technologies such as chemical modification of proteins, and structure observation technologies such as cryo-electron microscopy (Cryo-EM) have been well utilized. Some complex structures of ubiquitinated nucleosomes and effector proteins have been resolved, which revealed molecular mechanisms of ubiquitin regulated epigenetic process. This article summarizes the six complexes between ubiquitinated nucleosomes and effector proteins reported recently, comparatively analyzes the selective recognition and activation mechanism of effector proteins by ubiquitinated nucleosomes, and describes the key interaction sites on ubiquitinated nucleosomes.

Keywords: post-translational modification, ubiquitin, nucleosome, modification crosstalk, structural biology, chromatin

doi: [10.1360/SSC-2019-0170](https://doi.org/10.1360/SSC-2019-0170)