

植物同源结构域(PHD 结构域) ——组蛋白密码的解读器*

马红辉 方存磊 曾平耀**

(复旦大学生物医学研究院, 上海 200032)

摘要 植物同源结构域(plant homeodomain, PHD 结构域), 是真核生物中一种进化保守的锌指结构域。多种调控基因转录、细胞周期、凋亡的蛋白质含有 PHD 结构域。研究表明, PHD 结构域涉及多种功能, 包括蛋白质相互作用, 特别是同核小体组蛋白的作用。目前认为, 各种组蛋白修饰(包括甲基化、乙酰化、磷酸化、泛素化等)的模式和组合, 调节染色质状态和基因转录活性, 并提出了组蛋白密码理论。PHD 指结构域能特异性识别组蛋白的甲基化(修饰)密码, 可能是组蛋白密码的一种重要解读器。

关键词 PHD 结构域, 组蛋白密码, 解读器, 组蛋白甲基化

学科分类号 Q7

1 PHD 结构域的发现

1993 年, Schinder 首次发现了植物同源结构域(plant homeodomain, PHD 结构域)。他注意到在拟南芥蛋白 HAT3.1 和 HOXIA 内有一段富含半胱氨酸的保守序列, 这段序列同金属离子结合结构域(metal-binding domains)非常相似^[1]。进一步研究发现, PHD 指结构域是 14 种已知的锌指结构域(zinc-binding motif)中的一种, 存在于 400 多种真核生物蛋白质中, 在进化过程中高度保守^[2, 3]。

2 PHD 结构域的序列和结构特征

PHD 结构域属于“cross-brace”锌指蛋白家族(cross-brace zinc finger protein), 它由约 60 个氨基酸残基组成, 具有典型的 C4HC3(cys4-His-cys3)锌结合基序(zinc binding motif)^[3], 其半胱氨酸之间的间距相对保守^[4](图 1a), 更显著的是, 在最后 2 个半胱氨酸之前存在一个色氨酸或其他芳香族氨基酸残基^[5]。在 PHD 结构域内, 7 个半胱氨酸残基和 1 个组氨酸残基结合 2 个锌离子而使该结构域得以稳定, 并形成了 2 个环(Loop1 和 Loop2)^[6](图 1b)。

在各种 PHD 结构域蛋白中, 其 PHD 三维结构大致相同, 为球状结构域(globular domain)^[7]。但随不同 PHD 结构域氨基酸残基序列的差异, 其三维结构也会有细微差异, 正是这种差异性, 赋予了 PHD 结构域家族成员不同的生物学功能。如, ISWI/SNF-染色质重构复合物(ISWI/SNF-chromatin remodeling complex)的一个亚单位 ACF1 的 2 个 PHD 指结构域具有核小体结合活性, 可以结合到 4 个核心组蛋白的中心区域, 并增加 ISWI 引起的核小体滑行的效率^[8, 9]。转录共活化物(transcription coactivator)Pygopus 的 PHD 结构域识别并结合到 Legless/BCL9 的 E2 连接酶的区域, 这 2 个蛋白质间的相互作用是 Wnt 信号系统的一个重要步骤^[6, 10]。原癌蛋白 c-Cbl 和 MEKK1(MAP kinase/ERK kinase)的 PHD 结构域具有 E3 泛素连接酶活性, 介导 ERK1/2 的泛素化和降解^[11]。

* 国家自然科学基金资助(30671188)。

** 通讯联系人。Tel: 021-54237598, Fax: 021-54237876

E-mail: pingyaozeng@hotmail.com

收稿日期: 2007-09-25, 接受日期: 2007-12-27

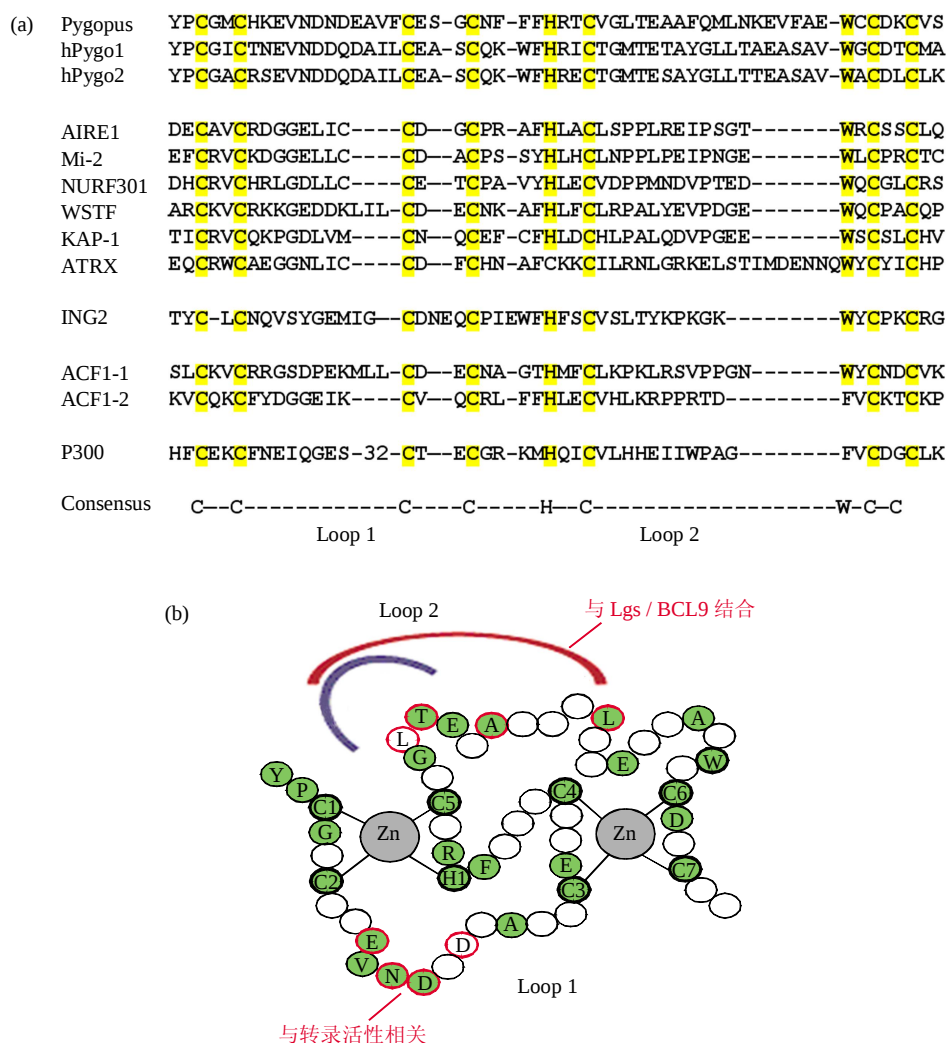


Fig. 1 The sequence signature of PHD fingers (a)^[49] and cross-brace model of the PHD finger of pygopus (b)^[6]

图 1 不同蛋白质 PHD 结构域的序列对比 (a)^[49] 及 pygopus 蛋白 PHD 结构域的 cross-brace 模型 (b)^[6]

(a)每个半胱氨酸的跨度都相对保守, Loop1 和 Loop2 中的氨基酸与 PHD 结构域功能高度相关。(b)该模型显示通过 2 个锌离子稳定的 cross-brace 拓扑结构, 绿色表示 pygopus 家族直系同源的保守残基, 红圈标记的 Loop1 和 Loop2 表示与功能相关的氨基酸, 其中 C、H、W 分别代表半胱氨酸, 组氨酸和色氨酸, 在 Loop2 中的蓝色弧形表示 c-Cbl 识别 E2 连接酶的区域。

3 PHD 结构域对组蛋白密码的解读功能

构成染色质的基本单位是核小体, 它由核心组蛋白八聚体和绕其约 2 周的 146 bp 双链 DNA 组成, 核心组蛋白八聚体中央是由组蛋白 H3 和 H4 形成的异四聚体, 侧翼是组蛋白 H2A 和 H2B 形成的异二聚体^[12]。每个组蛋白的氨基端约 20~35 个富含基本氨基酸的片段延伸到核小体的表面(在组蛋白 H2A, 其羧基端也有约 37 个氨基酸伸展到核小体外), 这些突出于核小体的部分, 称为“组蛋

白尾”(histone tail)^[13]。组蛋白尾在空间结构上相对可变, 并与核内各种蛋白质或酶直接接触而被修饰。组蛋白尾的共价修饰, 包括乙酰化、磷酸化、甲基化、泛素化等^[14](图 2), 影响染色质结构的重构和 DNA 与蛋白质的相互作用, 对染色质状态的形成和维持、基因转录表达调控、DNA 损伤反应和修复、细胞记忆、以及多余 DNA 的清除等多种细胞生命活动具有至关重要的作用, 从而提出了组蛋白密码(histone code)理论^[15]。

Fig. 2 Post-translational modifications of the core histones^[14]图 2 各种组蛋白尾巴的修饰^[14]

p=磷酸化, AC=乙酰化, Ub=泛素化, Me=甲基化.

一般说来, 组蛋白赖氨酸残基的乙酰化(acetylation)修饰, 无论发生在哪个位置都与基因表达活化相关. 组蛋白的甲基化修饰, 依其发生位点、状态(无甲基化、单甲基化、双甲基化、三甲基化)以及与其他修饰的联合谱式不同, 分别能激活或抑制基因的转录. 组蛋白修饰导致生物学效应是怎样产生的呢^[16]? 一种假说认为是通过下游效应蛋白(effector protein)^[17]特异的识别和解译这种修饰来完成的. 近来, 有关识别和结合特定修饰组蛋白赖氨酸残基的效应蛋白的鉴定, 使这一假说得到了有力的支持. 例如, 含 chromo 结构域的异染色质蛋白-1(heterochromatin protein-1, HP1)和 Polycomb 蛋白被证明能分别识别和结合 H3K9 和 H3K27, 含 bromo 结构域的蛋白质特异识别乙酰化的赖氨酸^[18]. 由于这些结构域特异性识别组蛋白密码(修饰), 并发挥生物效应, 所以称之为“组蛋白密码的解读器(histone code reader)”. 最近研究发现, 含 PHD 结构域的蛋白质能特异地识别结合不同甲基化状态的 H3K4^[19-21], 这一发现不仅揭示了 PHD 的新功能, 而且对组蛋白 H3K4 甲基化的功能研究提供了新的视角.

3.1 PHD 结构域——组蛋白密码解读器的发现

斯坦福大学的 Gozani 小组在对肿瘤抑制蛋白 ING(inhibitor of growth)家族的 PHD 结构域进行了研究, 通过 ChIP(chromatin immunoprecipitation)检测, 发现 ING PHD 结构域能紧密地同 H3K4me3 和 H3K4me2 结合. 进一步地研究发现, 在 DNA 损伤和遗传毒性损伤(genotoxic insults)反应时, ING2 PHD 结构域识别 H3K4me3 并将复合物

mSin3a-HDAC1 稳定在 cyclin D1 的启动子上, 从而抑制活性基因的转录^[22]. 此外, ING 家族其他成员, 如酵母 ING 家族的成员 Yng1、Yng2 和 Pho23 的 PHD 结构域都具有很强的 H3K4me2/3 结合活性^[19]. 洛克菲勒大学的 Allis 小组的研究也证明了 PHD 结构域作为组蛋白密码解读器的功能. 利用质谱和染色质免疫共沉淀技术, 他们筛选到一个与 H3K4me3 结合的 BPTF(bromodomain and PHD finger transcription factor), BPTF 同时包含 PHD 和 bromo 两个结构域. 它是 ATP 依赖的核小体重构因子 NURF(nucleosome remodelling factor)中最大的亚单位^[23,24]. 在发育过程中, 通过 BPTF 与 H3K4me3 的结合, NURF 复合物能维持 HOX 基因的表达模式, 去除 H3K4me3, 会引起 NURF 复合物的亚单位 BPTF 从染色质中释放出来, 导致 NURF 复合物不能募集相关的 ATPase——SNF2L(也称之为 ISWI 和 SMARX1)到 HOXC8 的启动子上. 缺失分析证明, BPTF/NURF301 中 bromo 结构域附近的 PHD 结构域在与 H3K4me3 结合时起决定作用^[20].

这两个小组的研究表明, PHD 结构域能特异识别并结合组蛋白 H3K4 的甲基化密码(修饰), 从而发挥基因转录调控功能. 那么 ING2 的 PHD 结构域和 BPTF 的 PHD 结构域在对 H3K4me3 识别的分子机理上有哪些异同呢? 相关的研究结果, 解答了这一问题.

3.2 PHD 结构域对 H3K4me3 识别和结合的分子机制

Kutateladze 小组利用核磁共振技术对 ING2 的 PHD 结构域作了详细的解析, 在 0.2Å 的分辨率

下, 他们发现该结构域 215 位酪氨酸(Tyr215, Y215)和 238 位色氨酸(Trp238, W238)两个芳香族氨基酸的侧链通过 H 键与 H3K4me3 相互作用, 将 Y215 突变为亮氨酸后, 结合活性几乎丧失, 而 W238 的突变导致不能区分组蛋白 4 位的赖氨酸(K4)和 2 位的精氨酸(R2)(图 3a)^[25]. 这种双芳香族氨基酸残基识别甲基化的赖氨酸的模式在以前研究 CHD1 的 chromo 结构域识别特性时也报道过^[26]. 虽然 ING2 的 PHD 结构域和 CHD1 的 chromo 结构

域结构上完全不相关, 但在识别结合 H3K4me3 时存在高度相似, 组蛋白 4 位的赖氨酸和 2 位的精氨酸占据邻近的两个沟, 并都是被酪氨酸残基分开. 与此同时, Allis 等对 NURF 复合物中 BPTF 的 PHD 结构域结合识别的分子机制也作了同样的研究, 发现与 ING2 PHD 的结构非常相似^[27], 唯一不同是在 BPTF 的 PHD 结构域中, 4 个芳香族氨基酸(Y10, Y17, Y23 和 W32)组成的笼样结构来识别结合 H3K4me3(图 3b)^[18].

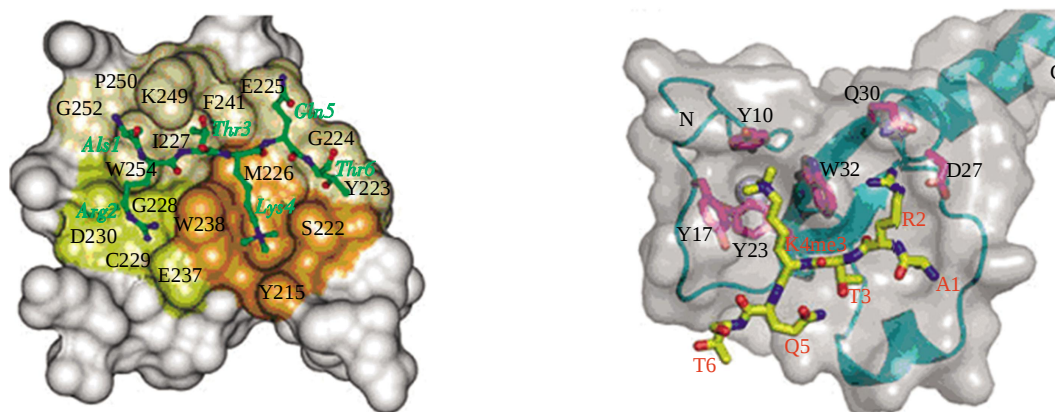


Fig. 3 Structure of ING2 (a) and BPTF (b) PHD finger in complex with a histone H3 peptide trimethylated at Lys4

图 3 ING2 (a) 和 BPTF (b) PHD 结合 H3K4me3 的结构

(a)H3 的 Lys4 和 Arg2 被 w238 分开, 分别占据两个结合沟(用棕色和黄色标记), 组蛋白的肽用球-棒模型显示, 绿, 红, 蓝分别代表 C, O, N 原子^[25]. (b)BPTF PHD 的 Y10, Y17, Y23, W32 四个芳香族氨基酸形成一个 H3K4me3 结合的笼, 黄, 红, 蓝分别代表 C, O, N 原子^[18].

3.3 PHD 结构域与其他密码解读结构域的组合

在组蛋白 H3K4 甲基化密码解读复合物中, 经常还有其他组蛋白密码识别的结构域, 它们可以在复合物的同一亚单位中, 也可以在复合物的不同亚单位中^[28]. 例如, PHD 结构域和 bromo 结构域共存于很多染色质相关的蛋白质中, 并在功能上相

互协调^[29, 30]. BPTF 的 bromo 结构域和 PHD 结构域能分别识别乙酰化和甲基化的组蛋白尾巴, 这表明在识别既甲基化又乙酰化(trimethyl-acetyl)的修饰模式时, 他们两者可能存在合作, 从而调节 BPTF 的功能^[31](图 4a). 此外, 酵母的 NuA4 乙酰转移酶复合物中的 2 个亚单位 Yng2 和 Eaf3, 分别含有 PHD

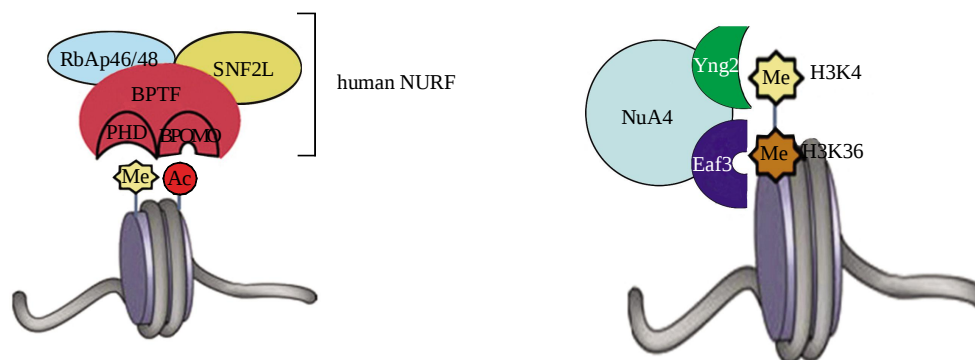


Fig. 4 BPTF of the NURF complex (a) and Yng2 and Eaf3 in NuA4 complex (b)

图 4 NURF 复合物中的 BPTF 蛋白 (a) 和 NuA4 复合物中的 Yng2 和 Eaf3 蛋白 (b)

(a) BPTF 中的 PHD 和 BROMO 结构域, 他们共存于 BPTF 中, 却分别与甲基化和乙酰化的组蛋白密码结合^[31]. (b) NuA4 复合物中的 Yng2 和 Eaf3 亚单位, Yng2 上的 PHD 结构域和 Eaf3 上的 chromo 结构域, 分别结合甲基化 H3K4 和 H3K36^[32].

结构域和 chromo 结构域, 它们分别识别组蛋白尾甲基化的 4 位(K4)和 36 位赖氨酸(K36)^[32] (图 4b)。最近还发现, Rpd3S HDAC 复合物(Rpd3S histone deacetylase complex)中的 Eaf3 和 Rco1 亚单位, 也分别含有 PHD 结构域和 chromo 结构域, 它们协同结合甲基化的 H3K36, 将 Rpd3S HDAC 复合物介导到转录后的染色质上^[33]。可见, 这种组合能使不同的组蛋白密码得到更精确的解读。

PHD 结构域蛋白 BHC80 是 LSD1 复合物(lysine specific demethylase 1 complex)中的一个亚单位, 研究发现, 其 PHD 结构域特异性识别结合无甲基化的 H3K4(H3K4me0)^[21]。LSD1 去甲基化修饰作用于组蛋白 H3K4me1/2 后形成 H3K4me0, BHC80 与其结合而参与 LSD1 的下游功能。晶体结构分析显示 BHC80 的 PHD 结构域, 与 ING2 和 BPTF 的 PHD 结构域具有高度相似的折叠, 但不形成笼样结构, 提示 PHD 作为结合模块在结构上的灵活性和柔韧性^[21]。此外, RAG1/2 V(D)J 重构酶的重要组分 RAG2, 也包含 PHD 结构域, 并通过它特异识别 H3K4me3, 调节抗原受体基因(antigen-receptor gene)V(D)J 的重排^[34]。

4 结语与展望

PHD 结构域广泛地存在于人类的基因组中, 利用 SMART 软件分析有 180 多种蛋白质含此结构域, 其中大部分蛋白质的功能尚未知。目前研究提示, PHD 结构域识别结合组蛋白甲基化密码, 参与核内不同的生物学进程, 如转录调控, 细胞周期, 凋亡等, 是组蛋白密码的一种解读器。因此, 广泛深入地对 PHD 结构域蛋白功能的鉴定, 将进一步阐明组蛋白密码的解读机制。

参 考 文 献

- Schindler U, Beckmann H, Cashmore A R. HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J*, 1993, 4(1): 137~150
- Breen T R, Chinwalla V, Harte P J. Trithorax is required to maintain engrailed expression in a subset of engrailed-expressing cells. *Mech Dev*, 1995, 52(1): 89~98
- Kaadige M R, Ayer D E. The polybasic region that follows the plant homeodomain zinc finger 1 of Pfl1 is necessary and sufficient for specific phosphoinositide binding. *J Biol Chem*, 2006, 281(39): 28831~28836
- Bienz M. The PHD finger, a nuclear protein-interaction domain. *Trends in Biochemical Sciences*, 2006, 31(1): 35~40
- Aasland R, Gibson T J, Stewart A F. The PHD finger: implications for chromatin mediated transcriptional regulation. *Trends Biochem Sci*, 1995, 20(2): 56~59
- Townsend F M, Thompson B, Bienz M, et al. Pygopus residues required for its binding to Legless are critical for transcription and development. *J Biol Chem*, 2004, 279(7): 5177~5183
- Nakamura Y, Umehara T, Hamana H, et al. Crystal structure analysis of the PHD domain of the transcription co-activator pygopus. *J Mol Biol*, 2007, 370(1): 80~92
- Vetter I, Ferreira R, Becker P B, et al. ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts. *EMBO J*, 2004, 23(20): 4029~4039
- Eberharder A, Ferrari S, Längst G, et al. Acf1, the largest subunit of CHRAC, regulates ISWI-induced nucleosome remodelling. *EMBO J*, 2001, 20(14): 3781~3788
- Kramps T, Peter O, Brunner E, et al. Wnt/wingless signaling requires BCL9-legless-mediated recruitment of pygopus to the nuclear b-catenin-TCF complex. *Cell*, 2002, 109(1): 47~60
- Kramps T, Peter O, Brunner E, et al. The PHD domain of MEKK1 acts as an E3 ubiquitinligase and mediates ubiquitination and degradation of ERK1/2. *Mol Cell*, 2002, 9(5): 945~956
- Kornberg R D, Lorch Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryotic chromosome. *Cell*, 1999, 98(3): 285~294
- Luger K, Richmond T J. The histone tails of the nucleosome. *Curr Opin Genet Dev*, 1998, 8(2): 140~146
- Peterson C L, Laniel M A. Histones and histone modifications. *Current Biology*, 2004, 14(14): R546~551
- Strahl B D, Allis C D. The language of covalent histone modifications. *Nature*, 2000, 403(6765): 41~45
- Mellor J. It Takes a PHD to Read the Histone Code. *Cell*, 2006, 126(1): 22~24
- Becker P B. A finger on the mark. *Nature*, 2006, 442(7098): 31~32
- Zhang Y. It take a PHD to interpret histone methylation. *Nature Structural Molecular Biology*, 2006, 13(7): 572~574
- Shi X B, Kachirskaja L, Walter K L, et al. Proteome-wide analysis in *saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J Biol Chem*, 2007, 282(4): 2450~2455
- Wysocka J, Swigut T, Xiao H, et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, 2006, 442(7098): 86~90
- Lan F, Collins R E, Cegli R D, et al. Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature*, 2007, 448(7154): 718~723
- Shi X B, Hong T, Walter K L, et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature*, 2006, 442(7098): 96~99
- Tsukiyama T, Wu C. Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*, 1995, 83(6): 1011~1020
- Barak O, Lazzaro M A, Lane W S, et al. Isolation of human NURF: a regulator of Engrailed gene expression. *EMBO J*, 2003, 22(22): 5177~5183

- 6089~6100
- 25 Pen a1 P V, Davrazou1 F, Shi X B, et al. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*, 2006, 442(7098): 100~103
 - 26 Flanagan J F, Mi L Z, et al. *Nature*, 2005, 438(7071): 1181~1185
 - 27 Haitao L, Serge I, Wooikoon W, et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*, 2006, 442(7098): 91~95
 - 28 Mellor J. Dynamic nucleosomes and gene transcription. *Trends Genet*, 2006, 22(6): 320~329
 - 29 Ragvin A, Valvatne H, Erdal S, et al. Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. *J Mol Biol*, 2004, 337(6): 773~788
 - 30 Zhou Y, Grummt I. The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. *Curr Biol*, 2005, 15(15): 1434~1438
 - 31 Ruthenburg A J, C. David Allis, et al. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Molecular Cell*, 2007, 25(1): 15~30
 - 32 Joshi A, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell*, 2005, 20(6): 971~978
 - 33 Li B, Gogol M, Carey M, et al. Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science*, 2007, 316(5827): 1050~1054
 - 34 Adam G, Matthews1 W, Kuo A J, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V (D)J recombination. *Nature*, 2007, 450(7172): 1~6

The PHD Finger: a Reader of The Histone Code*

MA Hong-Hui, FANG Cun-Lei, ZENG Ping-Yao**

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

Abstract The PHD finger is a Zn-binding domain found in all eukaryotic genomes, and typically show a C4HC3 signature. Notably, many if not all PHD fingers are found in nuclear proteins whose functions are associated with the regulation of transcription, cell cycle and apoptosis. Increasingly evidences suggest that the PHD finger has multiple functions, including the protein-protein interaction, especially interact with nucleosomes. The pattern and combination of histone modifications, for example, methylation, acetylation, phosphorylation and ubiquitination etc, have been believed to be an important regulator of gene expression and state of the chromatin, which have raised the histone code hypothesis. With the feature of specific recognizing methylated histone, the PHD finger may functions as an important reader of the histone code.

Key words PHD finger, histone code, reader, histone methylation

*This work was supported by a grant from The National Natural Science Foundation of China (30671188).

**Corresponding author.

Tel: 86-21-54237598, Fax: 86-21-54237876, E-mail: pingyaozeng@hotmail.com.

Received: September 25, 2007 Accepted: December 27, 2007