

# 红系 KLF 因子研究进展

夏芸, 钱若兰<sup>1\*</sup>

(华东师范大学生命科学学院, 上海 200062; <sup>1</sup>中国科学院生物化学与细胞生物学研究所, 上海 200031)

**摘要:**  $\beta$ -珠蛋白基因簇是真核生物基因组中颇具代表性的模型, 其表达具有红系组织专一性及发育时期特异性, 这些特异性的产生与红系专一及通用的反式作用因子与相应的顺式作用元件之间相互作用有关。综述了反式作用因子中 KLF 家族中近年来新发现的三个因子 EKLF、FKLF 和 FKLF-2, 介绍了它们的结构、功能及它们与其他反式作用因子间的相互作用。

**关键词:** 红系 KLF 因子; CACCC box; 蛋白因子的相互作用

**中图分类号:** Q291 **文献标识码:** A **文章编号:** 0253-9977(2004)02-149-04

有机体从一个受精卵发育至完整的生物个体, 严格依赖于各种特定基因的精确表达与调控。哺乳动物的珠蛋白基因家族为研究真核基因表达调控提供了独特且具代表性的模型系统。 $\alpha$ -、 $\beta$ -珠蛋白基因簇位于两个不同的染色体上, 各自排列成簇。在红系细胞中, 它们在特定的发育阶段严格顺序表达。

在红系细胞分化过程中, 大量红系组织专一的基因被激活。其中  $\beta$ -珠蛋白基因簇最受关注。这不仅是因为它为揭示珠蛋白基因转录及转录后调控机制提供了良好的模型, 还因为它与人类各种遗传性血红蛋白功能及珠蛋白基因表达紊乱所导致的疾病有关联。

在个体发育过程中,  $\beta$ -珠蛋白基因顺序表达, 经历了两次“开关”, 分别为从胚胎型到胎儿型及从胎儿型到成人型开关。胎儿型与成人型珠蛋白基因表达产物, 即血红蛋白也具有完全不同的生化属性, 这反映了个体在不同发育时期对氧气的不同需求。

$\beta$ -珠蛋白基因的表达调控涉及到顺式调控元件、细胞内反式作用因子、及它们之间的相互作用、染色质结构等多种因素的协调作用。本文仅关注近年来进行了较多研究的红系 KLF 家族反式调控因子, 及与之相关的主要顺式调控元件、和蛋白因子间的相互作用。

## 1 珠蛋白基因启动子及基因近侧端调控区

珠蛋白基因启动子包含许多在其他基因中发现的模体, 包括 TATA、CCAAT、和 CACCC。启动子内也含有一个或多个 GATA 结合位点, 这些位点由谱系专一的锌指蛋白 GATA 家族成员所识别。除了近侧端调控区, 在上游及下游的远侧端 DNA 序

列内, 还发现了与转录活动有关的调控元件<sup>[1]</sup>。

这些调控元件在珠蛋白基因的表达调控中起着不同的作用。珠蛋白基因调控的竞争模型阐明  $\gamma$ -珠蛋白基因在早期发育阶段通过竞争 LCR 的增强子活性来阻止  $\beta$ -珠蛋白基因的表达。人  $\gamma$ -珠蛋白基因 TATA 及 CACCC 元件在胚胎/胎儿期对  $\beta$ -珠蛋白基因的表达具有竞争作用, 但这两个元件所起的作用又不尽相同<sup>[2]</sup>。此外, CACCC 元件在胚胎发育过程中具有时期特异性作用, 这就表明会存在一个或多个因子, 就象 EKLF 结合在  $\beta$ -CACCC 盒上一样, 它/它们能特异结合在  $\gamma$ -珠蛋白基因的 CACCC 盒上。

每一个  $\beta$  类珠蛋白基因在保守的启动子序列中几乎都含有一个或两个 CACCC 盒, CACCC 盒对  $\beta$ -珠蛋白基因转录的重要性已由偶然地自然突变证明, 它的突变会导致地中海贫血症。CACCC 盒在  $\gamma$  基因表达中的作用已得到了证明。在 K562 细胞中, 缺失 CACCC 盒的  $\gamma$  基因启动子对该基因的激活作用降低了。其次, 携带一个缺失 CACCC 盒功能的  $\gamma$  启动子的转基因鼠在所有发育阶段  $\gamma$  基因的表达量均降低。

## 2 红系 KLF 家族转录调控因子

在众多的具有 Cys2-His2 型锌指结构的转录因子中, Krüppel 样因子(KLF)是其中的一个亚家族。该家族的一个显著特点是在近羧基端有三个串连的 C2H2 型锌指。这些锌指由两个半胱氨酸和两个组氨酸残基构成。与高度保守的羧基端相比, 各个

KLF 因子的氨基端则十分不同<sup>[3]</sup>。在哺乳动物中, KLF 调控不同细胞过程, 如细胞增殖和分化, 它们对早期胚胎发育也是至关重要的。这些蛋白通过结合相似的 GT 盒或 CACCC 盒来调控特定基因的表达, 它们对不同启动子的 GT 盒具有不同的亲和力。这些 KLF 因子或作为转录激活子, 或作为转录抑制子, 有的甚至既有激活子功能也有抑制子功能。KLF 家族的一些成员对造血及红系细胞分化具有重要的作用。另外, 在对非哺乳动物的 KLF 因子的研究中, 也有了一些进展。如分离了斑马鱼的 KLF 基因家族, 并对基因的结构与染色体定位及在胚胎发生过程中的表达进行了研究, 发现它们在造血、血管功能及鳍和表皮发育中可能具有作用<sup>[4]</sup>。

### 2.1 EKLF(Erythroid Krüppel Like Factor)

EKLF 同源于 Sp1, 具有三个 C<sub>2</sub>H<sub>2</sub> 锌指; 它能专一地结合具有 CCN CNC CCN 特征的 DNA 序列。在人体珠蛋白基因 β-LCR 和 α-HS40 调控元件的核心区域都存在着 EKLF 的结合位点。并通过它们来识别红系启动子中的 CACCC 盒<sup>[5, 6]</sup>。

EKLF 具有红系细胞表达的特异性, 主要在分化的红系组织, 如脾脏和骨髓中表达, 而在原始的红系组织, 如卵黄囊、肝脏中表达水平较低<sup>[7]</sup>。红系 β 与 γ 基因启动子的 CACCC 盒序列有着很大的区别, 人们发现 EKLF 激活人体成年型 β- 珠蛋白基因启动子的程度要比激活胎儿型 γ- 珠蛋白基因启动子强得多。表明 EKLF 在珠蛋白基因表达由胎儿型向成年型转换的过程中可能起着很重要的作用。EKLF 通过与 β- 珠蛋白基因 5'-DNA 序列内 CACCC 盒高亲和力的结合激活了 β- 珠蛋白基因的表达<sup>[8]</sup>, 它对 β- 珠蛋白基因的表达至关重要, 如果这一识别序列发生突变就会引起 β- 地中海贫血症。

### 2.2 FKLF<sup>[9]</sup>(Fetal Krüppel Like Factor)

目前已克隆了 FKLF 的 cDNA。FKLF 由 512 个氨基酸组成, 在近羧基末端有 3 个邻近的锌指, 是由 C-X<sub>4</sub>-C-X<sub>12</sub>-H-X<sub>3</sub>-H-X<sub>7</sub>-C-X<sub>4</sub>-C-X<sub>12</sub>-H-X<sub>3</sub>-H-X<sub>7</sub>-C-X<sub>2</sub>-C-X<sub>12</sub>-H-X<sub>3</sub>-H 组成(此处 X 代表任何氨基酸残基), 它与 Sp1、EKLF 及 KLF 家族其他蛋白的锌指结构相类同。根据锌指的氨基酸同源性, KLF 家族的蛋白进一步被分为三类: ①具有 70%~90% 氨基酸高度同源性的 Sp1 亚家族, 由 Sp1、Sp2、Sp3 及 Sp4 构成; ②由 EKLF、BTEB2、BKLF、LKLF 及 GKLF 构成了另一高度同源(50%~60%)的 EKLF 亚家族; ③由 BTEB、TIEG、FKLF/TIEG-2 和 FKLF-2 则构成的第三亚族, 以同源性 60%~70% 介于 Sp1 和 EKLF 亚家族之间。锌指是 FKLF 与 DNA 相互结合的区域。

FKLF 的特点是在氨基末端区含有两个酸性及两

个富含脯氨酸的区域。FKLF 主要的转录活性在氨基端。

FKLF 主要在红系细胞中表达。用 RT-PCR 方法在胎儿肝中可检测到 FKLF 的表达, 但在成年骨髓细胞中不能表达。

FKLF 能激活 γ 及 ε 珠蛋白基因启动子, 而对 β 珠蛋白启动子的激活能力相当低。FKLF 通过与 γ 珠蛋白基因启动子 CACCC 盒专一性结合来激活该启动子, FKLF 也能最低程度地激活其他含有 CACCC 调控元件或富含 GC 结合位点的红系启动子(GATA-1、glycophorin B、ferrochelatase、porphobilinogen deaminase, and 5-aminolevulinic synthase)。由此表明 FKLF 并不广泛地激活含 CACCC 模体的基因。

现有的文献表明 FKLF 是体内胚胎型及胎儿型珠蛋白基因表达的一个激活子, 主要激活 ε- 珠蛋白基因启动子和 γ- 珠蛋白基因启动子。

### 2.3 FKLF-2<sup>[10]</sup>

FKLF-2 是一个从鼠卵黄囊克隆获得的 KLF 锌指蛋白。推测有 289 个氨基酸, 在近羧基末端有三个连续锌指。其氨基端富含丙氨酸及脯氨酸, 以及一个富含丝氨酸的羧基末端区。

FKLF-2 锌指的结构为 C-X<sub>4</sub>-C-X<sub>12</sub>-H-X<sub>3</sub>-H-X<sub>7</sub>-C-X<sub>4</sub>-C-X<sub>12</sub>-H-X<sub>3</sub>-H-X<sub>7</sub>-C-X<sub>2</sub>-C-X<sub>12</sub>-H-X<sub>3</sub>-H (此处 X 代表任何氨基酸残基), 与 Sp1、EKLF 及这两个家族中的其他蛋白的锌指相似。FKLF-2 的锌指 motif 的同源性介于 Sp1(Sp1、Sp2、Sp3、Sp4) 及 EKLF(EKLF、BTEB2/KLF、LKLF、BKLF、GKLF/EZF、CPBP/Z19、UKLF 和 AP-2rep) 两家族的同源性之间, 因而它属于第三类 Sp1/EKLF 蛋白, 此类还包括 BTEB1、TIEG-1 和 FKLF/TIEG-2 蛋白质因子。

FKLF-2 的人体同源物在骨髓及横纹肌中表达, 而在被检测的 12 种其他人体组织中则不表达(胃, 甲状腺, 脊索, 淋巴结, 气管, 肾上腺, 脑, 胎盘, 肺, 肝, 肾和胰腺)。FKLF-2 在具有红系表型的细胞株中表达, 而在髓系及淋巴表型的细胞株中表达不恒定。鼠红白血病细胞株用诱导剂 hemin、δ-ALA 或 hemin + HMBA 诱导后, FKLF-2 mRNA 上调了, 这与 EKLF 和 FKLF 不同, 红系分化的诱导剂不能促进后两者的表达。FKLF-2 主要激活 γ- 珠蛋白基因, 而对 ε- 及 β- 珠蛋白基因激活能力较低。对 γ- 珠蛋白基因启动子的激活不依赖于 HS2 增强子的存在。FKLF-2 主要通过 γ- 珠蛋白基因启动子中 CACCC 盒相互作用来激活 γ- 珠蛋白基因启动子, 而通过与 TATA 盒或其周围的 DNA 序列的相互作用所起的激活功能则较低。FKLF-2 也能激活其他的红系专一的启动子(GATA-1、glycophorin B、ferrochelatase、porphobilinogen

deaminase, and 5-aminolevalinate synthase)。这些结果表明, FKLf-2 除了激活  $\beta$ - 类珠蛋白基因的表达, 可能还与红系细胞株内其他一些基因的转录激活有关。

FKLF-2 有许多特点使其区别于 FKLf。首先, 它主要激活  $\gamma$  基因启动子, 而 FKLf 则主要激活  $\epsilon$  基因启动子。其次, FKLf 的活性完全依赖于 -140 CACCC 盒, 而 FKLf-2 虽然主要通过 CACCC 盒来激活  $\gamma$ - 珠蛋白基因启动子, 但也可通过 TATA 盒或邻近序列在较低程度上激活  $\gamma$ - 基因的表达。再次, 红白血病细胞株 K562、HEL 及 MEL 细胞经 hemin、 $\delta$ -ALA 或 hemin + HMBA 诱导后, EKLF 和 FKLf 的表达没有上调, 而 FKLf-2 表达上调了。

### 3 调控因子间的相互作用与基因表达调控

$\beta$ - 类珠蛋白基因表达的发育调控需要每个珠蛋白基因近侧端调控序列与位于基因远侧端 DNA 序列内的座位控制区(LCR)相互作用。这些相互作用的分子机制仍然不太明了, 但人们认为是通过谱系专一的、或通用的、或时期专一的调控蛋白介导的。根据这个观点, 蛋白质-蛋白质间的相互作用能促进并稳定 DNA 成环, 由此使得 LCR 的特定区域与珠蛋白基因启动子及上游调控序列的接近。

然而, 基因表达的调控往往是多个因子共同参与的。这些因子除了与特定的 DNA 序列相互结合, 还能与特定的转录共激活子(Coactivator)/ 抑制子(Corepressor)发生相互作用, 形成激活或抑制复合物, 并通过它的“桥梁”作用来调控特定的基因。因而, 这就涉及到了反式作用因子间的相互作用。

在 KLF 因子中, Sp1 和 EKLF 这两个因子研究得较为深入。已知 GATA-1 能与 Sp1 和 EKLF 相互作用, 并增强它们的功能<sup>[11]</sup>。Sp1 富含谷氨酸的激活区能与 TF II D 复合物的 TAF110 亚单元直接接触并介导转录激活<sup>[12]</sup>。此外, 转录辅助因子复合物 CRSP(Cofactor required for Sp1 activation)也被证明是 Sp1 体外转录激活所必需的<sup>[13]</sup>。EKLF 在体内可与 CBP/p300 及 PCAF 结合, CBP/p300 能使 EKLF 或染色体的组蛋白发生乙酰化, 同时又招募了其他一些转录因子及染色质重组复合物, 来共同影响  $\beta$ - 珠蛋白基因的表达<sup>[14]</sup>。EKLF 也与 ERC-1(EKLF coactivator remodeling complex 1)相互作用。ERC-1 是一个 SWI/SNF 相关的染色质重组复合物, 对染色质 DNase I 超敏感性产生与  $\beta$ - 珠蛋白基因的活跃转录都是必需的<sup>[15]</sup>。人们也发现 EKLF 能与共抑制子 mSin3A 及 HDAC1 结合, 共同抑制转录<sup>[16]</sup>。另外, BKLF 可以在体外有效地竞争及关闭 EKLF 介导的  $\beta$ - 珠蛋白基因激活<sup>[17]</sup>, 以上这些研究都表明蛋白质间的相互作用对珠蛋白基因调控是极为重要的。

那么, 一个因子是如何与另一个或多个因子相互作用以建立和维持红系分化的程序? 在这个问题上人们又作了进一步的研究。

新近发现 FKLf-2 可与 CBP/p300 和 PCAF 相互作用。CBP/p300 和 PCAF 通过 FKLf-2 的锌指区与 FKLf-2 相互作用, 使 FKLf-2 乙酰化, 增强它与 DNA 的结合, 来共同提高 FKLf-2 的转录活性。虽然 CBP/p300 和 PCAF 都能促进 FKLf-2 与 DNA 结合, 但仅仅是 PCAF 的 HAT 区对激活 FKLf-2 的 DNA 结合及转录活性是必须的<sup>[18]</sup>。

CBP/p300 与 PCAF 是许多转录因子的共激活子, 它们与细胞发育和分化有关联<sup>[19-24]</sup>。CBP/p300 与 PCAF 本质上具有组蛋白乙酰转移酶活性。除了组蛋白, 许多转录因子也是乙酰化的底物。转录因子的乙酰化修饰已被证明在多个水平上具有激活功能, 包括与 DNA 结合、蛋白质因子的相互作用及其稳定性。FKLF-2 和 KLF 家族中的其他成员(如 EKLF)与 CBP/p300 或 PCAF 相互作用的方式不尽相同, 而且这些家族成员的活性由这些共激活子通过截然不同的机制来调控, 暗示了 CBP/p300 及 PCAF 在专一靶基因的激活中起着十分重要的作用。

### 4 小结与展望

本文讨论的三个转录因子, 它们均具有相同的锌指结构, 在珠蛋白基因表达调控中也起着相似的作用, 都能激活相应基因地表达, 然而, 它们又有各自的特点。FKLF 不像 EKLF, 后者只是激活  $\beta$ - 珠蛋白基因启动子, 而前者则不仅激活  $\epsilon$ - 与  $\gamma$ - 珠蛋白基因启动子, 它还能在 K562 细胞内较低水平上激活  $\beta$  基因启动子。不像 EKLF 只在成人红系细胞中表达, FKLf 主要在胚胎及胎儿红系细胞中表达, FKLf-2 在胎儿及成人的造血组织都有表达, 虽然对 FKLf、FKLF-2 已开始了研究, 但对于 FKLf 到底是激活  $\epsilon$  还是  $\gamma$  基因尚需进一步明了。假设 FKLf 激活  $\epsilon$  珠蛋白基因, 那么是否存在这样的因子, 如 CBP/p300 和 PCAF 与 EKLF、FKLF-2 的作用一样, 它/它们也能作用于 FKLf 呢? 若 FKLf、FKLF-2 可以激活  $\beta$ - 珠蛋白基因启动子, 即使激活水平很低, 那是否也是通过  $\beta$  基因启动子的 CACCC 盒呢? 在红系细胞中, 是否仍然有未知的调控因子存在, 它/它们对珠蛋白基因表达调控起着更为重要的作用呢? 这些问题我们尚无法知晓, 目前, 本实验室正在进行有关方面的探索, 希望能获得一些有益的结论。

### 参 考 文 献

- [1] STAMATOYANNOPOULOS G, NIENHUIS A. Hemo-

- globin switching in The Molecular Basis of Blood Diseases [M]. Saunders W B, Philadelphia, 1994, 107 — 155.
- [2] THANH G S, JOYCE A L. The human  $\gamma$ -globin TATA and CACCC elements have key, distinct roles in suppressing  $\beta$ -globin gene expression in embryonic/fetal development[J]. *J Biol Chem*, 2001, **276**: 41817 — 41824.
- [3] JEREMY T, MERLIN C. Mammalian Krüppel-like transcription factors: more than just a pretty finger[J]. *Trends Biochem Sci*, 1999, **24**: 236 — 240.
- [4] ANDREW C O, STEPHEN J R, BRENDA V, *et al.* The zebrafish klf gene family[J]. *Blood*, 2001, **98**: 1792 — 1801.
- [5] WADE P A, PRUSS D, WOLFFE A P. Histone acetylation: chromatin in action[J]. *Trends Biochem Sci*, 1997, **22**: 128 — 132.
- [6] WU C. Chromatin remodeling and the control of gene expression[J]. *J Biol Chem*, 1997, **272**: 28171 — 28174.
- [7] KULOZIK A E, *et al.* Thalassemia intermedia: moderate reduction of beta globin gene transcriptional activity by a novel mutation of the proximal CACCC promoter element [J]. *Blood*, 1991, **77**: 2054.
- [8] MILLER I J, BIEKER J J. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear factors [J]. *Mol Cell Biol*, 1993, **13**: 2776 — 2786.
- [9] HARUHIKO A, LI X S, STAMATOYANNOPOULOS G. FKLf, a novel Krüppel-like factor that activates human embryonic and fetal  $\beta$ -like globin genes[J]. *Mol Cell Biol*, 1999, **19**: 3571 — 3579.
- [10] HARUHIKO A, LI X S, STAMATOYANNOPOULOS G. FKLf2, a novel Krüppel-Like transcriptional factor that activates globin and other erythroid lineage genes[J]. *Blood*, 2000, **95**: 3578 — 3584.
- [11] MERIKA M, ORKIN S. Functional synergy and physical interaction of the Krüppel transcription factor GATA-1 with the Krüppel family proteins SP1 and EKLF[J]. *Mol Cell Biol*, 1995, **15**: 2437 — 2447.
- [12] GILL G, PASCAL E, TSENG Z H, *et al.* A glutamine-rich hydrophobic patch in transcription factor sp1 contacts the dTAF11110 component of the drosophila TFIID complex and mediates transcriptional activation[J]. *Proc Natl Acad Sci USA*, 1994, **91**: 192 — 196.
- [13] RYU S, ZHOU S, LADURNER A G, *et al.* The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1[J]. *Nature*, 1999, **397**: 446 — 450.
- [14] ZHANG W, BICKER J J. Acetylation and modulation of erythroid Krüppel-like factor(EKLF) activity by interaction with histone acetyltransferases[J]. *Proc Natl Acad Sci USA*, 1998, **95**: 9855 — 9860.
- [15] ARMSTRONG J A, BIEKER J J, EMERSON B M. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF *in vitro*[J]. *Cell*, 1998, **95**: 93 — 104.
- [16] CHEN X, BIEKER J J. Unanticipated repression function linked to erythroid Krüppel-like factor[J]. *Mol Cell Biol*, 2001, **21**: 3118 — 3125.
- [17] TURNER J, CROSSLEY M. Loning and characterization of mCtBP2, a co-repressor that associates with basic Krüppel-like factor and other mammalian transcriptional regulators [J]. *EMBO J*, 1998, **17**: 5129 — 5140.
- [18] CHAO Z S, KIMBERLY K, KEN M, *et al.* Functional interaction between coactivators CBP/p300, PCAF, and transcription factor FKLf2[J]. *J Biol Chem*, 2002, **277**: 7029 — 7036.
- [19] GREGORY R C, TAXMAN D J, SESHASAYEE D, *et al.* Functional interaction of GATA1 with erythroid Krüppel-like factor and Sp1 at defined erythroid promoters. *Blood*, 1996, **87**: 1793 — 1801.
- [20] GOODMAN R H, SMOLIK S. CBP/p300 in cell growth, transformation and development[J]. *Genes Dev*, 2000, **14**: 1553 — 1576.
- [21] JANKNOCHI R, HUNTER T. A growing coactivator network[J]. *Nature*, 1996, **383**: 22 — 23.
- [22] KOUZARIDES T. Acetylation: a regulatory modification to rival phosphorylation?[J]. *EMBO J*, 2000, **19**: 1176 — 1179.
- [23] SHIKAMA N, LYON J, LA THANGUE N B. The p300/CBP family: integrating signals with transcription factors and chromatin[J]. *Trends Cell Biol*, 1997, **7**: 230 — 236.
- [24] VO N, GOODMAN R H. CREB-binding protein and p300 in transcriptional regulation[J]. *J Biol Chem*, 2001, **276**: 13505 — 13508.

## Recent Advances in Studies of Erythroid Krüppel-like Factors

XIA Yun, QIAN Ruo Lan<sup>1\*</sup>

(School of Life Science, East China Normal University, Shanghai 200062, China; <sup>1</sup>Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai 200031, China)

**Abstract:** The  $\beta$ -globin gene family has been used as a model for study of gene expression and regulation in eukaryotic cells. These genes expressed specifically in erythroid tissue and differentially expressed in a precise temporal order during development. It has been shown that interaction of erythroid specific *trans*-acting factors with *cis*-acting elements may play an important role in regulation of globin gene expression. Here we mainly focus on introducing three erythroid specific factors of KLF family (EKLF, FKL and FKLf-2), including their structure, function and interaction with other *trans*-acting factors. Previous studies have revealed that they may be critical in the developmental regulation of globin gene expression.

**Key words:** erythroid KLF factors; CACCC box; proteins interaction

\*Corresponding author, E-mail: qianlab@sunm.ac.cn