

- Biochim Biophys Acta*, 1985; 813(1): 96
- 14 Lee P M, Cherry R J, Bächi T. Correlation of rotational mobility and flexibility of sendai virus spike glycoproteins with fusion activity. *Virology*, 1983; 128(1): 65
 - 15 Blatt E, Corin A F. The microsecond rotational motions of eosin-labelled fatty acids in multilamellar vesicles. *Biochim Biophys Acta*, 1986; 857(1): 85
 - 16 Tilley L, Sawyer W H, Morrison T R *et al.* Rotational diffusion of human lipoproteins and their receptors as determined by time-resolved phosphorescence anisotropy. *J. Biol Chem*, 1988; 263(33): 17541
 - 17 Wang J, Hogan M, August R H. DNA motions in the nucleosome core particle. *Proc Natl Acad Sci USA*, 1982; 79(19): 5896
 - 18 Hogan M E, Rooney T F, Austin R H. Evidence for kinks in DNA folding in the nucleosome. *Nature*, 1987; 328: 554
 - 19 Berkoff B, Hogan M, Legrange J *et al.* Dependence of oxygen quenching of intercalated methylene blue triplet lifetime on DNA base-pair composition. *Biopolymers*, 1986; 25(2): 307
 - 20 Austin R H, Karohl J, Jovin T M. Rotational diffusion of *Escherichia coli* RNA polymerase free and bound to deoxyribonucleic acid in non-specific complexes. *Biochemistry*, 1983; 22(13): 3082

螺旋-环区-螺旋蛋白质——DNA 结合蛋白的新类型

李建义 童坦君

(北京医科大学生化教研室,北京 100083)

提 要

最近发现了一种特殊的蛋白质结构域,它广泛地存在于动、植物体的 DNA 结合蛋白 (DBP) 中。此结构称为螺旋-环区-螺旋 (helix-loop-helix, HLH) 结构。c-myc 基因和 MyoD 基因的蛋白质产物有此结构。在增强子结合蛋白 (EBP) 中也发现了 HLH 结构,如 E2A 基因的产物——E12/E47。已报道的 20 多种 HLH 蛋白质几乎全与转录的调节和肿瘤的发生有关。

关键词 HLH 蛋白, DNA 结合蛋白 (DBP), c-myc 基因

目前关于 DNA 结合蛋白的研究已成为分子生物学的热点。DBP 可结合在 DNA 的某一特定部位,它不仅可以调节 DNA 的复制,重组和转录,还在染色质的解旋、盘绕和折叠过程中起重要作用。基因的活化蛋白和阻遏蛋白研究较为深入。这些蛋白质分子有以下几种特殊结构: a. 螺旋-转折-螺旋 (helix-turn-helix, HTH) 结构。它最初是在细菌的蛋白质中发现的^[1]。b. 锌指结构 (zinc fingers), 含有此结构的蛋白质是一类真核细胞基因的调节因子^[2], TFIIIA 中有 9 个锌指结构, c. 亮氨酸拉链 (leucine zipper, LZ), 此结构存在于 fos, Jun, myc 基因编码的蛋白质中^[1]。d. 螺旋-环区-螺旋结构^[3]。本文主要介绍 HLH 结构。

1 HLH 结构特点

HLH 结构由三部分组成,两端为 α -螺旋,中间是由一个或几个 β -转角组成的环区。大约有 60 个氨基酸残基^[4]。两个 α -螺旋含有许多高度保守的疏水氨基酸。 α -helix I 中有 12 个疏水氨基酸残基。在 α -螺旋的疏水侧, Leu 和 Phe 具有高度的保守性,与它们相连的氨基酸残基也是疏水氨基酸。 α -helix II 中有 13 个疏水氨基酸,其中有 5 个具有高度保守性^[5](见图 1)。

中间的环区一般含有多个阻碍 α -螺旋形

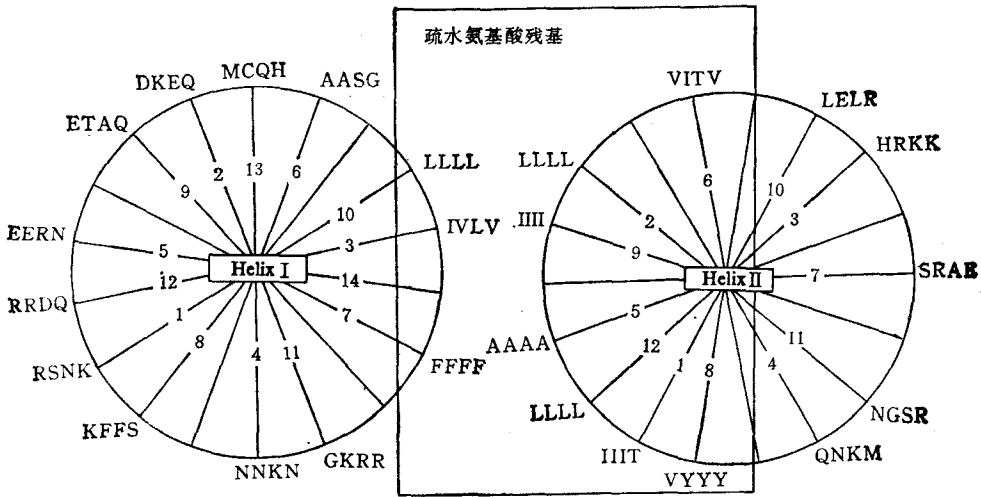


图 1 E12, c-myc, MyoD 和 achaete-scute 的氨基酸保守序列

E12, MyoD, c-myc 和 achaete-scute 的氨基酸序列按由左到右排列。两个 α -螺旋的疏水面已位于矩形内

成的氨基酸残基。多由 Gly、Pro、Asp、Asn 和 Ser 连接而成。如 c-myc 癌基因编码的蛋白质，其环区的氨基酸序列为 Pro-Glu-Leu-Glu-Asn-Glu。这种序列同样也存在于 E12 以及 MyoD 基因编码的蛋白质中^[5]。同族的 HLH 结构蛋白的环区的大小和组成在不同的种属是相同的，但不同族的就有很大的差异^[3]。

一般在 HLH 结构的上游有富含碱性氨基酸的区域 (basic region, BR)，大约由 10—20 个氨基酸组成。它可分为 3 个碱性氨基酸聚集片段: B₁, B₂, B₃，这种结构特点在 MyoD, Myogenin 和 Myf-5 基因编码的蛋白质中都是高度保守的^[6]。最近又发现 BR 缺乏或中断的一些 HLH 结构蛋白，它们是基因转录的负调节因子^[7-9]。这里必须指出 HLH 的存在也就是从氨基酸序列特征推断的，至今并没有真正的结构数据发表。

2 α -螺旋参与 HLH 蛋白二聚体的形成

HLH 结构蛋白质能与同家族的其它蛋白质分子结合形成同源或异源二聚体，同源二聚体与相应基因结合的能力比异源二聚体弱^[9]。

HLH 蛋白质可分为 3 类: A 类为分布广

泛的 HLH 蛋白,如: E12、E47。B 类是具有组织特异性的蛋白质,如 MyoD 的蛋白质产物。C 类主要指 myc 编码的蛋白质。A 类与 B 类可形成异源二聚体,而 C 类则不能与 A、B 类 HLH 蛋白结合成异源二聚体。Wood 和 Eisenman (1991 年)利用含有 c-myc 编码蛋白的 C 末端的标记融合蛋白分离出一种蛋白质,被命名为 Max^[10]。Max 也含有 BR 区、HLH 和 LZ 结构,它可与 3 类 myc 基因编码的蛋白质结合,但不能与其它 HLH 结构蛋白结合。在不表达 c-myc 的细胞中也有 Max 的存在,可见 Max 还可能行使不依赖于 c-myc 的功能,或许 Max 可与其它未知蛋白质相结合,来拮抗 myc-Max 异源二聚体的基因激活作用^[10]。

Murre 等人通过对 E12 与 MyoD 编码蛋白质结合的研究^[9],发现 α -螺旋直接参与 HLH 蛋白二聚体的形成。他们的实验证明, E12 的编码基因发生 Bal 31 区带缺失时, E12 的 α -helix II 的 C 末端就不能形成。MyoD/E12 异源二聚体也就不能形成。同样 α -helix I 的编码基因的 Bal 14 区带缺失后,也会影响二聚体的形成^[9]。两个 HLH 蛋白质分子的 α -螺旋上的疏水氨基酸残基相互靠近,形成疏水键,从而

形成二聚体。

3 HLH 结构蛋白的 DNA 结合区和基因激活区

Lesczynski 和 Rose 认为 HLH 蛋白的 α -螺旋参与二聚体的形成,而环区保守性差,他们认为环区是 HLH 蛋白的 DNA 结合部位^[9]。而 Murre 认为 α -螺旋在 HLH 蛋白与 DNA 结合的过程中起重要的作用^[9]。最近 Davis 等人通过基因定位突变的研究,表明 BR 才是 HLH 结构蛋白的 DNA 结合区^[6]。与 DNA 的结合之所以需要 HLH 结构,是因为 HLH 蛋白质只有形成二聚体之后才能与 DNA 结合。HLH 结构区与 BR 的关系十分密切,只要在两者的交接处插入几个氨基酸残基,就会影响二聚体的形成以及与 DNA 的结合。如果环区的长度和氨基酸序列发生变化,也会影响与 DNA 的结合能力。

HLH 结构蛋白质的碱性区带有大量的正电荷,在溶液中不能形成稳定的 α -螺旋结构。当与 DNA 相靠近时,其正电荷被 DNA 的磷酸根离子所中和,从而形成稳定的 α -螺旋结构,然后与 DNA 双螺旋的大沟相结合^[6]。许多 HLH 蛋白的异源二聚体可与 DNA 上的 E“Box” (CACGTG 序列)相结合^[5,10],在免疫球蛋白的重链、轻链基因和胰岛素基因的增强子上都含有 E“Box”,它还存在于肌肉肌酸激酶基因的启动子上。

HTH 结构蛋白作为一类转录调节因子,它不仅含有与 DNA 相结合的区域,还应含有转录激活区。myc 基因的编码蛋白质是通过 HLH 结构区与 Max 形成异源二聚体,二聚体又通过 BR 与 DNA 的特定区域相结合。而其转录激活作用完全取决于 myc 编码的蛋白质的 N 末端序列^[10]。N 末端序列只有与其它部分相连时才有细胞转化和转录激活作用。可见 HLH 结构蛋白的 DNA 结合区和激活区是密切相关的。

4 HLH 蛋白对基因表达的调控作用

在原核生物中,有 4 种具有 HTH 结构的蛋白质对基因转录起调控作用。 λ 阻遏蛋白和 Cro 蛋白可调节噬菌体的 λ 基因的表达。CAP (catabolite activator protein) 在缺乏葡萄糖的情况下可激活大肠杆菌的乳糖操纵子的结构基因。还有色氨酸操纵子阻遏蛋白也含有 HTH 结构。在真核细胞中,myc 基因编码的蛋白质以及 Id 蛋白^[7] (BR 缺失的一种 HLH 结构蛋白质)对基因表达的调控作用研究较深入。

Id 蛋白是由一类 cDNA 编码的具有 HLH 结构的 DNA 结合蛋白质,它不具有碱性区。Robert 和 Davis 证明,在体外 Id 至少可与 3 种 HLH 蛋白结合 (E12 E47 和 MyoD 的编码蛋白质),从而影响它们与 DNA 结合的能力。Id 是肌肉肌酸激酶基因表达的负调节因子^[7]。此外,Id 在红细胞系、肌细胞以及网织红细胞系的分化过程中都处于低表达状态。Id 可能对各类细胞系的其它 HLH 蛋白的基因激活作用起拮抗效应^[7]。在果蝇,extra-machroaete (emc) 编码的 HLH 蛋白没有 BR。hairy 编码的蛋白质由于 Pro 的存在,BR 出现中断。它们都对 achaete-scute 基因有负调节作用^[3]。可见 BR 缺失或中断的 HLH 蛋白可形成无功能的异源二聚体,从而对其它的 HLH 蛋白激活的基因起抑制作用。

人的 c-myc 基因的编码产物是一种分子

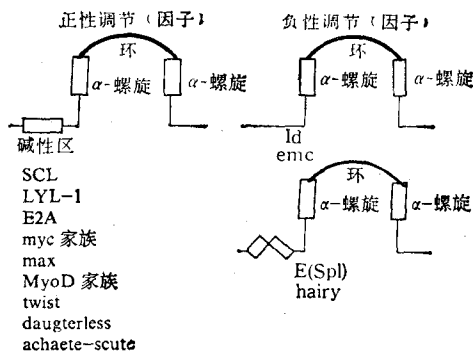


图 2 HLH 蛋白对基因表达的调控作用类型

正调节因子含有 BR 区,而负调节因子不含 BR 区,或者它们的 BR 区有中断现象

量为 65kD 的 DNA 结合蛋白 (p65^{c-myc})^[11]。在转录水平和转录过程之后，它都对基因的表达起调控作用。并且对不同的基因作用不同^[12]。p65^{c-myc} 对成纤维细胞的血纤溶酶原激活因子 I (plasminogen activator I, PAI-I) 编码基因起激活作用^[13]。而对神经母细胞瘤和黑色素瘤细胞的 HLA 基因起阻遏作用^[14]。p65^{c-myc} 对 PAI-I 编码基因的调节主要发生在转录之后^[12]。此外在一些非肿瘤细胞，如：正常的淋巴细胞和成纤维细胞，c-myc 的编码产物对于 c-myc 基因可产生自身反馈抑制^[12]。这种自身调节发生在转录起始水平，并且至少需要一种中介物或辅助因子来介导。但在肿瘤细胞中不存在这种自身负反馈^[12]。最近研究表明 c-myc 基因的编码蛋白可促进肿瘤细胞 c-myc 基因的复制和转录的起始^[15]。可见 c-myc 编码蛋白对不同基因的作用方式和作用时间有所不同。这可能由于 c-myc 编码蛋白具有多种功能，也可能由于所参与的中介物和辅助因子不同。

HLH 结构蛋白既能形成二种二聚体，又可在不同的时间对不同的基因表达产生正负调节作用，从而形成了真核细胞基因表达的调控系统。

5 与肿瘤的关系

在 Burkitt's 淋巴瘤和其它 B 淋巴细胞肿瘤中，c-myc 原癌基因可易位到免疫球蛋白的重链或轻链的调控区附近^[3]，这种易位并不改变 c-myc 的编码潜能。由于 c-myc 的编码蛋白可结合于 Ig 的重链及轻链基因增强子上的 E“Box”上^[3]，同时由于 c-myc 的直接损伤，其自身负反馈调节作用丧失，最终导致 c-myc 的高表达。在 Burkitt's 淋巴瘤细胞中还发现 c-myc 基因的第一个内含子的点突变。在肿瘤细胞，c-myc 的编码产物可以促进 c-myc 基因的复制，因而在小细胞肺癌和前髓细胞白血病可见到 c-myc 基因的扩增和高表达^[16]。其他 HLH 结构蛋白的编码基因 (SCL, LYL-1, E2A) 的易位也与肿瘤的发生有关^[17-19]。

综上所述，正常的 HLH 结构蛋白参与多种基因的调节作用。从而维持细胞正常的增殖、分化。当 HLH 基因发生突变、易位和扩增时，就可能致肿瘤的发生。在 myc 编码的蛋白质中即有 HLH 结构，又有 LZ 结构，二者的关系还不清楚。HLH 结构蛋白对正常细胞的基因表达的调控还应更深入地研究。

参 考 文 献

- 1 Landschuitz W H, Johnson P F, McKnight SL. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*, 1988; 240: 1759
- 2 Berg J M. Potential metal-binding domains in nucleic acid binding proteins. *Science*, 1986; 232: 485
- 3 Visvader J, Begley C G. Helix-loop-helix genes translocated in lymphoid leukemia. *TIBS*, 1991; 16: 330
- 4 Henthion P, Kadesch T, Kiledjian M. Two distinct transcription factors that bind the immunoglobulin enhancer mE₃/x E₂ motif. *Science*, 1990; 247: 467
- 5 Murre C, McCaw P S, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD and myc proteins. *Cell*, 1989; 56: 777
- 6 Davis R L, Lassar A B, Cheng PF *et al.* The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell*, 1990; 60: 733
- 7 Beuzra R, Davis R, Lockshon D. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell*, 1990; 61: 49
- 8 Jones N. Transcriptional regulation by dimerization: two sides to an incestuous relationship. *Cell*, 1990; 61: 9
- 9 Murre C, McCaw P S, Vaessin H *et al.* Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, 1989; 58: 537
- 10 Cole M D. Myc meets its max. *Cell*, 1991; 65: 715
- 11 Persson H, Leder P. Nuclear localization and DNA binding properties of a protein expressed by human c-myc oncogene. *Science*, 1984; 225: 718
- 12 Grignani F, Lombardi L, Inghirami G *et al.* Negative autoregulation of c-myc gene expression is inactivated in transformed cells. *EMBO J*, 1990; 12: 3913
- 13 Prendergast G C, Cole M D. Post transcription regulation of cellular gene expression by the c-myc oncogene. *Mol Cell Biol*, 1989; 9: 124
- 14 Verteeg R, Noordermeer I A, Margreest K W. c-myc down-regulates class I HLA expression in human melanomas. *EMBO J*, 1988; 7: 1023
- 15 Ariga H *et al.* DNA replication origin and transcriptional enhancer in c-myc gene share the c-myc proteins binding sequences. *EMBO J*, 1989; 8(13): 4273

(下转第 73 页)

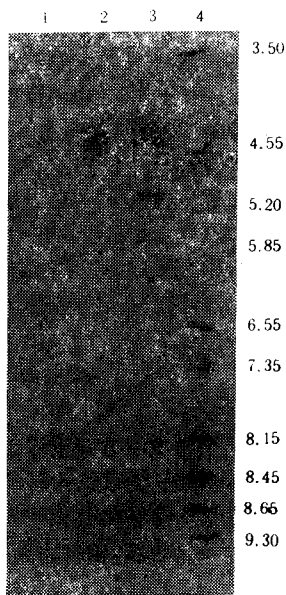


图5 等电聚焦电泳结果

1. PGFB-2 肽纯品 2, 3. PGFB-2 肽不纯部分
4. 等电聚焦电泳标准等电点

肽N端的前四个氨基酸进行了序列分析, 前四个氨基酸的顺序为甲硫氨酸、苏氨酸、苏氨酸、色氨酸, 这与编码该多肽的基因片段的 DNA 序列分析结果完全一致^[11], 从而也进一步证实了该多肽的纯度及正确性。

2.4 PGFB-2 肽对热、SDS 等的稳定性比较

蛋白质的稳定性是决定其应用前景的一个重要指标, PGFB-2 肽是由遗传工程构建的一种蛋白质, 所以, 有必要了解其稳定性如何。于是我们首先用热及 SDS 对 PGFB-2 肽进行处理, 然后用免疫酶联法测定其与 IgG 结合活性, 发现 PGFB-2 肽经 100℃ 处理 5min 后其活性基本无变化, 加入 SDS 后再煮沸 15min, 活性依然无明显下降, 这表明 PGFB-2 肽是一较稳定的多肽, 由此推测天然蛋白 G 中 Fc 结合区

结构域能形成一较稳定的结构。这与最近 Gronenborn 等^[12]报告的有关结果是一致的。

参 考 文 献

- 1 Guss B, Eliasson M, Olsson A *et al.* Structure of the IgG-binding regions of streptococcal protein G. *EMBO J*, 1986; 5: 1567
- 2 Bjorck L, Kronvall G. Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. *J Immunol*, 1984; 133: 969
- 3 Reis K J, Ayoub E M, Boyle M D P. Streptococcal Fc receptors; Isolation and partial characterization of the receptor from a group C streptococcus. *J Immunol*, 1984; 132: 3091
- 4 Reis K J, von Mering G O, Karis M A *et al.* Enzyme-labeled type III bacterial Fc receptors, a versatile tracer for immunoassay. *J Immunol Methods*, 1988; 107: 273
- 5 蔡仕英, 王园园, 姚志建. PCR 扩增的 Protein G IgG Fc 结合区单结构域基因片段的克隆. *中华流行病学杂志*, 1990; 11(特刊 6号): 100
- 6 郭燕捷, 姚志建. 几种微量测定蛋白质新方法的比较. *生物化学与生物物理进展*, 1986; (4): 67
- 7 张龙翔, 张庭芳, 李令媛. *生物化学实验方法和技术*. 湖北: 湖北科学技术出版社, 1987; 112—118
- 8 Burnette W N. "Western blotting". Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gel to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem*, 1981; 112: 195
- 9 周本正. *实用电泳及免疫电泳技术*. 湖北: 湖北科学技术出版社, 1988; 86, 104
- 10 Chang J Y, Brauer D, Wittmann Liebold B. Micro-sequence analysis of peptides and proteins using 4-NN-dimethylaminoazobenzyl 4'-isothiocyanate phenylisothiocyanate double coupling method. *FEBS Letter*, 1978; 93: 205
- 11 蔡仕英, 王园园, 姚志建. Protein G 中 Fc 段结合区的结构分析. *中国科学 B 辑*, 1991; (12): 1284
- 12 Gronenborn A M, Filipula D R, Essig N E *et al.* A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science*, 1991; 253: 675—661

(上接第 14 页)

- 16 Little C D, Nau M M, Cayney D N *et al.* Amplification and expression of the c-myc oncogene in human lung cancer cell lines. *Nature*, 1983; 306: 194
- 17 Begley C G, Aplan P D, Davey M P *et al.* Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor δ -chain diversity region and results in a previously unreported fusion transcript. *proc Natl Acad Sci USA*, 1989; 86: 2031
- 18 Mellentin J D, Smith S D, Cleary M L. Lyl-1, a novel gene altered by chromosomal translocation in T-cell leukemia, codes for a protein with helix-loop-helix DNA binding motif. *Cell*, 1989; 58: 77
- 19 Nourse J, Mellentin J D, Galili N *et al.* Chromosomal translocation t(1; 19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell*, 1990; 60: 935