

序列特异的三锌指多肽的构建及其在大肠杆菌中的表达

张书祥** 马清钧 赵志虎*

(军事医学科学院生物工程研究所 北京 100850)

摘 要 在获得单一锌指突变体的基础上,以小鼠转录因子 Zif268 的三锌指 DNA 结合区为模板,利用重叠(Over-lap)PCR 技术,获得了关键氨基酸位点同时突变的三锌指突变体 ZF123、2ZF123。ZF123、2ZF123 分别克隆进 pUC 18 质粒,序列测定正确后,以 pGEX-2T 为表达质粒,在大肠杆菌 JM109 中实现了功能性的表达。经 SDS-PAGE 分析,表达出了分子量 34.0kD 的融合蛋白,扫描分析其含量在 20%左右。菌体经超声波破碎后,对可溶性融合蛋白进行了纯化得到了游离的目的蛋白,为进一步的 DNA 结合特性分析、杂交转录因子的构建等奠定了基础。

关键词 锌指蛋白,重叠 PCR,蛋白表达纯化

中图分类号 Q78 文献标识码 A 文章编号 1000-3061(2001)04-0406-04

锌指结构是一种长约 30 氨基酸的最常见的核酸识别基序,整个肽段通过锌离子与半胱氨酸和/或组氨酸的螯合配位形成一种稳定的“指头”样高级结构。自 1985 年在非洲爪蟾转录因子 TFIIIA 中发现第一个锌指以来,目前已在 560 多种蛋白中发现 4100 多种不同的锌指序列,成为所有 DNA 结合蛋白中最庞大的一个家族^[1,2],此外锌指结构还是目前为止唯一的既可识别 DNA,又可识别 RNA 甚至还可以识别 DNA-RNA 杂交双链的蛋白元件,是研究蛋白-核酸特别是蛋白-DNA 相互作用的理想对象。早期的一些经典方法如 Dnase I 足迹法、甲基化保护实验以及晶体结构的分析等发现锌指蛋白以一种特异性的方式与 DNA 相互作用,每个锌指识别约 3 个碱基对的特异序列^[3,4],而且在这种 1:3 的识别过程中起关键作用的是几个保守的氨基酸位点,噬菌体展示技术则揭示改变这些位置上的氨基酸可以改变整个锌指的识别特异性^[5]。

但是以上这些实验都是一些体外实验,很难真实地反映体内锌指-DNA 相互作用规律,鉴于此,我们以小鼠转录因子 Zif268 的三锌指结合区^[4]为材料,建立了一种体内遗传筛选模型,利用上述模型,通过对单一锌指识别螺旋中几个保守的氨基酸位点进行随机突变与筛选,获得了一些识别特异序列如

5'-GCA、GAT、GCT、GCC-3'等的单锌指突变体(赵志虎博士论文)。本工作则在此基础上,通过重叠 PCR 对 Zif268 的三锌指结合区同时进行突变,实现三个单一锌指突变体的串联,以获得特异性识别 9bp 靶序列即 5'-GCAGCTGCT-3'、5'-GATGCTGCA-3'的三锌指多肽,从而为利用上述元件构建杂交转录因子、实现特异基因表达的人为干预奠定基础。

1 材料与方 法

1.1 材 料

1.1.1 菌株与质粒:大肠杆菌 DH5 α 、JM109;质粒 pUC18, pGEX-2T 本室保存, pBKS-Zif268,来源于 ATCC。

1.1.2 试剂:EcoRI、BamHI、T4 DNA 连接酶、Taq DNA 聚合酶、pfu DNA 聚合酶为 Promage、华美公司产品。蛋白标准为 Promage 公司产品,核酸标准、IPTG、dNTP、X-gal 等为华美公司产品。Sephrose 4B 为 Phamarical 公司产品。

1.2 方 法^[6,7]

1.2.1 重叠(Over-lap)PCR^[8]:PCR 引物设计与合成 为完成 Zif268 三锌指元件的 2 次同时突变,我们设计 10 条引物,其中 5'-框内加 BamHI 位点,3'-端添加 EcoRI 位点和终止密码 TAA/TGA。

收稿日期 2000-12-31,修回日期 2001-04-18。

基金项目 国家自然科学基金面上项目(39970162)。

* 通讯作者。Tel 86-10-66931809;Fax 86-10-68215721;E-mail zhaozh@nic.bmi.ac.cn

**现在安徽大学生命科学学院。

1. ZN1 : 5'-AAGGATCCCATGAACGCCCATATGCTTG-3'
2. GCA11 : 5'-CTTCGCCGCCATATCCGC ATCC-3'
3. GCA12 : 5'-GCGGATATGCGCGGAAGATCACCCGACTGAAAAGCGGCGATCGC-3'
4. GCT11 : 5'-CATGCGTAACTTCAGTACTAGTGGCGAACTTGTCCGCCACATCCGCACCCACAC-3'
5. GCT12 : 5'-CACTACTACTGAAGTTACGCATG CAG-3'
6. GAT : 5'-GATTTTGGTATGCCTCACCAGGTTACCCTCTGCGCAAACCTTCTCCAC-3'
7. GCT21 : 5'-CTTTCGCCCATATCCGC ATCCAC-3'
8. GCT22 : 5'-GCGGATATGCGCGACAAG CTCACCCGAGGTAGAAAAGCGGCGATC-3'
9. GCA : 5'-GATTTTGGTATGCCTCTCAGATCACCACTCTGGCAAACCTTCCAC-3'
10. ZC1 : 5'-AAGAATTCATTAGTCCCTTCGTCTTAATGGATTTGGTATGCCTC-3'

粗体部分为酶切位点,阴影部分为重叠互补区,下划线为突变碱基。以上引物由上海生工生物技术公司合成。

重叠 PCR

Zif268 的 DNA 结合区由三个前后串联锌指组成,以其为模板,通过 Over-lap PCR,可获得具有全新 DNA 结合特性的三锌指多肽。重叠 PCR 由 3 次 PCR 完成,以中间锌指 ZF2 为例。

第一次:在 50 μ L 的反应体积中,加入模板:天然 Zif268 10ng,引物 ZN1、GCT12 各 50pmol,dNTP100 μ mol/L,Taq2.5u,94 $^{\circ}$ C 30s,55 $^{\circ}$ C 30s,72 $^{\circ}$ C 45s 反应 30 循环,回收目的片段。

第二次:以 ZC1,GCT11 为引物,以天然 Zif268 基因为模板,回收目的片段,反应条件同反应 1。

第三次:取反应 1、2 目的产物各 10pmol 为模板,引物 ZN1,ZC1 各 50pmol,dNTP 100 μ mol/L,Tag 酶 2u,按照 94 $^{\circ}$ C 30s,50 $^{\circ}$ C 30s,72 $^{\circ}$ C 45s 的条件反应 10 个循环后,94 $^{\circ}$ C 30s,55 $^{\circ}$ C 30s,72 $^{\circ}$ C 45s 反应 20 个循环,回收目的片段,获得中间锌指发生突变的突变体 ZF2。

第一、二两个锌指发生突变的突变体 ZF12 的突变换以 ZF2 为模板,GCA11、GCA12 为引物,其它条件同 ZF1 的突变。3 个锌指均发生突变的突变体 ZF123 突变分两步进行,先以 ZF12 为模板,ZN1、GAT 为引物扩增,再以此扩增产物为模板,ZN1、ZC1 为引物完成全部三锌指的突变,ZZF123 的突变过程同 ZF123。

1.2.2 克隆质粒、表达质粒的构建:参见参考文献^[6]。

1.2.3 DNA 序列分析:由上海博亚生物公司操作。

1.2.4 菌体的诱导表达:挑选单克隆接种培养过夜按 2% 的比例接摇瓶培养(37 $^{\circ}$ C 250r/min)2~5h 至 OD₆₀₀ = 0.3~0.5,加入 10% IPTG 使其终浓度为 1mmol/L,诱导培养(30 $^{\circ}$ C 250r/min)4~6h,取 500 μ L 细菌培养物,12000r/min 离心收菌,沉淀加入 100 μ L

双蒸水和 100 μ L 2 \times 上样缓冲液,重悬后沸水浴 5min,室温下冷却后取 16 μ L,进行聚丙烯酰胺凝胶电泳。积沉胶浓度为 5%,分离胶浓度为 15%。

1.2.5 GST 融合蛋白的纯化:细菌培养物超声波破碎、离心。上清与准备好的 Sepharose 4B 充分混合,一部分转移到层析柱,pBS 洗涤后,用谷胱甘肽洗脱液洗脱。另一部分先用凝血酶切割,再洗脱分离。

2 结果

2.1 ZF123、ZZF123 克隆及序列测定

经过 Over-lap PCR,获突变产物 ZF123、ZZF123,与 pUC18 连接后,转入大肠杆菌 JM109 中。经序列测定,结果与我们设计的突变序列一致。ZF123、ZZF123、Zif268 序列比较见图 1。

Zif268:	5'-ATGCATGAACGCCCATATGCTTGCCCTGTCGAGTCTCTGCGATCGCCGCTTTTCT
ZF123:	5-----
ZZF123:	5-----
Zif268:	CGCTCGGATGAGCTTACCCGCCATATCCGCATCCACACAGGCCAGAAGCCCTTC
ZF123:	-AG---G---T---CG-----
ZZF123:	AC---G---GT-----
Zif268:	CAGTGTGCAATCTGCATGCGTAACTTCAGTCTGATGACCACTTACCAACCCAC
ZF123:	-----AC-----G-G-A---GT-CG---
ZZF123:	-----AC-----G-G-A---GT-CG---
Zif268:	ATCCGCACCCACACAGCGGAGAAGCCTTTTGCCCTGTGACATTTGTGGGAGGAAG
ZF123:	-----
ZZF123:	-----
Zif268:	TTTGCCAGGAGTGATGAACGCAAGAGGCATACCAAAATCCATTTAAGACAGAAG
ZF123:	-----C-----G-A-C-TGCT-----
ZZF123:	-----CA-----G---T-TG-G-----
Zif268:	GACAAGAAAGCAGACAAATAATGA-3'
ZF123:	-----3'
ZZF123:	-----3'

图 1 天然 Zif268 锌指结合区与突变体 ZF123、ZZF123 的序列比较

Fig.1 The sequence comparison of Zif268 DNA binding domain with the mutants ZF123、ZZF123

2.2 表达质粒的构建

取测序鉴定后的 pUC-ZF123,pUC-ZZF123 用 BamHI、EcoRI 双酶切,回收目的片段。与用同样两种限制酶处理后的 pGEX-2T 连接,完成表达质粒的构建。

2.3 表达质粒的鉴定

挑阳性克隆,质粒抽提,EcoRI、BamHI 双酶切

产物用 1.5% 琼脂糖凝胶电泳检测, 出现两条 DNA 带, 位置与预测一致, 表明获得正确的表达质粒。

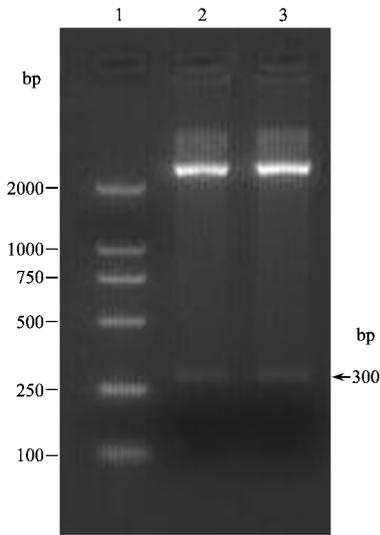


图 2 表达质粒双酶切分析

Fig.2 Restriction enzyme analysis of expression plasmids

1. DNA marker
2. pGEX-ZF123 digested by *EcoRI* and *BamHI* ;
3. pGEX-2ZF123 digested by *EcoRI* and *BamHI*

2.4 融合蛋白的表达

SDS-PAGE 结果见图 3。由图可见在分子量为 34.0kD 处有目的蛋白条带。经扫描分析表明: 重组蛋白约占菌体总蛋白的 20%。

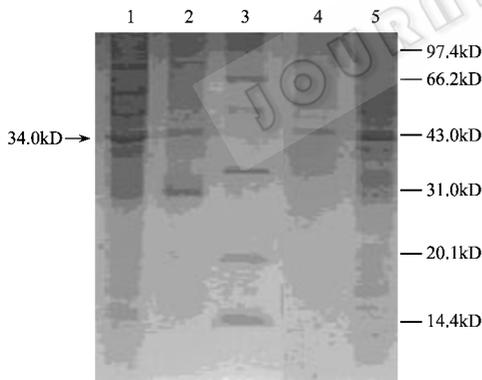


图 3 表达产物的 SDS-PAGE 分析

Fig.3 SDS-PAGE analysis of expression products

1. DH5α (pGEX-ZF123) cell extract without IPTG induction ;
2. DH5α (pGEX-ZF123) cell extract with IPTG induction ;
3. DH5α (pGEX-2ZF123) cell extract with IPTG induction ;
4. DH5α (pGEX-2T) cell extract with IPTG induction ;
5. Protein molecular weight marker

2.5 融合蛋白的纯化

上清用 Sepharose 4B 亲和层析。一部分用谷胱甘肽洗脱液洗脱得到纯化的融合蛋白 GST-ZF123, GST-2ZF123。另一部分用凝血酶切割融合蛋白后, 再行洗脱分离, 得到 GST 和游离的目的蛋白, 见图 4。

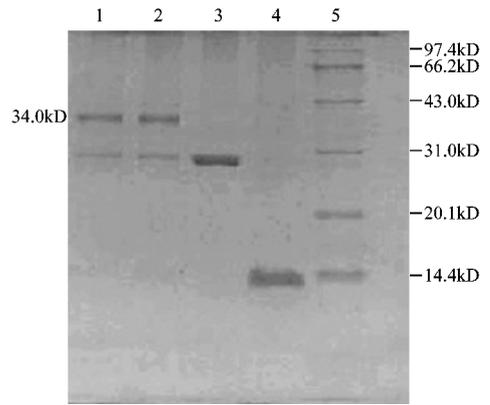


图 4 蛋白纯化产物的 SDS-PAGE

Fig.4 SDS-PAGE analysis of purified production

1. Purified GST-ZF123
2. Purified GST-2ZF123
3. Purified GST ;
4. Purified ZF123
5. Protein molecular weight marker

3 讨 论

外源基因在表达过程中, 表达载体的启动子强弱是一个重要因素。此外诱导后的培养温度、宿主菌以及外源基因密码子的组成等都会影响基因的表达^[9, 10]。综合各方面的考虑, 我们选择 pGEX-2T 为表达载体, IPTG 诱导终浓度为 1.0mmol/L, 诱导后培养温度 30℃, 诱导时间 5h, 结果目的蛋白占总蛋白 20% 左右, 与前期研究的天然锌指 Zif268 以及单锌指突变体表达水平相当。

Zif268 是锌指蛋白家族成员之一, 其 DNA 结合区属于 C₂H₂ 型经典锌指, 3 个串联的锌指结构识别靶 DNA 序列 5'-GCCGTGGCG-3'。与 DNA 相互作用时, 通过 α-螺旋-1、+3、+6 位点与 DNA 大槽上 3 碱基对特异性接触^[3, 11]。在利用锌指蛋白-DNA 相互作用的遗传筛选模型获得单一锌指突变体的基础上, 本文以天然 Zif268 三锌指基因为模板, 通过重叠 PCR 成功地实现了三锌指关键位置同时突变, 序列测定与设计完全一致。这表明, 本文所设计的引物及重叠 PCR 条件是可行的。根据设计, ZF123, 2ZF123 分别识别 5'-GATGCTGCA-3' 和 5'-GCAGCTGCT-3' 序列测定及表达成功为进一步开展其体内、体外 DNA 结合特性以及两者串联构建识别更长靶序列 5'-GCAGCTGCT-NNNN-GATGCTGCA-3' (HIV 特异的 DNA 序列) 的六锌指多肽奠定基础。

REFERENCES (参考文献)

- [1] Henikoff S, Greene EA, Pietrovski S *et al.* Gene families: the taxonomy of protein paralogs and chimeras. *Science*. 1997, **278**(5338):

- [2] Miller J ,McLachlan AD ,Klug A. Repetive zinc-binding domains in the protein transcription factor IIIA from xenopus oocytes. *EMBO J.* 1985 **4** (6) :1609 ~ 1614
- [3] Pavletich NP ,Pabo CO. Zinc finger-DNA recognition :crystal structure of a Zif268-DNA complex in 2.1 Å. *Science* ,1991 **252** (5007) :809 ~ 817
- [4] Christy B ,Daniel Nathan. DNA binding site of growth factor inducible protein Zif268. *Proc Natl Acad Sci USA* ,1989 **86** (22) :8737 ~ 8741
- [5] Choo Y ,Klug A. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc Natl Acad Sci USA* ,1994 **91** (23) :11168 ~ 11172
- [6] Sambrook J ,Fritsch E F ,Maniatis T. *Molecular Cloning :A Laboratory Manual*. 2nd New York :Cold Spring Harbor Laboratory Press ,1989
- [7] LU S D. *Current protocols for molecular biology* ,Beijing : Peking Union Medical College Press (中国协和医科大学出版社) ,1999
- [8] Yolov AA ,Shabarova ZA. Constructing of DNA by polymerase recombination. *Nucleic Acids Res.* ,1990 **18** (13) :3983 ~ 3986
- [9] Hockney RC. Recent developments in heterologous protein production in *Escherichia coli*. *Trends Biotechnol.* ,1994 **12** (11) :456 ~ 463
- [10] Ford KG ,Whitmarsh AJ ,Hornby DQ. Overexpression and purification of eukaryotic transcription factors as glutathione-S-transferase fusions in *E. coli*. *Methods Mol Biol.* ,1994 **30** :185 ~ 197
- [11] Pavletich NP ,Pabo CO. Crystal structure of a five-finger GLI-DNA complex at 2.1 Å. *Science* ,1993 **261** (5129) :1701 ~ 1707

Constructing and Expression of Three Zinc-fingers Peptide with Specific DNA Recognition Property in *Escherichia coli*

ZHANG Shu-Xiang MA Qing-Jun ZHAO Zhi-Hu*
(Beijing Institute of Biotechnology ,Beijing 100850 ,China)

Abstract For investigating the DNA binding property of classical zinc finger protein Zif268 ,an *in vivo* transcription interference experiment was once utilized to develop a genetic selection assay. By screening a library in which the key amino acids of the third zinc finger from Zif268 were randomized ,some single fingers with new binding specificity were obtained. In this study ,by combining the single fingers ,two three-finger peptides cDNA ZF123 and 2ZF123 were constructed by an over-lap PCR technique using the DNA binding domain of Zif268 as the template. After three times PCR ,the products were inserted into pUC18 for cloning. The ZF123 and 2ZF123 cDNA were also inserted into pGEX-2T for expression in *Escherichia coli* after sequencing confirmation. The result showed that the three-finger peptides were expressed at a high level in *E. coli* JM109. The fusion protein GST-ZF123/2ZF123 have the relative molecular weight of 34.0kD and consisted about 20% of the total soluble cell protein as detected by SDS-PAGE. After supersonic treatment ,the soluble part of the bacterial extract was purified. After two additional thrombin cleavage and Sepharose 4B affinity purification steps ,the free three-fingers peptide proteins were also obtained. The construction and obtaining of the three-fingers peptide cDNA and its products will facilitate the *in vivo* and *in vitro* DNA binding specificity study and the design of the hybrid transcription factors.

Key words zinc-finger protein , over-lap PCR , protein expression and purification

Received :December 31 2000

This work was supported by grant from the National Natural Science Foundation of China(39970162).

* Corresponding author. Tel 86-10-66931809 ,Fax 86-10-68215721 ,E-mail :zhaozh@nic. bmi. ac. cn