

一种特异识别 SV40 启动子的人工转录因子的构建

赵兴卉 朱旭东* 刘 娟 饶相君 黄培堂

(北京生物工程研究所 北京 100071)

摘 要 转录因子是真核表达调控中非常重要的一类反式作用因子,通常由 DNA 结合域与效应域两部分组成,而锌指结构是 DNA 结合域的常见组成单元。人工转录因子就是基于转录因子的这种结构特点,人为地选择针对特定序列的 DNA 结合域与具有特定作用的效应域组合而成。利用噬菌体展示技术,筛选到与 SV40 启动子上 9bp 序列特异结合的锌指结构,再连接 KOX1 的 KRAB 域构建了一种人工转录因子。转染实验表明它对 SV40 下游的报告基因的表达有很显著的抑制作用。

关键词 人工转录因子, 锌指, 噬菌体展示, KOX1-KRAB 结构域, SV40 启动子

中图分类号 Q756 文献标识码 A 文章编号 1000-3061(2003)05-0608-05

转录因子(Transcription factor, TF)是一类反式作用因子,能够与真核基因启动子中的顺式元件发生相互作用,因而在真核基因表达调控中至关重要^[1]。通常而言,天然的转录因子至少由两部分组成,即 DNA 结合域(DNA-binding domain)与效应域(Effector domain)。DNA 结合域通过对 DNA 上特异位点的识别与结合使转录因子精确地定位,从而使得效应域可以有有的放矢地针对特定区域起作用。常见的效应域有激活域(Activation domain)和抑制域(Repression domain)等等,分别对 DNA 的转录起到激活或抑制的作用^[2]。根据转录因子的这种结构特点,我们可以设计人工转录因子,特异地识别并结合我们感兴趣的靶序列,并依照我们的需要起到激活或抑制转录的作用。

天然转录因子中,最常见的 DNA 结合域是锌指结构。锌指结构是一种自然界广泛存在的蛋白质结构基序(Motif)^[3],每一个锌指单元由大约 30 个氨基酸组成,形成反平行的 β 折叠片与一个 α 螺旋,通常每一个锌指通过 α 螺旋(被称为识别螺旋)外侧的氨基酸来识别 DNA 双螺旋大沟中的相连的 3bp 序列,并可通过氢键作用与相应的碱基结合^[4]。实验表明每个 α 螺旋外侧有 5~7 个氨基酸位点与该锌指结合 DNA 的特异性及 kD 值有密切联系^[5,6]。多个锌指可以串联成簇,从而识别并结合较长的 DNA 序列

区域。

为了能够在理论上针对任意给定的 DNA 序列都能得到与之特异结合的锌指结构,我们采用噬菌体展示技术来构建锌指库以便筛选。目前已有多种构建用于锌指筛选的噬菌体展示文库的策略^[7-13]。我们采用的是“两部分互补法”(Bipartite-complementary method)^[9]。

常见的转录因子效应域有具有转录激活作用的 p65、VP16 以及在本实验中使用的有抑制转录作用的 KOX1 的 KRAB 域。KOX1 是一种重要的转录抑制因子,是 *Kox1* 基因的产物。经研究知 KOX1 的第 1~75 氨基酸就具有全部的转录抑制活性,这 75 个氨基酸是具有高度同源性的序列,称为 KRAB 域(Kruppel-associated box),几乎存在于所有真核生物中^[14]。

本文利用上述噬菌体展示锌指文库筛选到特异识别并结合 SV40 启动子的锌指,连接 KOX1 的 KRAB 域构建人工转录因子。通过活性的检测实验证实其对 SV40 启动子下游报告基因的表达抑制作用显著。

1 材料与方法

1.1 材料

1.1.1 载体、菌株与细胞株:质粒 pHB-1HScFv,噬

菌体 M13K07 为本室俞炜源教授和刘志刚博士惠赠,质粒 pGEM(R)-T Easy Vector、pGL3-Control、pRL-TK、pcDNA3.1(+)分别购自 Promega 与 Invitrogen 公司。大肠杆菌菌株 JM109、XL1-Blue,以及 HeLa 细胞株为本室保存。

1.1.2 酶与试剂: T4 多核苷酸激酶、T4 DNA 连接酶、限制酶、Taq DNA 聚合酶等分别购自 Promega、BioLab、TaKaRa、Sagon 等公司。ATP、dNTP、BSA、链亲和素、D-MEM 细胞培养基以及常规试剂为本室保存或购自 Sigma、Sangon、Promega、Gibco/BRL 公司。LipofectamineTM2000 购自 Invitrogen 公司。HRP-Anti-M13-Ab 为童贻刚教授惠赠。

1.1.3 试剂盒等: 质粒提取、胶回收、PCR 产物提纯试剂盒分别购自 Promega、Qiagene 及 Roche 公司。链亲和素包被的酶联板购自 Roche 公司。荧光素酶检测试剂盒 Dual-Luciferase[®] Reporter Assay System 购自 Promega 公司。

1.2 文库中锌指序列片段的合成与连接

1.2.1 片段的合成: 依照文献[9],设计普通序列 18 条,含有随机化位点的序列 108 条,由上海博亚公司合成。

1.2.2 文库锌指序列片段的连接: Lib12 与 Lib23 的锌指片段的加磷、退火、连接系常规操作,依照分子克隆实验指南(第二版)进行。

1.2.3 片段的 PCR 扩增: 设计引物 Lib-top: 5'-TCGCGGCCCCAGCCGGCCATGGCGGAAGAGA-3'(30bp) 与 Lib-end: 5'-TGTAGCGGCCGCTTCTGTCTTAAATG-3'(27bp),分别对 Lib12 与 Lib23 的锌指片段连接产物进行 PCR 扩增,并在片段两端分别引入 *Sfi* I 与 *Not* I 的酶切位点。PCR 的条件为: Taq 酶 61℃退火,延伸 30s,做 25 个循环。其他同常规设定。

1.3 噬菌体展示锌指库的构建

1.3.1 载体片段的制备: 质粒 pHB-1HScFv 用 *Not* I 与 *Sfi* I 消化,回收大片段(约 4.3kb)。

1.3.2 建库载体的构建: PCR 扩增之后的 Lib12 与 Lib23 的锌指片段同样用 *Not* I 与 *Sfi* I 消化,与上一步回收的载体片段连接,并乙醇沉淀浓缩去盐。

1.3.3 电转建库: 制备好的转化载体电转大肠杆菌 XL1-Blue 感受态,电转孵育后取部分菌液涂布平板以记数。其余的菌液 37℃孵育后,加终浓度 15% 的甘油,菌体冻存于 -70℃。

1.4 靶序列的确定与合成

1.4.1 在含有报告基因的质粒 pGL3-Control 的 SV40 启动子区域选择一个 9bp 序列 5'-

GCAGAGGCC-3',作为筛选的靶位。

1.4.2 依照文献[9],设计分别用于 Lib12 与 Lib23 筛选的靶序列,交上海博亚公司合成,并在其中加入生物素标记。

Screen12:

A 5'-biotin-ATTTAT GCGGAGGCC GAGGCC-3'

B 3'-TAAATA CGCCTCCGG CTCCGG-5'

Screen23:

A 5'-biotin-ATTTAT GCAGAGGCC GAGGCC-3'

B 3'-TAAATA CGTCTCCGG CTCCGG-5'

Screen:

A 5'-ATTTAT GCAGAGGCC GAGGCC-3'

B 5'-biotin-GCGGGCCTC GGCCTCTCG ATAAAT-3'

1.5 噬菌体展示锌指库的筛选与特异性鉴定

1.5.1 Lib12 与 Lib23 针对各自靶序列的筛选: 文库 Lib12 针对上述合成的片段 Screen12(AB 退火成双链),文库 Lib23 针对上述合成的片段 Screen23(AB 退火成双链)各进行四轮筛选。噬菌体拯救、效价、辅助噬菌体的培养等常规操作依照分子克隆实验指南(第二版)、刘志刚博士论文及文献[9]。

1.5.2 Lib12 与 Lib23 筛选到的产物的重组与再筛选: 重组得到的片段集合针对上述合成的片段 Screen(AB 退火成双链)再筛选。筛选、富集、重组以及 ELISA 检测方法参照文献[9]。

1.5.3 阳性克隆的特异性鉴定: 根据文献[9],合成 9 对生物素标记的序列,每一对也是由 A、B 两链退火而成,每条 A 链与 Screen 的 A 链只相差 9bp 靶位中的一个碱基,B 链亦然,并与 A 链配对。如 Screen-mutation1 两条链的序列是:

A 5'-ATTTAT ACAGAGGCC GAGGCC-3'

B 5'-biotin-GCGGGCCTC GGCCTCTCT ATAAAT-3'

其余类推。根据文献[9],用 ELISA 来鉴定这些阳性克隆的特异性。

1.6 人工转录因子效应域——Kox1 的 KRAB 域的合成与克隆

1.6.1 Kox1 的 KRAB 域分段合成: 通过 <http://srs.ebi.ac.uk> 核酸与蛋白质序列数据库查找到大肠杆菌中 *Kox1* 的 KRAB 域序列,根据人的密码子偏性对个别密码子稍做修改,使其更符合在人的细胞中表达。之后分 9 段合成。

1.6.2 合成片段的退火连接及 PCR 扩增、酶切鉴定: 加磷、退火、连接;之后使用 Taq DNA 聚合酶扩增后连入 T 载体,转化 *E. coli* JM109;挑克隆提质粒,酶切鉴定,均系常规操作,依照分子克隆实验指

南(第二版)进行。

1.6.3 序列测定 选酶切片段大小正确的质粒,送本所 DNA 序列分析中心测序。为了标记方便,在以后的实验中称此序列为 KOX1。

1.7 人工转录因子表达载体的构建

如图 1,我们设计的人工转录因子由我们筛选得到的锌指与 KOX1 的 KRAB 域(这里简记为 KOX1)两部分组成。由于锌指可以特异地识别并结合 SV40 启动子上的 9bp 序列,从而使得 KOX1 能够特异地作用于报告基因(荧光素酶基因)上游的 SV40 启动子,下调报告基因的表达。

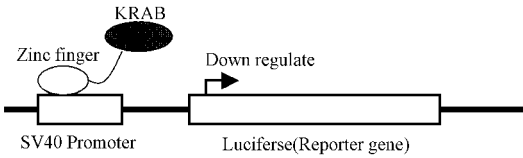


图 1 人工转录因子的作用机制示意图
Fig.1 Structure and function of designed artificial transcription factor

在本实验中,分别把人工转录因子全长(锌指部分 + KOX1 的 KRAB 域)人工转录因子 DNA 结合域(即锌指部分)以及人工转录因子效应结构域(即 KOX1 的 KRAB 域部分)克隆到真核表达载体 pcDNA3.1(+)中,并在其 5'-端 ATG 之后加入 NLS(核定位信号)。

1.7.1 设计引物 用 PCR 的方法在各片段中引入酶切位点以方便克隆并在恰当的位置引入 ATG 起始、终止密码子与 NLS^[16]。

1.7.2 将各片段连接入 pcDNA3.1(+)构建转染载体 测序。

1.8 人工转录因子的活性检测

1.8.1 Hela 细胞的培养 :Hela 细胞接种于 24 孔细胞板,用无抗生素、含 10% 小牛血清的 DMEM 常规培养。

1.8.2 人工转录因子的有效性检测 :待在各孔细胞的汇合度达到 90% 以上时,用 Lipofectamine 2000 转染,每孔都为质粒 pGL3-Control 与 pRL-TK 各 20 ng,与 100 ng 待测质粒共转染。每种待测质粒,重复 3 孔。用 Promega 公司的荧光素酶检测试剂盒(Dual-Luciferase Reporter Assay System)与荧光照度仪(Luminometer)来检测报告基因的表达情况。

1.8.3 人工转录因子的抑制效率检测 除对照外,每孔都为质粒 pGL3-Control 与 pRL-TK 各 20 ng 与一系列质量梯度的人工转录因子全长的表达载体共转染。

每种加量做 3 个孔作为重复实验。检测方法同上。

2 结 果

2.1 特异识别、结合 SV40 启动子上 9bp 序列的锌指的筛选与鉴定

2.1.1 噬菌体展示锌指库的库容估计 通过菌液稀释、平板记数,得到文库 Lib12 与文库 Lib23 的库容都为 $10^6 \sim 10^7$ 。

2.1.2 经过筛选、重组之后得到的锌指特异性鉴定:

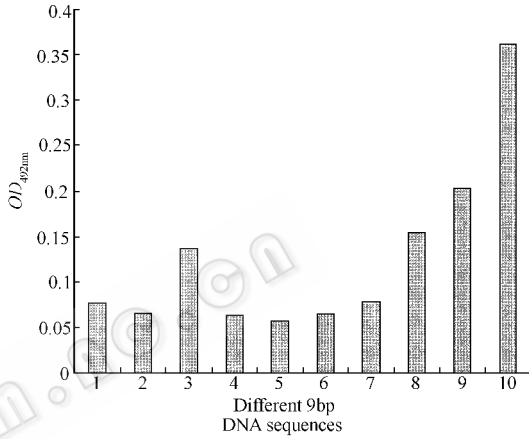


图 2 筛得锌指特异性的鉴定

Fig.2 Zinc finger binding specificity analyses by ELISA
1.5'-aCAGAGGCC-3' 2.5'-GtAGAGGCC-3' 3.5'-GCgGAGGCC-3' ;
4.5'-GCAaAGGCC-3' 5.5'-GCAGgGCC-3' 6.5'-GCAGAAgCC-3' ;
7.5'-GCAGAGaCC-3' 8.5'-GCAGAGGtC-3' 9.5'-GCAGAGGCl-3' ;
10.5'-GCAGAGGCC-3' (9bp sequence used as target in screen)
(1 ~ 9 are mutagenesis sequences each contains single transition mutation site compared with 10 shown in a lower case nucleotide)

从图中可以看出筛选到的这个锌指序列对于只改变一个碱基的靶序列的亲和力就有很明显的下降,表明这个锌指针对目的序列的特异性是非常高的。

2.1.3 筛得特异锌指的序列测定 测序结果显示筛选到的这个锌指的 3 个指头的 α 螺旋外侧-1 2 3 5, 6 位(这 5 个位点是文库的随机化部位)的氨基酸分别为 N, A, D, T, R(F1); R, D, N, S, R(F2); Q, A, S, I, N(F3)。其余框架部分与设计完全一致(测序图略)。

2.2 抑制因子 KOX1 的 KRAB 域克隆测序

测序结果与设计的序列完全吻合(测序图略)。

2.3 人工转录因子的构建

测序结果与设计的序列完全吻合(测序图略)。

2.4 转录因子活性检测

2.4.1 人工转录因子的有效性检测(图 3a):Relay 表

示荧光光度计的读数。由于本实验使用的是双荧光报告系统,引入了内标,从而去除了转染效率等因素的影响,使得 Rdy 的数值可以更加客观地反映报告基因(荧光素酶基因)的表达情况。从图 3 可以看出,共转染人工转录因子表达载体以后,报告基因的表达与对照(空白对照,以及只加入空 pcDNA3.1(+)的对照)相比都有显著的降低。只加入人工转录因子的抑制域,报告基因的表达量与只加入 pcDNA3.1(+)的情况差别不是很大;而只加入人工转录因子的 DNA 结合域,报告基因表达比只加入 pcDNA3.1(+)的一组只有轻微的降低。

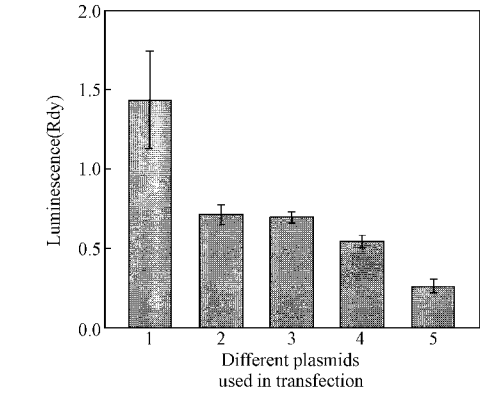


图 3 人工转录因子对荧光素酶表达的抑制作用
Fig.3 Suppression of the expression of luciferase as a reporter gene by artificial transcription factors

1. Blank 2. pcDNA3.1(+) 3. KOX1-KRAB domain only ;
4. Zinc finger only 5. Intact artificial transcription factor

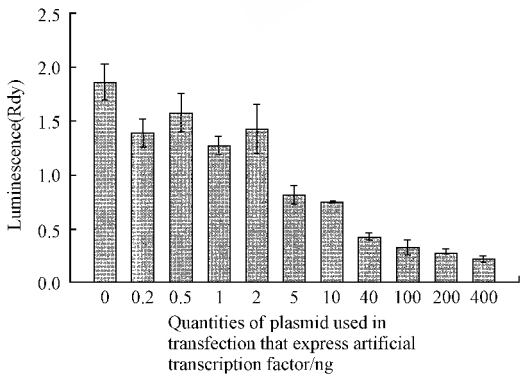


图 4 人工转录因子对荧光素酶表达的抑制作用的初步定量研究
Fig.4 The relation between repression effect and quantity of artificial transcription factor

2.4.2 人工转录因子的抑制效率检测(图 4):从图 4 可以看出,使用一系列质量梯度的人工转录因子全长的表达载体与报告基因载体共转染,发现报告基因的表达随着人工转录因子用量的增加而降低,

荧光素酶表达最高被抑制近 90%。

3 讨论

从荧光素酶的活性检测实验可以看到,人工转录因子可以显著地抑制报告基因的转录(可以达到近 90%)。而仅仅加入 KOX1-KABA 域却没有如此显著的抑制作用,说明锌指使得人工转录因子可以特异地靶向作用位点对于人工转录因子的作用是至关重要的。我们原先设想,只有锌指蛋白,不加抑制因子可能也有一定的抑制作用,因为锌指与 SV40 启动子特异地结合可能会造成空间的位阻作用而使得有激活作用的转录因子无法起作用;不过从实验的结果看,单纯的特异结合 SV40 启动子的锌指蛋白对报告基因的抑制作用是很有限制的。

由于构建了锌指库,我们可以针对感兴趣的序列筛选到与之特异结合的锌指,然后连接效应因子,构建可以针对我们感兴趣的基因起作用的人工转录因子。有研究表明合理设计的人工转录因子能够针对抑癌基因、原癌基因以及病毒的转录表达都起到相应的调控作用^[2,4,7-9],一个合适的人工转录因子甚至可以调节某个基因家族或是整条代谢通路中的一系列相关产物^[2],因此具有广泛的应用潜力。例如,在基因表达调控等方面的基础研究领域,人工转录因子可以通过与顺式元件的作用达到类似基因敲除的效果;在探讨染色体结构的变化与基因表达联系方面,人工转录因子也是一个很好的工具。另外,人工转录因子也可以被应用于肿瘤等疾病的基因治疗与抗病毒治疗等的应用研究中^[2,8,9]。

致谢 俞炜源教授、刘志刚博士提供了建库与拯救所需的噬菌体展示载体与辅助噬菌体,刘志刚博士还在建库方面给予了大量的帮助与建议;童贻刚教授提供了 HRP 标记抗 M13 抗体,在此表示衷心感谢。

REFERENCES(参考文献)

[1] Rolf O ,Rainer R ,Victor L. CTCF is a uniquely versatile transcription regulator linker to ipigenetics and disease. *Trends Genet* , 2001 ,17 (9) :520 - 527

[2] Roger R B ,Carlos F B. Engineering polydacyl zinc-finger transcription factors. *Nat Biotechnol* 2002 20(2) :135 - 141

[3] Miller J , McLachlan A D , Klug A A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* ,1985 4 :1609 - 1614

[4] Aaron K. Zinc finger peptides for the regulation of gene expression. *J Mol Biol* ,1999 293 215 - 218

- toward a therapeutic application. *Proc Natl Acad Sci USA*, 1995, **92** (1): 344 – 348
- [6] Martha L. B, Huang X H, Yen C *et al.* Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. *Proc Natl Acad Sci USA*, 2001, **98** (13): 7158 – 7163
- [7] David J S, Carlos F B. Design of novel sequence-specific DNA-binding proteins. *Curr Opin Chem Biol*, 2000, **4** (1): 34 – 39
- [8] Mark I, Aaron K, Yen C. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nat Biotechnol*, 2001, **19** (7): 656 – 660
- [9] Mark I, Yen C. Rapid, high-throughput engineering of sequence-specific zinc finger DNA-binding proteins. *Methods Enzymol*, 2001, **340**: 593 – 609
- [10] Mark I, Yen C. Engineered zinc finger proteins that respond to DNA modification by HaeIII and HhaI methyltransferase enzymes. *J Mol Biol*, 2000, **295**: 417 – 477
- [11] Michael M, Aaron K, Yen C. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger unit. *Proc Natl Acad Sci USA*, 2001, **98** (4): 1437 – 1441
- [12] Igor F, Roland E K, Ricarda F *et al.* A strategy of exon shuffling for making large peptide repertoires displayed on filamentous bacteriophage. *Proc Natl Acad Sci USA*, 1996, **93**: 7761 – 7766
- [13] Nikola P P, Carl O P. Zinc finger-DNA recognition: crystal structure of a zif268-DNA complex at 2.1 Å. *Science*, 1991, **252**: 809 – 817
- [14] Bellefroid E J, Poncelet D A, Lecocq P J *et al.* The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc Natl Acad Sci U S A*, 1991, **88** (9): 3608 – 3612

Construction of a SV40 Promoter Specific Artificial Transcription Factor

ZHAO Xing-Hui ZHU Xu-Dong* LIU Juan RAO Xiang-Jun HUANG Pei-Tang

(Beijing Institute of Biotechnology, Beijing 100071, China)

Abstract Transcriptions are regulated by transcription factors. Natural transcription factors usually consist of at least two functional domains: a DNA-binding domain and an effector domain. According to this, novel artificial transcription factors are designed to up or down regulate transcription and expression of a target gene. The Cys2-His2 zinc finger domain is a DNA-binding module that has been widely used as the DNA-binding domain in artificial transcription factors. Each zinc finger domain, which comprises about 30 amino acids that adopt a compact structure by chelating a zinc ion, typically functions by binding 3 base pairs of DNA sequence. Several zinc fingers linked together would bind proportionally longer DNA sequences.

According to the “bipartite complementary” library strategy, a pair of zinc finger phage display libraries were constructed. After construction of the libraries, a 9bp sequence (5'-GCAGAGGCC-3') on the promoter of SV40 was chosen as a target for next step. After parallel selection, PCR amplification, desired fragments recovery, re-ligation and additional rounds of selection, phage enzyme-linked ELISA experiments were performed to identify specific binding clones displaying the zinc fingers with predetermined sequence-specificity to our target sequence. Then two clones with strong ELISA signals were chosen to be tested for binding both to its full target site (5'-GCAGAGGCC-3') and to sites containing single transition mutations. The binding specificity of one of the two clones (clone 3) was shown to be fairly good.

The three-finger DNA-binding domain targeted to SV40 promoter, that is, zinc finger sequences on clone 3, was fused to KOX1 suppression domain KRAB and cloned into pcDNA3.1(+) (which expression product was artificial transcription factor). The zinc finger (which expression product was the DNA-binding domain of artificial transcription factor) and KRAB domain only (which expression product was effector domain of artificial transcription factor) were also cloned separately into the same expression vector. All constructs contained an N-terminal nuclear localization signal. Every of the vectors (including pcDNA3.1(+)) without inserting sequences were cotransfected with pGL3-Control and pRL-TK and the activity of luciferase was used to indicate the function of product from transfected expression vectors. Our artificial transcription factor was proved to repress the expression of reporter gene efficiently, while with only DNA-binding domain or effector domain the repression was not remarkable.

By adding different effector domains and changing the DNA-binding domain, artificial transcription factor would have a wide range of potential applications.

Key words artificial transcription factor, zinc finger, phage display, KOX1-KRAB domain, SV40 promoter