

缺失突变法研究尿激酶——单链抗体融合基因在大肠杆菌中的高表达*

王 翔 俞炜源

(军事医学科学院生物工程研究所 北京 100071)

摘 要 对抗人纤维蛋白二聚体单链抗体与 scu-PA-32K 融合基因的表达质粒进行内切酶切割、拼接和外切酶 Bal 31 消化,得到了该融合基因的一系列缺失突变体。通过对这些突变体表达情况的研究,将影响表达的关键因素定位于基因的 841~851 位,可能对应于两个连排的大肠杆菌稀有密码子 AGC(精氨酸),用 PCR 定位诱变法把这两个 AGG 均改造成 CGT 后,融合蛋白的表达量从改造前的 2%~3% 提高到了约 30%。

关键词 大肠杆菌,高表达,缺失突变体,稀有密码子,PCR 定位诱变

分类号 Q552 **文献标识码** A **文章编号** 1000-3061(1999)01-0023-27

过去的工作中曾得到了一株特异性的抗人纤维蛋白 D-二聚体的单链抗体的基因^[1],有望用于导向溶栓,将该基因与低分子量单链尿激酶(scu-PA-32K)基因融合后,对表达产物进行分析时,可测到两种蛋白的活性,但融合基因的表达量较低(约 2%~3%),这成了进一步研究的主要障碍,鉴于我们曾在大肠杆菌中表达过全长的尿激酶基因^[2],表达量也不高,因此影响表达的主要因素可能存在于尿激酶基因内部。本文通过构建各种缺失突变体搜索影响表达的关键原因,并以此为依据对基因进行改造并实现了高表达。

1 材料与方法

1.1 材料

1.1.1 菌种和质粒: *E. coli* XLI-blue 及 BL21(DE3) 本室保存。pET 3b,为带 T7 启动子的表达质粒,本所崔立斌博士惠赠。pET 3b/C119,是在 pET 3b 载体上构建的抗人纤维蛋白二聚体单链抗体与 scu-PA-32K 融合基因的表达质粒,本室构建。

1.1.2 工具酶及生化试剂: 内切酶 *EcoRI*, *PvuII*, *XbaI*, *PstI*, *HindIII*, *BamHI* 及 T4 DNA 连接酶、小牛肠碱性磷酸酶(CIAP)均购自 Promega 公司,核酸外切酶 Bal31 购自 New England Biolabs 公司,抗尿激酶 IgG 由本室制备,其它试剂为分析纯。

1.2 方法

1.2.1 融合基因酶切缺失突变体的构建: 以 pET 3b/C119 质粒为基础,利用融合基因内部的几个 *PvuII*、*EcoRI* 和 *BamHI* 位点,对整个基因进行剪切并拼接,得到 4 个初步的大范围的缺失突变体(B1~B4)。

* 国家高技术研究发展计划项目(No. 102-09-03-01)

收稿日期:1997-10-20,修回日期:1998-10-12。

1.2.2 初步缺失突变体的表达研究：将上述 4 个缺失突变体质粒转化 *E. coli* BL21 (DE3) 30℃ 培养至 $OD_{600} = 0.5$,用 1mmol/L IPTG 诱导 4h ,SDS-PAGE 检测。

1.2.3 Bal 31 外切缺失突变体库的建立：在初步结果的指导下 ,以缺失突变体 B2 为出发点 ,用 *Bam*HI 酶切将质粒线性化后用核酸外切酶 Bal 31 消化 ,于不同时间取出小份终止 ,合并后热灭活 Bal 31 ,再用 *Xba*I 酶切得到融合基因的缺失突变体库 ,将之重新克隆于 pET 3b 载体的 *Xba*I 和 *Bam*HI(补平)位点 ,形成缺失突变体的表达库。

1.2.4 外切缺失突变体的表达研究：通过酶切鉴定选出一系列长度的缺失突变体 ,先进行一轮较粗的表达比较研究 ,确定了表达量突变的区间后 ,再在此区间内精细选择几个质粒进行进一步的表达研究 ,以准确定位影响表达的因素的位置 ,表达条件同 1.2.2。

1.2.5 关键质粒的长度测量：将表达量突变前后的两个质粒双酶切后与分子量标准混合 ,在同一个道内电泳 ,以排除缓冲体系对泳动率的影响。另外 ,如果外切产物的最后一个碱基是 G 将使 *Bam*HI 位点恢复 ,必要时可用 *Bam*HI 酶切帮助判断。

1.2.6 PCR 定位诱变：除原有引物(scFv 上游引物和 scu-PA-32K 下游引物)外 ,合成一对诱变引物 ,将 844~849 位两个 AGG 密码子诱变成同义密码子 CGT ,方法见文献 [3]。

1.2.7 序列测定：使用 T7 Sequenase 2.0 试剂盒。

1.2.8 诱变前后的表达量比较研究：方法同 1.2.2。

1.2.9 Western blot 分析：诱变前后的表达产物经 SDS-PAGE 分离 ,半干法转移至硝酸纤维素膜上 ,在含 3% BSA 的 Tris-HCl(pH7.4)溶液中室温振洗 30min 封闭 ;然后与抗 UK 血清(1:100)稀释液振荡过夜(室温) ,0.01mol/L PBS-Tween 20 液洗涤 ;与抗 HRP-SPA 室温反应 30min 3-3 二氨基联苯胺(DAB)显色 ,2% 硫酸终止。

2 结 果

2.1 酶切缺失突变体的构建

融合基因表达质粒 pET 3b/C119 中的若干有用的酶切位点和初步构建的几个酶切缺失突变体如图 1 所示 ,碱基位置以融合基因的第一个 ATG 为 1。

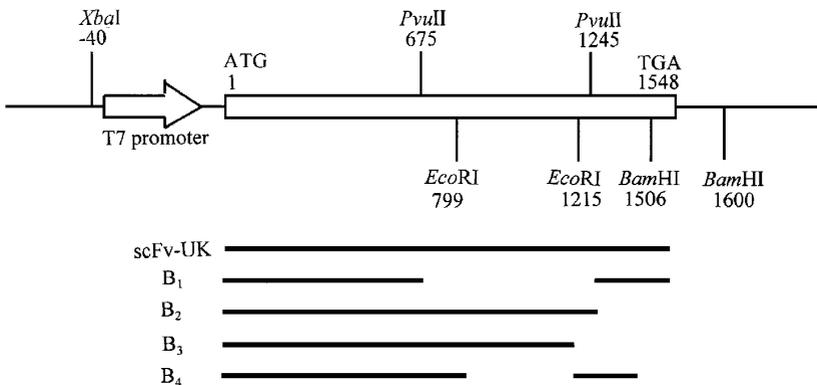


图 1 scFv-scu-PA-32K 融合基因酶切位点及酶切缺失突变体的翻译区段示意图

Fig.1 Useful restriction sites of scFv-scu-PA-32K fusion gene and

the translation regions of restriction deletion mutants.

几个酶切缺失突变体的构建方法如下：

B1 :用 *Pvu* II 切割除去 675~1245 段。

B2 :用 *Pvu* II 及 *Bam* HI 切割除去 1245~1600 段。

B3 :用 *Eco* RI 及 *Bam* HI 切割除去 1215~1600 段。

B4 :用 *Eco* RI 切割除去 799~1215 段。

其中 B2 缺失突变体的 *Pvu* II 与 *Bam* HI 补平后后端连接使 *Bam* HI 位点得以恢复 , 在此位点后 pET3b 载体上有 3 种读框的终止密码子 , 所以在构建缺失突变体库时 Bal31 消化的缺失片段可直接插入此位点 , 无需再加终止密码子 , 只是 3 种读框的翻译产物长度有区别。

gagCAGGATCCggctgcTAAcaagccccgaaaggaagctgagttggctgctgccaccgctgagcaaTAAcTAGcat
(大写字母依次显示连接点、-1、0 及 +1 读框的终止密码子)

上述几个切点中 , 两个 *Pvu* II 切点处于同一读框 , 两个 *Eco* RI 不在同一读框 , 经过酶切拼接后翻译情况有一定变化 , 全长基因及 4 个缺失突变体的翻译区段见图 1。

表达研究结果 , 缺失突变体 B1 及 B4 的表达量很高 , 如图 2 所示 , 表明影响表达的主要因素在 799~1215 范围内。

2.2 Bal31 外切缺失突变体库的构建及表达研究

以缺失突变体 B2 为进一步研究的出发点得到了 Bal31 外切缺失突变体的表达库 , 缺失突变体末端长度范围为 650~1245 , 从中初步选取一批质粒 C1~C8 , 它们的末端长度分别约为 700、800、850、900、1000、1100、1200 和 1245 , 表达结果如图 3 所示 , 与 C2 相比较 C3 的表达量显著下降。

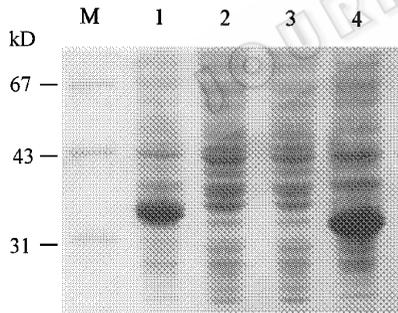


图 2 酶切缺失突变体的表达结果

Fig. 2 Expression of restriction mutants

M. Protein molecular weight market ; 1~4 B1~B4

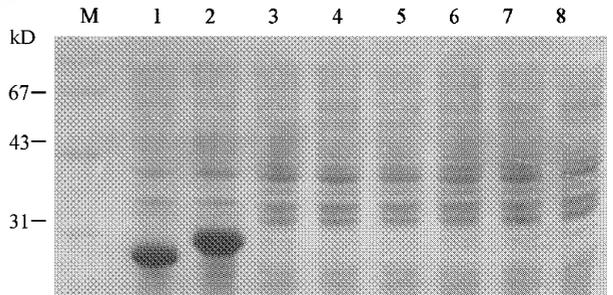


图 3 外切缺失突变体库在大肠杆菌中的初步表达研究

Fig. 3 Primary study of the expression of exo-deletion mutants in *E. coli*

M. Protein molecular weight market ; 1~8 C1~C8

据此结果 , 影响表达的因素大致可定位于第 800 个碱基以后的序列中 , 再在此范围内精细选取几个克隆 D1~D8 , 末端长度约为 780、810、830、840、850、870、890、900。表达结果如图 4 所示 , 其中 D2 由于终止密码子读框的关系 , 表达产物的分子量与 D4 相仿。表达量的突变区间应在 D4 之后 , 经精细测算其范围为 841~851 , 序列是 acAGGAGGcac , 844~849 位是两个连续的 AGG 密码子。

2.3 PCR 定位诱变及结果检测

根据以上研究结果,设计一对诱变引物如下:

引物 M1 5'-ACGACGGTAGATGGCCGCAAACC-3' 23 mer

引物 M2 5'-TGCGGCCATCTACAGGAGGCACCGGGGGGGCTC-3' 33 mer

用以上两个引物及原有的 scFv 上游引物、UK 下游引物一起,采用文献中的方法进行 PCR 定位诱变^[3]将 844~849 位的两个 AGG 密码子改造成了 CGT 密码子,经测序与设计相符,表达研究发现,改造后的目的基因表达量约占菌体总蛋白的 30%,比改造前提高了一个数量级(图 5),Western blot 的结果进一步证实了表达带是我们所需的产物(图 6)。

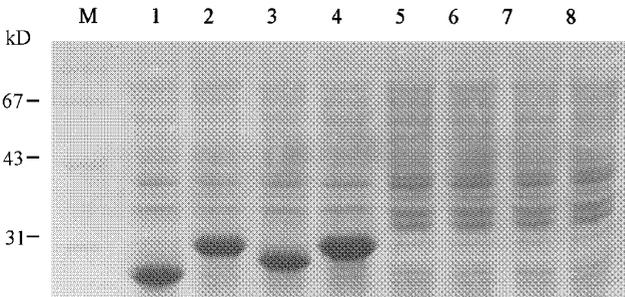


图 4 外切缺失突变体库在大肠杆菌中的精细表达研究

Fig. 4 Precise study of the expression of exo-deletion mutants in *E. coli*

M. Protein molecular weight market; 1~8. D1~D8

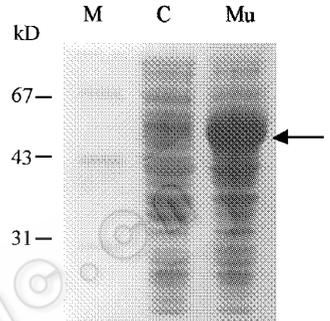


图 5 诱变改造前后的表达量变化

Fig. 5 Expression of the fusion gene before and after site-mutation
M. Protein molecular weight market; C. Control; Mu. Mutant. The arrow show the fusion protein

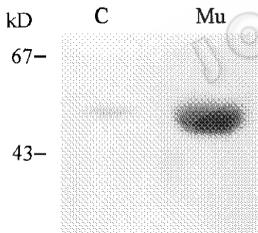


图 6 Western blot 结果

Fig. 6 Result of Western blot
C. Control; Mu. Mutant.

3 讨论

很多外源的真核基因可以很容易地在大肠杆菌中获得高表达,然而也有许多外源基因的表达难尽人意,从多年来的研究结果中已总结出了许多影响表达的因素,涉及宿主、转录、翻译、翻译后等多个方面^[4~6]。

此前,我们所表达的 scFv 可以达到很高的表达量,而融合基因在多种载体上的表达效果均不佳。我们曾在大肠杆菌中表达过全长的尿激酶基因,表达量也很低,据此可以初步推断融合基因的低表达主要由基因内部结构造成。于是,我们选用 Bal31

外切酶建立了一种专门用于搜索基因内部影响表达的未知因素的方法,此方法可一定程度上避免主观盲目性,该方法有可能用于解决大肠杆菌系统中的其它基因表达问题。

关于大肠杆菌慎用码对外源基因表达的影响,有过一些报道,连续的稀有密码子在一些事例中明显地影响表达,尤其是精氨酸密码子 AGG、AGA 和异亮氨酸密码子 AUA^[7,8]。本文的结果也证实了这一点,然而,也有许多含有串联的 AGG 或 AGA 密码子的外源基因可以在大肠杆菌中获得高表达,因此,串联的稀有密码子在什么情况下、如此影响基因的表达值得进一步深入研究。

参 考 文 献

- [1] 黄君健,俞炜源,黄翠芬. 生物技术通讯,1996,7(1):1~5.
[2] 胡宝成,李 军,俞炜源等. 生物工程学报,1993,9(2):122~127.
[3] 黄培堂,俞炜源编译. PCR 技术原理和应用. 北京:中国科学技术出版社,1990.
[4] M. L. Sprengart, H. P. Fatscher, E. Fuchs. *Nucleic Acids Res.*, 1990, 18(7):1719~1723.
[5] 隋广超,胡美浩. 生物化学与生物物理进展,1994,21(2):128~132.
[6] G. Gross, C. Mielke, I. Hollatz *et al.* *J. Biol Chem.*, 1990, 265(29):17627~17636.
[7] B. J. Del Tito Jr, J. M. Ward, J. Hodgson *et al.* *J. Bacteriol.*, 1995, 177(24):7086~7091.
[8] J. F. Kane. *Curr Opin Biotechnol.*, 1995, 6(5):494~500.

High Level Expression of the Deleted and Mutated Urokinase-scFv Fusion Gene in *Escherichia coli* *

Wang Xiang Yu Weiyuan

(*Institute of Biotechnology, Academy of Military Medical Sciences, Beijing 100071*)

Abstract The fusion gene of a specific anti-human fibrinogen D-dimer scFv and low molecular weight single chain urokinase (scu-PA-32K) was restricted, spliced and digested by exonuclease Bal31 to obtain a series of deletion mutants. Study of their expression in *E. coli* revealed that the key sequence which reducing its expression level resides in the range from 841 bp to 851 bp in which a tandem AGG codons (encoding arginine, rarely used in *E. coli*) exists. By means of PCR mediated site-mutation, we altered these two AGG codons to CGT codons which is more efficiently translated in *E. coli*, and the expression level turned out to be about 30% of the total bacterial proteins while that of nature gene to be 2%~3%.

Key words *E. coli*, high level expression, deletion mutant, rare codon, PCR mediated site-directed mutagenesis

* Supported by Project of Chinese National Programs for High Technology Research and Development (No. 102-09-03-01).