

## 人 $\alpha$ -肿瘤坏死因子基因的高效表达

吕兆海<sup>4\*</sup> 王子轩<sup>1</sup> 闵永洁<sup>2</sup> 包慧中<sup>3</sup> 刘新垣<sup>1</sup>

(中国科学院上海生物化学研究所 上海 200031)<sup>1</sup>

(复旦大学遗传所 上海 200433)<sup>2</sup>

(上海华新生物高技术有限公司 上海 200233)<sup>3</sup>

(广东省微生物研究所 广州 510070)<sup>4</sup>

**摘 要** 用人工合成的  $\alpha$ -肿瘤坏死因子 (TNF- $\alpha$ ) 基因构建了五个不同的表达质粒, 它们不同之处是 SD 序列与起始密码子 ATG 间距离 (D) 各异。计算机模拟计算出翻译起始区域 (TIR) 中二级结构的最小生成自由能, 以 (D) 为 6 个核苷酸时最小 (绝对值)。它的表达效率也最高, 产物 TNF- $\alpha$  可达菌体总蛋白的 60%。密码子的选用对表达效率有很大的影响, 故人工合成 TNF- $\alpha$  基因 (选用大肠杆菌偏爱的密码子) 的表达效率高于 sc-DNA (对部分密码子改造的半合成 cDNA)。

**关键词**  $\alpha$ -肿瘤坏死因子 (TNF- $\alpha$ ), 人工合成基因, 半合成基因, 高效表达, 翻译起始区域 (TIR)

肿瘤坏死因子是巨噬细胞或单核细胞激活后分泌的一种蛋白质。小鼠经卡介苗和脂多糖处理几小时后, 在血清中含有一种可以使动物体内的肿瘤发生出血性坏死的物质, 1975 年 Carswell 等将该物质命名为肿瘤坏死因子 (TNF, Tumour Necrosis Factor)<sup>[1]</sup>。以后发现淋巴毒素 Lymphotoxin (LT) 也属于这类物质, 将原来的 TNF 称为 TNF- $\alpha$ , 而将 LT 称为 TNF- $\beta$ 。TNF- $\alpha$  含 157 个氨基酸残基, 参与机体炎症与免疫反应, 在抗肿瘤与抗病毒活性方面, 与干扰素、白细胞介素等有协同作用, 因此肿瘤坏死因子可能有广泛的临床应用前景<sup>[2,3]</sup>。

本工作用人工合成的 TNF- $\alpha$  基因及 N-端部分氨基酸密码子作了更换的 cDNA, 构建了多个表达质粒, 在大肠杆菌中得到了高效表达。

### 1 材料和方法

#### 1.1 材料

pLY-4 是本组构建的高效表达载体, 含 P<sub>L</sub> 启动子, 温控阻遏蛋白 Clts857 基因。大肠杆菌 JF1125 由本组保存。限制性内切酶、DNA 修饰酶购自美国 Promega, GibcoBRL 或瑞典 Pharmacia 公司。聚合酶链反应 (PCR) 所需试剂购自美国 Perkin Elmer Cetus 公司。放射性同位素购自英国 Amersham 公司。DEAE-Sepharose FF, S-Sepharose FF 层

\* 先在中科院上海生物化学研究所进修, 后在上海华新生物高技术有限公司工作。  
本文于 1993 年 9 月 2 日收到。

析填料购自瑞典 Pharmacia 公司。其它化学试剂均选用国产分析产品。s-TNF（人工合成的肿瘤坏死基因）由闵永洁构建，sc-TNF（经修饰的  $\alpha$ -肿瘤坏死因子 cDNA，它的 5' 端有一小段序列为人工合成）由南京大学徐贤秀赠送。

## 1.2 方法

**1.2.1 寡聚核苷酸的合成及纯化：**用亚磷酸三酯法在 DNA 合成仪上完成寡聚核苷酸的合成。脱保护基后，溶解于 TE 缓冲液中，走 20% 聚丙烯酰胺-7M 尿素变性胶电泳分离，从电泳胶上切下所需条带，TE 浸泡过夜，浸出液用 Sephadex G-25 柱脱盐，冷冻浓缩，测定  $A_{260}$  值后备用。

**1.2.2 DNA 重组操作：**质粒提取，DNA 连接，内切酶水解鉴定等操作参照《分子克隆》<sup>[4]</sup>，DNA 序列测定用双脱氧末端终止法<sup>[5]</sup>，PCR 反应按 Perkin Elmer Cetus 公司的试剂盒说明书进行。

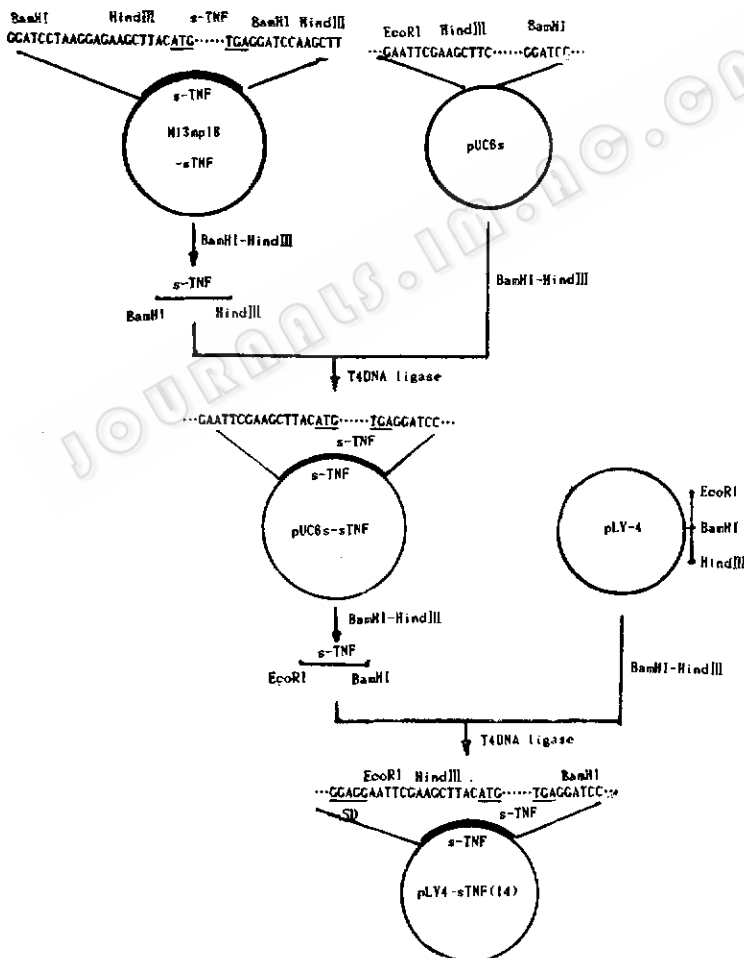


图 1 表达质粒 pLY4-sTNF (14) 的构建

Fig. 1 Construction of expression plasmid pLY4-sTNF (14)

1.2.3 最小生成自由能的计算: 按已有的设计程序, 模拟表达质粒中 0—80 个核苷酸间最稳定的二级结构以计算其最小生成自由能  $\Delta G^{\circ}_{T298}$  值。

1.2.4 工程菌摇瓶发酵: 挑单菌落接种 3ml LB (含氨基青霉素 100 $\mu$ g/ml), 30 $^{\circ}$ C 振荡培养过夜。以 1:100 比例转接发酵培养液 (M9 加 10%LB, 或 M9 加 10%CA, 或 YT 培养液), 30 $^{\circ}$ C 培养至 OD<sub>600</sub> 为 0.5, 42 $^{\circ}$ C 热诱导培养 3 小时。离心收集菌体, 取样走 SDS-PAGE 蛋白电泳, 方法参照 Laemmli 等人的条件<sup>(6)</sup>。蛋白电泳黑度扫描在 Shimadru CS-930 双波长扫描仪上进行。

1.2.5 重组 TNF 的细胞毒性活性测定: 参照 Carewell<sup>(3)</sup>和 Aggarwal<sup>(7)</sup>方法。小鼠成纤维细胞株 L929, 加丝裂霉素 C 共培养。加入待测样品, 继续培养 18 小时, 用中性红活性染料将细胞染色, 观察结果。

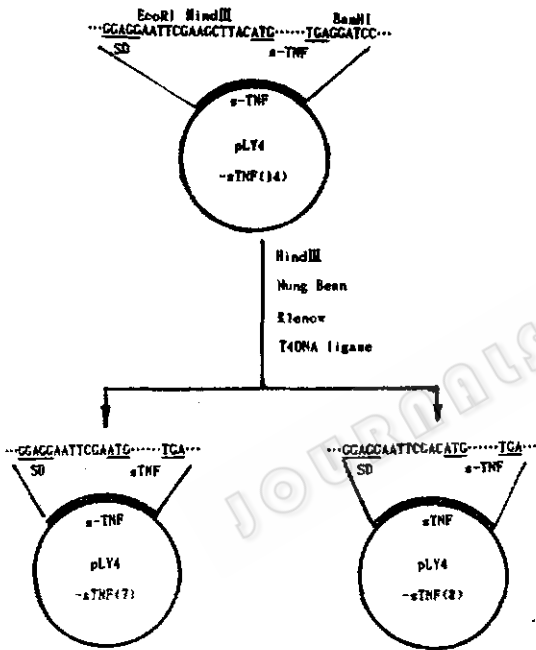


图 2 表达质粒 pLY4-sTNF (7), pLY4-sTNF (8) 的构建

Fig. 2 Construction of expression plasmid pLY44-scTNF (7) and pLY4-sTNF (8)

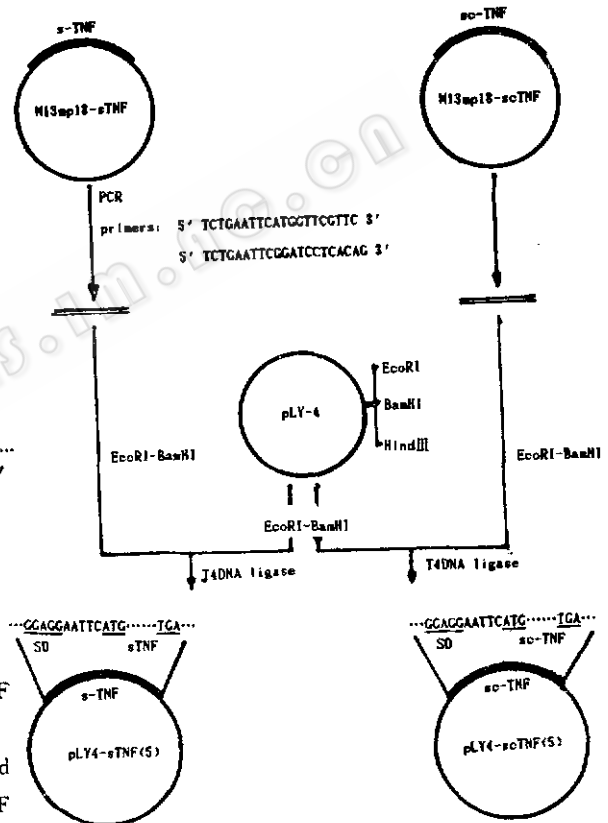


图 3 表达质粒 pLY4-sTNF (5), pLY4-scTNF (5) 的构建

Fig. 3 Construction of expression plasmid pLY4-sTNF (5) and pLY4-scTNF (5)

## 2 结 果

### 2.1 TNF- $\alpha$ 基因的鉴定

成熟 TNF 分子由 157 个氨基酸残基组

成, 采用经改造的肿瘤坏死因子 cDNA (简称 sc-TNF) 和人工全合成的 TNF 基因 (简称 s-TNF) 来构建 TNF 表达质粒。sc-TNF 只在原 cDNA 基础上改变其 N 端第 1、2、3、6 位氨基酸的密码子。s-TNF 的全长为 471 个碱基, 采用了大肠杆菌偏爱密码子。它与

cDNA 比较,共改变了 113 个碱基,更换了 102 个氨基酸的密码子。将 s-TNF 及 sc-TNF 基因与 M13mp18 重组,得到 M13mp18-sTNF, M13mp18-scTNF。经 DNA 序列测定两种基因完全正确。

## 2.2 不同 TNF 表达质粒的构建

为了研究 SD 序列与起始密码子 ATG 之间距离 (D) 对表达效率的影响,我们借用 pUC-6s 质粒,构建了一系列 SD-ATG 不同距离 (D) 的 TNF 表达质粒。质粒 pLY4-sTNF (14) (括号中数字表示 SD 序列与起始密码子 ATG 间的距离,核苷酸数目) 中, D 长度为 14 个核苷酸 (图 1)。而 pLY4-sTNF (7) (图 2)、pLY4-sTNF (8) (图 2)、pLY4-sTNF (5) (图 3)、pLY4-scTNF (5) (图 3) 和 pLY4-sTNF (6) (图 4) 中, SD-ATG 之间的距离 (D), 依次分别为 7、8、5、5 和 6 个核苷酸, 其中 pLY4-sTNF (5) 采用了人工全合成基因, 而 pLY4-scTNF (5) 则为 cDNA, 只是在基因编码区 N-端作了部分修饰而已。质粒构建过程详见图 1、2、3 和 4。

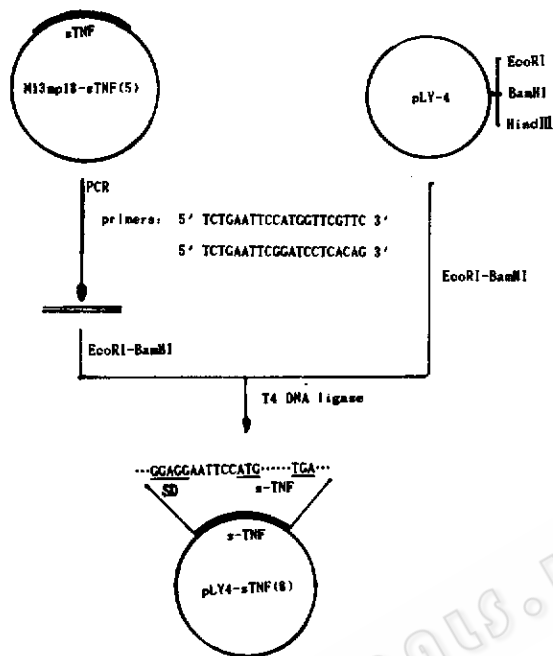


图 4 表达质粒 pLY4-sTNF (6) 的构建

Fig. 4 Construction of expression plasmid pLY4-sTNF (6)

## 2.3 人工全合成 TNF- $\alpha$ 基因在不同 TIR 载体中的表达效率

用 sTNF- $\alpha$  基因构建了五种不同的表达质粒, 它们的表达效率见图 5, 在 pLY4-sTNF (5), (6), (7), (8), (14) 5 种不同表达质粒中, 表达效率最高者为 pLY4-sTNF (5), pLY4-sTNF (6), 两者表达的 TNF- $\alpha$  最高均占菌体总蛋白 60%。其他三者 (D) 的长度增加, 效率也较低。用计算机模拟各表达质粒 TIR 区域 mRNA 的二级结构, 发现各质粒中, SD 序列与 ATG 在 TIR 区域中所处位置是相似的 (图未列出), 但计算出来的最小生成自由能 ( $\Delta G^{\circ}_{f_{298}}$ ) 不同。pLY4-sTNF (5), (6), (7), (8), (14) 的  $\Delta G^{\circ}_{f_{298}}$  值分别为 -32.44, -31.60, -40.60, -40.18, -39.89 kJ/mol, 其中 pLY4-sTNF (5)、(6) 两者的 TIR 区域中生成稳定的二级结构时, 放出的能量最小, 因而分子最不稳定, 这样有利于翻译的起始。所以 pLY4-sTNF (5), pLY4-sTNF (6) 表达效率较其它质粒高。

## 2.4 比较人工全合成 TNF- $\alpha$ 基因及半合成 TNF- $\alpha$ 基因的表达水平

为了研究全合成基因中大肠杆菌偏爱密码子在表达中的作用, 我们将 s-TNF- $\alpha$  及 sc-TNF- $\alpha$  基因分别装入同一种表达载体 (pLY-4) 中, 得到 pLY4-sTNF (5) 及 pLY4-scTNF (5), 两者的表达水平见图 6, sTNF- $\alpha$  的表达水平可达菌体总蛋白的 60%, 而 scTNF- $\alpha$  只有 20%, 左右相差甚远, 由此可见, 大肠杆菌偏爱密码子的选用, 在 TNF- $\alpha$

的高效表达中起了很重要的作用。



图 5 SD 与 ATG 间距离对 TNF- $\alpha$  表达影响  
Fig. 5 Effect of the distance between the SD and the initiation codon on TNF- $\alpha$  expression

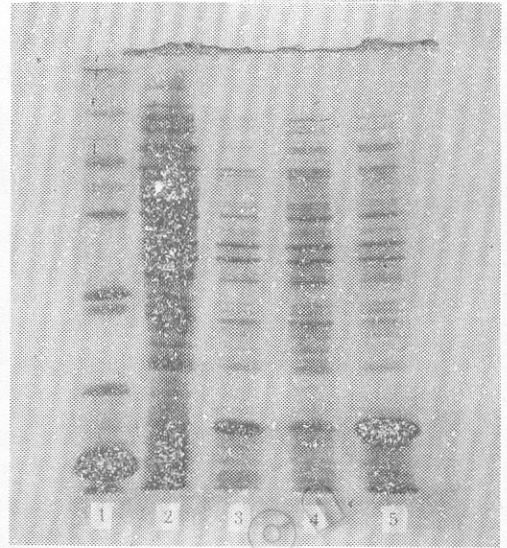


图 6 比较 s-TNF- $\alpha$  与 sc-TNF- $\alpha$  的表达  
Fig. 6 Comparison of expression for sTNF- $\alpha$  and scTNF- $\alpha$

### 3 讨 论

在大肠杆菌中高效表达外源基因，特别是人的基因，是生物工程所面临的一个大问题。为得到高表达，除选用强启动子外，通常考虑用大肠杆菌偏爱的密码子替换原来的密码子，以提高翻译效率。还有，翻译起始区域 (Translation Initiation Region, TIR) 也很重要，即 SD 序列与 ATG 距离要适当，并要尽量减少 mRNA 在该区域中形成不利于翻译起始的二级结构，使 SD 序列处于易被结合的状态<sup>[8,9]</sup>。

本实验选用了两种 TNF 基因，s-TNF 和 sc-TNF。sc-TNF 是在原 TNF 的 cDNA 基础上替换了 N 端的第 1, 2, 3, 6 号氨基酸密码子，以提高表达效率。而且实际表达水平仍低于全部人工合成的 TNF- $\alpha$  基因，因为后者采用了大肠杆菌偏爱的密码子。而且即使用人工合成的 TNF- $\alpha$  基因，其构建的质粒不同，表达效率也差别很大，pLY4-sTNF (5), (6), (7), (8), (14) 等质粒的表达效率的各不相同的原因：一是 TIR 中 SD 序列与 ATG 间距离 (D) 长度不同，二是 TIR 中二级结构的最小生成自由能绝对值高低不同之故。

采用 pLY4-sTNF (5) 或 (6) 转化受体菌发酵。由于表达效率很高，重组 TNF 在细菌内堆积形成不溶的包涵体，但仍有相当部分以溶解状态存在。改变发酵的条件，可以提高可溶性的 TNF 含量，约为包涵体中 TNF 含量的 2 倍。

TNF 有很好杀伤肿瘤细胞的作用，但毒性太大，很多人对 TNF- $\alpha$  分子作了改造。1988 年日本 Soma 等对 TNF- $\alpha$  作了改造，得到一种活性高，毒性低的 TNF- $\alpha$ <sup>[10]</sup>。1993 年常金丽报道，将 TNF- $\alpha$  的 N 端进行改造，可提高活性 100 多倍<sup>[11]</sup>，但目前他们尚无纯化

TNF- $\alpha$  的报告。

致谢:

郭海鸿、杨思远(上海华新生物高技术有限公司), 孙兰英、李孟津(中科院上海生化所), 王启松(复旦大学遗传所)等同志也参加了本文的工作, 特致谢意。

### 参 考 文 献

- [1] Carswell E A *et al.* Proc Natl Acad Sci USA, 1975, 726 : 3666.
- [2] Munher *et al.* Nature, 1986, 323 : 79.
- [3] Neta *et al.* J Immunol, 1988, 140 : 108.
- [4] Sambrook *et al.* Molecular Cloning, Cold Spring Harbor Lab. Press, 1989.
- [5] Sanger *et al.* Proc Natl Acad Sci USA, 1977, 74 : 5463.
- [6] Laemmli *et al.* Nature, 1970, 227 : 680.
- [7] Aggarwal *et al.* J Biol Chem, 1985, 260 : 2345.
- [8] Gold *et al.* Annu Rev Microbiol, 1981, 35 : 363.
- [9] Annu *et al.* The EMBO J, 1984, 3 : 623.
- [10] Soma *et al.* J Biol Response Mod, 1988, 7 : 587.
- [11] 常金丽等, 中国生物工程学会成立大会论文摘要, 1993, p. 53.

## High Level Expression of Human Tumor Necrosis Factor- $\alpha$

Lü Zhaohai<sup>4</sup> Wang Zixuan<sup>1</sup> Min Yongjie<sup>2</sup> Bao Huizhong<sup>3</sup> Liu Xinyuan<sup>1</sup>

(Shanghai Institute of Biochemistry, Academia Sinica, Shanghai 200031)<sup>1</sup>

(Institute of Genetics, Fudan University, Shanghai 200433)<sup>2</sup>

(Shanghai Huaxin High Biotechnology Inc. Shanghai 20033)<sup>3</sup>

(Guangdong Institute of Microbiology, Guangzhou 510070)<sup>4</sup>

**Abstract** Five expression plasmids for a total chemically synthesized gene of human tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) were constructed. There are two different points: 1. the distance (D) between the SD and the initiation codon ATG; 2. the energy- $\Delta G^{\circ}_{1298}$  for forming the stable secondary structure in the translation initiation region. The plasmid with a (D) of 6bp has the lowest  $\Delta G^{\circ}_{1298}$  (absolute value) and showed the highest expression level, up to 60% of the total bacterial proteins.

Codon usage will impose much influence on expression level. So it is reasonable and understandable that the total chemically synthesized TNF- $\alpha$  gene with the codons preferable to *E. coli* usage gave much higher expression level than the sc-DNA in which only partial N-terminal codons were modified.

**Key words** Human tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), total chemically synthesized gene, high level expression, translation initiation region (TIR)