

# 芽孢杆菌蛋白质分泌工程

郭兴华 门大鹏 贾士芳

(中国科学院微生物所, 北京)

无论真核生物还是原核生物, 每个细胞都至少能合成上千种蛋白质(酶)。作为理想的生物工程菌, 应能把这些蛋白质(酶)的大部分分泌到细胞外。芽孢菌是重要的工业生产用菌, 遗传背景也较清楚。芽孢菌细胞表层比大肠菌简单, 蛋白质易于向胞外分泌。芽孢菌的信号序列和蛋白质的分泌有密切的关系, 用芽孢菌的强启动子和适宜的信号序列能构成分泌运载体。原来不能向胞外分泌的外源基因产物, 若将该基因插入到分泌载体信号肽序列之后, 其产物即可分泌到细胞外。

## 一、枯草杆菌和大肠杆菌细胞表层的差异

大肠杆菌为革兰氏阴性菌, 细胞表面有三层生物膜<sup>[1,14]</sup>: 细胞质膜、肽糖苷、外膜。与物质分泌有关的是外膜和细胞质膜, 细胞质膜和肽糖苷之间叫细胞周质, 大肠杆菌的蛋白质(酶)能通过细胞质膜, 但不能通过外膜, 所以蛋白质的绝大部分都停留在周质里。

枯草杆菌是革兰氏阳性菌, 细胞没有外膜, 蛋白质如能通过细胞质膜, 就有可能渗透到体外。因此, 从细胞表层的结构来看, 枯草杆菌系统比大肠杆菌简单。适合于作为研究蛋白质分泌的模式。

## 二、芽孢杆菌分泌蛋白质(酶)的类型

革兰氏阳性菌分泌的蛋白质(酶)分为两类<sup>[1]</sup>: 一类为脂蛋白, 如蜡状芽孢杆菌分泌的  $\beta$ -内酰胺酶 III, 金黄色葡萄球菌分泌的  $\beta$ -内酰胺酶, 它们都跟膜结合在一起, 本身就是膜的一个组成部分, 这为工艺设计提供了一个可能位点。

另一类是非脂蛋白, 如蜡状芽孢杆菌分泌的  $\beta$ -内酰胺酶 I, 解淀粉芽孢杆菌和枯草杆菌分泌的  $\alpha$ -淀粉酶, 在它们肽链的 N-末端, 有一个信号肽序列, 这个序

列的开始是由甲硫氨酸, 然后是赖氨酸和精氨酸等极性氨基酸组成的亲水序列, 接着是亮氨酸、缬氨酸和丙氨酸等非极性氨基酸组成的疏水序列<sup>[1]</sup>。信号肽酶在酶解位点  $\left(-\text{信号肽}-\left[\begin{smallmatrix} \text{ALA} \\ \text{GLY} \end{smallmatrix}\right]-\downarrow-\text{分泌蛋白}-\right)$  酶解掉信号肽, 分泌蛋白(酶)才能分泌到胞外。我们讨论的是非脂蛋白分泌问题, 即如何构建一个分泌载体系统。

## 三、构建分泌载体的意义

基因工程发展到今天, 所遇到的主要问题之一, 是外源基因不能高效表达和得不到大量产物。为此, 利用强启动子和 SD box 序列组建了表达运载体<sup>[5,6]</sup>; 利用强启动子、SD box 和信号序列组建了分泌载体<sup>[10,11]</sup>。组建分泌载体的目的是使完整的外源基因在比较理想的受体菌——芽孢杆菌中得以表达, 并使基因产物向胞外分泌, 从而可得到较多的基因产物的量, 还可了解某些基因表达调控机制, 同时也为分子生物学及分子遗传学研究转译产物提供手段。

## 四、芽孢杆菌信号肽序列和蛋白质(酶)分泌的关系

芽孢杆菌是主要的工业生产菌之一。芽孢杆菌能分泌几十种酶蛋白, 如  $\alpha$ -、 $\beta$ -淀粉酶, 中性、碱性蛋白酶, 酯酶和蔗糖酶等, 数种抗生素, 如多粘杆菌素, 枯草杆菌肽等, 也是生物防治——杀虫毒素的产生菌, 如苏芸金芽孢杆菌, 球状芽孢杆菌和侧孢芽孢杆菌等。

为什么芽孢杆菌能向胞外分泌蛋白呢? 1975 年 Blobel<sup>[13]</sup> 提出了信号肽的假说, 1980 年 Inouye 又提出了环状模型, 同年 Steiner 还提出了  $\beta$ -折叠转位, 1982 年 Blobel 进一步完善了他的假说。研究证明假说基本上是正确的, 即细胞内蛋白质的前体 N-末端含有疏水氨基酸残基的信号序列(信号肽), 细胞内游离

的核糖体先合成信号肽,信号肽由于疏水性很快和细胞质膜结合,然后与膜结合的多核糖体合成分泌蛋白。信号肽能识别并和膜上受体蛋白结合,接着进行穿膜转位,信号肽被信号肽酶水解掉,使胞内蛋白变成胞外蛋白。蛋白中的信号肽是分泌的关键。蛋白质的穿膜转位是一个重要的生物现象。假如没有信号肽和穿膜转位现象,蛋白质就只能在胞内,不能向外分泌。

## 五、分泌载体和基因工程的关系

芽孢杆菌作为基因工程受体菌,有利因素在于:对它的遗传学研究比较深入;遗传物质的传递方式也多样化;大部分的种不致病;它是重要的工业生产用菌;能分泌多种蛋白质。

从事基因工程研究的目的之一,是利用基因工程菌生产人们所需要的产物;目的之二,是利用基因工程的技术研究分子生物学和分子遗传学。芽孢杆菌分泌的各种蛋白质都有特异性的信号肽。克隆得到的非芽孢杆菌外源基因(如革兰氏阴性菌基因)的信号肽是不能向外分泌基因产物的。所以构建分泌载体使外源基因插入到芽孢杆菌中某个基因所具有的强启动子和信号序列的后面,通过信号肽使外源基因产物分泌到培养液中,并利用分泌载体研究分子生物学和分子遗传学问题是非常重要的。

## 六、分泌运载体的研究进展

目前国外都在积极从事这方面的工作,他们已利用芽孢杆菌中两个具有强启动子的 $\beta$ -内酰胺酶基因<sup>[7,8]</sup>和 $\alpha$ -淀粉酶基因<sup>[11]</sup>的信号序列构建成了分泌载体。这两种基因都能向胞外分泌大量产物。在所构建的分泌载体中,都有启动子序列,RNA聚合酶的识别位点,Pribnow box 序列转译时核糖体的结合位点——Shine-Dalgarno (SD) 序列、起始密码 ATG 和信号序列等。

1980年 Brammer<sup>[4]</sup>报道了用 $\lambda$ 噬菌体克隆了地衣状芽孢杆菌 749/C 菌株的 $\beta$ -内酰胺酶基因。1981年 Gray 和 Chang<sup>[7]</sup>用 pSC 101 从同一菌株中构建了 pTB<sub>1</sub> 质粒,并对该质粒中的 $\beta$ -内酰胺酶基因进行了 DNA 序列分析。他们还用 pUB110 的复制子、pC 194 的氯霉素抗性基因、大肠杆菌 COLE1 质粒的复制子和地衣状芽孢杆菌 $\beta$ -内酰胺酶基因构建了一个双功能质粒 pOG2165,观察在大肠杆菌和枯草芽孢杆菌中的表达。同年 Neugebauer<sup>[9]</sup>也做了类似的报道。

地衣状芽孢杆菌 749/C 菌株 $\beta$ -内酰胺酶的前体有 307 个氨基酸残基<sup>[7,9]</sup>,其中 34 个为信号肽,信号肽氨基酸 1—12 为亲水区,13—18 为疏水区,(真核生物和其它原核生物的疏水区比较短),还有一个 29—34 的肽链,它和膜连接。相似的结构,在 $\alpha$ -淀粉酶基因中也存在<sup>[11]</sup>,它们含有以非极性的氨基酸残基为中心的 29—41 肽的信号序列。在 $\beta$ -内酰胺酶和 $\alpha$ -淀粉酶中,一般推测有两种信号肽参与蛋白质的降解。

革兰氏阴性菌和革兰氏阳性菌中的 $\beta$ -内酰胺酶的信号肽有很大不同<sup>[11]</sup>。前者的信号序列由 23 个氨基酸残基所组成,7 个氨基酸残基组成亲水部分,16 个氨基酸残基组成疏水部分。后者的信号肽由 34 个氨基酸残基组成,12 个氨基酸残基组成亲水部分,22 个氨基酸残基组成疏水部分。这种结构上的变化,反映了革兰氏阳性菌和革兰氏阴性菌分泌机制的不同。这种长的信号肽可能是对于正在生长的 $\beta$ -内酰胺酶分子附着在膜上是需要的,如同假说所述,蛋白质通过膜向外分泌时要形成 $\beta$ -折叠。

Gray 和 Chang<sup>[7]</sup>把 pOG 2165 分泌载体转到枯草杆菌,当菌为  $2 \times 10^6$  菌落单位/ml 时,测出有  $10 \mu\text{g/ml}$  的 $\beta$ -内酰胺酶,有 60—80% 的酶分泌到培养液中,这个水平代表  $10^6$  分子/菌落单位。可见,枯草杆菌能以很高的效率进行转录、转译/转位和转译后加工。

掘越弘毅<sup>[11]</sup>用嗜硷性芽孢杆菌 $\beta$ -内酰胺酶基因构建分泌性质粒,使大肠杆菌 HB101 分泌基因产物。他的做法是这样:用内切酶 Hind III 把嗜硷性蜡状芽孢杆菌 $\beta$ -内酰胺酶基因切下来,接到 pMB9 质粒上,转化到大肠杆菌 HB101 里,得到杂种质粒 pEAP2,克隆片段为 2,400bp,含有启动子、信号序列和结构基因,又用其它限制性内切酶把无用的片段切掉,得到 pEAP3 质粒,片段为 1,200bp。含 pEAP3 的 HB101,培养 30 小时后,有 80% 的产物分泌到胞外,然而含 pBR322 质粒的 HB101 并不分泌或很少分泌 $\beta$ -内酰胺酶。pEAP3 之所以能使产物分泌,是因为它能使大肠杆菌外膜透性增大。有意义的是:不但 $\beta$ -内酰胺酶向外分泌,而且原来不能分泌的碱性磷酸酯酶也同时分泌出来。分泌的关键,可能是嗜硷性芽孢杆菌 $\beta$ -内酰胺酶基因在起作用。

Lovett 实验室用 pUB110 和短小芽孢杆菌染色体上抗氯霉素的基因组建了能克隆启动子的 pPL603 质粒,他们用这个质粒克隆了噬菌体 SPO2 的启动子,叫 pPL608。他们又把小鼠的二氢叶酸还原酶基因和大肠杆菌色氨酸基因插入到 pPL608 的 SPO2 启动子后,结果这两个基因的产物在枯草杆菌中都得到了表达<sup>[6,13]</sup>。

Palva 等用 pUB110 和解淀粉芽孢杆菌 $\alpha$ -淀粉酶基因构建了质粒 pKTH10,经过切割加工建成 pKTH29 (含有 $\alpha$ -淀粉酶的启动子、信号序列及部分结构基因),又用外核酸酶处理 pKTH29 上 $\alpha$ -淀粉酶的部分结构基因,希望只留信号肽序列,然后插入 $\lambda$  Hind III 的人工接头。转化得到大约 120 个转化子,通过 DNA 序列分析确定了其中 30 个 Hind III 人工接头的位置。图 1 是由长短不同的信号序列构建的分泌载体的一部分。

由 pBR322 和 pC194 融合而组建的 pHV33,被 ECOR1 切开后,再用外核酸酶 III 切去编码 $\beta$ -内酰

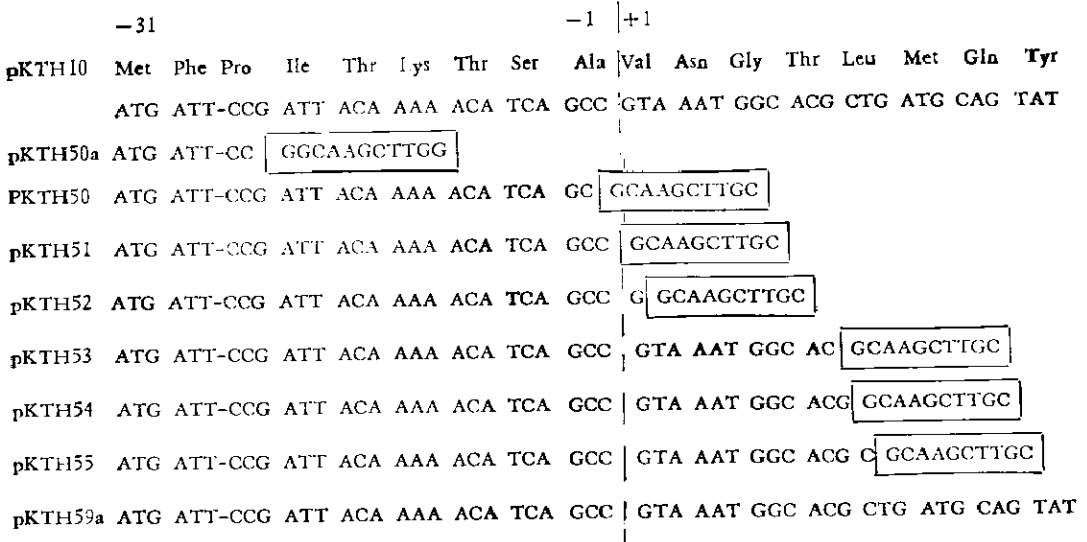


图1 由长短不同的信号序列构建的分泌载体<sup>[12]</sup>

表 1  $\beta$ -内酰胺酶在枯草杆菌不同分泌载体中的酶活性

分泌载体	连接位置 ( $\alpha$ -淀粉酶中的 氨基酸)	$\beta$ -内酰胺酶	
		分泌的	束缚于细胞膜的
pKTH50a	- 7	30	220
pKTH50	- 1	2030	80
pKTH53	+ 4	1850	75
pKTH59a	+ 14	450	35

胺酶(来自大肠杆菌)信号序列,进一步修饰后,在两端加上 Hind III 人工接头,把这种来自大肠杆菌只有结构基因而无信号序列的  $\beta$ -内酰胺酶基因连接到上述载体上,使原来不能在革兰氏阳性菌枯草杆菌中表达的  $\beta$ -内酰胺酶基因得以表达和分泌,如表 1 和图 2。在 pKTH50a、pKTH50、pKTH53 和 pKTH59a 等分泌载体中,对  $\alpha$ -淀粉酶/ $\beta$ -内酰胺基因连接处做了 DNA 序列分析。只要是含  $\alpha$ -淀粉酶基因全部信号序列的都能有效地向培养液中分泌  $\beta$ -内酰胺酶,然而缺乏信号序列中 19 个核苷酸(7 个氨基酸)的  $\beta$ -内酰胺酶的大部分是与细胞膜结合在一起的。由表 1 看出,质粒 pKTH50 和 53 有 95% 的酶分泌, pKTH50a 的酶有 90% 在膜上,酶活约为 10%。这些事实说明,信号肽后边的氨基酸序列的裂解位点对于枯草杆菌信号肽酶的裂解可能是严格的、特异的。另外, pKT50a 的酶大部分束缚于膜上,不能加工,说明淀粉酶信号序列羧基端 7 个氨基酸对加工是重要的。

把人  $\alpha$ -2 干扰素基因连接到 pKTH53 分泌运载体上,在培养液中得到具有生物活性的干扰素,最大活性为 200,000 IU/mL<sup>[12]</sup>。这说明其它真核生物基因如

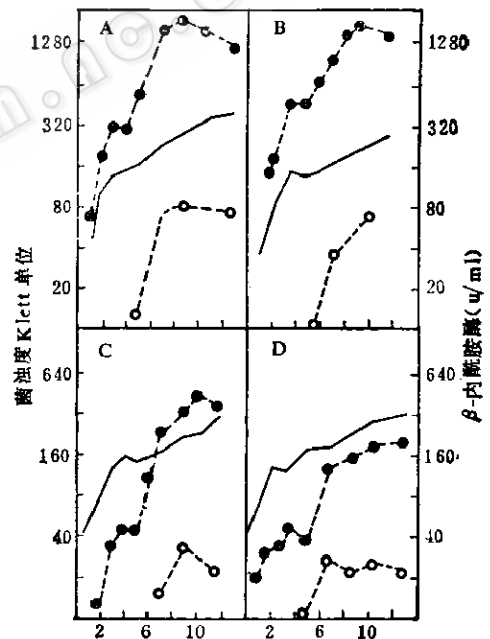


图2  $\beta$ -内酰胺酶基因插入不同分泌载体后测定的酶活曲线<sup>[13]</sup>

- A: pKTH50 (Hind III 人工接头在-1 的位置)  
B: pKTH53 (Hind III 在+4 的位置)  
C: pKTH59a (Hind III 在+14 的位置)  
D: pKTH50a (Hind III 在-7 的位置)
- .....● 上清液中  $\beta$ -内酰胺酶活性  
○.....○ 超声波处理后细胞中  $\beta$ -内酰胺酶活性  
—— 超声度

(下转第 225 页)

(上接第 217 页)

果接在信号肽后面,基因的产物也可能分泌到胞外。

有关分泌工程设计的范围很广,从微生物角度考虑,有酵母、丝状真菌、放线菌和细菌之分,而细菌中又有革兰氏阳性和革兰氏阴性之分。从工程设计的内容来讲,有的用诱变等方法来研究膜的渗透类型;有的试图把功能性的酶或功能性的片段放在细胞表面发展成为一个生化催化床;有的通过已经了解的膜渗透机制来设计产物的分泌;有的通过往培养基中加青霉素和/或表面活性剂促使菌体内产物的分泌;更多的是如本文所述的用蛋白质的信号肽和信号肽酶来研究分泌。限于篇幅,本文着重谈的是芽孢杆菌中构建分泌载体。

微生物蛋白质分泌工程的设计是基因工程或者说生物工程的一个重要组成部分。是当前急待解决的问题。对它的研究,不仅有利于微生物工业发酵的生产;而且也有利于高等动植物细胞分泌机制的探讨,所以这项工作的本身是实际性较强,理论性较深的课题。相信不久的将来会有更大的突破。

### 参 考 文 献

[1] 掘越弘毅:现代化学,154:26,1984。

- [2] 吴明:生物科学动态,2:30,1984。
- [3] Blobel, G. and B. Dobberstein: *J. Cell Biol.*, **67**: 852, 1975.
- [4] Brammar, W. J. et al.: *Mol. Gen. Genet.*, **178**: 217, 1980.
- [5] David, S. et al.: *Molecular cloning and Gene Regulation in Bacilli* edited by A. T. Ganesan et al., pp. 311, 1982.
- [6] Donna, M. et al.: *Ibid*, pp. 91, 1982.
- [7] Gray, O. and S. Chang: *J. Bacteriol.*, **145**(1): 422, 1981.
- [8] Kroyer, J. and S. Chang: *Gene*, **15**: 343, 1981.
- [9] Neugebauer, K. et al.: *Nucleic Acids Res.* **9**(11): 2577, 1981.
- [10] Palva, I.: *Gene*, **19**: 81, 1982.
- [11] Palva, I.: *Proc. Natl. Acad. Sci. USA.* **79**: 5582, 1982.
- [12] Palva, I. et al.: *Proceedings of the IVth. International Symposium on Genetics of Industrial Microorganisms* pp. 287, 1982.
- [13] Ronald, G. et al.: *Gene*, **22**: 47, 1983.
- [14] Stryer, L.: *Biochemistry* pp. 744, 1981.
- [15] Takeich, Y. et al.: *Agric. Biol. Chem.*, **47**(1): 159, 1983.