



大肠杆菌色氨酸操纵子的结构与克隆

唐任天

(中国科学院微生物研究所,北京)

色氨酸(Trp)具有重要生理活性,是哺乳动物的必需氨基酸之一。Yanofsky 小组对大肠杆菌 Trp 合成酶的研究,首次确证了核苷酸三联密码与多肽链氨基酸的对应关系。Monod 等发现大肠杆菌 Trp 合成酶的合成被培养基中的 Trp 所抑制,而且过量 Trp 对邻接的五个 Trp 基因都起着调节作用,由此提出了 Trp 基因组成一个操纵子的著名设想。以上两项工作在分子生物学领域具有十分重要的意义,促使人们对 Trp 操纵子及其基因调控机制进行了更为深入的研究,并试图以基因工程手段获得工业化生产的 Trp 产生菌。本文尝试对 Trp 操纵子的结构、基因调控及基因分离的主要研究成果作一简要介绍。

一、 Trp 操纵子的组织结构^[1-4]

Trp 操纵子包括调节区、结构基因与终止区三部份。五个结构基因与它们编码的酶及代谢步骤的关系见图 1。 trpE 和 trpD 分别编码邻氨基苯甲酸合成酶(ASase)成份 1 和 2。两成份结合后方具有 ASase 活性,而成份 2 单独具有磷酸核糖转移酶(PRTase)活性。

性。 trpC 编码吲哚甘油磷酸合成酶(IGPSase),催化反应 3、4。 trpB 和 trpA 分别编码色氨酸合成酶(TSase)的 B、A 多肽链,两肽链结合才具有正常 TSase 活性,催化反应 5。

五个相邻的结构基因 EDCBA 共长 6.7kb。从另一方面看,结构基因编码的多肽分子量依次为 60、65、45、45、29($\times 10^3$),按每个氨基酸残基分子量 110 计算,则相应为 1.64、1.77、1.23、1.23、0.8(kb),共约 6.67kb。通过 DNA-RNA 杂交,测定出 mRNA^{Trp} 之沉降速度为 33 S,这相当于 6.7 kb,也与上述数据一致。而操纵子本身在 25 分钟处。

大肠杆菌 Trp 终止子 trpC 座落在 trpA 之后仅 36 个核苷酸处。在其后约 250 bp 处另有一个终止子 trpT 。这样,大肠杆菌 Trp 操纵子全长约 7.1kb。

二、 Trp 操纵子的基因调控^[1,3,5-9]

Trp 操纵子的基因调控是负反馈调节的典型例子。在 Trp 操纵子首端有启动基因 P(在结构基因 trpD 处也可能存在内启动基因)。RNA 多聚酶在 P

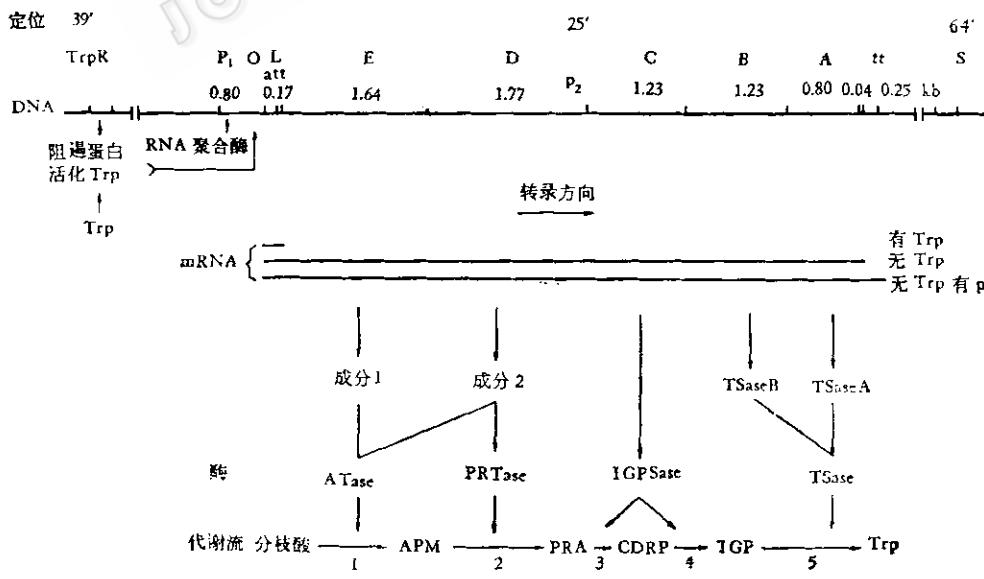


图 1 色氨酸操纵子及其编码的酶

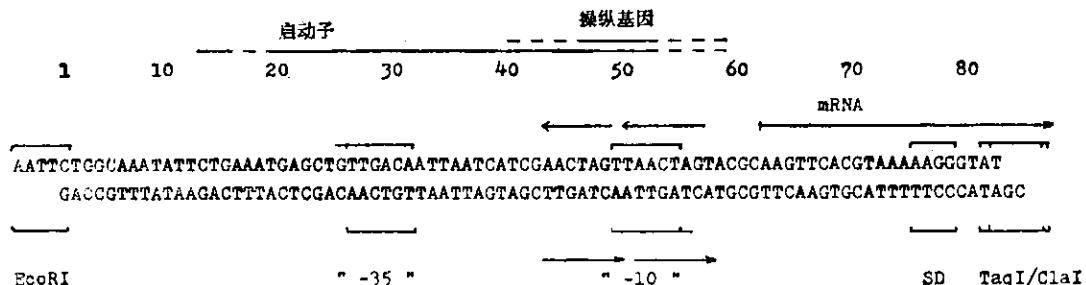


图 2 Trp P-O 区的结构

处附着便能启动结构基因的依次转录。假使 P 发生负突变，便无法启动转录，所以转录速度首先取决于 P。在 P 前有调节基因 R，它产生无活性的阻遏蛋白。阻遏蛋白是一种变构蛋白，与辅阻遏物 Trp 结合之后，结构改变，便能使复合物与位于 P 后的操纵基因 O 上的特殊核苷酸序列相结合（图 2）。即活性阻遏蛋白具有识别 O 的能力，从而关闭整套结构基因。不过若核糖体事先结合时则仍能往下转录，说明活性阻遏蛋白只是关闭 O 区，使结合在 P 上的核糖体无法通过。

若调节基因 R 发生 trpR 突变，不能产生阻遏蛋白，或者不能被活化，则结构基因一直打开。trpR⁻ 对 trpR⁺ 是隐性，在这种突变株中 Trp 合成酶系的合成是组成型的。若发生 trpR^S 突变，将产生即使没有 Trp 也具有活性的阻遏蛋白，O 区便一直关闭，trpR^S 对 trpR⁺ 为显性。另一类突变 trpO^c，使 O 区丧失与活性阻遏蛋白结合的能力，即 O 区一直开启，Trp 合成酶系的形成也是组成型的。

实际上，Trp 本身并不能与阻遏蛋白结合。它须先经色氨酰-tRNA 合成酶活化，后者由 trpS 基因编码。若编码该酶的基因突变，过量终产物 Trp 便不能成为辅阻遏物，也就不能起到抑制 Trp 继续合成的作用。

以上是启动转录的负反馈调节的主要机制。除此之外还需要在终止转录上加以调节，才能得到符合要求的产物。结构基因前往往有衰减区，其功能是减弱启动讯号，以对转录作更精细的调节（图 3）。在一定条件下 RNA 聚合酶到此中止转录，而形成 132bp 外加 8bpU，再接上 22bp 的 mRNA。在 Trp 存在下指令一条 14 个氨基酸多肽链的合成，而无 Trp 时则合成 12 肽。

大肠杆菌 Trp 操纵子体内转录高频率地终止在称为 trp^t 的位点上。许多原核细胞终止子的初级结构显示出有同一性，转录以几个尿嘧啶核苷酸残基终结，在终止位点前有一个二重对称的富 GC 区。然而 trp^t 只有 25% 的终止效力。当不存在 P 因子时，转录部份在 trp^t 处停止，但大部份通读到末端。存在 P 蛋

白时则大部份越过 trp^t 而终止在 trp^{t'} 处。体外转录表明，这两个终止子无论是串连结构，或者分别克隆都能独立运转。

三、Trp 操纵子的基因分离与重组^[10]

在噬菌体 φ80，及 λ 转导的大肠杆菌之染色体中，虽然 Trp 操纵子两侧的序列不尽相同，但因都没有限制性内切酶 EcoRI、BamHI 及 SmaI 的作用位点，而不能将完整的 Trp 操纵子切下。可是噬菌体 trp EA51 的 Trp 操纵子紧靠左右两侧的碱基序列却可被 SmaI 裂解。为此先用 EcoRI 消化得到 16 kb 片段，再用 SmaI 进一步得到 8.7 kb 片段，然后与 pBR313 重组。选择 trp^t Ap^r Tc^r，得到 pHp3。

利用 HpaI 位点正好在 Trp 操纵子启动基因中的特点，可以判断 Trp 操纵子在杂合质粒上的方向。pHP3 的 8.7 kb 片段有 4 个 HpaI 切点。在一定条件下 RNA 多聚酶能选择性地结合到启动子上，而使启动基因免受 HpaI 的裂解。对有无 RNA 多聚酶存在时的 HpaI 酶切图谱进行比较，发现 0.23 kb 片段被 6.1 kb 片段代替，后者是 0.23 kb 与 5.9 kb 两片段连接而成的。因此 Trp 启动基因正好位于 HpaI 切点上，而这一位点处于 0.23 和 5.9 kb 片段之间，并且转录为顺时针方向。

四、结构基因的分离^[10,11]

将 pHp3 用 SmaI 消化，连上 EcoRI 接头 CCGA-AT TCGG，与 pBR345 一起用 EcoRI 消化。pBR345 很小，但有复制区且不带抗药性标记。结果分离到携带 pHp12 及 pHp39 的 trp^t 转化菌。改造 pHp12 得到 pEP121 (12.5 kb)，带 Ap^r，有完整的 trpA 和 E，并具有 7 个内切酶单切点。由于该质粒缺失 trpC，trpB、D 也不完整，因此可用于克隆外源 trp B、C、D。

pHP39 为 9.8 kb。经改造得到 pHp3921 与 pHp3923。pHP3921 为 2.8 kb，有完整的 trpE、A，有 6 个单裂解点。pHP3923 为 7.4 kb，有 5 个单切点，有完整的 trpA、B，可用于克隆外源 trpC、D 或 E 基因。

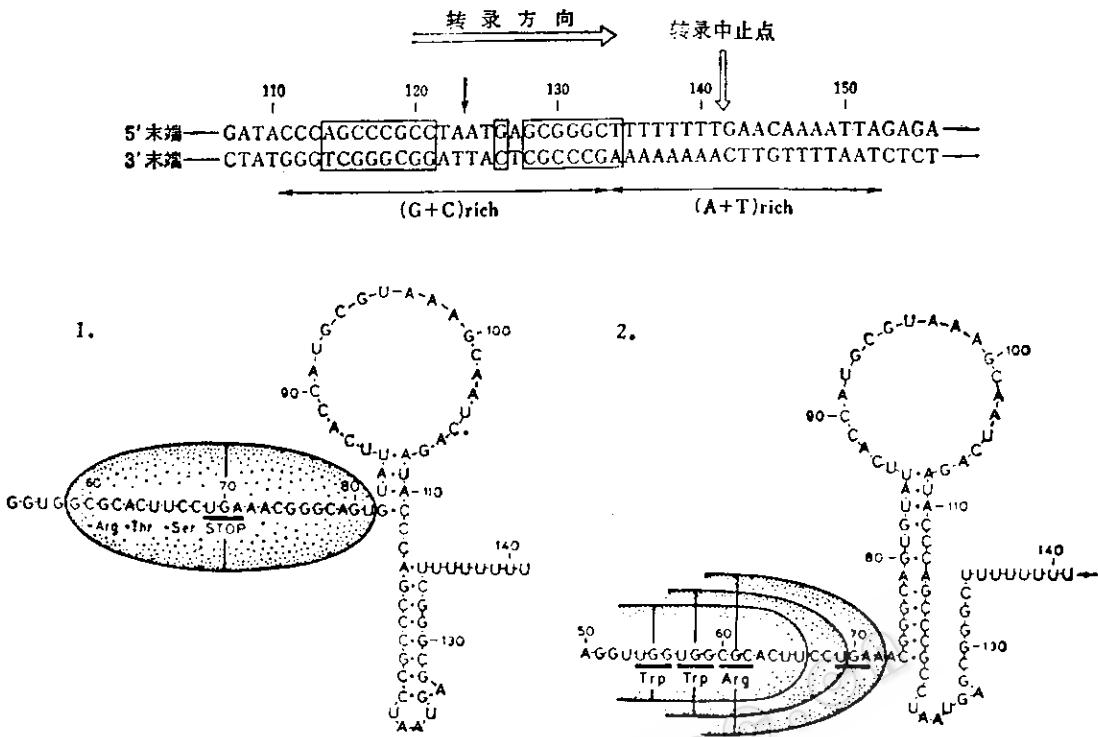


图3 大肠杆菌衰减区的结构与作用机制

上. 衰减区的DNA初级结构

图中→示出迴文结构的对称轴

下. 衰减区mRNA二级结构的变化与核糖体位置的关系

1. 低浓度Trp存在下,衰减区的富GC前后序列之间形成氢键

2. 无Trp、精氨酸时,衰减区的富GC段迴文结构间的氢键解开,核糖体向前移动进行翻译

五、Trp启动基因的克隆^[7,11,12]

pHP3被HindIII消化改造成pEP3012、pEP3015。这些质粒在trpA之前缺失启动子，而带一个EcoRI切点，可用trpA作为选择标记来克隆启动子。例如将pEP3012的EcoRI-HindIII片段用来自pHP39的3.2kb代替，便得到pEP292。测定表明pEP3012的trpA已处于pEP39的trp启动子的控制之下。又如移去pEP3015的EcoRI-HpaI片段，插入pHP3的EcoRI-SalI0.65kb片段，所得到的pEP320中trpA能够表达，表明插入片段含有启动子。

pBR322的HindIII位点居于Tc^r基因启动子内，因此插有外源片段时一般会使Tc基因失活，而变为Tc^s，而如果插有外源片段后Tc基因依然能够表达，说明已经克隆了一个外源启动子。取大肠杆菌5.4kb HindIII段，插到pBR322的HindIII切口中。得到的ptrpED₁，以EcoRI消化后再用外切酶III消化，并以S₁核酸酶形成钝末端，再用T₄DNA连接酶连接。转化W3110 trp O^EΔ1菌，从中分离到ptrpED₁-1质粒(6.7kb, Ap^r Tc^r)。该转化菌在过量Trp存在时即已

能耐23μg/ml的Tc，无Trp时更增加2.5倍，为56μg/ml。由于pBR322的Tc启动基因已遭破坏，说明现在已插入了一个外源trp启动子，并且有效地控制着Tc基因的表达，而这种控制受到Trp的调节。

六、色氨酸操纵基因^[4,11]

由于trpO与trpP相互重叠，因此Russell等将trpP-O区域一道分离出来进行研究。首先从大肠杆菌W3110ΔLC145中分离出带有trpP-O区的噬菌体λtrpΔLC145，取其EcoRI片段克隆到pMK16中而构成pGB62。从中取HpaII段经AluI消化，得到41bp。将其插入到pBR322的AluI位点，构成了pDR系列质粒。

pDR11含有1份方向正确的41bp段，转化菌呈现trpO^r表型，方向相反的pDR12则无trpO活性。判断片段方向借助于BspRI-HpaI双酶消化法。PvuII将BspRI片段裂解为118和422bp两块碎片，HpaI将41bp分为13和28bp两段。因此若将PvuII位点上插有41bp的BspRI片段以HpaI消化，插入方向正确者将产生435和146bp两个碎片。

Pvu II 裂解能把 41bp 从质粒中释放出来。将 41bp 片段以 T_4 DNA 连接酶处理，然后插到 pBR 322 的 **Pvu II** 位点中，可以得到多重 41 bp 的衍生质粒。从 β -内酰胺酶活性看，这些质粒都有类似的拷贝数，并不因多重 **trp P-O** 片段的插入而受影响。

具有 **trpO** 活性的 41bp 片段，除了在 -4 和 -18 处有差异外，从 -2 到 -21 的碱基呈现出完美的对称。若从 **Hpa I** 处切开，将左右两半各自连接，并插到 pBR 322 的 **Hpa I** 处。得到的 pDR 36 (**O 右 O 左**) 有与 pDR 11 或 pDR 12 相类似的 β -半乳糖苷酶活性，但 pDR63 (**O 左 O 右**) 却测不到活性，而且转化菌长得又慢又小。为进一步剖析 41 bp 片段中哪一部份比较重要，将其用 **Rsa I** 裂解，并插到 pBR322 的 **Sma I** 处。结果这个新的序列没有 **trpO** 活性，从而表明从 -6 到 -3 附近的序列对于 **trpO** 是重要的。一般说来，**trpO** 至少需要 -20 到 -3 这一段序列，才能被 **Trp**-阻遏物识别。

七、终止子的克隆^[1,2]

转录的终止，其实质是 RNA 聚合酶对 DNA 模板上译成密码的讯号的识别，因此基因表达也在终止区受到调节。

pEP165 的 **Tc** 结构基因处于 **trpP** 的控制之下，在 **Tc** 结构基因之前有个 **HindIII** 位点，可用于克隆终止讯号。假如把 pEP162 的 2.9 kb 之 **HindIII-HindIII** 片段插到 pEP165 的 **HindIII** 位点中，转化菌便变得对 **Tc** 敏感，说明 **Tc** 结构基因的表达被阻断；也就是说，一个 **trp t** 已被克隆到 **Tc** 结构基因之前。

然而 **trp t** 只有 25% 的终止效力，证明这一点的是下面的实验。从 pLD 102 分离出 **HinfI-HaeIII** 的 380bp，与 pBR322 的 **EcoRI** 片段重组成 pTP119。从 **trpA** 后 387bp 处切取包含 500bp 的片段插到 pTP119 的 **EcoRI** 位点中产生 pWU5。在缺乏 **P** 蛋白时，部份转录在 **trp t** 处中止，产生 245bp mRNA；其余通读到末端，产生 545bp。当存在 **P** 因子时 **trp t** 仍有 25% 终止效力，但其余并不通读到末端，由此发现了依赖 **P** 的第二个终止子 **trp t'**。

把 pWU5 的 **HincII** 的 200 bp 片段插到 pTP 119 中，便产生只有 **trp t'** 的 pWU15。同样，将 λ RT38 的 **HaeIII** 的 315bp 片段插到 pTP119 的 **EcoRI** 位点，就产生只有 **trp t'** 的 pWU11。体外转录表明，**trp t** 或 **trp t'** 均有终止功能，证明这两个终止子在体外无论是串连结构，或者分别克隆都独自发挥作用。通过 DNA 序列分析，以及对转录生成的 mRNA 的消化图谱的分析，将 **trp t'** 定位于 **trp t** 后大约 250bp 处。

八、**Trp** 基因的表达^[4,11-18]

已经报导了许多氨基酸合成基因之分离、克隆，并以 DNA-RNA 杂交法证明了重组基因已经表达。然而检验表达与否，最根本的是看有无相应产物形成。

获得一个强有力的启动子，往往是研究者所盼望的。除了将 **trpP** 与 **Tc** 或 **Lac** 融合的报导外，Windass 等人工化学合成了含有大肠杆菌 **trpP-O** 序列，以及编码 **trp** 操纵子的第一个 SD 序列的 82 bp DNA 片段。该片段两端为 **EcoRI** 和 **Clal/TaqI** 粘性末端。将其克隆到源于 pBR322 的 pAT153 中，改造成 pSTP 2。然后与载有合成的干扰素 $\alpha-1$ 基因的 pJDW22 重组成 pJDW30。去掉小片段后借助合成的寡核苷酸使 **trpP** 与 **IFN** 融合。这样就在 **IFN- $\alpha 1$** 启动子位点上产生一个完全天然的、高度有效的细菌翻译起始讯号。携带这些重组质粒 pIFS 201—204 的 *E. coli* 转化菌几乎组型地表达 **IFN- $\alpha 1$** 活性，而且其效率明显地比以 **Lac UV5** 启动子为基础的系统高。

从结构基因的表达看，当 λ 噬菌体 PL 插入 **EcoRI**、**BamHI** 和 **Sall** 位点时，**trpA** 基因产物分别达到细胞总蛋白量的 2.0、6.6 及 1.2%。以携带整个 **trp** 操纵子的 **ColE1** 质粒，或以 λ **trp EDC** 噬菌体转染的细胞，其 **trp E** 产物则占 12%。前述以 pBR322 为载体，将含有 **trp P** 及 **trp E** 的 *E. coli* 之 **HindIII** 片段克隆进去，得到 p**trp E-D**。₁₀ 含该质粒的转化菌在加有 **Trp** 的培养基内时，**trp E** 基因产物很少。但加入 3 β -吲哚丙烯酸时能解除阻遏。其后 1—2 小时 **trp E** 蛋白合成达到最高峰，至少增加 50 倍，占总蛋白的 30%，是新合成蛋白的 55%。

相应多肽的产生并不意味着一定能够过量积累相应的氨基酸，但在这方面亦进行了探索。作者构建了一些质粒，其中有些似乎含有 **trpAB** 基因。在适当条件下，转化菌能转化吲哚生成 **Trp**，其转化率几达 100%（未发表资料）。

在从邻氨基苯甲酸生成 **Trp** 方面，有人利用 pUM 101(1—2 个拷贝/染色体) 及抗 **Trp**、抗反馈控制的受体菌；而第一个借助于基因工程技术实现 **Trp** 小型发酵生产的是日本 Shuichi Aiba 等。他们发现 TSase 活性的增加与质粒拷贝数的变化相一致。但拷贝数太高时将招致不稳定，尽管此时 TSase 增加会使生长速率降低。所以他们选择了质粒 pSC101 **trp-115** (**Tc^rtrp^r**)，拷贝数为 5—10 个/染色体。培养基成份 (%) 为：葡萄糖 5，酪蛋白水解物 1, NH_4Cl 0.3, KH_2PO_4 0.3, K_2HPO_4 0.7, 邻氨基苯甲酸 0.05, $MgSO_4 \cdot 7H_2O$ 0.02, $FeSO_4 \cdot 7H_2O$ 0.001, 四环素 0.001。自动流加氨水维持 pH 7.0。在邻氨基苯甲酸降到 0.03% 时，以 5mg/dl/h 速度补加至原值。通气量 1vvm，搅拌 500rpm, 37℃。

菌体生长分为两阶段，早期消耗氨基氮，后期以氨基氮源。比生长率较大，在培养 3—6 小时为 0.56/h，最大细胞浓度为干细胞 6.9 g/l。直到葡萄糖耗尽之前，**Trp** 不断产生，邻氨基苯甲酸不断消耗，而一旦葡萄糖耗尽，**Trp** 亦停止产生。培养 24 小时，**Trp** 积累

量达 5.5g/l, 生产速率为 0.229g/l·h。这一生产速度在 Trp 发酵生产的所有报道中是最高的。

从上面的简述可以看到, 利用分子生物学和基因工程技术已经对 Trp 操纵子的结构、调控、转录及表达等问题作了比较深入的研究, 借助强有力的 trpP 使真核干扰素表达, 以及克隆出 trp 操纵子产生 Trp, 是其中比较突出的成果。在克隆外源基因时, 多利用抗生素抗性作选择标记, 而利用氨基酸作标记的质粒不会产生安全问题, 则在工业生产意义上具有其明显的优点, 是值得借鉴的。不过, 目前的研究涉及调节机制的尚不多。按笔者的体验, 仅仅克隆天然基因是难以获得过量产物的。因此克隆 trpR、S 并加以改造, 应是下一步研究的重点。如此, 有可能获得积累纯氨基酸的生产菌株。

参 考 文 献

- [1] Stent, G. S.: *Molecular Genetics*, by Freeman, W. H. and company, San Francisco, 1971.
- [2] Smith-Keary, P. F.: *Genetic Structure and Function*, The Macmillan press LTD, 1975, p 234.
- [3] Umbarger, J. E.: *Ann. Rev. Biochem.*, **47**: 574—99, 1978.
- [4] Windass, J. D. et al.: *Nucleic Acids Research*, **10** (21): 6639—6657, 1982.
- [5] Platt, T.: *Cold Spring Harbor, NY*, 1978, p449.
- [6] Craaford, I. et al.: *Annu. Rev. Biochem.*, **49**: 163—95, 1980.
- [7] Oppenheim, D. S. et al.: *J. Mol. Biol.* **144**: 133—42, 1980.
- [8] Gunsalus, R. P. et al.: *Proc. Natl. Acad. Sci. USA*, **77**: 7117—7121, 1980.
- [9] 今本文男: 蛋白质核酸酶素, **27** (9): 1215—1226, 1982.
- [10] Enger-Valk, B. E. et al.: *Gene*, **15**: 297—305, 1981.
- [11] Russell, D. R.: *Gene*, **17**: 9—18, 1982.
- [12] Hallewell, R. A.: *Gene*, **9**: 27—47, 1980.
- [13] Wu, A. M. et al.: *Proc. Natl. Acad. Sci. USA*, **78** (5): 2913—2917, 1981.
- [14] Setlow, J. K. and Alexandra Hollaender: *Genetic Engineering, Principles and Methods*. New York and London, Vol 2, p145—155, 1980.
- [15] Walz, A. et al.: *Proc. Natl. Acad. Sci. USA*, **75** (12): 6172—6176, 1978.
- [16] Tribe, D. E. et al.: *Appl. Environ. Microbiol.* **38**: 181—190, 1977.
- [17] Shiba Aiba et al.: *Appl. Environ. Microbiol.* **43** (2): 289—297, 1982.
- [18] 佐野考之輔: *発酵と工業*, **41**(4): 275—282, 1983.