[19]中华人民共和国国家知识产权局

[51]Int. Cl[7]
C12N 1/00
C12N 5/10 C12N 15/11
C12N 15/12 C12N 15/16
C12N 15/18 C12N 15/29
C12N 15/31 C12N 15/63
C12P 21/06

# [12]发明专利申请公开说明书

[21]申请号 98803491.3

[43]公开日 2000年7月19日

[11]公开号 CN 1260831A

[22]申请日 1998.3.21 [21]申请号 98803491.3

[30]优先权

[32]1997.3.21 [33]US [31]08/821,559

[86]国际申请 PCT/US98/05655 1998.3.21

[87]国际公布 WO98/42822 英 1998.10.1

[85]进入国家阶段日期 1999.9.20

[71]申请人 柏内布拉斯加公司

地址 美国内布拉斯加

[72]发明人 夏远南

[74]专利代理机构 中国国际贸易促进委员会专利商标事务所

代理人 郭建新

权利要求书3页 说明书34页 附图页数9页

[54]发明名称 小球藻病毒启动子

[57]摘要

本发明提供了得自小球藻病毒的新的启动子序列。本发明包括含有本发明启动子序列的基因构建体，所述启动子序列与编码结构基因的DNA序列可操作地相连。本发明还提供了表达由本发明基因构建体中的结构基因编码的产物的载体和宿主细胞以及被与启动子可操作相连的异源基因转化的细胞。

ISSN1008-4274

# 权　利　要　求　书

1. 根据 SEQ ID No. 1 的分离的核酸序列。

2. 根据 SEQ ID No. 2 的分离的核酸序列。

3. 根据 SEQ ID No. 3 的分离的核酸序列。

4. 根据 SEQ ID No. 4 的分离的核酸序列。

5. 根据 SEQ ID No. 5 的分离的核酸序列。

6. 根据 SEQ ID No. 6 的分离的核酸序列。

7. 根据 SEQ ID No. 7 的分离的核酸序列。

8. 含有 SEQ ID Nos. 1, 2, 3, 4, 5, 6 或 7 的核酸序列的启动子。

9. 含有与结构基因可操作相连的启动子的基因构建体，所述启动子含有 SEQ ID Nos. 1, 2, 3, 4, 5, 6 或 7 的核酸序列。

10. 权利要求 9 的基因构建体，其中所述结构基因编码碳酸酐酶。

11. 权利要求 9 的基因构建体，其中所述碳酸酐酶是人碳酸酐酶。

12. 权利要求 9 的基因构建体，其中所述结构基因编码胰高血糖素样肽-1(GLP-1)，生长激素释放因子(GRF)或甲状旁腺激素(PTH)。

13. 含有基因构建体的宿主细胞，所述基因构建体含有与结构基因可操作相连的启动子，所述启动子含有 SEQ ID Nos. 1, 2, 3, 4, 5, 6 或 7 的核苷酸序列。

14. 权利要求 13 的宿主细胞，其中所述宿主细胞是原核细胞。

15. 权利要求 14 的宿主细胞，其中所述原核细胞是大肠杆菌。

16. 权利要求 14 的宿主细胞，其中所述大肠杆菌包括 HB101, JM109, BL21 或 HMS174。

17. 生产蛋白质组分的方法，所述方法包括:

用核酸序列转化宿主细胞，所述基因构建体含有与异源 DNA 序列可操作相连的启动子，所述启动子含有 SEQ ID Nos. 1, 2, 3, 4, 5, 6 或 7 的核苷酸序列; 和

在宿主细胞中表达异源 DNA 序列以产生蛋白质组分。

18. 生产蛋白质组分的方法，所述方法包括:

在宿主细胞中表达核酸序列，其中所述表达受与异源核酸序列可操作相连的启动子驱动，所述启动子含有 SEQ ID Nos. 1，2，3，4，5，6 或 7 的核苷酸序列。

19. 核酸序列，它含有:

含有 SEQ ID No. 1，2，3，4，5，6 或 7 所示核酸序列的启动子;

*lac* 操纵基因;

和结构基因;

其中启动子，操纵基因和结构基因可操作地相连。

20. 权利要求 1 的核酸序列，其中启动子含有 SEQ ID No. 15 所示的核酸序列。

21. 含有 SEQ ID No. 1，2，3，4，5，6 或 7 所示核酸序列的诱导型启动子。

22. 含有与 *lac* 操纵基因可操作相连的 SEQ ID No. 1，2，3，4，5，6 或 7 所示核酸序列的诱导型启动子。

23. 权利要求 22 的诱导型启动子，它含有 SEQ ID No. 15 所示的核酸序列和 *lac* 操纵基因。

24. 含有权利要求 22 的诱导型启动子的宿主细胞。

25. 权利要求 24 的宿主细胞，其中宿主细胞是植物细胞。

26. 权利要求 24 的宿主细胞，其中宿主细胞是原核细胞。

27. 权利要求 24 的宿主细胞，其中宿主细胞是大肠杆菌细胞。

28. 核酸序列，它含有:

小球藻启动子;

*lac* 操纵子; 和

结构基因，其中启动子，操纵子和结构基因可操作地相连。

29. 权利要求 28 的核酸序列，其中结构基因含有编码胰高血糖素样肽 1，生长激素释放因子，甲状旁腺激素，碳酸酐酶，β-半乳糖苷酶，氯霉素乙酰转移酶或谷胱甘肽乙酰转移酶的核酸序列。

30. 在其基因组中携有权利要求 28 的核酸序列的植物。

31. 在其基因组中携有权利要求 22 的诱导型启动子的植物。

32. 在宿主细胞中产生蛋白质的方法，所述方法包括下列步骤:

用含有小球藻启动子、*lac* 操纵基因和编码蛋白质的结构基因的核酸序列转化宿主细胞；和

在足以使结构基因表达为蛋白质的条件下保温经转化的细胞。

33. 权利要求 32 的方法，其中所述经转化的细胞在 21 至 37℃的温度范围内保温。

34. 权利要求 33 的方法，其中所述经转化的细胞在低于 37℃的温度下保温。

35. 权利要求 32 的方法，其中所述经转化的细胞在 21 至 30℃的温度范围内保温。

36. 权利要求 28 的启动子，该启动子在 21 至 37℃的温度范围内起作用。

37. 权利要求 36 的启动子，该启动子在低于 37℃的温度下起作用。

# 说　　明　　书

## 小球藻病毒启动子

本发明涉及分离自小球藻(Chlorella)病毒的新启动子，所述新启动子可用于在宿主细胞中表达异源基因。

基因工程可从一个生物体中分离结构基因并在不同的生物体中表达该基因。基因的表达包括核酸转录为 mRNA 和 mRNA 翻译成蛋白质。为了在新的生物体中表达结构基因，必须将基因与适当位置处的传递基因转录信号的调节序列连接。调节序列一般包括结构基因上游的启动子序列。启动子序列是指导结构基因转录的 DNA 序列。结构基因的核酸序列被转录成信使 RNA(mRNA)，然后被翻译成特定多肽或蛋白质的特征性氨基酸序列。启动子序列一般位于基因的 5'区域，结构基因转录起始位点的上游。

启动子可以是诱导型的，也可以是组成型的。诱导型启动子对诱导剂的反应是活性增加，从而使可操作相连的编码序列的转录速率增加。与之形成对照的是，在组成型启动子的控制下，基因的转录速率不受调节。然而，值得注意的是通过加入操纵基因序列可使组成型启动子变为诱导型启动子。例如，在 T7 噬菌体启动子中加入 *lac* 操纵基因，使其由组成型启动子转变为能被 IPTG 诱导的启动子(Rosenberg 等，美国专利 4,952,496)。

尽管不受诱导剂的控制，一些组成型启动子可提供比其它启动子水平更高的转录。提供高水平基因转录的高活性启动子在基因产物的商业生产中具有显著的优点。

通常，启动子在其天然宿主体外指导转录的能力各不相同。另外，特定启动子的转录速率也可随启动子所作用的特定宿主的不同而变化。因此，需要能在广范围宿主细胞中促进高水平转录的新启动子。

小球藻病毒是一组感染某些单细胞的，真核的，小球藻样绿藻系的病毒(Van Etten, 1995, Mol. Cells, 5:99-106; Van Etten 等，

1991，微生物学评论，55:586-620)。这些病毒是已知的最大的和最复杂的病毒，一般为直径 150-190nm 的多角体，其中含有大于 300kb 的双链 DNA。小球藻病毒基因组具有编码几百个基因产物的潜力。

已分离出小球藻病毒甲基转移酶启动子，该启动子显示出在原核和真核宿主细胞系统中都可起作用。这些甲基转移酶启动子在一些细菌和高等植物细胞中能很好地起作用。例见美国专利 5,563,328 和 Mitra 等，1994，植物分子生物学，26:85-893("Mitra")。

本发明提供了分离自小球藻病毒的新的启动子序列，该启动子可在宽温度范围，如约 21℃至 37℃，在原核或真核细胞中诱导高水平的基因表达。

本发明提供了分离自小球藻病毒的新的启动子序列(SEQ ID NOS:1-7)。本发明包括含有本发明启动子序列的基因构建体，所述启动子序列与结构基因的 DNA 序列可操作相连。本发明还提供了表达由本发明基因构建体的结构基因编码的产物的载体和宿主细胞，以及被与启动子可操作相连的异源基因转化的细胞。

本发明的启动子包括诱导型小球藻启动子，例如通过与操纵基因序列融合可使该启动子变为诱导型启动子。在一个优选的实施方案中，本发明的小球藻启动子与 *lac* 操纵基因融合。本发明的小球藻启动子在宽温度范围内，如约 21℃至约 37℃，显示出很强的启动子活性。

在一个实施方案中，结构基因是编码用于在宿主细胞中产生的蛋白质的非小球藻病毒 DNA 序列。根据此实施方案，非小球藻病毒 DNA 序列可以是编码肽，蛋白质，激素，酶等的任何适当的结构基因。适当结构基因的例子包括胰高血糖素样肽 1(GLP-1)，生长激素释放因子(GRF)，甲状旁腺激素(PTH)，互连肽，酰胺化密码子序列，碳酸酐酶，β-半乳糖苷酶，氯霉素乙酰转移酶(CAT)，谷胱甘肽乙酰转移酶等。

本发明的基因构建体被导入适当的用于表达基因产物的宿主细胞。直接转化宿主细胞，或通过载体进行转化。在一个实施方案中，

用于例如大肠杆菌菌株 HB101 或 JM109 的原核细胞的适当载体是质粒 pKK232-8。

本发明的启动子可用于多种原核和真核宿主细胞。在一个实施方案中，在包括烟草，小麦和其它植物的植物中使用本发明的启动子促进高水平的基因表达。

本发明还提供了生产蛋白质组分的方法。根据此实施方案，在被本发明基因构建体转化的宿主细胞中产生了蛋白质产物。该基因构建体包括与编码有待在宿主细胞中产生的蛋白质的结构基因可操作相连的本发明的启动子序列。本发明也提供了筛选和分离具有强转录特性的启动子序列（包括下文所示小球藻病毒启动子的截短形式）的方法。

图 1 图解说明了 pKK232-8 质粒图谱。

图 2 图解说明了使用与编码 CAT 蛋白的异源 DNA 序列相连的 7 个小球藻病毒启动子的基因构建体。

图 3 是用与启动子序列 *cvp*-10，*cvp*-13，*cvp*-15 和 *cvp*-16 可操作相连的 CAT 基因转化的氯霉素抗性大肠杆菌的 CAT 活性比较。

图 4 是转化至大肠杆菌 HB101 并在多种浓度氯霉素的存在下生长的小球藻病毒启动子 *cvp*-13 和 *tac* 启动子的启动子活性比较。

图 5 图解说明了 pBN115-glp 质粒图谱。

图 6 是显示不同温度下，由 pBN115-glp 和 pBN115-glp/tac 表达 GLP-1 的 SDS-PAGE 凝胶。

图 7 是含有 3 拷贝 GLP-1 和 YX16 启动子或 tac 启动子的表达盒所表达蛋白质的 SDS 聚丙烯酰胺凝胶照片，分别为经诱导和未经诱导，可溶和不可溶的蛋白质，于 27℃和 37℃。

图 8 是由 8 拷贝 GLP-1 和 YX15，YX16 或 tac 启动子形成的表达盒所表达蛋白质的 SDS 聚丙烯酰胺凝胶照片，分别为经诱导和未经诱导，可溶和不可溶的蛋白质，于 27℃和 37℃。

图 9 图解说明了 pBN951nk2[GLP(7-36)AFAM]8HAE(tac)的质粒图谱。

本发明涉及分离自几个小球藻病毒基因组的新的核酸序列，所述

分离的核酸序列作为转录启动子起作用[SEQ ID NOS:1-7]。本发明的启动子序列与结构基因可操作相连以指导结构基因在原核或真核细胞中的转录。相对于天然或其它非天然启动子而言，本发明的启动子序列提供了高水平的基因表达，这一点如下文实施例的细菌宿主细胞中所示。

本发明的启动子序列与结构基因序列可操作相连，形成“基因构建体”或“表达盒”。在典型的实施方案中，基因构建体的结构基因序列是异源序列。本文所用的“异源序列”是不同于小球藻病毒中与启动子序列可操作相连之序列的 DNA 序列。基因构建体也优选包括增强子，标记物，聚腺苷酸化序列或其它调节核酸序列。当欲产生分泌蛋白质时，结构基因的编码序列优选包括编码信号肽的核酸序列。

本发明的基因构建体被用于在宿主细胞中表达由结构基因编码的蛋白质产物。使用已知方法将基因构建体直接地，或经由载体掺入宿主细胞。蛋白质产物保持细胞内后-表达或者当基因构建体中含有编码信号肽的核酸序列时，所述蛋白质产物被分泌到细胞外。

小球藻病毒启动子序列

下列核酸序列分离自小球藻病毒，并包括指导结构基因转录的启动子序列。

*cvp*-1 [SEQ ID NO: 1]

CCCGGGGATC GCAGGGCATG GGCATTAAAA GAACTTTATG GAATCAAAAA
TCTTAGTGAA TTTCCACCAC AGGTATATAG TCTTCAGGAC GCTAACGATG
ATATCAACGA TTGTATCAAA GGTTATCGTT TGAGGCACTC ATATCAGGTA
GTTTCTACAC AGAAACTTGA ACAACGCCTG GGAAAAGATC CTGAGCATAG
TAACTTATAT ACTAGCAGAT GTTGTAACGA TGCTTTATAT GAATATGAAT
TAGCACAACG ACAACTACAA AAACAACTTG ATGAATTTGA CGAAGATGGG
TATGATTTTT TTCAGGCACG TATAAATACA TTAGATCCGT CGACCTGCAG
CCAAGCTT

*cvp*-3 [SEQ ID NO: 2]

CCCGGGGATC TAATTCAGGG TGCGAATTTC TTGAACATCA AAGGTCTGTT
GGACGTTTTG TGTGCAGCGG TTGCTGATCG CATTGAATCC ATCAATAAAC
AGATTGGGGT AAATATCAAA CCCAGTTAGT CGGACATTAG AAGGATTTGT
GAGACCACCA CATCCAACGA CACCTAATGG TGTTGTGAAT GATATATTAG
AAATGTTACT TATCATTGAT ATTTGCATAA CACCATTTCC CTTTGCTTGA
TTTCTACCTA TACTAATTGA TTGTATTGTA GTGCACGCGG CGTACTTACT
TGTATTTGCC GTCTCAGACG TGCTTGATAA TAGTGTGGAA CTCGAGTATG
ATCCGTCGAC CTGCAGCCAA GCTT

*cvp*-6 [SEQ ID NO: 3]

CCCGGGGATC ATCGAAAGCA ACTGCCGCAT TCGAAACTTC GACTGCCTCG
TTATAAAGGT TAGTGAAAGC CATTGTATGT TATTACTGAG TTATTTAATT
TAGCTTGCTT AAATGCTTAT CGTGTTGATA TGATAAATGA CAAATGATAC
GCTGTATCAA CATCTCAAAA GATTAATACG AAGATCCGTC GACCTGCAGC
CAAGCTT

*cvp*-10 [SEQ ID NO: 4]

CCCGGGGATC GTTTCTCAGG GCGTCCGGGA GCATATTTCA GACTTGTCCA
GCCGTATGAG CATCACGTGC GCGTTCCTAG CAAGAGCGTG TACGTATATT
CTTTCGCTCT AGAAGATGCA GATTCGAGAC AACCGAATGG ATCGAATCTA
TTTGTACCCC GATATATATA GAATCTAGTC TAAACAAAAC GACCGCGGCT
CTTGCCAATA AATGTGACGC AATTAACGCA TTCGTGAATG ATGACTTGTC
CGCCCCGGTT CTTGACATTC TAAAAAAATG TGGAGTATCC TCGATCCGTC
GACCTGCAGC CAAGCTT

*cvp*-13 [SEQ ID NO: 5]

CCCGGGGATC TGCGTATTGC GGGACTTTTG AGCATTTTCC AGAACGGATT
GCCGGGACGT ATACTGAACC TCCAGTCCCT TTGCTCGTCG TATTTCCCAT
AATATACATA TACACTATTT TAATTATTTA CACCGGTTGT TGCTGAGTGA
TACAATGCAA ATTCCCTCCA CCGAGGAGGA TCGCGAACTG TCCAAATGTC
TTCTTTCTGC AGCTCCATAC GGAGTCGTTA GGAAACATTC ACTTAATTAT
AGGATCCGTC GACCTGCAGC CAAGCTT

*cvp*-15 [SEQ ID NO: 6]

CCCGGGGATC AGGCCTCGCT TATAAATATG GTATTGATGT ACTTGCCGGT
GTGATTGACT CAGATTACAG AGGAGAGTTG AAAGCAATCC TTTACAATAC
TACAGAACGT GACTATATTA TCAAAAAAGG CGATCAGCCA AGCTTCGTCG
ACCTGCGATC CGTCGACCTG CAGCCAAGCT T

*cvp*-16 [SEQ ID NO: 7]

CCCGGGGATC GCAAAACTCA CAGTCAACAA ACCAAAACAC GGAATGAAGA
AAGGAGAAAC TGTGATCATG TGGCAACAAG ATGGAGGTGT CATAGACTAC
ATTTACCCTC CCTCTGATCA TCGAAAGCAA CTGCCGCATT CGAAACTTCG
ACTGCCTCGT TATAAAGGTT AGTGAAAGCC ATTGTATGTT ATTACTGAGT
TATTTAATTT AGCTTGCTTA AATGCTTATC GTGTTGATAT GATAAATGAC
AAATGATACG CTGTATCAAC ATCTCAAAAG ATTAATACGA AGATCCGTCG
ACCTGCAGCC AAGCTT

下文实施例中更全面地描述了发现这些新启动子序列的方法。简而言之，由5株小球藻病毒，CA-4B，A1-1A，PBCV-1，SC-1A和NC-1A的病毒基因组产生限制性DNA片段(Van Etten，1991，微生物学评论，55:586-620)。使用已知方法，通过鸟枪克隆法将限制性片段插入质粒pKK232-8(Sambrook等，1989，分子克隆)。

质粒载体pKK232-8含有无启动子的氯霉素乙酰转移酶(CAT)基因，在其上游含有多克隆位点以插入DNA限制性片段。然后，用克隆的pKK232-8转化大肠杆菌，用氯霉素抗性筛选携有启动子序列的转化子。为了得到高活性的启动子，使用生长培养基中增加浓度的氯霉素进一步筛选氯霉素抗性转化子。

本文所用的启动子"强度"指的是由启动子指导的转录水平。"强"启动子提供了比弱启动子更高水平的转录。因此，术语"强启动子"可以与"高活性启动子"互换使用。强启动子对于基因产物的商业生产特别有用。

应理解启动子功能可能不需要SEQ ID NOS:1-7中每一个所述的完整核酸序列。使用上文和下文实施例中所述的方法，可进一步地限制，如截短或修饰本发明的启动子序列，并进行筛选以提炼出有活性的启动子区域。

制备基因构建体

根据本发明，基因构建体包括至少一种与转录控制区可操作相连的结构基因编码序列。转录控制区包括启动子和其它调节元件，如增强子，调节元件，聚腺苷酸化序列，转录起始区和转录终止序列。SEQ ID NOS:1-7各包括启动子序列，也可包括其它调节元件。将启动子序列与结构基因序列可操作相连的方法是已知的，例见Itakuri等，1977，科学，198:1056-1063。

本发明基因构建体的结构基因一般编码蛋白质或多肽产物。使用已知方法可将任何已知的或后来发现的编码所需产物的结构基因与本发明的启动子序列可操作相连。适用于与本发明启动子一起使用的已

知结构基因的例子包括编码下列物质的那些核酸序列：胰高血糖素样肽 1(GLP-1)，生长激素释放因子(GRF)，甲状旁腺激素(PTH)，碳酸酐酶，β-半乳糖苷酶，氯霉素乙酰转移酶(CAT)，谷胱甘肽乙酰转移酶，互连肽，酰胺化序列，以及其它结构基因。

因此，在一个实施方案中，本发明的启动子序列与编码碳酸酐酶，如人碳酸酐酶的 DNA 序列可操作相连。在另一个实施方案中，本发明的启动子序列与编码多拷贝所需蛋白质的 DNA 序列相连。美国专利 5,595,887 中公开了胰高血糖素样肽－1(GLP-1)的适当多拷贝结构基因的例子。

基因转化方法

一旦形成基因构建体，直接将它导入宿主细胞，或亚克隆至适当载体中以转化宿主细胞。

转化细胞的方法是已知的，优选的方法随被转化宿主细胞的类型而变化。本文所用的“宿主细胞”指的是最终表达基因构建体的结构基因的细胞。对于包括细菌，酵母和动物宿主细胞的原核和真核宿主细胞而言，优选的转化方法包括冻融法，氯化钙沉淀法，磷酸钙沉淀法，质粒，原生质体转化，脂质体指导的转化，电穿孔和其它已知的转化方法。

对于植物细胞而言，优选的转化方法包括农杆菌(Agrobacterium)指导的转化，电穿孔，微粒轰击，原生质体融合，这些方法和其它已知转化方法的结合。

宿主细胞

本发明的启动子可用于多种宿主细胞，包括原核和真核宿主细胞。用于表达本发明基因构建体的适当细菌宿主细胞包括大肠杆菌(Escherichia coli)，枯草芽孢杆菌(Bacillus subtilis)和链霉菌属(Streptomyces)。本发明的启动子也可用于植物和动物细胞，包括酵母细胞，如烟草，小麦和其它植物细胞，胚胎，肿瘤和其它动物细

胞等。

将基因构建体携至宿主细胞的适当载体系统包括例如质粒，病毒，噬菌体和酵母人工染色体(YAC)。如下文实施例中所述，适于用本发明基因构建体转化细菌宿主细胞的适当质粒包括 pKK232-8 或 pB0304。

将外源基因导入植物的方法是已知的，可使用该方法将本发明的基因构建体插入植物宿主，所述方法包括生物和物理植物转化法，例见 Miki 等，1993，“将外源 DNA 导入植物的方法”，植物分子生物学和生物技术方法，Glick 和 Thompson 编，CRC 出版公司，Boca Raton，p67-88。选定的方法随宿主植物的不同而不同，包括化学转染法如磷酸钙，微生物指导的基因转移如农杆菌(Horsch 等，科学，227:1229-31，1985)，电穿孔，显微注射，和 biolistic 轰击。

植物细胞的表达载体和体外培养方法或植物的组织转化和再生是已知的，可以获得的，例见 Gruber 等，1993，“植物转化载体”，植物分子生物学和生物技术方法，Glick 和 Thompson 编，CRC 出版公司，Boca Raton，p89-119。

实施例

参照下列实施例更全面地描述本发明，但实施例不以任何方式限制本发明。

实施例 1

产生病毒 DNA 片断

提取自 5 株小球藻病毒 CA-4B；A1-1A；PBCV-1；SC-1A 和 NC-1A 的病毒 DNA 由 Nebraska 大学的 Van Etten 博士提供。合并提取的病毒 DNA(各 1.5μg)，通过在(100μl 体积的)100mM NaCl，10mM Bis Tris Propane-HCl，10mM $MgCl_2$，1mM 二硫苏糖醇和 100μg/ml BSA 中用 *Sau*3A1(New England Biolabs，Beverly，MA)消化产生了病毒 DNA 片断。将混合物于 37℃保温 120 分钟，通过 10mM EDTA 终止反应。用

乙醇从消化混合物中沉淀出 *Sau*3A1 片断，并用乙醇洗涤。

实施例 2

将片断克隆至 pKK232-8

通过 Sambrook 等在 1989，分子克隆中所述的鸟枪克隆法将按实施例 1 所述产生的 *Sau*3A1 病毒片断克隆至质粒 pKK232-8 中。质粒 pKK232-8 购自 Pharmacia(Piscataway, NJ)。pKK232-8 的图谱示于图 1。按 Sambrook 等（文献同上）所述或根据厂商的指导进行 DNA 的制备，补平反应和连接。

将小球藻病毒 DNA 片断与 1μg 预先经 *BamH*I 和牛小肠碱性磷酸酶(CIP)处理过的 pKK232-8 连接，产生重组质粒 pCVP-1; pCVP-3; pCVP-6; pCVP-10; pCVP-13; pCVP-15;和 pCVP-16。这 7 个与异源 CAT 基因可操作相连的小球藻病毒启动子示于图 2。(缩写: Sm, *Sma*I; Sa, *Sal*I H, *Hind*III; CAT, 氯霉素乙酰转移酶基因)，牛小肠碱性磷酸酶购自 Promega(Madison, WI)。T4 DNA 连接酶购自 BRL (Rockville, MD)。

pKK232-8 是 pBR322 的衍生物，它含有无启动子的氯霉素酰基转移酶(CAT)基因，在其上游含有多克隆位点以插入 DNA 限制性片段。被 pKK232-8 转化的大肠杆菌细胞对氨苄青霉素具有抗性，但对氯霉素敏感，除非含有启动子的 DNA 片断被插入 CAT 基因的上游以诱导 CAT 的表达。如果插入这种启动子，携有重组质粒的细胞即表达 CAT，从而获得氯霉素抗性。

pKK232-8 被设计成降低氯霉素抗性背景并增加筛选强启动子的能力。为了降低背景，CAT 基因的侧翼为有效的转录终止子，该转录终止子阻断质粒上存在的其它启动子转录 CAT 基因。在多克隆位点和 CAT 基因起始密码子之间的所有 3 个阅读框中导入翻译终止密码子以防止从克隆的启动子片断可能导入的任何 ATG 起始密码子开始翻译读。CAT 基因也含有其自身的核糖体-结合信号和 ATG 起始密码子，能有效地翻译 CAT mRNA。

实施例 3

转化大肠杆菌和选择转化子

按 Sambrook 等，1989(文献同上)所述进行大肠杆菌 HB101 和 JM109 的转化。所述大肠杆菌菌株购自 Promega(Madison, WI)。

筛选经转化的大肠杆菌的可显示被 pKK232-8 转化的氨苄青霉素抗性和显示启动子插入的氯霉素抗性。通过在递增量的氯霉素存在下测定细胞生长来估计插入的启动子的强度。

在含有 30μg/ml 氨苄青霉素和多种浓度氯霉素(5, 10, 20, 30, 100μg/ml)的 Luria 肉汤(LB)平板上分离阳性菌落。选择抗 100μg/ml 氯霉素的大肠杆菌菌落，接种于含 200, 400, 600, 700, 900μg/ml 氯霉素的 LB 培养基中。通过测定培养物的 $O.D._{600}$ 来监测细胞生长。

得到几千个抗 30μg/ml 氯霉素的转化子。抗氯霉素的转化子数目随抗生素浓度的增加而显著降低。仅有约 500 个转化子抗 100μg/ml 氯霉素，仅有 36 个转化子显示出对 500μg/ml 氯霉素的抗性。被未插入启动子的对照 pKK232-8 转化的细胞对浓度大于 5μg/ml 的氯霉素没有抗性。

将显示出对 500μg/ml 氯霉素具有抗性的 36 个转化子进一步暴露于 500, 700 和 900μg/ml 氯霉素中。7 个转化子在含有 700μg/ml 氯霉素的 LB 培养基中显示出正常生长，在 900μg/ml 氯霉素存在下显示出较缓慢的生长(图 2)。

实施例 4

比较由小球藻病毒启动子诱导的 CAT 活性

比较由实施例 3 氯霉素抗性小球藻病毒启动子转化子表达的 CAT 活性。在 600μg/ml 氯霉素存在下，使 7 个氯霉素抗性大肠杆菌转化子中的 4 个生长(CVP-10; CVP-13; CVP-15 和 CVP-16)。通过测定不同时间的细胞密度($O.D._{600}$)来监测细胞生长。如图 3 所示，4 个转化子显示出类似的细胞生长，这表明各个培养物中表达了类似水平的

CAT 活性。

实施例 5

纯化质粒 DNA

使用 Wizard 微量制备质粒 DNA 纯化试剂盒(Promega 公司，Madison, WI)或 Qiagen DNA 纯化试剂盒(Qiagen 公司，Chatsworth, CA)从显示出最高水平氯霉素抗性(700 至 900μg/ml)的那些菌落中纯化质粒 DNA。通过限制性核酸内切酶消化对质粒进行分析，结果表明所有质粒都携有 DNA 插入片段。

用 *Sma*I 和 *Hind*III 消化质粒，并在 7.5%聚丙烯酰胺凝胶或 0.8%琼脂糖凝胶上电泳。如图 2 所示，电泳凝胶显示出大小为 100 至 400bp 的 7 个病毒启动子片段。

实施例 6

测定启动子片段的序列

用 *Sma*I 和 *Hind*III 从 pKK232-8 载体上切下 7 个含启动子的病毒 DNA 片段(*cvp*-1, *cvp*-3, *cvp*-6, *cvp*-10, *cvp*-13, *cvp*-15 和 *cvp*-16)并亚克隆至 pUC19 或 pBluescriptSK(+)II 以供测序。由 Nebraska Lincoln 大学(UNL)测序实验室使用自动 LICOR 测序仪进行 DNA 测序。使用 Sanger 双脱氧链终止法在两个方向测定序列。

所有 7 个小球藻病毒片段的序列分析揭示出 *Sau*3A1 位点位于小球藻病毒插入物的每个末端，并侧翼于 pKK232-8 载体的多克隆位点。7 个启动子片段 DNA 序列的大小与通过聚丙烯酰胺凝胶电泳测定的限制性片段的大小相符。

| SEQ ID NO: | 小球藻病毒 DNA 插入物 |
|---|---|
| 1 | cvp-1 |
| 2 | cvp-3 |
| 3 | cvp-6 |
| 4 | cvp-10 |
| 5 | cvp-13 |
| 6 | cvp-15 |
| 7 | cvp-16 |

已知的大肠杆菌启动子序列含有 3 个必需的元件：两个六聚体序列(-35 区域和-10 区域)和这两个区域之间的 16-18bp 间隔区。使用 *lac*, *lac*UV5, *trp*, *tac* 和 PL 启动子共有序列作为参照，将-35 和-10 区域推定序列对 7 个病毒启动子中的每一个进行排布。如下表 I 所示，这些六聚体序列与已知大肠杆菌启动子的共有序列或者相同，或者仅稍有不同。例如，*cvp*-10(TTGACA)或 *cvp*-15(TTTACA)的-35 序列区域与 *trp*(TTGACA)或 *lac*(TATGTT)中的相同。

本发明 7 个分离的病毒启动子中的 6 个在推定的-35 和-10 区域之间具有 16-18bp 间隔区。*cvp*-13 病毒启动子和 *lac* 启动子都具有相同的-35 六聚体序列(TTTACA)，-35 和-10 区域之间完全相同的间隔区(18bp)，非常相似的-10 六聚体序列 (*cvp*-13 为 TACAAT，*lac* 为 TATAAT)。

| 启动子 | -35 位 | -10 位 | +位 | |
|---|---|---|---|---|
| 共有序列 | TTGACA | TATAAT | | SEQ.ID NO: |
| *cvp*-1 | CAAAAACAACTTGATGA ATTTGACGAAGATGGGTATGATTTTTTTCAGGC | | | 10 |
| *cvp*-3 | TCAGACGTGCTTGATAATAG TGTGGAACTCGAGTATGATCCG TCGACCT | | | 11 |
| *cvp*-6 | GCTTATCGTGTTGATATGATAAATGACAAAT GATACGCTGTATCAACA | | | 12 |
| *cvp*-10 | GCCCCGGTTCTTGACATTCTAAAAAAATGTGGAGTATCCTCGATCCGTCGA | | | 13 |
| *cvp*-13 | ATTTTAATTATTTACACCGGTTGTTGCTGAGTGATACAATGCAAATTCCCT | | | 14 |
| *cvp*-15 | AAAGCAATCCTTTACAATAC TACAGAACGTGACTATATTATCAAAAAAGG | | | 15 |
| *cvp*-16 | ACTGCCTCGTTATAAAGGTTAGTGAAAGCCATTGTATGTTATTACTGAGTT | | | 16 |
| *lac* | CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTG TGGAATT | | | 17 |
| *lac*UV5 | CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTG TGGAATT | | | 18 |
| *trp* | AAATGAGCTGTTGACAATTAATCATCGAACTA GTTAACTAGTACGCAAGT | | | 19 |
| *tac* | AAATGAGCTGTTGACAATTAATCAT CGGCTCGTATAATGTG TGGAATT | | | 20 |
| *PL* | TCTGGCGGTGTTGACATAAATACCACT GGGGGTGATACTGAG CACATCA | | | 21 |

黑体“A”表示转录起始位点。缩写：*lac*, *lac*启动子；*lac*UV5, *lac*UV5启动子； *trp*, *trp*启动子； *tac*, *tac*启动子； *PL*, 噬菌体λ *PL*启动子。

实施例 7

比较 *cvp*-13 和 *tac* 的启动子活性

tac 启动子是很强的大肠杆菌启动子(de Boer 等, 1983, Proc. Natl. Acad. Sci. 80:21-25)，已在研究和工业中被广泛用于基因表达。为了在相同检测系统中将病毒启动子 cvp-13 的启动子活性和 tac 的启动子活性进行比较，使用 pKK232-8 构建 tac 的启动子质粒。

将含有 tac 序列的两个互补寡聚体，

5′-GGGAAATGAGCTGTTGACAATTAATCATGGCTCGTATAATGTGTGGAAGCTT-3′

[SEQ ID NO: 8]

和

5′-CCCTTTACTCGACAACTGTTAATTAGTAGCCGAGCATATTACACACCTTCG-3′

[SEQ ID NO: 9]

退火并插入 pKK232-8 中 CAT 基因上游，*Sma*I 和 *Hind*III 之间。将所得质粒 pTAC-cat 转化至大肠杆菌 HB101 中。在 100, 300 或 600μg/ml 氯霉素存在下培养含有 pTAC-cat 或 pCVP-13 的细胞。通过测定不同时间的细胞生长密度来监测抗生素抗性。如图 4 所示，在所有 3 个氯霉素浓度水平下，*cvp*-启动子显示出比 *tac* 启动子更高的启动子活性。

实施例 8

构建受调节的病毒启动子

分离自小球藻病毒基因组的 7 个病毒启动子片段在大肠杆菌中显示出强启动子活性。然而，它们的启动子活性不受调节，因此，只能组成型指导蛋白质表达。已为几个启动子开发了通过融合启动子与 lac 操纵基因将组成型启动子转变为诱导型启动子的方法，所述启动子如核糖体 RNA 基因启动子(Brosius 和 Holy, 1984, PNAS USA 81:6929-6933)和脂蛋白基因启动子(Nakamura 等, 1982, EMBO J. 1:771-775)。这些经修饰的杂合启动子受 *lac*I 基因产物，*lac* 阻遏物的阻遏，所述阻遏物与 *lac* 操纵基因特异性的结合抑制转录的起始并且通过如异丙基 b-D-硫代半乳糖苷(IPTG)这样的诱导物而解除阻遏，从而防止阻遏蛋白与 *lac* 操纵基因结合。

将 lac 操纵基因与小球藻病毒启动子 cvp-1, cvp-15 和 cvp-16 融合以评估小球藻病毒启动子是否可变为诱导型启动子。由 Operon Technology 公司(Alameda, CA)化学合成了含有 lac 操纵基因序列和适当限制性克隆位点的两个互补寡核苷酸。这些寡核苷酸的序列如下:

5'-TCGACCGGGTACCTCGCGAGGAATTGTGAGCGGATAACAATTCCTCTAGA-3'

[SEQ ID NO: 22]

5'-AGCTTCTAGAGGAATTGTTATCCGCTCACAATTCCTCGCGAGGTACCCGG-3'

[SEQ ID NO: 23]

变性和退火之后，将所得寡核苷酸双链体插入 pKK232-8 上 *Sal*I 和 *Hind*III 位点处的病毒启动子下游。将在 CAT 基因上游含有病毒启动子/操纵基因杂合体的 pKK232-8 克隆转化至缺乏 *lac*I 功能的大肠杆菌 HB101。通过测定经转化细胞对氯霉素的抗性而检测到与缺乏 *lac* 操纵基因的病毒启动子类似水平的启动子活性。这表明通过与 *lac* 操纵基因融合而操作病毒启动子不会导致启动子强度的降低。与 *lac* 操纵基因融合的启动子 cvp-1，cvp-15 和 cvp-16 分别被称为 YX-1，YX-15 和 YX-16。

实施例 9

由 YX-15 表达载体表达胰高血糖素样肽 (GLP-1)

A. 构建 YX-15 表达载体，pBN115-glp

为了将病毒启动子(YX-1，YX-15 和 YX-16)亚克隆至表达载体(需要限制性位点 BglII 和 XbaI 以供克隆)，使用由 Operon Technology 公司合成的下列寡核苷酸进行聚合酶链反应(PCR)以修饰启动子片段：

5'-GGGTTCGAAGATCTAATTCCCGGGGATC-3' [SEQ ID NO: 24] 和

5'-TGCATTTCTAGAATTGTGAATTGTTATCCGCTCA-3' [SEQ ID NO: 25].

质粒 pGEX-2T 购自 Pharmacia Biotech 公司(piscataway，NJ)，该质粒携有 pBR322 复制起点，*lac*Iq 基因，氨苄青霉素抗性基因，和 *tac* 启动子/谷胱甘肽 S-转移酶(GST)融合基因表达盒。首先通过用 FspI 和 SmaI 消化除去 1.2kb 的 FspI-SmaI 片段以缺失 pGEX-2T 上的 *tac*/GST 表达盒，然后用含有 BglII/XbaI/NheI/XhoI 位点，可用作克隆位点的 DNA 片段取而代之。所得质粒被称为 pBN115。

GLP-1 是 29 个残基的激素肽。构建合成的 8 拷贝 GLP-1 基因并在

大肠杆菌中表达之。为了检测在 YX-15 控制下的 GLP-1 表达，通过将含有 YX-15 的 BglII-XbaI 片段与含有核糖体结合位点和 8 拷贝 GLP-1 基因的 XbaI-XhoI 片段融合，构建 YX-15/GLP-1 表达盒。然后，将 YX-15/GLP-1 表达盒插入 pBN115 唯一的 BglII 和 XhoI 位点处，产生 pBN115-glp(图 5)。通过用含有 *lac* 操纵基因的 *tac* 启动子替代 pBN115-glp 上 BglII 和 XbaI 位点处的 YX-15，构建 pBN115-glp/tac 以用作对照表达载体。*tac* 启动子也可被 IPTG 严格调节。

B. 表达 GLP-1

在氨苄青霉素的存在下，将 p'BN115-glp 和 pBN115-glp/tac 转化至大肠杆菌宿主菌株 HMS174(Novagen 公司，Milwaukee，WI)。在 37℃至 21℃下，用 0.1 至 2mM IPTG 诱导 GLP-1 的表达。诱导 15 小时之后收获细胞。制备总细胞蛋白供 SDS-PAGE 分析(图 6)。使用 Personal Color Scanner 和 Adobe photoshop 和 NIH Image Analysis 计算机程序，通过凝胶扫描来评价 GLP-1 的产生。如表 1 所示，凝胶扫描揭示出 37℃下，pBN115-glp 表达 GLP-1 的水平大于总细胞蛋白的 46%，而 pBN115-glp/tac 表达 GLP-1 的水平大于总细胞蛋白的 35%。1 O.D.$_{600}$ 细胞中，pBN115-glp 产生的 GLP-1 比 pBN115-glp/tac 产生的多 40%。此结果清楚地表明 YX-15 作为很强的启动子在大肠杆菌中产生高水平重组蛋白的潜在应用。

表 1

大肠杆菌中 YX-15 指导的 GLP-1 表达

| 诱导 | GLP-1 蛋白带强度的读数 | |
|---|---|---|
| | *tac* | YX-15 |
| 未诱导 | 84(7%) | 385(35%) |
| 37℃ | 1089(100%) | 1544(141%) |
| 27℃ | 55(5%) | 1248(114%) |
| 21℃ | ND | 830(76%) |

每次处理时从 0.25 O.D.$_{600}$ 细胞中提取总蛋白并在 16% PAGE 凝胶上分析(图 6)。用考马斯蓝染色后，使用 Personal Color Scanner 和 Adobe photoshop 和 NIH Image Analysis 计算机程序扫描蛋白质带的强度。圆括号中的百分数表示相对于 37℃下由 tac 启动子驱动的 GLP-1 表达水平的 GLP-1 表达水平。

C. 通过 IPTG 诱导

尽管在没有 IPTG 的情况下可检测到携有 pBN115-glp 的 HMS174 中 GLP-1 的基本表达水平，但加入 IPTG 后，GLP-1 表达水平增加了 4 倍。与之形成对照的是，IPTG 诱导 pBN115-glp/tac 产生的 GLP-1 增加了 13 倍。这表明 pBN115-glp 的 GLP-1 表达确实对 IPTG 诱导作出反应。然而，对 pBN115-glp 的 GLP-1 表达所作的调节不如对 pBN115-glp/tac 的严格。

37℃下，检测多种浓度(0，0.25，0.5，1.0 和 2.0mM)的 IPTG 对 GLP-1 表达的诱导能力。在 IPTG 浓度范围为 2mM 至 0.1mM 时，pBN115-glp 产生了类似高水平的 GLP-1。这表明低至 0.1mM 的 IPTG 浓度足以诱导 GLP-1 超量表达。

在所有受试的 IPTG 浓度下，pBN115-glp 产生的 GLP-1 比 pBN115-glp/tac 产生的 GLP-1 始终要高，这说明 YX-15 的启动子活性比 *tac* 高。

D. 温度的影响

培养温度对微生物的行为有很大的影响。通常使用 37℃作为诱导温度以在大肠杆菌中产生所需的重组蛋白，而在某些情况下，较低的温度在产生活性蛋白质和避免形成无活性的蛋白质聚集体方面显示出优点。然而，在较低的温度下，作为生物化学反应的转录和翻译的速率大大降低，因此，目的蛋白质的终产量降低。因此，使用在较低温度下具有高活性的启动子对于产生高水平的蛋白质是必需的。在本发明之前，这种大肠杆菌启动子还不能得到。如 tac 启动子和 T7 启动子

这样的常用启动子在 37℃下活性很高，但在较低的温度下活性弱得多，下文中有关 *tac* 启动子的数据可以说明此现象。

在 37℃至 21℃的温度范围内，比较 YX-15 控制 GLP-1 产生的启动子活性。表 1 所示的结果(上文)证明 YX-15 启动子在 37℃至 21℃的温度范围能驱动高水平 GLP-1 的表达。27℃下，使用 YX-15 启动子的表达水平甚至比 37℃时使用 *tac* 启动子的表达水平稍高。27℃和 21℃时，YX-15 驱动的 GLP-1 表达水平降低，与 37℃时相比，分别降低 20%和 50%。与之形成对照的是，使用 *tac* 时，27℃与 37℃相比，表达降低了 95%。21℃时未检测到 *tac* 启动子的蛋白质表达。

实施例 10

病毒启动子在植物中的功能

为了检测分离的病毒启动子是否能在植物中起作用，将 7 个启动子片段插入β-葡糖醛酸酶(GUS)表达载体 pBI221(Clotech Laboratories, Inc., Palo Alto, CA)中的 GUS 基因上游以替代 CaMV 35S 启动子。通过微粒轰击法将所得质粒导入烟草叶和小麦不成熟的胚中。然后通过计数受试外植体上蓝色 GUS 焦点的形成，组化检测经转化外植体的 GUS 表达情况。用含有 CaMV 35S 启动子的 pBI221 转化阳性对照外植体。

如表 2 所示，用启动子 cvp-15 观察到 GUS 瞬时测定的阳性结果，但用另外 6 个启动子未观察到阳性结果。

表 2

| 载体 | GUS 焦点平均范围/外植体/植物* | |
|---|---|---|
| | 烟草 | 小麦 |
| pBI221(CaMV 35S) | 2.6-9.5 | 7.6-17.2 |
| pBI221-YX15 | 0.7-1.5 | 0.4-1.6 |

*n=3 次实验

实施例 11

用 pYX16 低温表达 GLP 构建体

上文实施例 9 中所述的表达质粒 pBN115-glp(8 拷贝 GLP-1)使用 YX15 启动子(cvp-15 加上 *lac* 操纵基因)驱动 8 拷贝 GLP-1(7-36)表达盒的表达。通过用 YX-16 启动子的片段替代 pBN115 载体中 BglII/XbaI 启动子片段，构建含有 YX16 启动子(cvp-16 加上 *lac* 操纵基因)和 3 拷贝 GLP-1(7-36)的第二个质粒。通过 XbaI-XhoI 片段的取代用 3 拷贝 GLP-1 盒取代 8 拷贝 GLP-1 盒。此构建体 pGEX1nk2(3 拷贝 GLP)YX16 含有与 pBN115 相同的质粒骨架，不同之处是启动子是 YX16 而不是 YX15。

使用 pBN53 表达质粒制备驱动 8 拷贝 GLP-1(7-36)表达的 YX16 启动子构建体。YX16 启动子作为 BglII/XbaI 片段被克隆至此构建体。所得质粒 pBN531nk28 拷贝 GLP YX16 为四环素抗性，并含有 lacIq 基因以阻遏未诱导的表达。pBN95 质粒与 pBN53 质粒相同，不同的是它含有 lacI 基因而不是 lacIq 基因。这一变化导致 pBN95 中 lacI 基因产物较低水平的表达，因此，使用此载体可较低水平地阻遏未诱导的蛋白质水平。

使用上文实施例 9 中所述的方法和材料将表达载体转化至大肠杆菌细胞。按实施例 9 中所述，在 27℃和 37℃下用 IPTG 诱导 GLP-1 的表达。

把诱导受 tac 或 YX16 启动子控制的 3 拷贝 GLP-1(7-36)和 8 拷贝 GLP-1(7-36)之后从细胞培养物中取出的蛋白质样品在 16% TrisTricine SDS-PAGE 凝胶上分离，并用考马斯蓝染色。

图 7 所示的数据阐明 YX16 启动子在 27℃下能较好地发挥作用，而 tac 启动子在此温度下无活性。GLP 前体肽于 27℃是可溶性的，而于 37℃是不可溶的。

图 8 所示的数据阐明 YX15 和 YX16 启动子在 27℃下都能发挥作用，而 tac 启动子在此温度下无活性。数据表明 YX16 启动子在低温(27

℃)下发挥作用，具有两个独立的靶肽。

上述说明书，实施例和数据提供了制备和使用本发明组分的完整描述。由于在不违背本发明精神和范围的前提下可想出很多本发明的实施方案，本发明包含在下文所附的权利要求书中。

# 序列表

(1) 一般资料:

(i) 申请人: BIONEBRASKA, InC.

(ii) 发明题目: 小球藻病毒启动子

(iii) 序列数: 25

(iv) 通讯地址:

(A) 收件人: Merchant, Gould, Smith, Edell, Welter & Schmidt

(B) 街道: 3100 Norwest Center, 90 South 7th Street

(C) 城市: Minneapolis

(D) 州: MN

(E) 国家: 美国

(F) 邮政编码: 55402

(v) 计算机可读形式:

(A) 介质类型: 磁盘

(B) 计算机: IBM 兼容机

(C) 操作系统: DOS

(D) 软件: FastSEQ for Windows Version 2.0

(vi) 目前的申请资料:

(A) 申请号: (新提交)

(B) 申请日: 19-MAR-1998

(C) 分类号:

(vii) 在先申请资料:

(A) 申请号: 08/821,559

(B) 申请日: 21-MAR-1997

(viii) 代理人/律师资料:

(A) 姓名: Kettelberger, Denise M

(B) 登记号: 33,924

(C) 资料/文档号: 8648.63WOI1

(ix) 通讯资料:

(A) 电话: 612-332-5300

(B) 传真: 612-332-9081

(C) 电报:

(2) SEQ ID NO: 1 的资料:

(i) 序列特征:

(A) 长度: 358 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 1:

```
CCCGGGGATC GCAGGGCATG GGCATTAAAA GAACTTTATG GAATCAAAAA TCTTAGTGAA     60
TTTCCACCAC AGGTATATAG TCTTCAGGAC GCTAACGATG ATATCAACGA TTGTATCAAA    120
GGTTATCGTT TGAGGCACTC ATATCAGGTA GTTTCTACAC AGAAACTTGA ACAACGCCTG    180
GGAAAAGATC CTGAGCATAG TAACTTATAT ACTAGCAGAT GTTGTAACGA TGCTTTATAT    240
GAATATGAAT TAGCACAACG ACAACTACAA AAACAACTTG ATGAATTTGA CGAAGATGGG    300
TATGATTTTT TTCAGGCACG TATAAATACA TTAGATCCGT CGACCTGCAG CCAAGCTT      358
```

(2) SEQ ID NO: 2 的资料:

(i) 序列特征:

(A) 长度: 374 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 2:

```
CCCGGGGATC TAATTCAGGG TGCGAATTTC TTGAACATCA AAGGTCTGTT GGACGTTTTG     60
TGTGCAGCGG TTGCTGATCG CATTGAATCC ATCAATAAAC AGATTGGGGT AAATATCAAA    120
CCCAGTTAGT CGGACATTAG AAGGATTTGT GAGACCACCA CATCCAACGA CACCTAATGG    180
TGTTGTGAAT GATATATTAG AAATGTTACT TATCATTGAT ATTTGCATAA CACCATTTCC    240
CTTTGCTTGA TTTCTACCTA TACTAATTGA TTGTATTGTA GTGCACGCGG CGTACTTACT    300
TGTATTTGCC GTCTCAGACG TGCTTGATAA TAGTGTGGAA CTCGAGTATG ATCCGTCGAC    360
CTGCAGCCAA GCTT                                                      374
```

(2) SEQ ID NO: 3 的资料:

(i) 序列特征:

(A) 长度: 207 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 3:

```
CCCGGGGATC ATCGAAAGCA ACTGCCGCAT TCGAAACTTC GACTGCCTCG TTATAAAGGT     60
TAGTGAAAGC CATTGTATGT TATTACTGAG TTATTTAATT TAGCTTGCTT AAATGCTTAT    120
CGTGTTGATA TGATAAATGA CAAATGATAC GCTGTATCAA CATCTCAAAA GATTAATACG    180
AAGATCCGTC GACCTGCAGC CAAGCTT                                        207
```

(2) SEQ ID NO: 4 的资料:

(i) 序列特征:

(A) 长度: 317 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 4:

```
CCCGGGGATC GTTTCTCAGG GCGTCCGGGA GCATATTTCA GACTTGTCCA GCCGTATGAG     60
CATCACGTGC GCGTTCCTAG CAAGAGCGTG TACGTATATT CTTTCGCTCT AGAAGATGCA    120
GATTCGAGAC AACCGAATGG ATCGAATCTA TTTGTACCCC GATATATATA GAATCTAGTC    180
TAAACAAAAC GACCGCGGCT CTTGCCAATA AATGTGACGC AATTAACGCA TTCGTGAATG    240
ATGACTTGTC CGCCCCGGTT CTTGACATTC TAAAAAAATG TGGAGTATCC TCGATCCGTC    300
GACCTGCAGC CAAGCTT                                                   317
```

(2) SEQ ID NO: 5 的资料:

(i) 序列特征:

(A) 长度: 277 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 5:

```
CCCGGGGATC TGCGTATTGC GGGACTTTTG AGCATTTTCC AGAACGGATT GCCGGGACGT     60
ATACTGAACC TCCAGTCCCT TTGCTCGTCG TATTTCCCAT AATATACATA TACACTATTT    120
TAATTATTTA CACCGGTTGT TGCTGAGTGA TACAATGCAA ATTCCCTCCA CCGAGGAGGA    180
TCGCGAACTG TCCAAATGTC TTCTTTCTGC AGCTCCATAC GGAGTCGTTA GGAAACATTC    240
ACTTAATTAT AGGATCCGTC GACCTGCAGC CAAGCTT                              277
```

(2) SEQ ID NO: 6 的资料:

(i) 序列特征:

(A) 长度: 181 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 6:

```
CCCGGGGATC AGGCCTCGCT TATAAATATG GTATTGATGT ACTTGCCGGT GTGATTGACT     60
CAGATTACAG AGGAGAGTTG AAAGCAATCC TTTACAATAC TACAGAACGT GACTATATTA    120
TCAAAAAAGG CGATCAGCCA AGCTTCGTCG ACCTGCGATC CGTCGACCTG CAGCCAAGCT    180
T                                                                    181
```

(2) SEQ ID NO: 7 的资料:

(i) 序列特征:

(A) 长度: 316 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 7:

```
CCCGGGGATC GCAAAACTCA CAGTCAACAA ACCAAAACAC GGAATGAAGA AAGGAGAAAC     60
TGTGATCATG TGGCAACAAG ATGGAGGTGT CATAGACTAC ATTTACCCTC CCTCTGATCA    120
TCGAAAGCAA CTGCCGCATT CGAAACTTCG ACTGCCTCGT TATAAAGGTT AGTGAAAGCC    180
ATTGTATGTT ATTACTGAGT TATTTAATTT AGCTTGCTTA AATGCTTATC GTGTTGATAT    240
GATAAATGAC AAATGATACG CTGTATCAAC ATCTCAAAAG ATTAATACGA AGATCCGTCG    300
ACCTGCAGCC AAGCTT                                                    316
```

(2) SEQ ID NO: 8 的资料:

(i) 序列特征:

(A) 长度: 52 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 8:

GGGAAATGAG CTGTTGACAA TTAATCATGG CTCGTATAAT GTGTGGAAGC TT 52

(2) SEQ ID NO: 9 的资料:

(i) 序列特征:

(A) 长度: 51 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 9:

CCCTTTACTC GACAACTGTT AATTAGTAGC CGAGCATATT ACACACCTTC G 51

(2) SEQ ID NO: 10 的资料:

(i) 序列特征:

(A) 长度: 49 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 10:

CAAAAACAAC TTGATGATTT GACGAAGATG GGTATGATTT TTTTCAGGC 49

(2) SEQ ID NO: 11 的资料:

(i) 序列特征:

(A) 长度: 49 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 11:

TCAGACGTGC TTGATAATAG TGTGGAACTC GAGTATGATC CGTCGACCT 49

(2) SEQ ID NO: 12 的资料:

(i) 序列特征:

(A) 长度: 48 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 12:

GCTTATCGTG TTGATATGAT AAATGACAAA TGATACGCTG TATCAACA 48

(2) SEQ ID NO: 13 的资料:

(i) 序列特征:

(A) 长度: 51 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 13:

GCCCCGGTTC TTGACATTCT AAAAAAATGT GGAGTATCCT CGATCCGTCG A 51

(2) SEQ ID NO: 14 的资料:

(i) 序列特征:

(A) 长度: 51 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 14:

ATTTTAATTA TTTACACCGG TTGTTGCTGA GTGATACAAT GCAAATTCCC T 51

(2) SEQ ID NO: 15 的资料:

(i) 序列特征:

(A) 长度: 50 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 15:

AAAGCAATCC TTTACAATAC TACAGAACGT GACTATATTA TCAAAAAAGG 50

(2) SEQ ID NO: 16 的资料:

(i) 序列特征:

(A) 长度: 51 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 16:

ACTGCCTCGT TATAAAGGTT AGTGAAAGCC ATTGTATGTT ATTACTGAGT T 51

(2) SEQ ID NO: 17 的资料:

(i) 序列特征:

(A) 长度: 50 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 17:

CACCCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATGTT GTGTGGAATT 50

(2) SEQ ID NO: 18 的资料:

(i) 序列特征:

(A) 长度: 50 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 18:

CACCCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATAAT GTGTGGAATT 50

(2) SEQ ID NO: 19 的资料:

(i) 序列特征:

(A) 长度: 50 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 19:

AAATGAGCTG TTGACAATTA ATCATCGAAC TAGTTAACTA GTACGCAAGT 50

(2) SEQ ID NO: 20 的资料:

(i) 序列特征:

(A) 长度: 48 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 20:

AAATGAGCTG TTGACAATTA ATCATCGGCT CGTATAATGT GTGGAATT 48

(2) SEQ ID NO: 21 的资料:

(i) 序列特征:

(A) 长度: 49 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 21:

TCTGGCGGTG TTGACATAAA TACCACTGGG GGTGATACTG AGCACATCA 49

(2) SEQ ID NO: 22 的资料:

(i) 序列特征:

(A) 长度: 50 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 22:

TCGACCGGGT ACCTCGCGAG GAATTGTGAG CGGATAACAA TTCCTCTAGA 50

(2) SEQ ID NO: 23 的资料:

(i) 序列特征:

(A) 长度: 50 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 23:

AGCTTCTAGA GGAATTGTTA TCCGCTCACA ATTCCTCGCG AGGTACCCGG 50

(2) SEQ ID NO: 24 的资料:

(i) 序列特征:

(A) 长度: 28 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 24:

GGGTTCGAAG ATCTAATTCC CGGGGATC 28

(2) SEQ ID NO: 25 的资料:

(i) 序列特征:

(A) 长度: 34 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线性

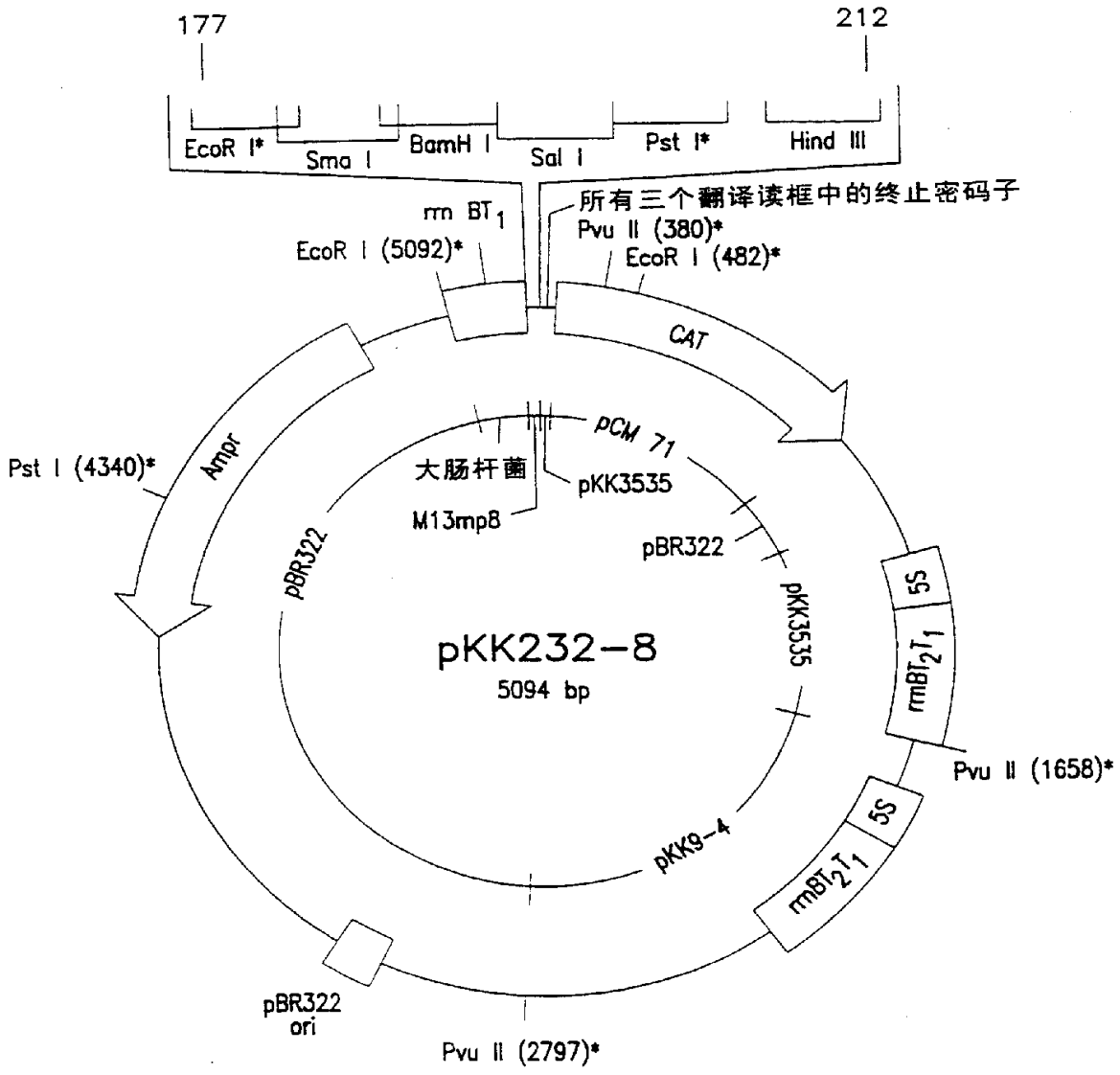
(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 25:

TGCATTTCTA GAATTGTGAA TTGTTATCCG CTCA 34

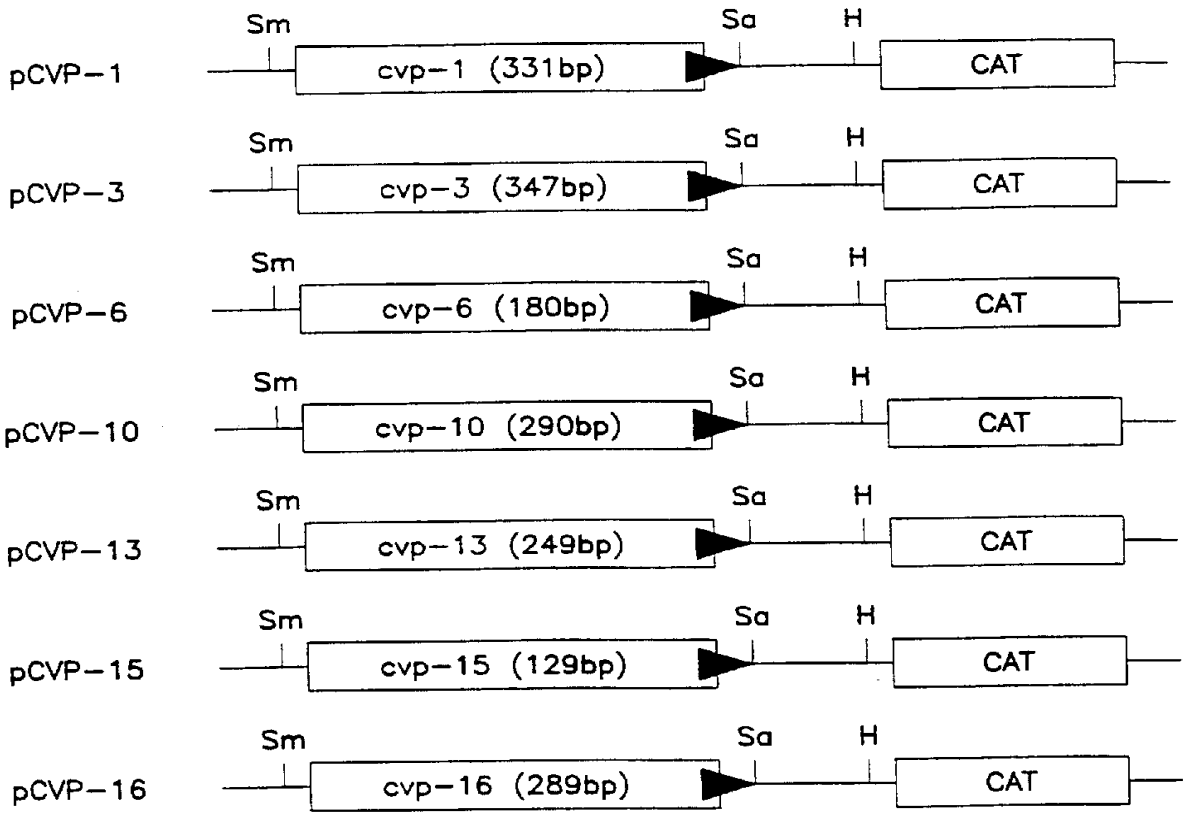
# 说 明 书 附 图

## 图 1




* 不是唯一的位点

# 图 2


pCVP-1
Sm
cvp-1 (331bp)
Sa
H
CAT
pCVP-3
Sm
cvp-3 (347bp)
Sa
H
CAT
pCVP-6
Sm
cvp-6 (180bp)
Sa
H
CAT
pCVP-10
Sm
cvp-10 (290bp)
Sa
H
CAT
pCVP-13
Sm
cvp-13 (249bp)
Sa
H
CAT
pCVP-15
Sm
cvp-15 (129bp)
Sa
H
CAT
pCVP-16
Sm
cvp-16 (289bp)
Sa
H
CAT

# 图 3

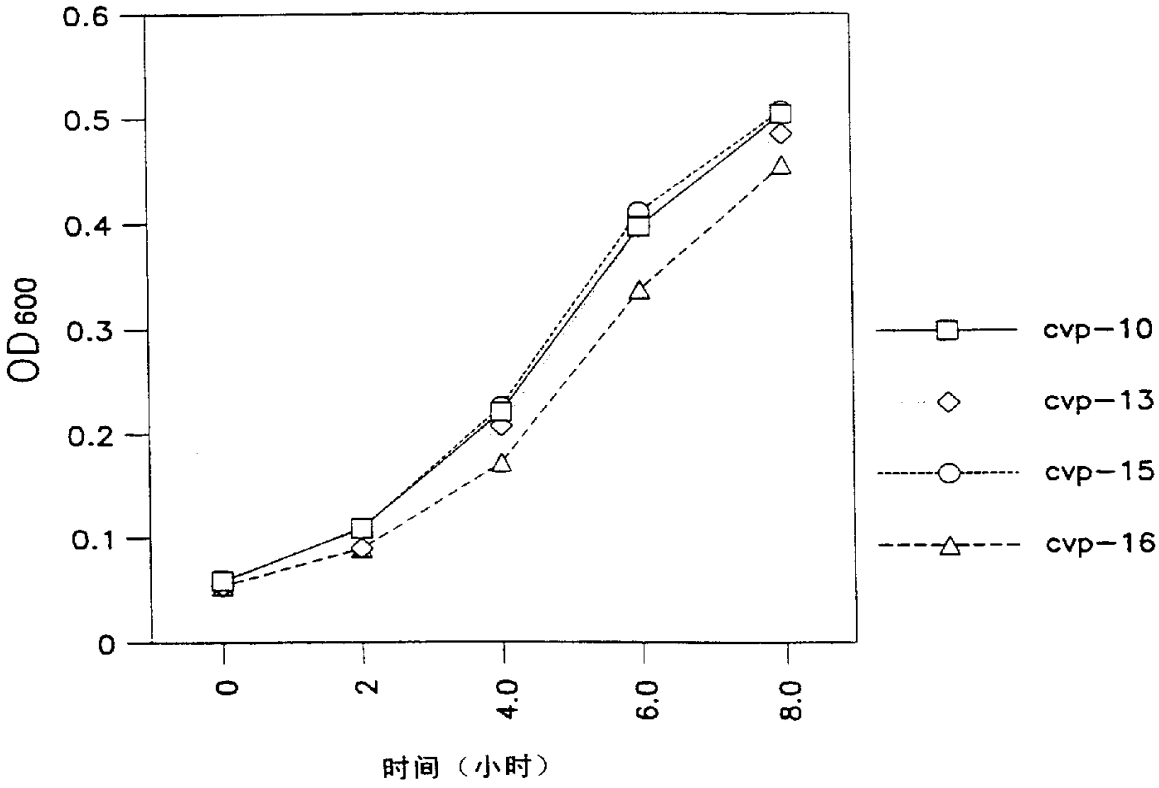
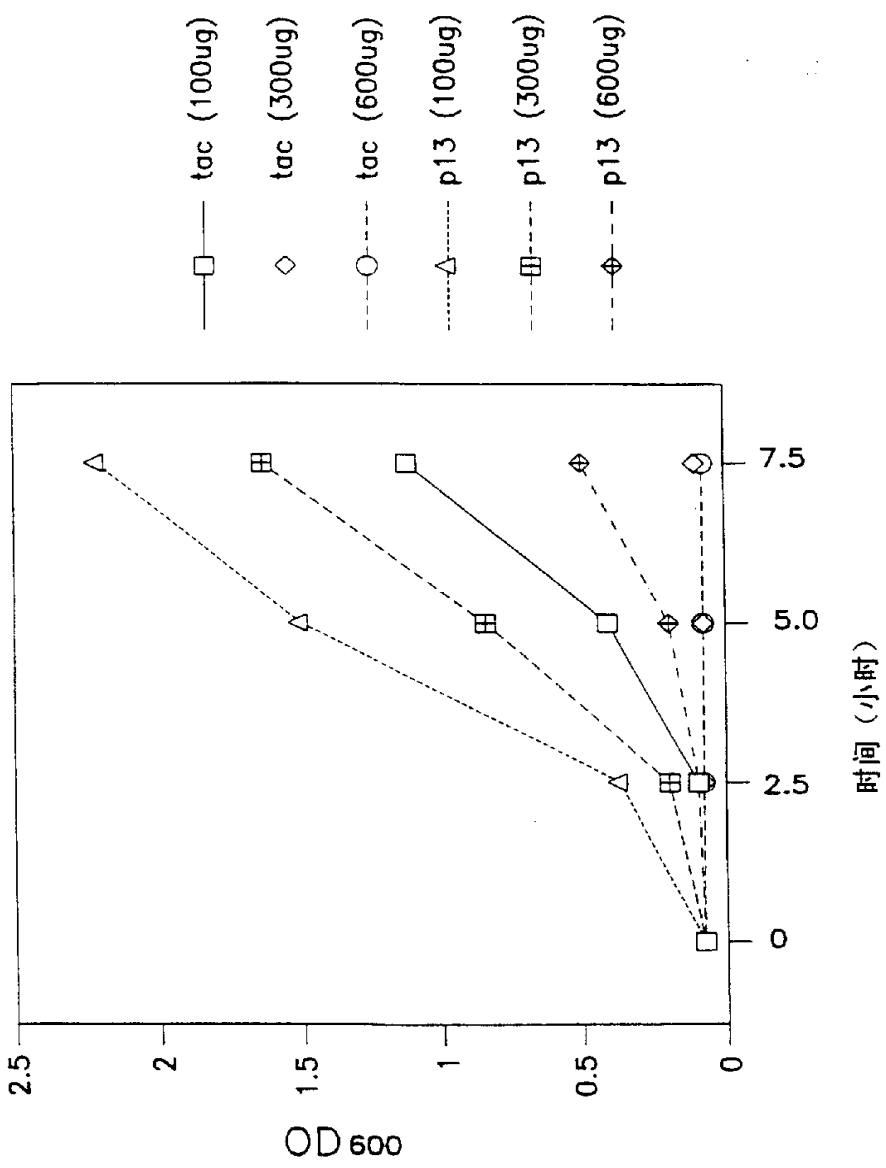
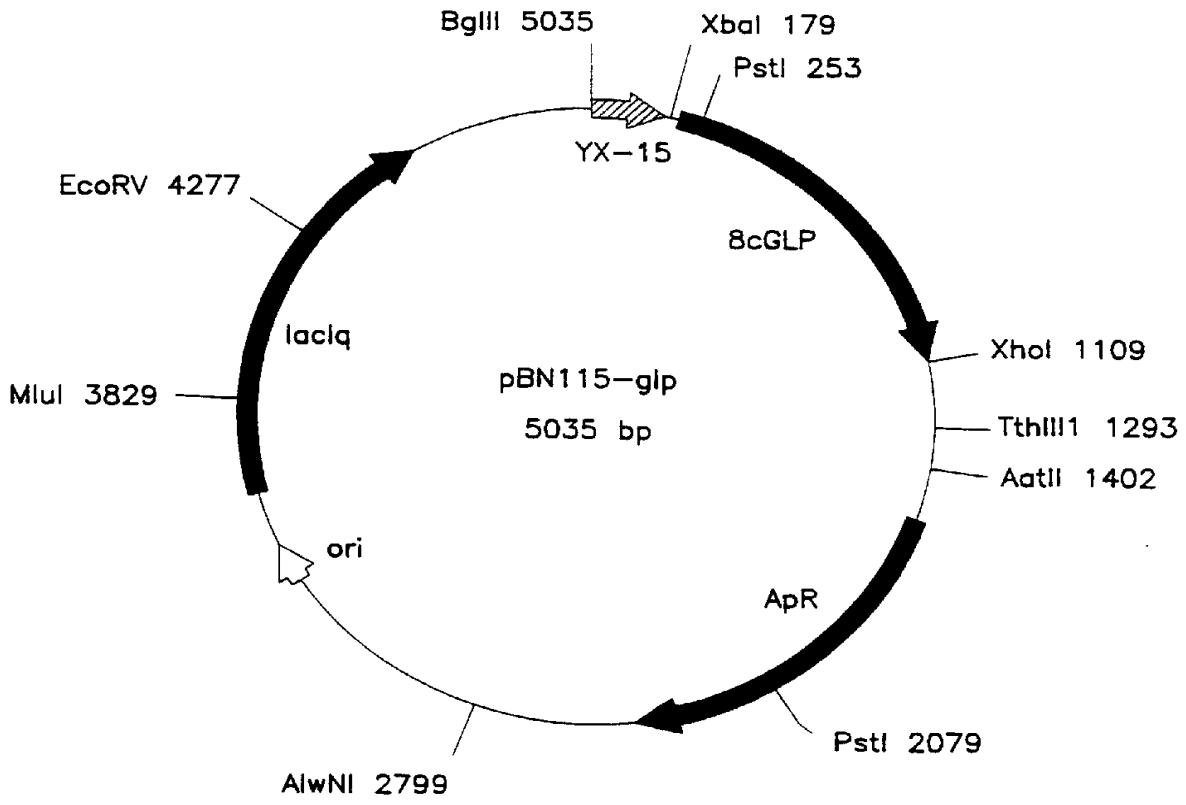

0.6
0.5
0.4
0.3
0.2
0.1
0
OD600
0
2
4.0
6.0
8.0
时间（小时）
cvp-10
cvp-13
cvp-15
cvp-16

图 4


tac (100ug)
tac (300ug)
tac (600ug)
p13 (100ug)
p13 (300ug)
p13 (600ug)
OD 600
2.5
2
1.5
1
0.5
0
7.5
5.0
2.5
0
时间（小时）

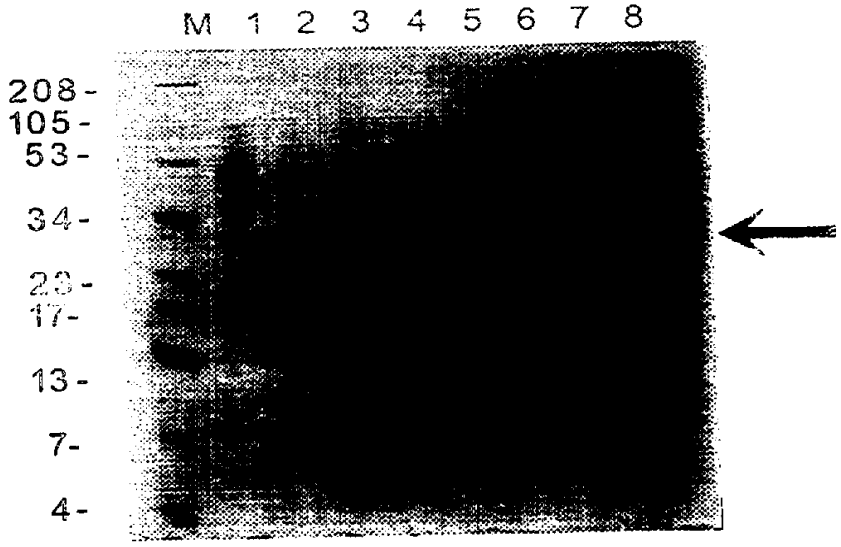
# 图 5

pBN115-glp 表达载体的质粒图谱。载体质粒的构建描述于说明书中。符号：ori，pBR322 复制起点；lacIq，lacI 阻遏基因；$Ap^R$，氨苄青霉素抗性基因；8cGLP，8 拷贝 GLP-1 基因；YX-15，YX-15 启动子。



# 图 6

比较不同温度下 pBN115-glp 和 pBN115-glp/tac 表达的 GLP-1。于 37℃(泳道 2 和 6)，27℃(泳道 3 和 7)或 21℃(泳道 4 和 8)，在缺乏(泳道 1 和 5)和存在(泳道 2-4 和 6-8)1mM IPTG 时，使含有 pBN115-glp(泳道 1-4)和 pBN115-glp/tac(泳道 5-8)的大肠杆菌 HMS174 生长 15 小时。然后通过离心和超声处理收获细胞。在 100mM Tris，100mM Tricine 和 0.1% SDS(pH8.3)中，16% SDS-聚丙烯酰胺凝胶上分析得自 0.1 $O.D._{600}$ 单位的细胞的总细胞裂解物。用考马斯蓝染色凝胶。箭头标出 GLP-1 多蛋白带。GLP-1 蛋白的相对量概述于表 1。泳道 M 表示以千道尔顿计的分子量标准。

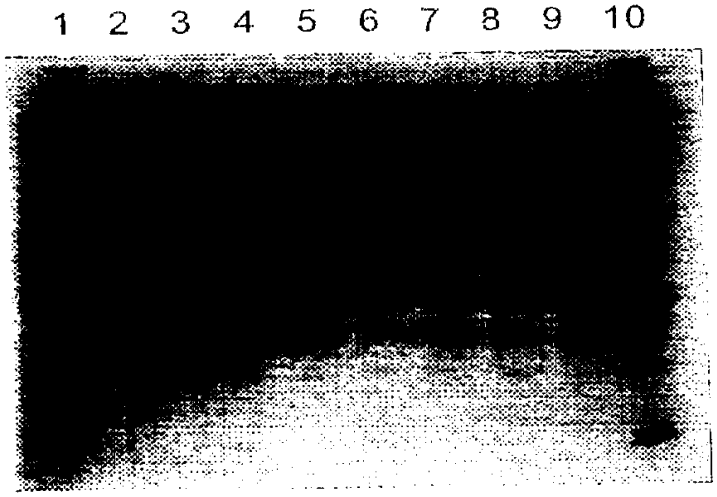


# 图 7

3 拷贝 GLP(7-36)蛋白的 Tac 和 YX16 表达

泳道 1: Novex MW 标记物
泳道 2: pGEXlnk2(3 拷贝 GLP)YX16 未诱导的总蛋白质，37℃
泳道 3: pGEXlnk2(3 拷贝 GLP)YX16 诱导的总蛋白质，37℃
泳道 4: pGEXlnk2(3 拷贝 GLP)YX16 诱导的总蛋白质，27℃
泳道 5: pGEXlnk2(3 拷贝 GLP)YX16 诱导的可溶性蛋白质，27℃
泳道 6: pBN95lnk2(3 拷贝 GLP)Tac@未诱导的总蛋白质，37℃
泳道 7: pBN95lnk2(3 拷贝 GLP)Tac@诱导的总蛋白质，37℃
泳道 8: pBN95lnk2(3 拷贝 GLP)Tac@诱导的总蛋白质，27℃
泳道 9: pBN95lnk2(3 拷贝 GLP)Tac@诱导的可溶性蛋白质，27℃

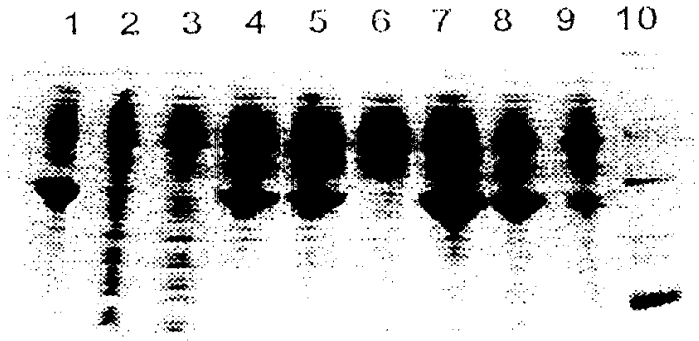
@该质粒的完整描述是 pBN95lnk2[GLP(7-36)AFAM]8HAE(tac)。泳道 2-5 中 3 拷贝 GLP pGEX 质粒中的融合蛋白相同。



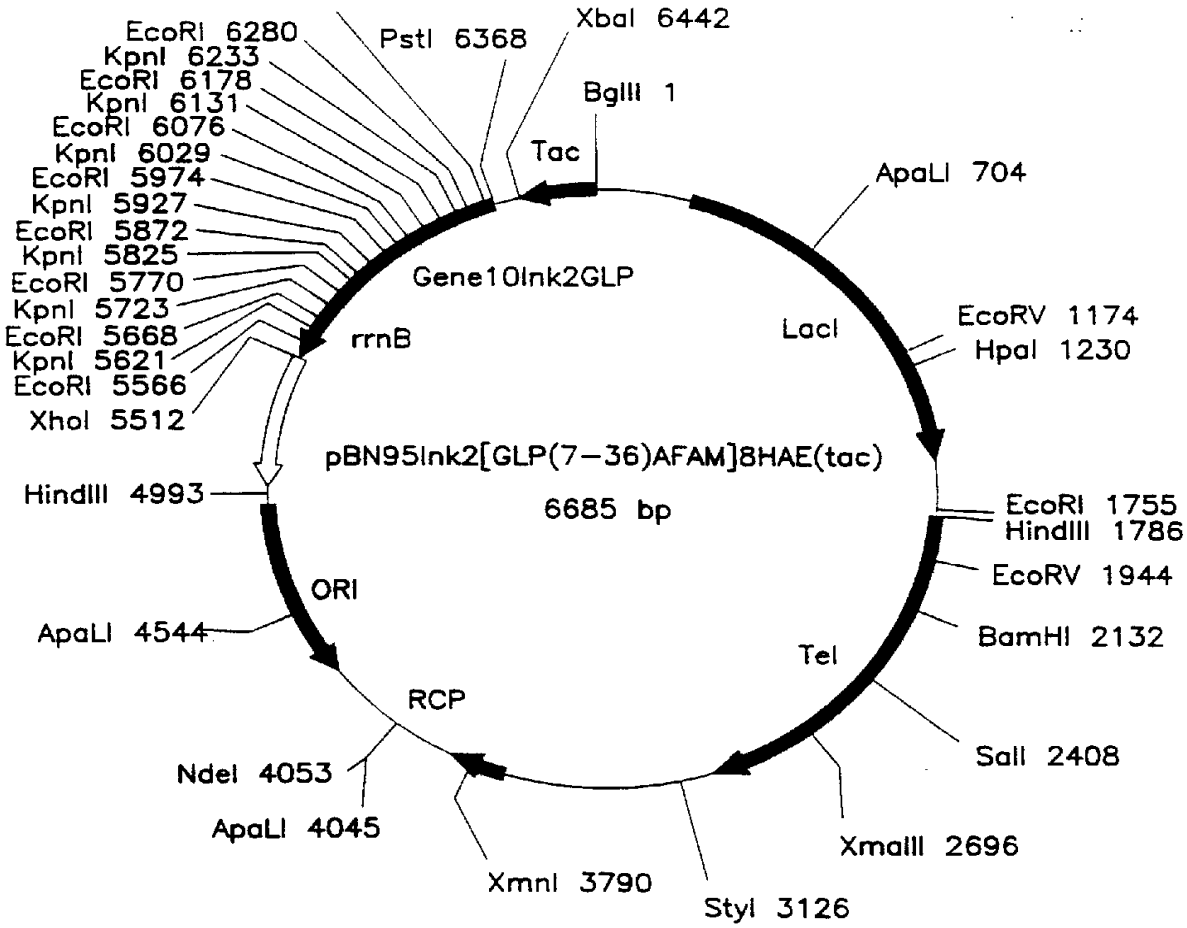
# 图 8

8 拷贝 GLP(7-36)蛋白的 Tac, YX15 和 YX16 表达

泳道 1:pBN95lnk2[GLP(7-36)AFAM]8HAE(tac)诱导的总蛋白质,37℃
泳道 2:pBN95lnk2[GLP(7-36)AFAM]8HAE(tac)诱导的总蛋白质,27℃
泳道 3:pBN95lnk2[GLP(7-36)AFAM]8HAE(tac)诱导的可溶性蛋白质,27℃
泳道 4:pBN53lnk2[GLP(7-36)AFAM]8HAE(YX16)诱导的总蛋白质,37℃
泳道 5:pBN53lnk2[GLP(7-36)AFAM]8HAE(YX16)诱导的总蛋白质,27℃
泳道 6:pBN53lnk2[GLP(7-36)AFAM]8HAE(YX16)诱导的可溶性蛋白质,27℃
泳道 7:pBN115-glp(8 拷贝 GLP)(YX15)诱导的总蛋白质, 37℃
泳道 8:pBN115-glp(8 拷贝 GLP)(YX15)诱导的总蛋白质, 27℃
泳道 9:pBN115-glp(8 拷贝 GLP)(YX15)诱导的可溶性蛋白质, 27℃
泳道 10:Novex MW 标记物



# 图 9




质粒名称： pBN95Ink2[GLP(7−36)AFAM]8HAE(tac)
质粒大小：6685 bp
构建者：
构建日期： 6/13/97
注释/参考文献：