[19] 中华人民共和国国家知识产权局

[12] 发明专利说明书

[21] ZL 专利号 97190017.5

[51] Int. Cl[7]
C12N 15/62
C12N 15/31 C12N 15/48
C07K 14/22 C07K 14/16
C07K 16/12 C07K 1/22
A61K 39/095 C12N 15/70
C12N 1/21
//(C12N1/21,C12R1:19)

[43] 授权公告日 2003 年 4 月 30 日

[11] 授权公告号 CN 1107115C

[22] 申请日 1997.1.17 [21] 申请号 97190017.5

[30] 优先权

[32] 1996. 1.17 [33] CU [31] 10/1996

[86] 国际申请 PCT/CU97/00001 1997.1.17

[87] 国际公布 WO97/26359 西 1997.7.24

[85] 进入国家阶段日期 1997.9.17

[71] 专利权人 基因工程和生物技术中心

地址 古巴哈瓦那

[72] 发明人 C·A·杜阿特·卡诺
G·E·奎伦·奈托
A·阿尔瓦里茨·阿克斯塔
E·L·卡皮奥·莫诺茨
D·奎恩塔纳·瓦茨奎兹
C·E·格梅茨·罗德里奎茨
R·C·西尔瓦·罗德里奎茨
C·纳扎贝尔·加尔韦茨
M·J·里尔·安古洛
A·M·马蒂恩·杜恩

[56] 参考文献

US-5286484-A 1994.02.15 A61K39/02

审查员 李 博

[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 陈文平

权利要求书 2 页 说明书 21 页 附图 12 页

[54] 发明名称 作为融合蛋白的异源抗原的表达系统

[57] 摘要

本发明涉及生物技术和遗传工程领域，更具体地说，本发明涉及使用重组 DNA 技术经与细菌肽融合在微生物宿主中表达异源蛋白质的方法。 本发明提供一种在大肠杆菌中以融合多肽表达外源蛋白的方法，以获得高纯度、商用量且适用于人类疫苗制剂的蛋白。 为此主要使用衍生自脑膜炎奈瑟氏球菌 B：4：P1.15 P64K 抗原头 47 个氨基酸的稳定化序列。 具体讲，使用了含所述序列、大肠杆菌色氨酸启动子控制下和噬菌体 T4 转录终止子的质粒，其中含有供同框克隆编码目的多肽之 DNA 片段的限制位点。 本发明方法可应用于药物工业、诊断系统的开发、疫苗制剂、和任何需要在大肠杆菌中获得高产量融合多肽形式的外源蛋白的情形。

ISSN1008-4274

1. 含有源于脑膜炎奈瑟氏球菌B:4:P1.15的P64K抗原N-末端头47个氨基酸的稳定子肽和与其融合的异源蛋白的融合蛋白。

2. 按照权利要求1的融合蛋白，其中所说的异源蛋白质是脑膜炎奈瑟氏球菌的外膜蛋白质。

3. 按照权利要求2的融合蛋白，其中所说的异源蛋白质是脑膜炎奈瑟氏球菌的Opc(5c)蛋白质或者是脑膜炎奈瑟氏球菌的*PorA*蛋白质。

4. 按照权利要求1的融合蛋白，其中所说的异源蛋白质是多表位多肽，其包括HIV-1 gp 120 蛋白质可变区3(V3)中心部分的几个拷贝。

5. 按照权利要求4的融合蛋白，其中所说的多表位多肽是多肽TAB4和/或TAB9。

6. 在大肠杆菌中产生作为融合多肽的异源蛋白质的方法，其中将来源于脑膜炎奈瑟氏球菌B:4:P1.15的P64K抗原N-末端头47个氨基酸的肽用作表达所说的异源蛋白质的稳定子，将所说稳定子的特异性单克隆抗体用于免疫检测融合到所说稳定子上的任何蛋白质。

7. 按照权利要求6的方法，其中所说的稳定子肽的氨基酸序列基本上相应于序列识别号6的序列：

10 20 30 40 47

MVDKRMALVE LKVPDIGGHE NVDIIAVEVN VGDTIAVDDT LITLDLE

8. 按照权利要求6的方法，其中用于纯化融合蛋白的单克隆抗体被命名为448/30/7。

9. 按照权利要求6的方法，其中所表达的异源蛋白质是可以在疫苗制剂中用作免疫原的任何蛋白质。

10. 按照权利要求6的方法，其中所说的表达的异源蛋白质是多表位多肽TAB4和/或TAB9，或脑膜炎奈瑟氏球菌的外膜蛋白Opc(5c)或PorA。

11. 单克隆抗体448/30/7，其对来源于脑膜炎奈瑟氏球菌B:4:P1.15的P64K抗原N-末端头47个氨基酸的稳定子肽是特异性的，并且用于免疫

检测和纯化融合到所说稳定子肽上的任何蛋白质。

12. 融合蛋白在大肠杆菌中的表达载体，该载体包含编码稳定子肽的序列，所说稳定子肽来源于脑膜炎奈瑟氏球菌 B:4:P1.15 的 P64K 抗原 N-末端头 47 个氨基酸，所说序列在大肠杆菌色氨酸启动子的控制下，其后是用于同框克隆编码兴趣肽之 DNA 片段的限制性位点 XbaI、EcoRV 和 BamHI，和作为转录终止信号的噬菌体 T4 终止子，以及作为选择性标记的氨苄青霉素抗性基因。

13. 按照权利要求 12 的表达载体，其用于同框克隆编码脑膜炎奈瑟氏球菌外膜蛋白的 DNA 片段。

14. 按照权利要求 13 的表达载体，其被命名为 pM-80 pM-82。

15. 按照权利要求 12 的表达载体，其用于同框中克隆编码多表位多肽的 DNA 片段，所说的的多表位多肽包括属于 HIV-1 gp 120 蛋白质可变区 3(V3)中心部分的拷贝。

16. 按照权利要求 15 的表达载体，其被命名为 pTAB4 和 pTAB9。

17. 一种大肠杆菌重组菌株，其是用按照权利要求 12 到 16 任一之表达载体转化任何大肠杆菌 K12 菌株产生的。

18. 权利要求 1 到 5 的融合蛋白在用于人或动物的疫苗制剂中的用途。

19. 包含权利要求 1 到 5 任一之融合蛋白和合适佐剂的疫苗制剂。

作为融合蛋白的异源抗原的表达系统

技术领域

本发明涉及生物技术和遗传工程领域，更具体地说，本发明涉及使用重组 DNA 技术经与细菌肽融合在微生物宿主中表达异源蛋白质的方法。

现有技术

用重组DNA技术在大肠杆菌中产生任何来源的蛋白质的有效性已有广泛说明。为此，尽管由于克隆和表达的各种基因通常需要单个对待，需要新的变体，但已开发了大量载体(Denhardt, D.T.和 Colasanti, 载体杂志（Vectors），Butterworths, Stoneham, M, pp. 179,1987 和 Lukacsovich, T.等, 生物技术杂志 (Journal of Biotechnology)，13, 243,1990)。

由于可以获得高的表达水平，胞内合成已用于在大肠杆菌中获得异源多肽(Goeddel, D.V, 酶学方法 (Methods Enzymol.)，185, 3-7, 1990)。然而，对宿主蛋白酶的敏感性和表达的蛋白质的毒性之类的因素可以明显降低所说的水平，而不依赖于高效调节序列的使用(Reads, C. A.和Saier, M. H.，细菌学杂志 (J. Bacteriol)，153, 685-692; Gwyn, G. W.，膜蛋白表达系统: 用户指南 (Membrane Protein Expression Systems: A User's Guide)，波特兰出版社, 伦敦, 英国, 29-82)。

在合适的载体中，与编码宿主细胞中稳定多肽的核酸序列同读框克隆编码兴趣蛋白质的核苷酸序列，导致在细胞质中表达被称为融合蛋白的杂合产物 (Marston, F.A.O.，生物化学杂志 (Biochem J.) 240, 11986)。这种多肽一般对宿主的蛋白酶解降解较不敏感，或者由于形成包含体毒性较低。这导致比不使用稳定子肽时所获多肽较高的表达水平(Itakura, K.等，科学（Science），198, 1056,1977)。此外，如果可以使用不同的用于重组产物随后复性的方法，这类表达会有利于简化初始纯化步骤 (Fischer, B.，Sumner, I.和Goodenough, P.，生物技术和生物工程 (Biotechnol. Bioeng.，41, 3-13)。

在大肠杆菌中许多重组蛋白质的表达期间，包含体是不溶性蛋白质聚集体，在

胞质溶胶中表现为电子致密体 (Rinas, Or.和Bailey, 应用微生物学和生物技术杂志 (Appl. Microbiol. Biotechnol.), 37, 609-614)。它们是部分折叠的多肽之间相互作用的结果，由于其中疏水残基对溶剂的暴露，其聚集体是热力学上有利的(Kiefhaber, T., Rudolph, R.等, 生物技术 (Biotechnology), 9, 825-829)。由于含丰富的二硫键-形成氨基酸(半胱氨酸)或β转角-形成氨基酸(脯氨酸)，许多真核蛋白质在细菌胞质溶胶中的缓慢折叠使得它们作为稳定化肽广泛使用。为这一目的，前者的例子使用具有抗体结合活性的多肽，它们来源于人奶脂肪球蛋白(HMFG)(按照国际专利申请PCT No. WO 9207939 A2 920514); 免疫球蛋白恒定区(如欧洲专利申请No. EP 0464533 A1 920108 中描述的); 人血管生成素(欧洲专利申请 No. EP 0423641 A2 910424)、生长激素(EP 0429586 A1 910605), 谷胱甘肽-S-转移酶(WO 8809372 A1 881201)和猪腺苷酸激酶(EP 0423641 A2 910424 和EP 0412526 A2 910213)。

然而，如果融合蛋白是疫苗候选者，则使用构成融合蛋白显著部分的稳定子多肽具有某些缺点，因为外源序列的存在可以改变B和T细胞表位的天然顺序 (Denton G., Hudecz, F., Kajtar, J.等, 肽研究 (Pepticle Research), 7, 258-264)或者它们被抗原呈递细胞的加工(Del Val, M., Schlicht, H., Ruppert, T.,等, 细胞 (cell), 66, 1145-1153), 甚至经特异表位抑制现象严重影响候选者的免疫原性(Etlinger, H., 现代免疫学 (Immunol. Today), 13, 52-55)。

由于以上提及的现象，一些研究组试图确定仍然使表达稳定化的小片段。例如，德国专利申请 35 41 856 A1(Hoechst AG)报道了使用稳定子肽以获得在大肠杆菌中合成的可溶形式的融合蛋白的可能性，所述肽由人类蛋白质白介素(IL-2)N-末端头 95 个氨基酸构成。类似地，在同一公司的欧洲专利申请 0416 673 A2 和 229 998 中，使用了与所说蛋白质的头58 个或38 个氨基酸一致的稳定子肽。在欧洲专利416 673 B1 中，也使用了 IL-2 的头58 个氨基酸，为了这一目的，在使用人血清白蛋白N-末端片段(欧洲专利申请 EP 0423641 A1 920212)、结缔组织的激活剂肽 III (WO 9013647 A1 901115)、人类激肽释放酶片段(EP 0381433 A1900808)的情况下，采用了类似的策略。这些发明给出了以前的问题的解决办法，但是所获得的融合多肽不能包含到用于人类的疫苗制剂中，因为其中存在与人类蛋白质同源的或相同的序列，有诱导自身免疫疾病的可能性。

也已成功地开发了供胞内表达的使用细菌来源(并且由此没有与人源抗原的交叉反应性)的稳定子多肽的替代方案。一种对这一目的最有用的蛋白质是大肠杆菌β-半乳糖苷酶(Itakura, K.等, 科学, 198, 1056 1977)或其部分(Hoechst AG 公司的德国专利申请 EP 0235754 A2 870909) 。这个系统的主要缺点是这种蛋白质大, 这使得所需肽仅代表总杂合蛋白的一小部分(Flowers, N.等, 应用微生物学和生物技术 25, 267 1986; G oeddel, D.V. 等, P.N.A.S. USA, 76, 106)。类似问题出现在破伤风类毒素 C 片段和假单胞菌外毒素的应用中(国际专利申请PCT WO 9403615 A1 940217和欧洲专利申请 EP 0369316 A2 900523)。非常合适的表达变体是使用具有大肠杆菌 tiorredoxin 的融合物 (PCT 专利申请 WO 9402502 A1 940203), 为了有利于纯化, 其利用了经渗透压从细胞释放的性质(Elyaagoubi, A., Kohiyama, M., Richarme, G., 细菌学杂志, 176, 7074-7078)。然而, 这一方式对包含体的获得是不成功的, 因为通过这一步骤, 包含体不能被游离出来。

许多这样的问题已由模拟融合蛋白的设计解决。在这些设计中, 稳定子肽与兴趣蛋白质靠允许两者独立折叠的间隔分离开, 其氨基酸序列使其对特异性内肽酶的进攻敏感。如果有识别所选稳定子的配体, 则经亲和色谱法纯化融合多肽和经用不同的蛋白酶处理从稳定子最终分离它是可能的(Cress, D., Shultz, J.和 Breitlow, S., Promega you Note, 42, 2-7)。另一个优点是在接下来的纯化的中间步骤中利用这一分子相互作用, 而不需要每种所表达蛋白质的抗体。这一点的一个众所周知的实例是利用组氨酸 (His) 与系统中某些金属(如镍(Ni)和锌(Zn))的亲和性, 按照 Upjoho 公司的 PCT 专利申请 WO 9115589 A1 911017 中的描述, 所述系统由具有 6 个串联 His 的稳定子和镍螯合物亲和基质组成。尽管如此, 这一种类的表达系统不是在所有情况下都起作用, 因为除了其它原因外, 兴趣蛋白质可以具有所选蛋白酶的限制性位点, 或折叠后有暴露于溶剂的间隔(Uhlen, M.和 Moks, T., 酶学方法. 185, 129-140; Cress, D., Shultz, J.和 Breitlow, S., Promega you Note, 42, 2-7), 以致干扰稳定子和亲和基质之间的结合(新英格兰生物实验室, The NEB Transcript, 3, 14), 或者只是其纯化需要影响它的生物活性的条件。由于这些理由, 需要不同的变体, 因为各种要表达的蛋白质代

表特定的情形。为了这一目的，已经开发出基于下列蛋白质的稳定子肽：大肠杆菌麦芽糖结合蛋白(MalE)，其对直链淀粉树脂具有亲和性(欧洲专利申请 EP 0426787 A1 910515)；氯霉素乙酰转移酶(欧洲专利申请 EP 0131363 A1 850116)或谷胱甘肽-S-转移酶(Amrad 公司的欧洲专利申请 EP 0293249 A1 88130)，其可用固定化底物获得；金黄色葡萄球菌蛋白质 A(按照专利申请 PCT WO 9109946 A1 910711)；或薛曼纳氏丙酸杆菌转羧基酶复合物的 12.5 kDa 亚单位，其在体内生物素化，并使得可以进行基于抗生物素蛋白的生物素亲和性的纯化 (Cress, D., Shultz, J.和 Breitlow, S., Promega Notes, 42, 2-7, 专利申请 EP 0472658 Al 920304 或 WO 9014431 Al 901129)。

欧洲专利申请 EP 0472658 Al 920304 或 WO 9014431 Al 901129 中描述的方法具有特别意义，该方法是由生物技术研究与发展公司与美国伊利诺州大学一同开发的。在这一申请中，描述了一种表达系统，其使用二氢硫辛酰胺乙酰转移酶硫辛酸结合域(EC 2.3.1.12)，也被称为大肠杆菌丙酮酸脱氢酶复合物 E2 亚单位。这一结构域通过添加硫辛酸分子到一个赖氨酸的氮原子上被体内翻译后修饰(Guest, J.R., Angler, J.S.和 Russell, G.C., Ann . N. Y. Acad. Sci., 573, 76-99)，其被用来纯化和鉴别融合蛋白，其中使用仅识别硫辛酸化结构域的抗体。

然而，这一方法具有许多弊端。首先，已知包含硫辛酸结合结构域的蛋白质的超量表达会超过细胞硫辛酸化容量，结果产生非硫辛酸化结构域(Thousands, J.S.和 Guest, J.R., 生物化学杂志， 245, 869-874; All, S.T.和 Guest , J.R. , 生物化学杂志 , 271, 139-145) 或辛酸化结构域(Ali, S.T., Moir, A.J., Ashton, P.R.等，分子微生物, 4, 943-950; Dardel, F., Packman, L.C.和 Perham, R.N., FEBS Lett. 295, 13-16)，这可以导致免疫亲和纯化降低产量。在第二方面，有一组推定由自身免疫引起的疾病，其共同特征是存在在这些结构域中特异性识别硫辛酸的抗体。它们主要是早期胆汁性肝硬变，以发炎和肝内胆汁导管进行性阻碍为特征的一种慢性病(Tuaillon, N., Andre, C., Briand, J.P.等，免疫学杂志, 148, 445-450)；和由卤烷(一种通过形成三氟乙酰赖氨酸衍生某些蛋白质的广泛使用的麻醉剂)引起的肝炎(Gut,

J., Christen, Or., Frey, N.等, 毒理学, 97, 199-224)。患这种疾病的患者的血清识别所说复合物，其分子结构与二氢硫辛酰胺乙酰转移酶中的硫辛酸相似(Gut, J., Christen, Or., Frey, N.等, 毒理学, 97, 199-224)。鉴于这一理由，如果包含硫辛酸的融合蛋白构成用于人的疫苗候选者，则最好避免在这样的肽中存在硫辛酸。

发明描述

本发明的目的是提供一种在大肠杆菌中高水平地表达作为融合多肽的异源蛋白质的方法，其基于使用来源于脑膜炎奈瑟氏球菌(*Neisseria meningitidis*) B:4:P1.15(欧洲专利申请 0 474 313 A2)的P64K抗原头47 个氨基酸的稳定子序列，所说稳定子赋予它们被作为包含体表达的能力。尽管存在与部分二氢硫辛酰胺乙酰转移酶硫辛酸结合域的同源性，所说的序列已经由基因工程操作消除了自身修饰的可能性，并且具有低免疫原性的优点。这一方法也包括使用特异性地识别被论及的稳定子的单克隆抗体，使得可以进行任何融合到所述稳定子上的蛋白质的免疫检测。

更具体地说，在本发明中，使用了作为表达载体的重组质粒，该质粒携带在大肠杆菌色氨酸启动子(ptrip)控制下的所述序列，其后是限制性位点 XbaI， EcoRV 和 BamHI 。这些使得可以在同读框中克隆编码兴趣多肽的DNA 片段。这一载体也包含噬菌体 T4 基因 32 的转录终止子和作为选择性标记的氨苄青霉素抗性基因。

这一方法也使得在用于人类的疫苗制剂中包含所获得的融合多肽成为可能；并且所采用的稳定子肽的性质使得可以产生针对与其结合的外源蛋白质或多表位肽的保护性免疫反应。

基因工程操作和使用与部分二氢硫辛酰胺乙酰转移酶硫辛酸结合结构域的同源稳定子肽，以便经重组 DNA 技术在大肠杆菌中产生融合蛋白构成了本发明的新颖性。更具体地说，为了前述的目的，使用来源于脑膜炎奈瑟氏球菌 B:4:P1.15 的P64K抗原头47 个氨基酸的稳定子肽以及特异性识别所述稳定子的单克隆抗体构成了本发明的新颖性。

实施例:

实施例 1：

脑膜炎奈瑟氏球菌 LpdA 抗原(P64K LpdA)是 594 个氨基酸的蛋白质，属于二氢硫辛酰胺脱氢酶(EC 1.8.1.4)家族，更具体地说，属于其中的一个新亚组，以在其 N-末端部分具有硫辛酸结合结构域(其类似于存在于二氢硫辛酰胺乙酰转移酶中的结构域)为特征(Kruger, N., Oppermann, F. B. , Lorenzl, H.和 Steinbüchel, A.，细菌学杂志, 176, 3614-3630, 1994; Hein, S. 和 Steinbüchel, A.,细菌学杂志, 176, 4394-4408, 1994)。LpdA 蛋白质已在大肠杆菌中克隆和超量表达，在其 N-末端具有 5 个附加的氨基酸(MLDKR)(欧洲专利申请 0474 313 A2；图 1)，尽管名称 LpdA 和 P64K 是等价的，但将利用 P64K 指重组蛋白质。

为了确定所说抗原的不同片段的免疫原性和分析利用较低免疫原性稳定子肽的可能性，通过评价多克隆抗-P64K 血清对合成肽的反应性定位存在于 P64K 中的 B 细胞表位。

为了这一目的，经在丁基-TSK 中的疏水性色谱和凝胶-过滤纯化 P64K 蛋白质(欧洲专利申请 0 474 313 A2)；经三氯乙酸(TCA)沉淀变性，用 NaOH 中和，并且经凝胶-过滤色谱在磷酸盐缓冲液中平衡。将这一制剂经皮下路线以 20μg 剂量辅以 2μg 氢氧化铝(0 天)免疫 30 只小鼠 Balb/C，7 和 21 天后，用相同的抗原加强免疫。在第一次抽取之后 28 天收集血清。合并所获得的血清，将形成的混合物等分并贮存在-20 ℃下。

此外，按照制造者提供的说明，以 96 孔平板方式使用用于固相合成的市售试剂盒(Multipin 肽合成系统，Chairon Mimotope Pty.有限公司，美国)合成了具有 20 个氨基酸(a.a.) 的 59 种肽(各覆盖完整的重组蛋白质序列并有 10 个 a.a.重叠)。随后，将这些从蛋白质的 N-末端编号。采用 1:2000 相同的稀度测定血清抗 P64K 针对这些肽的反应性，使用的免疫测定的方式与由前述市售试剂盒的制造者推荐的相同。

结果在图 2 中显示，其中给出了每种肽的吸光值。很明显，头 110 个氨基酸(由肽 1 到肽 11 代表)形成差的致免疫区段(尽管免疫原变性，其甚至能够暴露隐藏的表位)。这一区段基本上包括硫辛酸结合结构域和富含脯氨酸和丙氨酸的间隔(其将所述结构域连接到蛋白质的其余部分上)。这一结果说明，所述的稳定子肽(或者其衍生物片段)可以有效地用作稳定子肽，这是由

于对与其融合的多肽的免疫原性影响小。如果融合多肽构成疫苗候选者，这一优点尤其重要。

实施例 2：

为了在大肠杆菌中通过将不同的异源蛋白质融合进脑膜炎奈瑟氏球菌 B:4:P1.15 的 P64K 抗原的硫辛酸结合结构域来表达异源蛋白质，构建表达载体 pM-83，其中引入了编码来源于所说蛋白质的头 47 个氨基酸的稳定子肽的序列(序列识别号：1)。在大肠杆菌色氨酸启动子的控制下克隆该序列，引入的序列包括作为转录终止信号的噬菌体 T4 终止子和作为选择性标记的氨苄青霉素抗性基因。

为了获得 pM-83 表达载体，首先从质粒 pM-6 经聚合酶链反应(PCR)(Randall, K.等，科学，42394, 487-491, 1988)扩增稳定子肽，所述质粒携带编码 P64K 抗原的核苷酸序列(欧洲专利申请 0 474 313 A2，图 1)。为了这一目的，使用寡核苷酸引物 1573 和 1575，其在扩增的 DNA 片段中引入相应于所述 DNA 编码的稳定子的氨基和羧基末端的 NcoI 和 XbaI 限制性位点：

NcoI

1573： 5'TTCCATGGTAGATAAAAG 3' (序列识别号:2)

XbaI

1575： 5'TTTCTAGATCCAAAGTAA 3' (序列识别号：3)

由形成的稳定子编码的氨基酸序列示于图 3 中(序列识别号：6)。NcoI 限制性位点的引入使亮氨酸 2 改变成缬氨酸；引物 1517 消除序列 ETD(45-47 位)，在它的位置上引入序列 DLE。以这种方式，硫辛酸的结合赖氨酸(48 位)不形成稳定子的一部分。其邻近区(在这些结构域中其是高度保守的(Russell, G.C., Guest, J.R., 生物化学和生物物理报告, 1076, 225-232, 1991))被改变。所有这一切保证消除含有这些结构域的融合蛋白的翻译后被硫辛酸化的可能性，以及在用这些蛋白质免疫期间产生与存在于早期胆汁性肝硬变患者中的那些有类似特异性的自身抗体(Tuaillon, N., Andre, C., Briand, J.P.等，免疫学杂志,148, 445-450) 的可能性。

通过克隆这一片段(序列识别号：5) 于质粒 pILM-29(Guillén, G., Loyal,

M., Alvarez, A.等，生物技术学报，15, 97-106, 1995)中构建质粒 pM-83，所述片段已用 XbaI /NcoI 事先消化。pILM29 质粒包含与人 IL-2 头 58 个氨基酸一致的稳定子肽融合的脑膜炎奈瑟氏球菌 Opc(5c)基因，因而这种克隆过程是除去 IL-2 片段并使 Opc 融合到 P64K 蛋白质稳定子上(图 4)。用酶 XbaI 和 BamHI 从形成的质粒(命名为 pM-80)切除 Opc 基因，在其位置上克隆由寡核苷酸 1576 和 1577 杂交形成的接头，这一过程在稳定子片段的 3'末端引入限制性位点 XbaI, EcoRV 和 BamHI：

1576 5' CTAGATTTGATATCAG 3' (序列识别号: 7)

1577 3' TAAACTATAGTCCTAG 5' (序列识别号: 8)

这一质粒命名为 pM-83(图 4)。按照 Sanger, F.等(PNAS, USA, 74: 54631977)的方法经 DNA 测序证实了所有 DNA 片段和寡核苷酸的插入以及正确读框的保持。

实施例 3：

如果形成的融合蛋白是疫苗候选者，重要的是所说稳定子不含有与人类蛋白质高度同源的区。通过在核苷酸序列数据库 EMBL 数据库 v.38 (Curl, C.M., Fuchs, R., Higgins, D.G.等，Nucl. Acids Beast. 21, 2967-2971, 1993)和氨基酸序列的 SWISS-PROT 数据库 v.38 (Bairoch, A.和 Boeckmann, B., Nucl. Acids Beast. 21, 3093-3096, 1993)中的同源性检索进行了 pM-83 稳定子肽(实施例 2)与人类蛋白质的类似性研究。使用的均是 1994 年 3 月版本的两个 BLAST 程序 (Altschul, S.F., Gish, W., Miller, W., Myers, And.W.和 Lipman, D.J., 分子生物学杂志, 215:403-410, 1990): BLASTP，其将一个氨基酸序列与蛋白质序列库(base)比较(在这种情况下是 SWISS-PROT)；和 TBLASTN，其将氨基酸序列与所有两个方向上的所有核苷酸序列库中读框中的翻译产物比较(在这种情况下是 EMBL 数据库)；在两种情况下均使用了辅助基质 PAM 120 [Dayhoff, M. O., Schwartz, R.M.和 Orcutt, B.B., in: Dayhoff, M.Or.(of.), 蛋白质序列和结构图谱, 5, 增补本 3, 345-352, Natn. Biomed. Beast. Found., 华盛顿, 1978]。

可以从图 6 和图 5 看出结果，其中显示了 BLASTP 和 TBLASTN 各自的结果单(为了更好地理解，省去了原核生物或低等真核生物的同源序列。

很明显，没有人类蛋白质或来源于其它哺乳动物的蛋白质表现出与 P64K 来源的稳定子的明显的类似性；因为由两种算法检测的同源性(在人和大鼠丙酮酸激酶中；和人和犬粘蛋白的 C-末端中)存在最高的偶然出现的可能性(在维涅兰德固氮菌二氢硫辛酰胺乙酰转移酶的情况下作为比较点的相同的可能性是 3.7×10^{-5})。

从所有以上所述可以得出结论，在疫苗候选者中使用所说的稳定子是绝对有把握的。

实施例 4：

通过比较融合到人白介素-2(IL-2)头 22 个或 58 个氨基酸的几种蛋白质(一种常用于该目的的融合肽)或按照实施例 2 所述修饰的融合到 P64K 抗原头 47 个氨基酸上的几种蛋白质的表达，评价了 pM-83 中的稳定子允许高水平和以包含体形式胞内合成的能力。

为了这一目的，将编码脑膜炎奈瑟氏球菌 B:4:P1.15 PorA 和 Opc 外膜蛋白的基因克隆进载体 pFP15 (hIL2-58 ； 欧洲专利 416 673 B1)或 pM-83(P64K-47)； 并将编码多表位肽(MEP)(包含人类免疫缺损病毒几种分离物的致免疫区)的基因克隆进载体 pISL31 (hIL2-22, Castellanos-Sierra, L.R., Hardy, E., Ubieta, R.,等， 提交的论文) 或 pM-83 。所形成的表达质粒是： pILM-28 (IL2-58 PorA; Guillén, G., Alvarez, A., Lion, L.,等， 494-498 in: Conde-González, C. J., Morse, S., Rice, P.等(编者)., 奈瑟氏球菌科的病理学和免疫学, Institute de Salud Pública Nacional, Cuernavaca, Mexico, 1994), pM-82 (P64K-47 PorA; Niebla, O., Alvarez, A., González, S. 等， 85-86 in: Evans, J.S., Yost, S.和 Maiden, M.C.J.等(编者), 奈瑟氏球菌属 94: 第 IX 次国际致病奈瑟氏球菌属会议论文集, Winchester, 英国, 1994), pILM-29 (IL2-58 Opc; Guillén, G., Leal, M., Alvarez, A.等， 生物技术学报, 15, 97-106, 1995), pM-80 (实施例 2, 图 4), pTAB4 (IL2-22 + MEP) 和 pTAB9 (P64K-47 MEP)。

TAB 4 和 TAB 9 蛋白质是多表位多肽(MEP)，其包括 HIV-1 gp120 蛋白质可变区 3(V3)中心部分的几个拷贝。为了构建这些 MEP，选择了下列分离物的 V3 区的 15 个中心氨基酸：

LR150: SRGIRIGPGRAILAT (序列识别号: 9)

JY1: RQSTPIGLGQALYTT (序列识别号: 10)

RF: RKSITKGPGRVIYAT (序列识别号: 11)

MN: RKRIHIGPGRAFYTT (序列识别号: 12)

BRVA: RKRITMGPGRVYYTT (序列识别号: 13)

IIIB: SIRIQRGPGRAFVTI (序列识别号: 14)

这些区由序列为 AGGGA(序列识别号: 17)的五个氨基酸的间隔肽结合。

为了达到这一目的，用化学合成获得编码由间隔肽结合的 V3 表位的 DNA 序列(序列识别号: 21)，并且在色氨酸启动子控制下克隆，融合进人 IL-2 的头 22 个氨基酸(图 7)。通过用酶 ScaI 和 HindIII 消化从形成的质粒(命名为 pTAB3)切除包含 MEP 基因、色氨酸启动子以及 T4 终止子的片段，并克隆进 pUC19 (Yanisch-Perron, C.等, 1985, 基因 33, 103-119)以获得 pTAB4(图 7)。最后，构建 pTAB9，通过用酶 NcoI 和 XbaI 消化消除编码来源于人 IL-2 的稳定子序列，在它们的位置上克隆由聚合酶链反应(PCR)获得的编码 P64K 抗原的头 47 个氨基酸的片段(如实施例 2 所述)。形成的序列如图 8 所示，其组成如图 9 所示。

用于所有这些质粒的宿主大肠杆菌是 W3110(Hill, C.W.,和 Hamish, B,W. 美国科学院学报, 78, 7069, 1981; Jensen, K. F., 细菌学杂志, 175, 3401-3407, 1993)(用于 pILM-28, pILM-29 , pM-80 和 pM-82); 和 W3110 trpA905(用于 pTAB4 或 pTAB9)。所有这些均获得了表达: 将培养物接种到 5ml 具有氨苄青霉素(Ap)50μg/ml 和色氨酸(W)100μg/ml 的 LB 培养基(Sambrook, J., Fritsch, Y. F. 和 Maniatis, T.,分子克隆: 实验室手册, 冷泉港实验室出版社, 1989, 纽约, 美国)上。于 37 ℃生长 12 小时。用所述培养物接种 50ml LB-Ap 培养基 (pTAB4 和 pTAB9)或由以下物质组成的限定培养基: M9 盐 (Miller, J.H., 分子遗传学实验, 冷泉港实验室出版社, 1972, 纽约, 美国), 葡萄糖 1%, 酪蛋白水解产物 1%, $CaCl_2$ 0.1 mM, $MgCl_2$ 1mM 和 Ap 50 ug/mL(pILM-28, pILM-29, pM-80, pM-82), 于 37 ℃和 250 r.p.m 下生长 12 小时。此后，取得总蛋白样品，经变性聚丙烯酰

胺凝胶电泳 (SDS-PAGE, Laemmli, O. K., 自然, 277, 680, 1970s)和用考马斯亮蓝 R-250 染色分析。在激光 Bromma-LKB 密度计中分析表达百分比。通过用溶菌酶和超声波结合处理裂解细胞确定它们的细胞定位。然后经离心将不溶性蛋白质和可溶性蛋白质分离。不溶性蛋白质用作估价其作为包含体的表达水平，因为在这种情况下，它们能表现出所说行为(与膜或肽聚糖结合)的其它条件是靠不住的。

总的结果可以从图 10A 中看出。在所有情况下，在来源于 P64K 的稳定子下的表达与在融合进 IL-2 肽时获得的表达是可比较的：这是就异源蛋白质/总细胞蛋白质之比而言(参见有关 MEP 的图 10B)，其证实了 pM-83 用作表达融合肽的载体的能力。值得注意的是如果它们不是融合的，则这些多肽难于在大肠杆菌中表达，这是因为它们的小的尺寸和对宿主蛋白酶的敏感性(如 MEP)或者蛋白质 PorA 和一般细菌 porins 的毒性(Carbonetti, N.H.和 Sparling, P.F.; 美国科学院学报 U.S.A., 84, 9084-9088)。在所有的情况下，所述产物以包含体获得，如由 pTAB9 所例证的(图 10C)。

总之，得出以下结论是可能的，与其它系统(欧洲专利申请 0 416 673 A2 和 229 998, Hoechst AG; 欧洲专利申请 0 416 673 B1; Castellanos-Sierra, L.R. , Hardy , E. , Ubieta, R. ,等，提交的原稿)比较，使用来源于脑膜炎奈瑟氏球菌 P64K 抗原头 47 个氨基酸的稳定子 (P64K-47)导致有效地以包含体表达异源蛋白质，并且由于不存在与人源抗原明显的同源性，具有产物可直接使用的额外的优点。

实施例 5：

特异性识别稳定子之配体(例如抗体，酶促辅因子等)的可获得性是任何重组蛋白质的表达中的一个所希望的特征。这是因为如果所述配体固定化在色谱树脂上，则以上特征使得可以例如设计亲和纯化的有效方案；并且在抗体的情况下甚至随后通过免疫技术纯化的中间步骤不依赖于所表达异源蛋白质的性质。

为了获得针对这一稳定子的单克隆抗体(Mab)，用蛋白质 TAB13 (序列识别号：20) 免疫小鼠达到了这一目标。 TAB13 是一种来源于 TAB9 的 MEP，后者不同于前者之处在于存在两个附加的 V3 共有区(图 9B):

-C6： TSITIGPGQVFYRTG (序列识别号: 15)

-C8： RQRTSIGQGQALYTT (序列识别号: 16)

用类似于有关 TAB4 和 TAB9 所描述的方式表达(实施例 4)和纯化(实施例 6)这一 MEP。

然后，经皮下路线用 3 剂 20μg TAB13 辅以氢氧化铝以 15 天间隔免疫小鼠 Balb/c。在最后一次用药后 20 天，经腹膜内路线用在磷酸盐缓冲液中的 20μg 相同的抗原加强免疫所说动物。将脾细胞与骨髓瘤 X63 Ag8 653 融合，分离形成的杂交瘤，按照已经建立的方法(Gavilondo, J. V. (编者)，单克隆抗体：理论和实践, Elfos Scientiae, 1995, 哈瓦那，古巴)检查。

经 ELISA(用 MEP TAB13， P64K 蛋白质或代表 TAB13 中存在的不同 V3 区的合成肽涂布平板)，评价由分离的杂交瘤分泌的抗体的反应性。在获得的总共 18 个阳性克隆中，其中一个(命名为 448/30/7)识别 TAB13 及 64K，但不识别来源于 gp120 的肽。

采用不同的样品(融合到来源于 P64K 的稳定子上的异源蛋白质(P64K-47)或相同的融合蛋白质或者融合到 IL-2 的头 58 个氨基酸上的蛋白质(IL2-58))经 Western 印迹法测定这一 MAb 对 pM-83 稳定子肽的特异性和其用于包含它的蛋白质的免疫检测的可能性。为此，用下列质粒转化(Sambrook, J., Fritsch, Y. F. 和 Maniatis, T.,分子克隆：实验室手册，冷泉港实验室出版社，1989，纽约，美国)大肠杆菌菌株 MM294： pILM-28(IL2-58+*porA*)、pM-82(P64K-47+*porA*)、pTAB13(P64K-47 + MEP)、pM-6(P64K)和 pFP15 (IL-2)。也使用表达质粒 pM-134，该质粒含有 P64K 的头 120 个氨基酸，其包括在如原有质粒中的相同的调节信号控制下的硫辛酸的结合结构域。用引物 1573(序列识别号: 2)和 2192(序列识别号: 4)经 PCR 扩增这一区段，用酶 NcoI 和 BamHI 消化，在同样消化过的质粒 pFP15(参见实施例 4)中克隆。在有关 pTAB4 和 pTAB9 的实施例 4 中限定的生长条件下获得这些转化体的表达。

图 11 显示了获得的结果，可以看出由于与质粒 pM-6， pM-82， pTAB13 和 pM134 样品的反应性(尽管所有这些蛋白质抗原性不同)，Mab 448/30/7 识别在稳定子 P64K-47 内的很可能为线性的表位，这一实验表明

这一反应性绝不是由于融合到稳定子上的蛋白质(例如质粒 pILM-28 和 pM-82, 在不同的稳定子下两者都携带基因 *porA*)引起的。这证明了这一 Mab 识别的特异性。

总之，由稳定子 P64K-47 形成的表达系统，含有它的质粒和 Mab 448/30/7 ，使得可以以包含体形式有效合成大量不同蛋白质，它们的检测不需要预先获得针对所表达的各种肽的免疫探针。

实施例 6：

在选择疫苗候选者表达系统时，对融合到稳定子上的多肽的免疫应答无有害影响是要考虑的一个重要因素。基于稳定子 P64K-47 表达系统的一个明显优点是其保证以上所述因素的降低的免疫原性(实施例 1)。然而，仍通过比较针对存在于 MEP TAB4(IL2-22)和 TAB9(P64K-47)中的 V3 区不同肽的抗体反应定量评价了在稳定子 P64K-47 对融合蛋白免疫反应的影响。

为了表达和纯化 TAB 4 和 TAB9 ，如实施例 4 所述获得菌株 W3110 trpA905 + pTAB4 和 W3110 trpA905+ pTAB9 的生物量。在苯甲基磺酰氟(PMSF)和非离子洗涤剂 Tritón-X-100 存在下用溶菌酶和超声波共同处理裂解生物量。经分级离心获得包含体，用阳离子移变剂（ caotrophic agent ）和洗涤剂(TAB4: 1. 尿素 4M Tritón-X-100 1%， 2. 尿素 8M 。TAB9: 1. 尿素 8M Tritón-X-100 1%， 2. 盐酸胍 6M)对包含体的两个连续的洗涤循环部分纯化并溶解 MEP 。通过在高效液相色谱(HPLC)， C4 VYDAC 柱中通过 20-80%乙腈梯度洗脱最终纯化所获得的上清液，纯度达约 90%。

以每毫克蛋白质 60 毫克佐剂的比例将纯化的重组蛋白质辅以氢氧化铝凝胶。将该制剂用于经皮下路线以 200μg/剂免疫 5 组兔。使用充分涂布有用于免疫的 MEP 或涂布有相应于存在于其上的各 V3 区的肽的 96 孔聚苯乙烯平板(高结合， Costar, USA)经 ELISA 评价免疫反应。用高于原血清混合物两倍的吸收值的各血清的最大稀释度计算滴度。所有血清均以双份分析。

所获得的值(图 12)表明，在不同的 IL2-22+MEP(TAB4)和 P64K-47+MEP(TAB9)之间针对 V3 区的滴度是类似的。尽管对 TAB9 的肽识别频率

稍高些，但这种差别在统计学上是不显著的($p<0.05$)。总之，异源蛋白质的免疫原性受稳定子 P64K-47 影响很小，并且可与现行使用的其它表达系统相比。

附图描述：

图 1： 编码 P64K 的基因 *IpdA* 基因的核苷酸序列。斜体显示插入质粒 pM-6 的 *IpdA* 原本缺乏的序列(欧洲专利申请 0474 313 A2)。

图 2： 抗 P64K 肽小鼠多克隆血清的反应性。选择被认为阳性的最小值为 0.4 光密度单位。

图 3： 稳定子的氨基酸序列，从质粒 pM-6 经 PCR 扩增的 DNA 序列推定。下划线序列相应于寡核苷酸引物。

图 4： 构建质粒 pM-83 的策略。

图 5：采用 BLASTP 程序进行的稳定子序列(“查找”)和 SWISS-PROT 库中存在的序列(“目标”)之间的同源性检索结果。仅显示了人蛋白质和哺乳动物蛋白质相应的结果。P(N)代表在由随机序列构成的库内发现 N 等同排列的几率。同源性的显著性随着 P(N)值减小。相同的残基用它的单字母代码表示。保守取代用“+”表示，没有显示差异。

图 6：采用 TBLASTN 程序进行的稳定子序列(“查找”)和 EMBL 数据库序列的所有可能的翻译产物(“目标”)之间的同源性检索结果。仅显示了人蛋白质和哺乳动物蛋白质相应的结果。P(N)代表在由随机序列构成的库内发现 N 等同排列的几率。同源性的显著性随着 P(N)值减小。相同的残基用它的单字母代码表示。保守取代用“+”表示，没有显示差异。

图 7： 构建质粒 pTAB4 和 pTAB9 的策略。

图 8： MEP TAB9 的核苷酸和氨基酸序列。

图 9： A： MEP TAB4 和 TAB9 的总的结构。B： MEP TAB13 的总的结构。

图 10：基因 *porA* 、Opc 和 MEP 在来源于人 IL-2 或来源于 P64K 抗原头 47 个氨基酸的稳定子下表达的比较。

A：比较表，hIL2-58 指人类 IL-2 头 58 个氨基酸，hIL2-22 指头 22 个氨基酸，P64K-4 指来源于 P64K 抗原头 47 个氨基酸的稳定子。

B：在质粒 TAB 4 和 TAB 9 中 MEP 表达的 SDS-PAGE 比较分析。A 道：分子量标记；B 道：菌株 W3110 *trpA905* 的总蛋白质；C 道：W3110 *trpA905* + pTAB4 的总蛋白质；D 道：纯化的 TAB4；E 道：W3110 *trpA905*+ pTAB9 的总蛋白质；F：纯化的 TAB9。

C：以包含体表达的 TAB9。A：样品的可溶性蛋白质。B：不溶性蛋白质或膜蛋白质。

图 11：采用 Mab448/30/7 和大肠杆菌 MM294 总蛋白质样品进行的 Western 印迹分析，所述大肠杆菌已由以下质粒转化：1: 阴性对照，2: pM-6 (P64K), 3: pM-82 (P64K-47 + *porA*), 4: pTAB13 (P64K-47 + MEP), 5: pFP15 (IL-2), 6: pM-134 (P64K-120) , 7: pILM-28 (IL2-58 + *porA*)。左边列出了分子量标志。

图 12：用 TAB4 和 TAB9 免疫的兔 ELISA 滴度值倒数。MG：抗 V3 滴度倒数的几何平均值；R：与 V3 肽的百分反应性。

序 列 表

序列识别号: 1

序列类型: 氨基酸

长度: 47 个氨基酸

分子类型: 蛋白质片段

性质: 脑膜炎奈瑟氏球菌重组蛋白质 P64k 的头 47 个氨基酸

MLDKRMALVELKVPDIGGHENVDIIAVEVNVGDTIAVDDTLITLET
D 44

序列识别号: 2

序列类型: 核苷酸

长度: 29 个碱基

分子类型: 合成寡核苷酸

性质: 脑膜炎奈瑟氏球菌重组蛋白质 P64k 的头 44 个氨基酸的 PCR 扩增的 5' 引物 No.1573

TTCCATGGTAGATAAAAGAATGGCTTTAG 29

序列识别号: 3

序列类型: 核苷酸

长度: 29 个碱基

分子类型: 合成寡核苷酸

性质: 脑膜炎奈瑟氏球菌重组蛋白质 P64k 的头 44 个氨基酸的 PCR 扩增的 3'引物 No.1575

TTTCTAGATCCAAAGTAATC AG GGTATCG 29

序列识别号: 4

序列类型: 核苷酸

长度: 26 个碱基

分子类型： 合成寡核苷酸

性质：脑膜炎奈瑟氏球菌重组蛋白质 P64k 的头 120 个氨基酸的 PCR 扩增的 3' 引物 No.2192

GGCGGTTCTGCCGATTAAGG ATCCGA 26

序列识别号：5

序列类型： 核苷酸

长度： 146 个碱基对

分子类型： PCR 扩增片段

性质： 脑膜炎奈瑟氏球菌重组蛋白质 P64k 的头 47 个氨基酸的衍生片段。限制性位点 NcoI(3-8 位)和 XbaI(139-144 位)由 PCR 引入，其引起这一片段核苷酸序列上的变化

TTCCATGGTAGATAAAAGAATGGCTTTAGTTGAATTGAAAGTGCCCGACATTGGCGGACA 60

CGAAAATGTAGATATTATCGCGGTTGAAGTAAACGTGGGCGACACTATTGCTGTGGACGA 120

TACCCTGATTACTTTGGATCTAGAAA 146

序列识别号： 6

序列类型： 氨基酸

长度： 47 个氨基酸

分子类型： 来源于脑膜炎奈瑟氏球菌重组蛋白质 P64k 的头 47 个氨基酸的稳定子片段

性质： 与 P64k 相比，该片段具有下列变化

L2→V2 ； E45→D45 ； T46→L46 ； D47→E47

MVDKRMALVELKVPDIGGHENVDIIAVEVNVGDTIAVDDTLITLDL E 47

序列识别号：7

序列类型： 核苷酸

长度： 16 个碱基

分子类型：合成寡核苷酸

CTAGATTTGATATCAG 16

序列识别号：8

序列类型： 核苷酸

长度： 16 个碱基

分子类型： 合成寡核苷酸

GATCCTGATATCAAAT 16

序列识别号：9

序列类型：氨基酸

长度： 15 个氨基酸

分子类型： HIV-1 LR150 分离物 gp120 蛋白质的 V3 环的中心区

SRGIRIGPGRAILAT 15

序列识别号：10

序列类型：氨基酸

长度： 15 个氨基酸

分子类型： HIV-1 JY1 分离物 gp120 蛋白质的 V3 环的中心区

RQSTPIGLGQALYTT 15

序列识别号：11

序列类型：氨基酸

长度： 15 个氨基酸

分子类型： HIV-1 RF 分离物 gp120 蛋白质的 V3 环的中心区

RKSITKGPGRVIYAT 15

序列识别号：12

序列类型：氨基酸

长度： 15个氨基酸

分子类型： HIV-1 MN分离物 gp120蛋白质的V3环的中心区

RKRIHIGPGRAFYTT 15

序列识别号：13

序列类型：氨基酸

长度： 15个氨基酸

分子类型： HIV-1 BRVA分离物 gp120蛋白质的V3环的中心区

RKRITMGPGRVYYTT 15

序列识别号：14

序列类型：氨基酸

长度： 15个氨基酸

分子类型： HIV-1 IIIB分离物 gp120蛋白质的V3环的中心区

SIRIQRGPGRAFVTI 15

序列识别号：15

序列类型：氨基酸

长度： 15个氨基酸

分子类型： MEP TAB13内7位上，不同HIV-1分离物gp120 蛋白质V3 环中心区的共有序列

TSITIGPGQVFYRTG 15

序列识别号：16

序列类型：氨基酸

长度： 15个氨基酸

分子类型： MEP TAB13内8位上，不同HIV-1分离物gp120 蛋白质V3 环中心区的共有序列

RQRTSIGQGQALYTT 15

序列识别号: 17

序列类型: 氨基酸

长度: 5个氨基酸

分子类型: 在MEP TAB3、TAB4、TAB5和TAB13中分开V3表位的柔性接头

AGGGA 5

序列识别号: 18

序列类型: 氨基酸

长度: 141个氨基酸

分子类型: 多表位多肽(MEP)TAB4。

MAPTSSSTAQTQLQLEHLLLDLQIFLSRGIRIGPGRAILATAGGGARQSTPIGLGGALYT 60

TAGGGARKSITKGPGRVIYATAGGGARKRIHIGPGRAFYTTAGGGARKRITMGPGRVYYT 120

TAGGGASIRIQRGPGRAFVTI 141

序列识别号: 19

序列类型: 氨基酸

长度: 162个氨基酸

分子类型: 多表位多肽(MEP)TAB9

MVDKRMALVELKVPDIGGHENVDIIAVEVNVGDTIAVDDTLITLDLDSRGIRIGPGRAIL 60

ATAGGGARQSTPIGLGGALYTTAGGGARKSITKGPGRVIYATAGGGARKRIHIGPGRAFY 120

TTAGGGARKRITMGPGRVYYTTAGGGASIRIQRGPGRAFVTI 162

序列识别号: 20

序列类型: 氨基酸

长度： 202 个氨基酸

分子类型： 多表位多肽(MEP)TAB13

MVDKRMALVELKVPDIGGHENVDIIAVEVNVGDTIAVDDTLITLDLDSRGIRIGPGRAIL 60

ATAGGGAROSTPIGLGOALYTTAGGGARKSITKGPGRVIYATAGGGARKRIHIGPGRAFY 120

TTAGGGARKRITMGPGRVYYTTAGGGARORTSIGOGOALYTTAGGGATSITIGPGOVFYR 180

TGAGGGASIRIQRGPGRAFVTI 202

序列识别号： 21

序列类型： 核苷酸

长度： 368 个碱基对

分子类型： 合成寡核苷酸

性质：编码在 MEP TAB9 中由 AGGGA 间隔的 V3 表位的核苷酸片段。引入了限制性位点 XbaI(1 至 6 位)和 BamHI(363 至 368 位)

TCTAGACTCGAGAGGCATTCGTATCGGCCCAGGTCGCGCAATTTTAGCAACAGCTGGCGG 60

TGGCGCACGTCAATCTACCCCTATTGGTTTAGGTCAGGCTCTGTATACGACTGCCGGCGG 120

TGGTGCGCGCAAAAGTATCACCAAGGGTCCAGGCCGCGTCATTTACGCCACCGCGGGCGG 180

CGGTGCCCGTAAGCGTATCCACATTGGCCCAGGCCGTGCATTCTATACTACAGCAGGTGG 240

TGGCGCACGTAAACGCATCACTATGGGTCCTGGTCGCGTCTATTACACGACCGCTGGCGG 300

CGGTGCTAGCATTCGCATCCAACGCGGCCCTGGTCGTGCATTTGTGACCATATGATAACG 360

CGGGATCC 368

图 1

```
        10         20         30         40         50         60         70
ATGCTAGATA AAAGATGGC TTTAGTTGAA TTGAAAGTGC CCGACATTGG CGGACACGAA AATGTAGATA
        80         90        100        110        120        130        140
TTATCGCGGT TGAAGTAAAC GTGGGCGACA CTATTGCTGT GGACGATACC CTGATTACTT TGGAAACCGA
       150        160        170        180        190        200        210
TAAAGCGACT ATGGACGTAC CTGCTGAAGT TGCAGGCGTA GTCAAAGAAG TTAAAGTTAA AGTCGGCGAC
       220        230        240        250        260        270        280
AAAATCTCTG AAGGTGGTTT GATTGTCGTC GTTGAAGCTG AAGGCACGGC AGCCGCTCCT AAAGCCGAAG
       290        300        310        320        330        340        350
CGGCTGCCGC CCCGGCGCAA GAAGCCCCTA AAGCTGCCGC TCCTGCTCCG CAAGCCGCGC AATTCGGCGG
       360        370        380        390        400        410        420
TTCTGCCGAT GCCGAGTACG ACGTGGTCGT ATTGGGTGGC GGTCCCGGCG GTTACTCCGC TGCATTTGCC
       430        440        450        460        470        480        490
GCTGCCGATG AAGGCTTGAA AGTCGCCATC GTCGAACGTT ACAAAACTTT GGGCGGCGTT TGCCTGAACG
       500        510        520        530        540        550        560
TCGGCTGTAT CCCTTCCAAA GCCTTGTTGC ACAATGCCGC CGTTATCGAC GAAGTGCGCC ACTTGGCTGC
       570        580        590        600        610        620        630
CAACGGTATC AAATACCCCG AGCCGGAACT CGACATCGAT ATGCTTCGCG CCTACAAAGA CGGCGTAGTT
       640        650        660        670        680        690        700
TCCCGCCTCA CGGGCGGTTT GGCAGGTATG GCGAAAAGCC GTAAAGTGGA CGTTATCCAA GGCGACGGGC
       710        720        730        740        750        760        770
AATTCTTAGA TCCGCACCAC TTGGAAGTGT CGCTGACTGC CGGCGACGCG TACGAACAGG CAGCCCCTAC
       780        790        800        810        820        830        840
CGGCGAGAAA AAAATCGTTG CCTTCAAAAA CTGTATCATT GCAGCAGGCA GCCGCGTAAC CAAACTGCCT
       850        860        870        880        890        900        910
TTCATTCCTG AAGATCCGCA CATCATCGAT TCCAGCGGCG CATTGGCTCT GAAAGAAGTA CCGGGCAAAC
       920        930        940        950        960        970        980
TGCTGATTAT CGGCGGCGGC ATTATCAGCC TCGAGATGGG TACGGTTTAC AGCACGCTGG GTTCGCGTTT
       990       1000       1010       1020       1030       1040       1050
GGATGTGGTT GAAATGATGG ACGGCCTGAT GCAAGGCGCA GACCGCGATT TGGTAAAAGT ATGGCAAAAA
      1060       1070       1080       1090       1100       1110       1120
CAAAACGAAT ACCGTTTTGA CAACATTATG GTCAACACCA AAACCGTTGC AGTTGAGCCG AAAGAAGACG
      1130       1140       1150       1160       1170       1180       1190
GCGTTTACGT TACCTTTGAA GGCGCGAACG CGCCTAAAGA GCCGCAACGC TACGATGCCG TATTGGTTGC
      1200       1210       1220       1230       1240       1250       1260
CGCCGGCCGC GCGCCCAACG GCAAACTCAT CAGCGCGGAA AAAGCAGGCG TTGCCGTAAC CGATCGCGGC
      1270       1280       1290       1300       1310       1320       1330
TTCATCGAAG TGGACAAACA AATGCGTACC AATGTGCCGC ACATCTACGC CATCGGCGAC ATCGTCGGTC
      1340       1350       1360       1370       1380       1390       1400
AGCCGATGTT GGCGCACAAA GCCGTTCACG AAGGCCACGT TGCCGCCGAA AACTGCGCCG GCCACAAAGC
      1410       1420       1430       1440       1450       1460       1470
CTACTTCGAC GCACGCGTGA TTCCGGGCGT TGCCTACACT TCCCCCGAAG TGGCGTGGGT GGGCGAAACC
      1480       1490       1500       1510       1520       1530       1540
GAACTGTCCG CCAAAGCCTC CGGCCGCAAA ATCACCAAAG CCAACTTCCC GTGGGCGGCT TCCGGCCGTG
      1550       1560       1570       1580       1590       1600       1610
CGATTGCCAA CGGTTGCGAC AACGGCTTTA CCAAGCTGAT TTTTGATGCC GAAACCGGCC GCATCATCGG
      1620       1630       1640       1650       1660       1670       1680
CGGCGGCATT GTCGGTCCGA ACGGTGGCGA TATGATCGGC GAAGTCTGCC TTGCCATCGA AATGGGCTGC
      1690       1700       1710       1720       1730       1740       1750
GACGCGGCAG ACATCGGCAA AACCATCCAC CCGCACCCGA CCTTGGGCGA ATCCATCGGT ATGGCGGCGG
      1760       1770       1780       1790    1797
AAGTGGCATT GGGTACTTGT ACCGACCTGC CTCCGCAAAA GAAAAAA
```

图 2

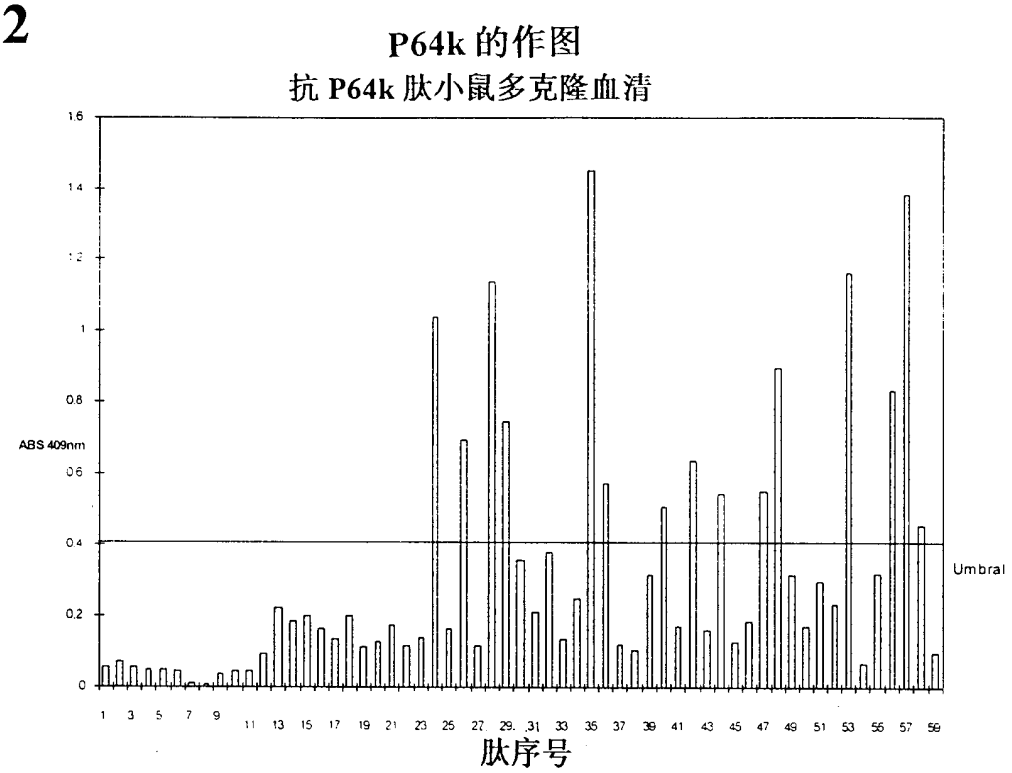

P64k 的作图
抗 P64k 肽小鼠多克隆血清
ABS 409nm
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Umbral
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 56 57 59
肽序号

图 3

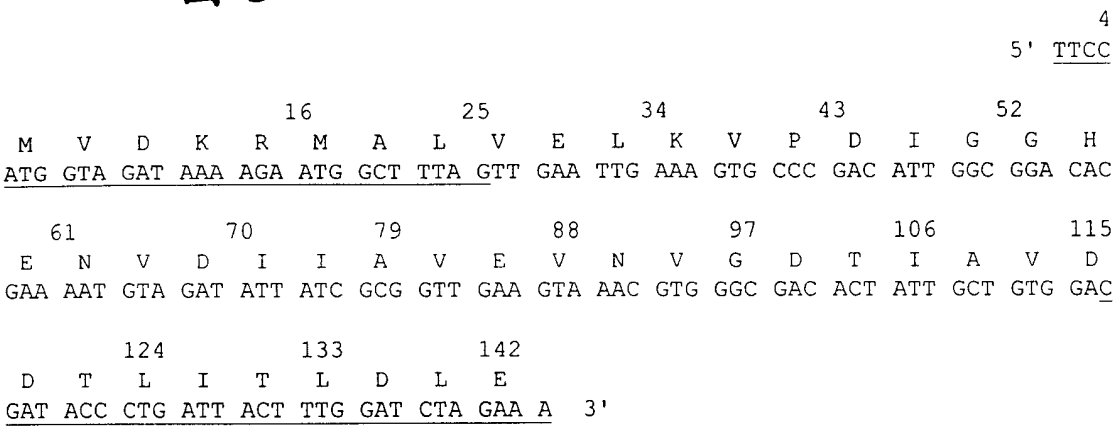

4
5' TTCC
16 25 34 43 52
M V D K R M A L V E L K V P D I G G H
ATG GTA GAT AAA AGA ATG GCT TTA GTT GAA TTG AAA GTG CCC GAC ATT GGC GGA CAC
61 70 79 88 97 106 115
E N V D I I A V E V N V G D T I A V D
GAA AAT GTA GAT ATT ATC GCG GTT GAA GTA AAC GTG GGC GAC ACT ATT GCT GTG GAC
124 133 142
D T L I T L D L E
GAT ACC CTG ATT ACT TTG GAT CTA GAA A 3'

图 4

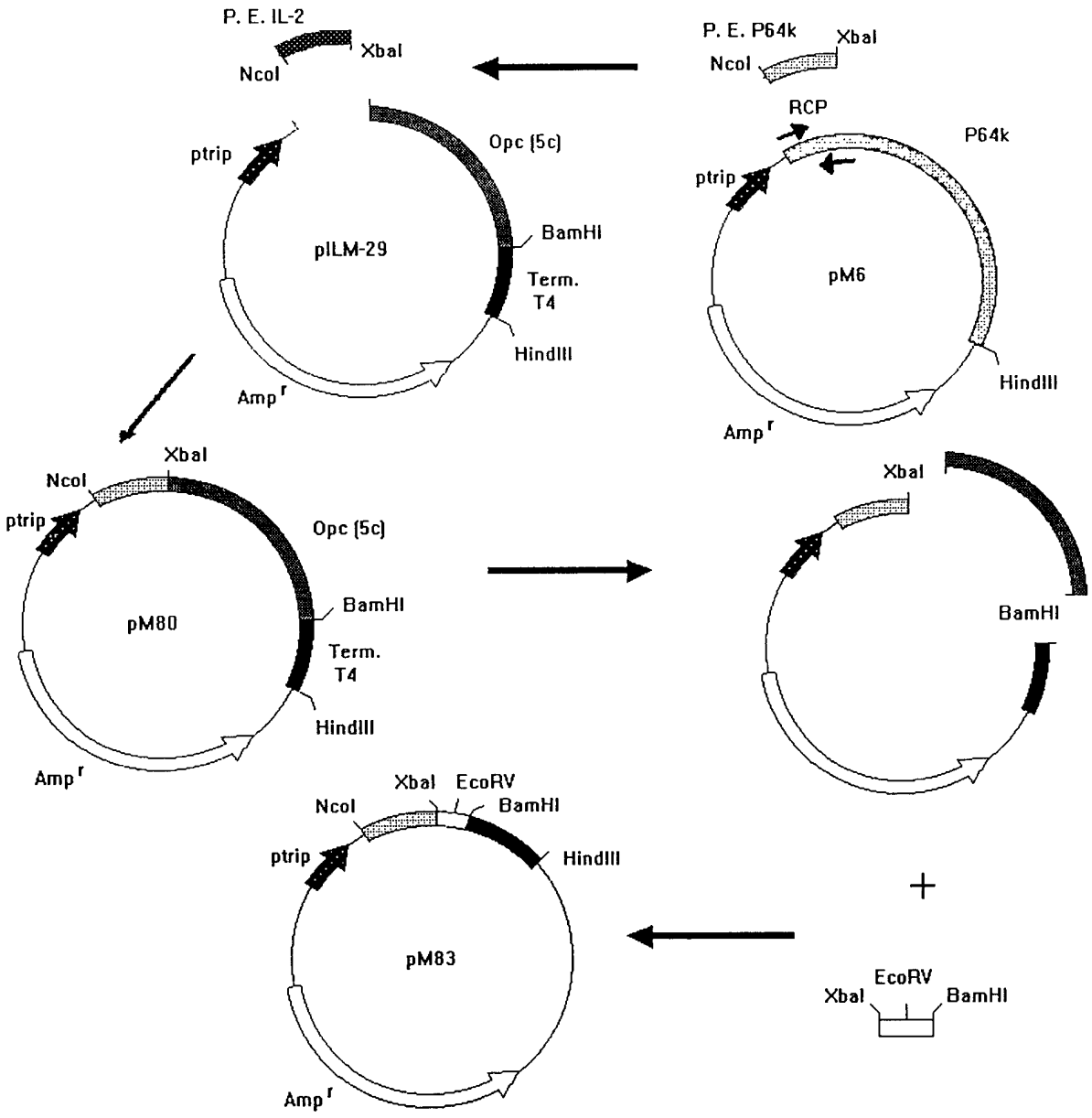

P. E. IL-2
XbaI
NcoI
ptrip
Opc (5c)
pILM-29
BamHI
Term.
T4
HindIII
Amp r
P. E. P64k
XbaI
NcoI
RCP
P64k
ptrip
pM6
HindIII
Amp r
XbaI
NcoI
ptrip
Opc (5c)
pM80
BamHI
Term.
T4
HindIII
Amp r
XbaI
BamHI
XbaI
EcoRV
NcoI
BamHI
ptrip
HindIII
pM83
Amp r
XbaI
EcoRV
BamHI

图 5

产生高分区段对的序列		高得分	最小泊松几率 P(N)	N
KPY1 人丙酮酸激酶, M1 (肌肉) 同工酶	(EC 2.7.1 …	51	0.98	1
KPY1 大鼠丙酮酸激酶,M1 (肌肉)同工酶	(EC 2.7.1 …	51	0.98	1
KPY2 人丙酮酸激酶, M2 同工酶	(EC 2.7.1.40)	51	0.98	1
KPY2 大鼠丙酮酸激酶,M2 同工酶	(EC 2.7.1.40)	51	0.98	1

>KPY1_人丙酮酸激酶, M1 (肌肉)同工酶 IEC 2.7.1.40)(胞质甲状腺激素结合蛋白) 长度 = 530

查找: 29 VNVGDTIA VDDTLITLDL 46
V+VG I VDD LI+L++
目标: 167 VEVGSKIY VDDGLISLQV 184

>KPY1_大鼠丙酮酸激酶, M1 (肌肉)同工酶(EC 2.7.1.40).
长度 = 530

查找: 29 VNVGDTIA VDDTLITLDL 46
V+VG I VDD LI+L++
目标: 167 VEVGSKIY VDDGLISLQV l84

>KPY2_人丙酮酸激酶, M2 同工酶(EC 2.7.l.40).
长度 = 530

查找: 29 VNVGDTIA VDDTLITLDL 46
V+VG I VDD LI+L++
目标: 167 VEVGSKIY VDDGLISLQV 184

>KPY2_大鼠丙酮酸激酶, M2 同工酶(EC 2.7.1.40).
长度 = 530

查找: 29 VNVGDTIA VDDTLITLDL 46
V+VG I VDD LI+L++
目标: 167 VEVGSKIY VDDGLISLQV 184

图 6

产生高分区段对的序列		读框	高得分	最小泊松几率 P(N)	N
CFMUCIN	犬属家族(克隆 pCTM-A)粘蛋白 c-末端...	-2	62	0.30	1
HS8671	EST02755 Home Sapiens... cDNA 克隆 HFBCA72 sim	-2	61	0.40	1

>CFMUCIN 犬属家族(克隆 pCTM-A)粘蛋白 c-末端 RNA, 3'末端
长度 = 1733

查找: 8 LV ELK VPD IGGH EN VDI I A
L E+ VPD H V+++ A
目标: 1015 LR EVQ VPD RKLH KG VQLL A

V E VN VGDTIA VDD 39
E + +G+++ VDD
G E LG IGEALQ VDD 920

>HS8~71 类似于粘蛋白 CTM-A 的 EST02755Home sapiens cDNA 克隆 HFBCA72 长度 = 286

查找: 8 LV ELK VPDIGGHEN VD IIAV
L E+ VPD HE V++++
目标: 240 LR EVQ VPDRKLHEG VQLLSG

EVNVGDTIAVDD 39
E+ VG VDD
EZGVGKXFQVDD 145

图 7

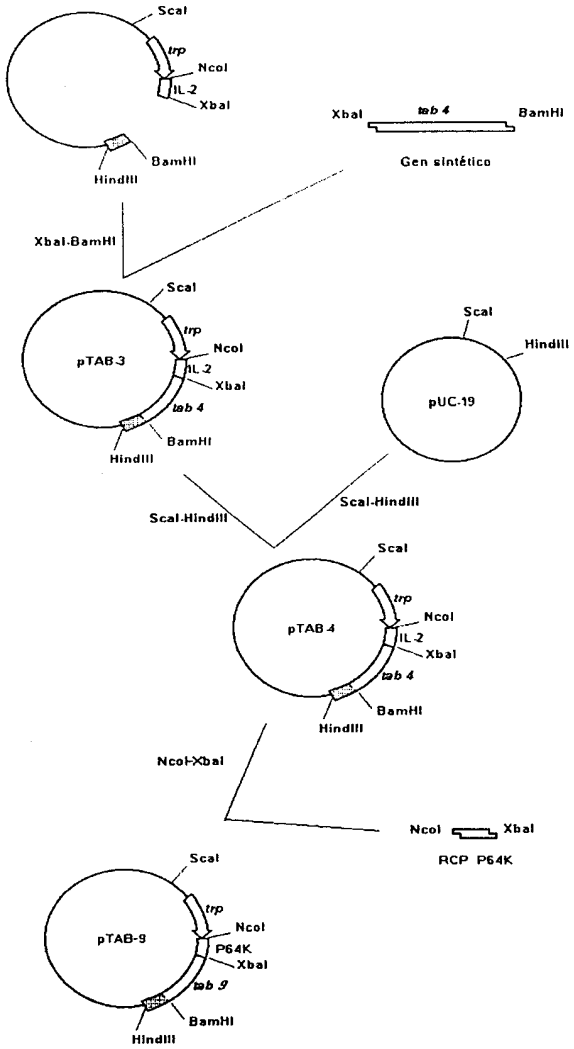

ScaI
trp
NcoI
IL-2
XbaI
BamHI
HindIII
XbaI
tab 4
BamHI
Gen sintético
XbaI-BamHI
pTAB-3
pUC-19
ScaI-HindIII
pTAB-4
NcoI-XbaI
NcoI
XbaI
RCP P64K
pTAB-9
P64K
tab 9

图 8

```
            12          21          30          39          48          57
 M   V   D   K   R   M   A   L   V   E   L   K   V   P   D   I   G   G   H
ATG GTA GAT AAA AGA ATG GCT TTA GTT GAA TTG AAA GTG CCC GAC ATT GGC GGA CAC
            69          78          87          96          105         114
 E   N   V   D   I   I   A   V   E   V   N   V   G   D   T   I   A   V   D
AA AAT GTA GAT ATT ATC GCG GTT GAA GTA AAC GTG GGC GAC ACT ATT GCT GTG GAC
            126         135         144         153         162         171
 D   T   L   I   T   L   D   L   D   S   R   G   I   R   I   G   P   G   R
GAT ACC CTG ATT ACT TTG GAT CTA GAC TCG AGA GGC ATT CGT ATC GGC CCA GGT CGC
            183         192         201         210         219         228
 A   I   L   A   T   A   G   G   G   A   R   Q   S   T   P   I   G   L   G
GCA ATT TTA GCA ACA GCT GGC GGT GGC GCA CGT CAA TCT ACC CCT ATT GGT TTA GGT
            240         249         258         267         276         285
 G   A   L   Y   T   T   A   G   G   G   A   R   K   S   I   T   K   G   P
CAG GCT CTG TAT ACG ACT GCC GGC GGT GGT GCG CGC AAA AGT ATC ACC AAG GGT CCA
            297         306         315         324         333         342
 G   R   V   I   Y   A   T   A   G   G   G   A   R   K   R   I   H   I   G
GGC CGC GTC ATT TAC GCC ACC GCG GGC GGC GGT GCC CGT AAG CGT ATC CAC ATT GGC
            354         363         372         381         390         399
 P   G   R   A   F   Y   T   T   A   G   G   G   A   R   K   R   I   T   M
CCA GGC CGT GCA TTC TAT ACT ACA GCA GGT GGT GGC GCA CGT AAA CGC ATC ACT ATG
            411         420         429         438         447         456
 G   P   G   R   V   Y   Y   T   T   A   G   G   G   A   S   I   R   I   Q
GGT CCT GGT CGC GTC TAT TAC ACG ACC GCT GGC GGC GGT GCT AGC ATT CGC ATC CAA
            468         477         486         495
 R   G   P   G   R   A   F   V   T   I   *
CGC GGC CCT GGT CGT GCA TTT GTG ACC ATA TGA
```

图 9

A


Est.
LR150
JY1
RF
MN
BRVA
IIIB
XhoI
BamHI
NcoI
XbaI
NdeI

B

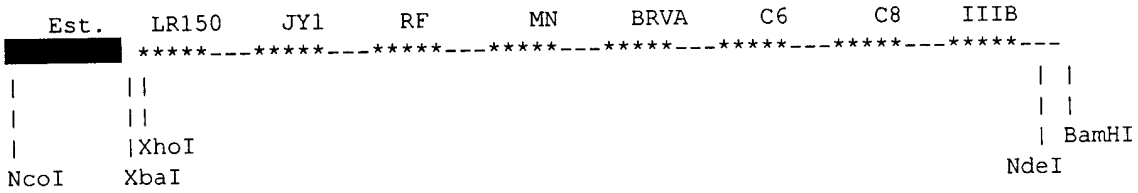
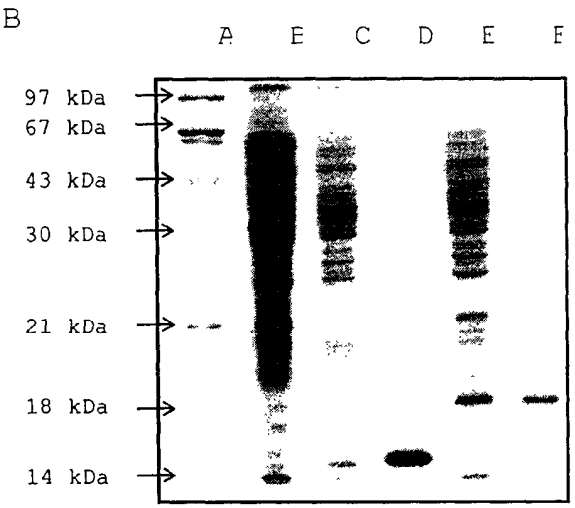

Est.
LR150
JY1
RF
MN
BRVA
C6
C8
IIIB
XhoI
BamHI
NcoI
XbaI
NdeI

图 10

A

基因	稳定子	质粒	培养基	表达百分率
porA	hIL2-58	pILM-28	M9	32
	P64k-47	pM-82	M9	34
opc	hIL2-58	pILM-29	M9	25
	P64k-47	pM-80	M9	20
TAB	hIL2-22	pTAB4	LB	5
	P64k-47	pTAB9	LB	10

B



C

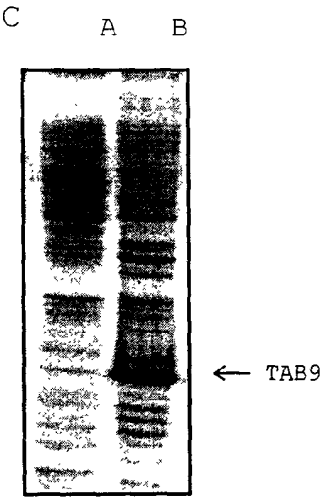


图.11

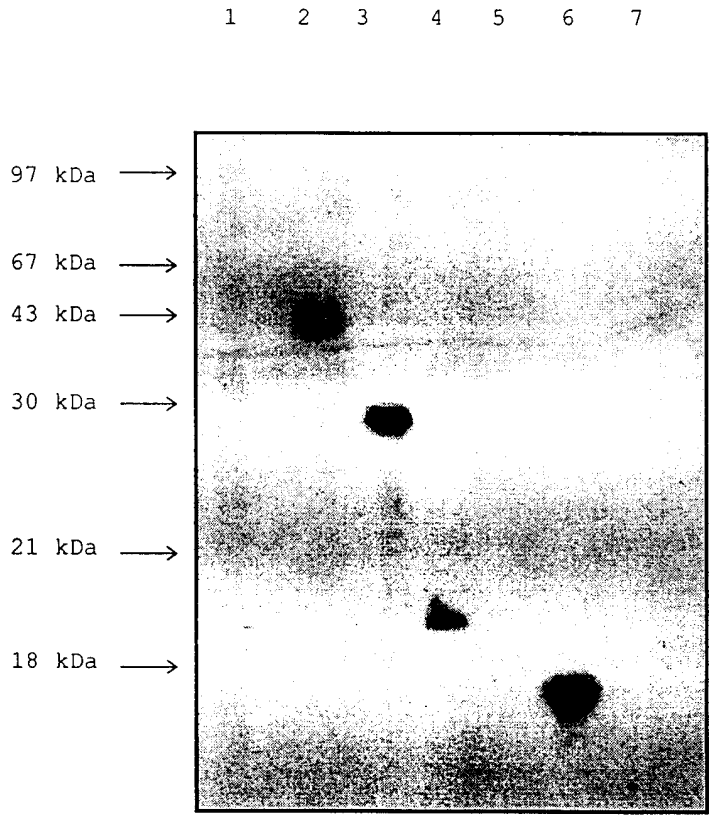


图.12

	TAB 4				TAB 9			
兔 #	12166	5725	5340	2310	1	2	3	10
TAB	20480	10240	10240	81920	20480	20480	51200	51200
	0	0	0	0	0	0		
LR150	<100	<100	<100	6400	400	<100	800	<100
JY1	200	<100	<100	12800	1600	6400	3200	<100
RF	6400	<100	3200	800	<100	200	3200	800
MN	200	<100	<100	1600	1600	3200	6400	<100
BRVA	<100	6400	400	3200	800	400	6400	1600
IIIB	<100	<100	<100	<100	800	<100	800	<100
GM	1820				1416			
R%	45.8%				75%			