[19]中华人民共和国专利局

[51]Int.Cl[6]
C12N 15/11
C12N 15/55 C12N 15/63
C12N 9/16
//(C12N15/11,C12R1:425)

[12] 发明专利申请公开说明书

[21] 申请号 96102722.3

[43]公开日 1997 年 1 月 29 日

[11] 公开号 CN 1141342A

[22]申请日 96.3.25

[30]优先权

[32]95.3.23 [33]JP[31]63772/95

[71]申请人 田边制药株式会社

地址 日本大阪府

[72]发明人 柴谷武尔 赤塚浩之
河合衣里

[74]专利代理机构 中原信达知识产权代理有限责任公司

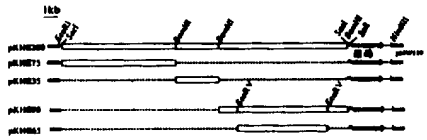
代理人 张天舒

权利要求书 1 页 说明书 19 页 附图页数 1 页

[54]发明名称 参与酯酶分泌机制的基因

[57]摘要

编码参与来自沙雷氏菌属微生物的酯酶的分泌机制的多肽的基因，包括通过将该基因插入一载体质粒所制得的质粒的重组质粒，用该重组质粒转化的微生物，以及一包括在一培养基中培养上述转化微生物及收集在细胞内外产生的酯酶的用于酯酶产生的方法。该转化微生物有显著优越的向胞外分泌酯酶的能力。



# 权 利 要 求 书

1. 一编码参与来自沙雷氏菌属的微生物的酯酶的分泌机制的多肽的基因。

2. 如权利要求1所述的具有示于序列号1中的DNA序列的基因。

3. 如权利要求1所述的包括一编码示于序列号2，3或4的多肽的DNA的基因。

4. 包括通过将权利要求1或权利要求2 所述的基因插入载体质粒所得的质粒的重组质粒。

5. 用权利要求4所述的重组质粒转化的微生物。

6. 通过用一包括通过同时将编码来自沙雷氏菌属微生物的酯酶的基因和编码参与来自沙雷氏菌属微生物的酯酶的分泌机制的多肽的基因插入一载体质粒所得质粒的重组质粒转化一宿主微生物所得的转化微生物。

7. 通过用一包括通过将编码来自沙雷氏菌属微生物的酯酶的基因插入一载体质粒所得的质粒的重组质粒和一包括通过将一编码参与来自沙雷氏菌属微生物的酯酶的分泌机制的多肽的基因插入一载体质粒所得的质粒的重组质粒转化一宿主微生物所得的转化微生物。

8. 如权利要求6或权利要求7所述的微生物，其中之宿主微生物为一沙雷氏菌属或大肠杆菌属的微生物。

9. 生产酯酶的方法，包括在一培养基中培养权利要求6 或权利要求7 所述的转化子以及收集在细胞内或在培养液中积累的酯酶。

# 说　明　书

## 参与酯酶分泌机制的基因

本发明涉及一编码一参与沙雷氏菌属的微生物的酯酶分泌机制的多肽的基因DNA，一包含该基因的重组质粒，一用此重组质粒转化的新型微生物，一同时含有上述重组质粒和带有编码一酯酶的基因的重组质粒的新型微生物，以及一通过培养这些新型微生物进行酯酶生产的过程。

近来，人们经常尝试将一些酶如酯酶用于水解反应。为此目的人们发现了各种各样的酯酶，如：来源于动物(如猪的肝脏、猪的胰脏等)的酯酶以及由微生物如球形节细菌(Arthrobacter globiformis)、白地霉(Geotrichum candidum)、Candida Cylindracea、萤光假单胞菌(Pseudomonas fluorescens)等所产生的酯酶。另外，众所周知由灵杆菌产生的酯酶是经重组DNA技术制备的。

然而，已知的酯酶和已知的生产酯酶的方法都有各种各样的问题。例如来源于动物的酯酶都很贵，来源于微生物的酯酶则在酶的活性、稳定性或特异性等方面有缺点或者是生产酯酶的微生物没有高的生产率。另外，即使利用基因重组的微生物，仅仅通过扩增酯酶基因将酯酶产率提高到适用于工业生产的水平也不一定容易。此外，即使提高了酯酶的产率，酯酶的胞外分泌有时也会不足，因此就要求复杂的程序回收此酶。

各种考察的结果表明本发明者已成功地获得了编码一参与沙雷氏菌属的微生物的酯酶分泌机制的多肽的基因(DNA)，并已进一步发现经用含该DNA 的重组质粒转化该微生物可显著提高此酯酶产生微生物分泌酯酶的能力，基于这些新的发现本发明已经成功地完成了。

本发明的目的之一是提供编码一参与沙雷氏菌属的微生物的酯酶分泌机制的多肽的基因。本发明的另一目的是提供一通

过将该基因插入一载体质粒而形成的重组质粒并进一步提供一含该重组质粒的转化子。本发明的又一目的是提供一通过培养该转化子并收集转化子细胞内外积累的酯酶以生产酯酶的方法。

附图简要说明

图1列示了从实施例2的重组细胞中分离出的质粒pPKHE200，pPKHE75，pPKHE35，pPKHE90和 pPKHE65 的限制性内切酶图谱。

此编码一参与酯酶分泌机制的多肽的基因(此后称为"酯酶分泌基因")为一长4,466个碱基对的双链DNA，序列中有三个开放的阅读框架(ORF)，尤其含有示于后面的序列身份号1中的核苷酸序列。

此酯酶分泌基因的供体微生物包括属于有酯酶生产能力的沙雷氏菌属的任何微生物如：灵杆菌Sr41(FERM BP-487)，液化沙雷氏菌ATCC 27592，灵杆菌ATCC 13880，灵杆菌ATCC 14764，灵杆菌ATCC 19180，灵杆菌ATCC 21074，灵杆菌ATCC 27117，灵杆菌ATCC 21212等。

插入酯酶分泌基因的载体质粒包括在被转化的细胞中可复制的任何质粒。 其中，该质粒最好是那些拷贝数在1至几千、带有一对抗生素如氯卡青霉素、卡那霉素、氯霉素的抗性标记并带有适当的启动子如lac，tac或trp的质粒。此外， 载体质粒可进一步携带一质粒稳定基因如par和parB。

这些载体质粒包括如pLG339[Gene，Vol.18，332(1982)]，pBR322[Gene，Vol.2，95(1977)]，pUC18[Gene，Vol.33，103(1985)]，pUC19[Gene，Vol.33，103(1985)]，pHSG298[Gene，Vol.61，63(1987)]，pHSG299[Gene，Vol.61，63(1987)]及类似的其他质粒。

上述这些质粒可购得，也可通过一常规方法如"廓清裂解物片"(参见Yasuyuki Takagi，"Procedure for Experiment in Genetic Engineering"，125页，Kodansha出版，1980) 或

"碱裂解片"(参见Maniatis等, Molecular Cloning", 368页, 美国科尔德斯普林港实验室(1982)] 从含这些质粒的微生物中获得。

宿主微生物 (既用于质粒的重组也用于所期望的酯酶的表达) 包括任何可用重组质粒转化且可复制此质粒、并可表达此质粒上的基因且产生一有功能的蛋白质的微生物。这些宿主微生物包括如属于沙雷氏菌属或大肠杆菌属的微生物，尤其是灵杆菌 Sr41 及由其衍生的各种突变株如灵杆菌 M-1(FERM BP-4068)，灵杆菌TT392 (参见Journal of Bacteriology, Vol. 161, 1(1985)] 或大肠杆菌 K12 DH5 [参见 Maniatis 等，"Molecular Cloning", 第二版，A10，美国科尔德斯普林港实验室(1989)]。

另外，携带一插入了沙雷氏菌属的酯酶基因的重组质粒的任何其它微生物也可用为宿主微生物，如灵杆菌TA5025 (FERM BP-4067)。

含酯酶分泌基因的染色体DNA 可用常规方法很容易地从含该基因的微生物获得，如将微生物细胞用溶菌酶处理再用表面活性剂(如十二烷基磺酸钠、月桂酰肌氨酸钠等)处理，然后用有机溶剂(如苯酚、氯仿、二乙酯等)抽提以除去蛋白，再用乙醇沉淀DNA[参见Journal of Molecular Biology, Vol. 3, 208 (1961) 及 Biochemica et Biophysica Acta, Vol. 72, 619 (1963)]。

由含酯酶分泌基因的染色体DNA和载体质粒DNA组成的重组质粒可用常规方法很容易地制备。如用适当的限制性内切酶(如EcoRI, BamHI, HindIII, SalI, SacI等)切染色体DNA 和质粒DNA，然后用DNA连接酶(如T4 DNA连接酶，E. coli DNA连接酶等)处理产物，若有必要，先用末端转移酶或DNA多聚酶处理，再用 DNA 连接酶进行处理 (参见 Methods in Enzymology, Vol. 68, 41(1979)，及 Yasuyuki Takagi, Procedure for Experiment in Genetic Engineering", 135页, Kodansha 出

版，1980]。

可按如下所述从经上述程序获得的重组质粒的混合物中筛选含酯酶分泌基因的目的重组质粒。

用含酯酶分泌基因的重组质粒转化限制内切酶缺陷且有酯酶生产力的微生物细胞，如携带含酯酶基因的重组质粒（如示于Japanese Patent First Publication (Kokai) No. 344891/1993的pLIPE111等）的大肠杆菌K12 DH5 菌株，然后将此转化细胞涂布到细胞可生长的含乳化的甘油三酯的琼脂培养基上，如含经聚氧化乙烯十六烷醇醚(Brij58)乳化的三丁酸甘油酯的营养琼脂培养基，培养基中同时也加有规定浓度的抗生素。30~37℃保温1~2天后，分离在其周围形成大的清亮环带的转化子克隆。

上面重组质粒对宿主微生物的导入可用一常规方法实现，如在低温下用$CaCl_2$ 水溶液处理宿主细胞以提高其通透性然后再将重组质粒导入（参见 Journal of Molecular Biology, Vol. 53, 159 (1970)]，或用电穿孔的方法。

将参与酯酶分泌机制的DNA 和酯酶基因一起插入同一个载体质粒所得的重组质粒转化微生物细胞（如大肠杆菌K12 DH5等），然后用与上述相同的方式处理所得到的转化细胞也可筛选所期待的转化子。

然后用"碱裂解片"从转化子中提取质粒以获得重组质粒，即通过将灵杆菌的酯酶基因插入载体质粒所得到的质粒。

为了提高转化效率，在限制性内切酶缺陷的且与用于表达酯酶的宿主微生物同种的微生物中修饰由此获得的重组质粒。也就是说当所用的宿主微生物是灵杆菌Sr41时，上述质粒被导入灵杆菌TT392中，此为一限制酶缺陷菌株。提取经如此修饰的重组质粒。然后将被修饰的质粒导入宿主微生物中以获得适于目的酯酶生产的目的转化子。

按Takagi和Kizumi的方法[参见Journal of Bacteriology, Vol. 161, 1 (1985)]可容易地将重组质粒导入宿主微生物中

，质粒的提取可按碱裂解法等完成。此外，通过分离表达抗生素抗性的菌落也可获得目的转化子。用于该转化的宿主微生物包括上面提及的微生物但最好是有高的酯酶产率的菌株。

用上述方法得到的被转化的微生物如灵杆菌TA5030和灵杆菌TBS90。灵杆菌TA5030是通过用将一含酯酶基因的约2.6kb的SalI-BstPI DNA片段和一含酯酶分泌基因的约6.5kb 的EcoRV-EcoRV DNA 片段插入载体 pMW119 所得的重组质粒转化灵杆菌Sr41而得的；灵杆菌TBS90是通过用将一含酯酶基因的约2.6kb的SalI-BstPI DNA片段和一含酯酶分泌基因的约9.0kb的BamHI-SacI DNA 片段插入载体 pMW119 所得的重组质粒转化灵杆菌Sr41而得的。这些转化株都有与宿主微生物灵杆菌Sr41相同的形态特征。

通过将微生物在培养基中培养并收集微生物细胞内外的酯酶可实现用上述被转化的微生物生产酯酶。

用于酯酶生产的培养基包括微生物可在其中生长的任何常规培养基。合适的培养基含有碳源如糖类(如葡萄糖、蔗糖、糖蜜等)、有机酸(如延胡索酸、柠檬酸等)或氨基酸(如丙氨酸、谷氨酸、天冬酰胺等)及氮源如无机铵盐(如硫酸铵、氯化铵等)、脲、蛋白胨、玉米浆、酵母抽提物、酪蛋白水解物等。依培养基之不同，碳源通常占1~15%(W/V)，氮源则为0.1~2.0%(W/V)。培养也可可有可无地进一步含有适量的无机盐(如磷酸盐、镁盐、钾盐、钙盐等)和/或金属离子 (如铁、锰、铜、锌等)。在合成培养基的情况下，还可进一步含维生素或氨基酸，以及酯酶生产的诱导剂(如植物油、表面活性剂等)、消泡剂、适于在微生物中稳定质粒的抗生素。培养基的pH值最好调至5~8。

此被转化的微生物的培养可用常规方法实现。例如，将微生物接种于培养基中，并经振荡培养、通气培养、持续培养、不间断培养等进行培养。虽然依培养基种类和培养方法的不同培养条件可有所改变，但可以是适于微生物生长的任意条件，

通常初始pH值为5~8，于20~40℃培养1~2天。

用常规方法收集在培养细胞内外产生的酯酶，例如，含于培养基中的酯酶可用无机盐盐析、有机溶剂沉淀、用离子交换树脂和各种柱色谱进行吸附或解吸附、凝胶过滤、使用蛋白质沉淀剂或这些方法的组合等方法来收集。在细胞内积累的酯酶可先用物理方法和摩擦破碎法(Dyno Mill) 或化学方法如溶菌酶处理来破碎细胞，然后再用上述方法收集细胞抽提物中的酯酶来得到。

本发明的酯酶分泌基因并不局限于此处详细说明中所明确揭示的那些DNA序列，且包括有通过如插入、缺失或置换在序列中进行修饰所获得的DNA序列的任意基因。也就是说，酯酶分泌基因可通过利用以此处明确揭示的编码酯酶基因的DNA 序列为基础设计的突变的DNA合成引物或利用化学诱变剂如甲酸、亚硫酸肼直接在试管中进行人工修饰。另外，也可用NTG 或紫外线处理酯酶生产菌株获得突变基因。

实施例

此以下面的例子来举例说明本发明，但不应认为本发明局限于此。

在实施例中酯酶活性用常规方法进行测量(利用日本Dainippon Pharmaceutical有限公司生产的脂酶试剂盒S，其中底物为三丁酸二巯基丙醇酯)，也可用另一方法测量，它包括将产物置于含橄榄油(作为底物)的酶促反应液中，pH8.0，37℃反应20分钟，然后测量形成的脂肪酸的量，其中酯酶活性单位以每分钟形成的脂肪酸的微摩尔数来表示。

此外，例子中所用的培养基描述如下，其中除非特别说明，'%'是重量/体积百分比(W/V%)。

LB培养基：1.0%的细菌胰蛋白胨(Difco公司生产)，0.5%的酵母抽提物(Difco公司生产)及0.5%的氯化钠。

LBG平板培养基：1.0%的细菌胰蛋白胨(Difco公司生产)，

0.5%的酵母抽提物(Difco公司生产),及0.5%的氯化钠及1.0%的Gellan树胶(日本Wako Pure Chemical Industries公司生产)。

含三丁酸甘油酯的LBG平板培养基:含0.5%(v/v)三丁酸甘油酯,0.5%聚氧化乙烯十六烷醇醚及0.005%氯苄青霉素的LBG平板培养基。

酯酶生产培养基:1.0%的葡聚糖,2.0%的酵母,0.2%的硫酸铵,0.1%的磷酸二氢钾,0.05%的七水硫酸镁,0.01%的二水氯化钙,0.001%的七水硫酸亚铁,0.5%的Tween 80及0.1%的亮色刺桐。

实施例1

(1) 含酯酶分泌基因的染色体DNA的制备:

在LB培养基(200ml)中通气振荡培养灵杆菌Sr41(FERM BP-487),30℃过夜,然后离心收集细胞,将细胞悬于0.9%的氯化钠水溶液(200ml)中,离心以洗涤细胞,洗后的细胞悬于含200mg溶菌酶的50mM Tris- 盐酸 -50mM 乙二胺匹乙酸二钠的水溶液(pH7.5, 200ml)中,混合物在室温下静置1小时。

以0.5%的终浓度向混合物中加入十二烷基磺酸钠,再加入蛋白酶K,混合物于50℃温和地摇动3小时以裂解细胞,用等体积的被10mM Tris-盐酸-1mM 乙二胺匹乙酸二钠水溶液(pH8.0)(此后称之为"TE")饱和的苯酚抽提混合物二次,再用TE 饱和的苯酚和氯仿的等体积混合物抽提二次,所得水相用乙醇沉淀。沉淀溶于TE中以制备含2.5mg带酯酶分泌基因的染色体DNA的TE溶液。

(2) 重组质粒DNA的制备:

用限制性内切酶SacI 彻底消化上述(1) 中制备的染色体DNA(20μg),用TE饱和的苯酚和氯仿的等体积混合物抽提二次,向混合物中加0.5倍体积的7.5M乙酸铵,用2 倍体积的乙醇沉淀以回收DNA。

将带有来自灵杆菌Sr41的酯酶基因的约2.6kb的基因组DNA

片段(SalI-BstPI片段)插入载体质粒pMWE119 中以制备质粒载体 pMWE121。将被上述相同的限制性内切酶彻底消化的质粒载体pMWE121用碱性磷酸酶(日本Takara Shuzo有限公司生产,0.4单位)在56℃处理1小时以去磷酸化, 产物用被TE饱和的苯酚和氯仿的等体积混合物抽提二次,再向其中加入0.5倍体积7.5M的乙酸铵, 然后用2倍体积的乙醇沉淀以回收载体质粒DNA。将此去磷酸化的质粒载体与上述染色体DNA(3μg)混合, 用DNA连接试剂盒(Takara Shuzo公司生产)在4℃连接16 小时以得重组质粒DNA。

(3) 用重组质粒转化及用大肠杆菌宿主构建克隆库:

用Hanahan的方法[参见Journal of Molecular Biology, Vol. 166,557(1983)]处理大肠杆菌DH5细胞, 并向其中加入含上述(2)中制备的质粒DNA的反应混合物, 使混合物转变实现。将如此处理的细胞涂布于含氯苄青霉素(50μg/ml)的LBG 平板培养基上, 37℃过夜培养以产生含插入了灵杆菌Sr41的染色体DNA片段的重组质粒的转化子(约50,000株)。

(4) 含酯酶分泌基因的转化子菌株的分离及确认:

当酯酶产生菌接种于含三丁酸甘油酯的LBG 平板培养基上时,产生到培养基中的酯酶会降解三丁酸甘油酯以产生脂肪酸,籍此菌落周围的甘油三酯乳剂便被转变, 在菌落周围会形成一清亮的环圈,用此现象可实现转化子的筛选。

将上述(3)得到的转化子(约50,000株) 接种于含三丁酸甘油酯的LBG平板培养基,37℃过夜培养,由此可从SacI的DNA库中分离出形成清亮环带的菌株。

此清亮环带的形成将被归因于通过导入含酯酶分泌基因的重组质粒提高了酯酶的胞外分泌能力, 因为当灵杆菌Sr41接种到含三丁酸甘油酯的LBG平板培养基上并于37 ℃过夜培养时可观察到清亮的环带的形成; 而当未被转化的大肠杆菌DH5 和携带载体质粒pMWE121的大肠杆菌DH5被接种到含三丁酸甘油酯的LBG平板上并于37℃过夜培养时,观察不到清亮环带的形成。

此外，将可形成清亮环带的转化子在含氯苄青霉素(200μg/ml)的LB培养基(60ml)中通气振荡培养，37℃过夜，然后用常规方法测量培养混合物上清的酯酶活性，结果可观察到34，800单位的酯酶活性。

另外，将上述培养混合物的上清走SDS 聚丙烯酰胺凝胶电泳，用兔的抗灵杆菌产生的酯酶的抗酯酶抗体通过Western杂交分析结果。结果产物显示出一条新带，此带在仅含载体质粒的大肠杆菌DH5的产物显示不出，且此带与从灵杆菌Sr41的培养物上清中得到的纯化的标准酯酶位于相同位置。

实施例2

质粒的分析

用常规方法[参见Maniatis等，"Molecular Cloning"，368页，美国科尔德斯普林港实验室(1982)]从例1-(4)的转化子细胞中制备质粒DNA，并用各种限制性内切酶作酶切然后走琼脂糖凝胶电泳 。结果证实除了含有来自灵杆菌的酯酶基因的约2.6kb的SacI-BstPI DNA片段外，此质粒(此后称为"pKHE200")还含有一约20.0kb的SacI DNA片段。

该pKHE200中的约20.0kb的SacI DNA 片段的限制性内切酶图谱示于所附的图1中。

用各种限制性内切酶切质粒pKHE200的约20.0kb的SacI 片段，并将每个片段亚克隆到质粒载体pMWE121上，然后用此重组质粒转化大肠杆菌DH5， 以上述同样方式在含三丁酸甘油酯的LBG平板培养基上培养转化子，并确定是否有清亮的环带形成。

每个DNA片段的限制性内切酶图谱示于所附的图1中。对于用含约9.0kb的BamHI-SacI DNA片段的质粒(pKHE90)和含约6.5kb的EcoRV-EcoRV DNA片段的质粒(pKHE65)转化的大肠杆菌DH5，可观察到清亮的环带的形成及酯酶的活性(用常规方法测量)。然后对于用含约7.5kb的SacI-BamHI DNA片段的质粒(pKHE75)和含约3.5kb的BamHI-BamHI DNA片段的质粒(pKHE35)转化的大

肠杆菌DH5，既观察不到清亮环带的形成也观测不到酯酶的活性(用常规方法测量)。

从上面的结果可以发现来自灵杆菌Sr41的酯酶分泌基因位于约6.5kb的EcoRV-EcoRV DNA片段上。

用质粒pKHE65转化的大肠杆菌DH5菌株已于1996年2月9 日交付工业科学和技术代理处国家生命科学和人类技术协会保管，登记号为FERM BP-5385(称为"大肠杆菌TA5029")。

实施例3

克隆DNA的分析

[核苷酸序列的确定]

用Kilobase Deletion试剂盒(Takara Shuzo公司生产) 处理重组质粒pKHE65以制备各种缺失质粒。将这些质粒与一引物退火，按照Sanger等的双脱氧链终止法 [Proc. Natl. Acad. Sci. USA, Vol.74,5463(1977)]用DNA多聚酶的Klenow 片段合成互补链[用 $\alpha$-$^{32}$P) dCTP (14.8×10$^{6}$Bq/pmol, 74×10$^{4}$Bq) 标记]，根据在8%的脲修饰的聚丙烯酰胺凝胶上电泳和放射自显影的数据确定其核苷酸序列。

结果发现来自灵杆菌Sr41的酯酶分泌基因的DNA 序列从起始密码子GTG到密码子BAG由4,466个碱基对组成，示于序列身份号1中，它包含参与酯酶分泌机制的三种多肽的DNA区域(ORF)。(推测这三个ORF构成了一对儿操纵子)。

此外，由每个ORF 编码的多肽的氨基酸序列分别示于序列身份号2,3和4中。

实施例4

酯酶高产菌株的制备：

将质粒pKHE65导入限制内切酶缺陷的菌株灵杆菌TT392 中以得到转化菌株。用"碱裂解法"从被转化的菌株中提取被灵杆菌修饰了的pKHE65质粒DNA。 然后通过电穿孔法用上述获得的

质粒DNA转化灵杆菌Sr41细胞以产生转化子灵杆菌TA5030。将此转化子(一接种环)接种到含氯苄青霉素(500μg/ml)的酯酶生产培养基中，30℃往复振荡培养20小时(摇幅7厘米，120 转/分)。离心培养液以获得酯酶活性约$3.5\times10^5$单位/ml的上清液(按常规方法测量)。此菌株的酯酶产率比宿主菌灵杆菌Sr41高约10倍，比含重组质粒(用作载体质粒)的灵杆菌Sr41高约2倍。

用带有本发明的酯酶分泌基因的重组质粒转化的微生物有显著优越的酯酶胞外分泌能力。因此，当一带有含酯酶基因的重组质粒的酯酶生产微生物被进一步用含本发明的酯酶分泌基因的重组质粒转化时，可获得一在酯酶产量和酯酶胞外分泌方面都有优越性状的微生物，培养该转化子可在工业水平上生产所期望的酯酶。

# 序列表

序列身份号：1
长度：4,466碱基对
类型：核苷酸
链型：双链
局部解剖：线形
分子类型：基因DNA
株：灵杆菌 Sr41
序列

```
GTG AAT CAA TTT ATC CCG CGC AAC GAA ATT GCG GAT GTT ATA CGT ACA   48
CGC AGC AAA GTC TTC TGG ACC GTT GGT ATA TTT ACT GCG TTT ATT AAC   96
CTG TTA ATG CTG GTT CCT TCC ATT TAT ATG CTC CAG GTT TAC GAC CGG  144
GTG CTG CCT TCG CGC AAT GAA ATC ACG CTG TTA ATG CTG ACG CTG ATC  192
ATG CTG GGC ATG TTC GGC ATG ATG TCG CTG TTG GAA TAC GTG CGC AGC  240
ATG GTG GTG ATC CGC ATC GGC AGC CAG CTG GAT ATG CGT CTC AAC ACG  288
CGA GTC TAT ACC GCG GCC TAC GAA GCG AAT CTG AAA AAC GGT TCG TCT  336
GAC GCC GGT CAG ATG CTG AGC GAT TTG ACC AAT CTG CGC CAA TTC CTC  384
ACC GGT AGC GCG CTG TTC GCC TTC TTT GAT GCG CCG TGG TTT CCG ATC  432
TAT CTG TTG GTG ATA TTC CTC TTT AAC CCT TGG TTG GGC CTT TTC GCC  480
CTG GTC GGT GCG CTG TTG CTG ATC GCA TTG GCG GTA ATC AAT GAA GTG  528
GTT TCG AAA AAG CCG CTG GGA GAA GCC AGC AAG CTG TCG ATC ATG TCA  576
GGT AAT TTG GCC AGC ACC AAT CTG CGA AAT GCC GAA GTG ATC GAG GCT  624
TTG GGG ATG TTG CCT AAC CTG AAA CGC CGG TGG TTC GGT CTG CAC CAG  672
CGG TTC TTG AAC AGC CAA CGC ATC GCC AGC GAA CGC GCA TCG CGG GTC  720
ACG TCA ATC ACC AAG TTC GTG CGT ATG TCG CTG CAG TCC TTA GTG TTG  768
GGC CTG GGG GGA TGG TTG GCG ATT GAT GGG CAC ATC ACG CCC GGC ATG  816
ATG ATC GCC GGT TCT ATA TTG ATG GGG CGA ACG TTG GCG CCG ATC GAG  864
CAG GTC ATT AAC GTT TGG AAA AGC TAT AGC GCG GCC AAA CTT TCT TAT  912
GGC CGC TTG GTC AAG CTG CTG GAA ACG CAT CCG CAG CGT GGT ACC GGC  960
ATG TCG CTG CCG CGT CCG GAA GGT GTG CTC TCC GTA GAA GGC GTG ACC 1008
GCC ACG CCT CCG GGA TCG AAA GGG GAT GCG GTG CTG CAT AAC GTA AGT 1056
TTT GCC ATT CAA CCC GGC GAT GTG CTG GGG ATT ATC GGT CCG AGT GCG 1104
TCG GGC AAA TCA ACA TTG GCG CGC TTA CTG GTC GGT ATT TGG CCT GTG 1152
AGC GAA GGG ATA GTG CGG TTG GAT AAT GCC GAC ATC TAC CAG TGG AAC 1200
AAA GAC GAA CTG GGG CCC TAT ATC GGC TAT CTG CCG CAG GAC ATC GAG 1248
```

```
TTG TTC GCC GGC ACT ATC GCC GAG AAC ATC GCT CGC TTT AAC GAC ATC  1296
GAT TCA GAG AAA GTG ATT GAG GCT GCC AAG CTG GCT GGT GTG CAT GAA  1344
CTG ATC CTG CGT TTC CCT AAC GGT TAC GAT TCG GTG ATC GGC AAC GGT  1392
GGT GCA GGG TTG TCC GGC GGG CAG AAG CAA CGT ATC GGC CTG GCG CGG  1440
GCA TTG TAT GGC GAT CCC GCG TTG GTG GTG TTG GAT GAG CCT AAC TCC  1488
AAC CTG GAT GAT GCC GGC GAG AAA GCG TTG AAC CAG GCC ATC ATG TTC  1536
CTT AAA CAG CGT AAT AAG ACG GTG GTC CTG ATC ACT CAC CGC ACC AAT  1584
CTG CTG TCG ATG ACC AGC AAG CTG TTG CTG TTG GTT AAC GGG AAC GTC  1632
AAT GCA TTC GGC CCA ACG CAG CAG GTG CTG CAG GCG TTG GCG AAT GCG  1680
CAA AAA GCG CAG GTG CCT CCG CAG GCG GTG CGT GCG GTG AAC TCC GAG  1728
CCG GAT GAA GGC GAA ATC CCT AAA ACT CAA ATT AAT TAA GCCGTGAACT   1777
TGCCCGGCGG CGCTTTTGCG TCGCCGACAG TCAAAGGAGT TGGT ATG TCT ACG CAT 1833
ATT GGC GAG CCG CAA GAC TCG TAT ACT GAA GAG ATC CCA CAA GAT GAA  1881
CGG CGG TTT ACC CGT ATG GGG TGG CTG GTG GTC GGG ATC GGT CTG TTC  1929
GGG TTT TTA GCC TGG GCG GCC TTT GCG CCG TTG GAT AAA GGG GTG GCG  1977
TCG CCG GGA TCG GTA ACC GTT TCC GGC AAC CGC AAA ACG GTG CAG GCC  2025
CCG GCC AGC GGC ATC ATT AAG AAT ATT GCG GTC AGA GAT GGC GAC AAA  2073
GTG AAA GCC GGT GAG GTG CTG GTG CAG CTC AGC CAG GTG CAG GCT CAA  2121
GCT CAG GTT GAT TCG CTG CGG GAT CAG TAC TAC ACC ACG CTG GCG ACA  2169
GAA GGG CGC TTG CTG GCA GAA CGC GAT GGG TTG AGC ATA GTG ACT TTC  2217
TCA CCC ATT TTG GAC GCG GTG AAA GAT AAA CCT CGC GTG GCA GAA ATC  2265
ATT GCA TTG CAA ACG CAG CTG TTC GCC TCC CGC CGC CAA GCG CTG CAA  2313
AGT GAA ATC GAC GGC TAT AAG CAG TCA ATG GAC GGA ATC CGT TTC CAA  2361
TTA AAA GGA CTG CAG GAT TCG CGC GGT AAC AAA CAG ATC CAG CTT TCC  2409
AGC CTG CGT GAG CAG ATG AAC AGC ATG AAG CAG TTG GCG GCG GAC GGT  2457
TAC CTA CCG CGT AAC CGT TAC CTG GAA GTG CAG CGC CAG TTT GCC GAG  2505
GTA AAT AGC AGC ATT GAT GAA ACG GTG GGG CGG ATT GGC CAA TTG CAA  2553
AAG CAG TTG CTG GAA TCA CAG CAA CGC ATC GAT CAG CGT TTC GCC GAC  2601
TAC CAG CGC GAA GTC AGA ACG CAG CTG GCG CAA ACT CAA ATG GAC GCC  2649
AGC GAA TTC CGC AAC AAG CTG CAA ATG GCC GAT TTC GAT CTG GGC AAC  2697
ACC GCC ATC ACC TCA CCG GTG GAC GGC ACC GTG GTT GGA TTG AAT ATC  2745
TTC ACT CAG GGG GGC GTC GTG GGA GCG GGT GAC CAC CTG ATG GAC GTT  2793
GTG CCC AGC CAG GCG ACT TTG GTG GTG GAT TCT CGC CTC AAA GTC GAC  2841
CTG TTC GAT AAG GTG TAC AAC GGG TTG CCG GTG GAT CTG ATG TTT ACC  2889
GCC TTC AAC CAA AAC AAA ACC CCG AAA ATT CCG GGA ACC GTC ACC TTG  2937
```

GTT TCC GCC GAC CGC CTG GTC GAC AAA GCC AAT GGC GAA CCT TAC TAC 2985
CAG ATG CAG GTC ACG GTC TCG CCG GAG GGC ATG AAA ATG CTC AGT GGC 3033
GAG GAC ATC AAG CCG GGG ATG CCG GTG GAG GTG TTC GTG AAA ACG GGG 3081
TCG CGC TCG CTG TTG AGC TAT CTG TTT AAA CCT ATT TTG GAT CGC GCT 3129
CAT ACT TCA TTA ACC GAG GAA TAA TT TTG ATT CAT TCA AAA CGA CAG 3176
GCT GCC GGT CTG GTT ATC GGC ACC CTT TTG TTT GCG ATG TCT GCG CCG 3224
GTT TAT TCG ATA GGG ATT TTA GAC GCA TAT TCG CTG GCA TTA GAA AAG 3272
GAC CCG ACC TTT CGG GCG GCT ATA AAA GAG AAA GAA GCG GGA GAT GAA 3320
AAC GAA AAT ATC GGC AGG GCA GGG CTG CTG CCG AAG GTA TCG CTG AAC 3368
TAC CAG AAT TCG CCG CGC AAC TGG CAA ACT CAG AAG TAC CCG CAA AGC 3416
GAC TTT TTC GGC AAT GTT TCG GAG GTT ACC CGG CGG CAG CAA TAT CGC 3464
AGC TAT TCC AGT TCG ATC ACC TTG ACG CAG CCG CTG TTC GAT TAT GAA 3512
GCT TAC GCC AGG TAC AAA GCC GGC GTG GCG CAG ACC ATG ATG TCG GAC 3560
GAG ACG TAT CGC GGT AAG TTG CTG GAT TTG GCG GTG AGG GTG ATT AAC 3608
GCC TAT GTC GAA GTG GCT TAT TCC AAG GAT CAA ATC GCG TTG GCC GAA 3656
GCT CAA AAG GCG GCT TAC AAG GAA CAG CTG ACT TTG AAC GAT CGC CTG 3704
ATG AGC GCC GGT GAA GGT ACC ATT ACC GAC GTA TCC GAA ACT CAG GCG 3752
CGC TAT AGC CTG GCT GAA GCA CAG GTG ATA GAA GCG CGC GAT GCA CTG 3800
GAT GCC GCA CAG CGT GAA TTG GAA GTA ATT ATC GGC ATG CCG CTG AAC 3848
CAA CTG GAT GAA TTG CAG GTC TTG CGG CCG GGT AAA TTC AAA GTG GCG 3896
CCG TTA ATC CCG TCC AAG TTC GAA GAG TGG CAA AAG ATC GCG TTG GAG 3944
AAC AAC CCT GTA TTG GCC GCC TCG CGT CAT GGC GTG GAT GCT GCT AAG 3992
TAT GAT GTC GAA AGG AAA CGG GCT GGC TTT ATG CCA CAG GTT CAG CTG 4040
TAT GCT TCG CAT TCG GAA AAC GAC GCC AGC AGC GAC AAC ACG GTT AAC 4088
CAG AAA TAC CGT ACT GAC AGC ATT GGC GTG CAG GTC AGC ATG CCT ATT 4135
TAT TCC GGT GGC GGT GTT TCG GCA TCG ACC CGC CAG GCA GCG GCG CGT 4184
TAC GGG CAA GCG ATG TAT GAA ATG GAT GCG CAA ACG GGC ACC ACG CTC 4232
AAC GAT CTG CGC AAA CAG TAC AAT TTG TGT ATT AGC AGC AGC GCT AAA 4280
GTG GCG GCC TAT GAA CTG GCG GTT CAA TCG GCG ACG ACC CAG GTG ACG 4328
GCG ACC CGG CAA AGC GTG CTG GCT GGG CAA CGT GTC AAC GTC GAT GTG 4376
CTC AAT GCC GAA CAG CAG CTC TAT AGC GCA CAG GCG ATT TTG GCC TCT 4424
GCT AAA TAC ACT TAT ATC AAA TCC TGG GAT CAC CCT ATT GAG 4466

序列身份号：2
长度：588氨基酸
类型：氨基酸
局部解剖：线形
分子类型：肽
序列

```
Val Asn Gln Phe Ile Pro Arg Asn Glu Ile Ala Asp Val Ile Arg Thr
 1               5                  10                  15
Arg Ser Lys Val Phe Trp Thr Val Gly Ile Phe Thr Ala Phe Ile Asn
            20                  25                  30
Leu Leu Met Leu Val Pro Ser Ile Tyr Met Leu Gln Val Tyr Asp Arg
        35                  40                  45
Val Leu Pro Ser Arg Asn Glu Ile Thr Leu Leu Met Leu Thr Leu Ile
    50                  55                  60
Met Leu Gly Met Phe Gly Met Met Ser Leu Leu Glu Tyr Val Arg Ser
65                  70                  75                  80
Met Val Val Ile Arg Ile Gly Ser Gln Leu Asp Met Arg Leu Asn Thr
                85                  90                  95
Arg Val Tyr Thr Ala Ala Tyr Glu Ala Asn Leu Lys Asn Gly Ser Ser
            100                 105                 110
Asp Ala Gly Gln Met Leu Ser Asp Leu Thr Asn Leu Arg Gln Phe Leu
        115                 120                 125
Thr Gly Ser Ala Leu Phe Ala Phe Phe Asp Ala Pro Tyr Phe Pro Ile
    130                 135                 140
Tyr Leu Leu Val Ile Phe Leu Phe Asn Pro Trp Leu Gly Leu Phe Ala
145                 150                 155                 160
Leu Val Gly Ala Leu Leu Leu Ile Ala Leu Ala Val Ile Asn Glu Val
                165                 170                 175
Val Ser Lys Lys Pro Leu Gly Glu Ala Ser Lys Leu Ser Ile Met Ser
            180                 185                 190
Gly Asn Leu Ala Ser Thr Asn Leu Arg Asn Ala Glu Val Ile Glu Ala
        195                 200                 205
Leu Gly Met Leu Pro Asn Leu Lys Arg Arg Trp Phe Gly Leu His Gln
    210                 215                 220
Arg Phe Leu Asn Ser Gln Arg Ile Ala Ser Glu Arg Ala Ser Arg Val
225                 230                 235                 240
Thr Ser Ile Thr Lys Phe Val Arg Met Ser Leu Gln Ser Leu Val Leu
```

```
                245                 250                 255
Gly Leu Gly Gly Trp Leu Ala Ile Asp Gly His Ile Thr Pro Gly Met
            260                 265                 270
Met Ile Ala Gly Ser Ile Leu Met Gly Arg Thr Leu Ala Pro Ile Glu
        275                 280                 285
Gln Val Ile Asn Val Trp Lys Ser Tyr Ser Ala Ala Lys Leu Ser Tyr
    290                 295                 300
Gly Arg Leu Val Lys Leu Leu Glu Thr His Pro Gln Arg Gly Thr Gly
305                 310                 315                 320
Met Ser Leu Pro Arg Pro Glu Gly Val Leu Ser Val Glu Gly Val Thr
                325                 330                 335
Ala Thr Pro Pro Gly Ser Lys Gly Asp Ala Val Leu His Asn Val Ser
            340                 345                 350
Phe Ala Ile Gln Pro Gly Asp Val Leu Gly Ile Ile Gly Pro Ser Ala
        355                 360                 365
Ser Gly Lys Ser Thr Leu Ala Arg Leu Leu Val Gly Ile Trp Pro Val
    370                 375                 380
Ser Glu Gly Ile Val Arg Leu Asp Asn Ala Asp Ile Tyr Gln Trp Asn
385                 390                 395                 400
Lys Asp Glu Leu Gly Pro Tyr Ile Gly Tyr Leu Pro Gln Asp Ile Glu
                405                 410                 415
Leu Phe Ala Gly Thr Ile Ala Glu Asn Ile Ala Arg Phe Asn Asp Ile
            420                 425                 430
Asp Ser Glu Lys Val Ile Glu Ala Ala Lys Leu Ala Gly Val His Glu
        435                 440                 445
Leu Ile Leu Arg Phe Pro Asn Gly Tyr Asp Ser Val Ile Gly Asn Gly
    450                 455                 460
Gly Ala Gly Leu Ser Gly Gly Gln Lys Gln Arg Ile Gly Leu Ala Arg
465                 470                 475                 480
Ala Leu Tyr Gly Asp Pro Ala Leu Val Val Leu Asp Glu Pro Asn Ser
                485                 490                 495
Asn Leu Asp Asp Ala Gly Glu Lys Ala Leu Asn Gln Ala Ile Met Phe
            500                 505                 510
Leu Lys Gln Arg Asn Lys Thr Val Val Leu Ile Thr His Arg Thr Asn
        515                 520                 525
```

```
Leu Leu Ser Met Thr Ser Lys Leu Leu Leu Leu Val Asn Gly Asn Val
    530                 535                 540
Asn Ala Phe Gly Pro Thr Gln Gln Val Leu Gln Ala Leu Ala Asn Ala
545                 550                 555                 560
Gln Lys Ala Gln Val Pro Pro Gln Ala Val Arg Ala Val Asn Ser Glu
                565                 570                 575
Pro Asp Glu Gly Glu Ile Pro Lys Thr Gln Ile Asn
            580                 585
```

序列身份号：3
长度：443氨基酸
类型：氨基酸
局部解剖：线形
分子类型：肽
序列

```
Met Ser Thr His Ile Gly Glu Pro Gln Asp Ser Tyr Thr Glu Glu Ile
 1               5                  10                  15
Pro Gln Asp Glu Arg Arg Phe Thr Arg Met Gly Trp Leu Val Val Gly
            20                  25                  30
Ile Gly Leu Phe Gly Phe Leu Ala Trp Ala Ala Phe Ala Pro Leu Asp
        35                  40                  45
Lys Gly Val Ala Ser Pro Gly Ser Val Thr Val Ser Gly Asn Arg Lys
    50                  55                  60
Thr Val Gln Ala Pro Ala Ser Gly Ile Ile Lys Asn Ile Ala Val Arg
65                  70                  75                  80
Asp Gly Asp Lys Val Lys Ala Gly Glu Val Leu Vla Gln Leu Ser Gln
                85                  90                  95
Val Gln Ala Gln Ala Gln Val Asp Ser Leu Arg Asp Gln Tyr Tyr Thr
            100                 105                 110
Thr Leu Ala Thr Glu Gly Arg Leu Leu Ala Glu Arg Asp Gly Leu Ser
        115                 120                 125
Ile Val Thr Pro Ser Phe Ile Leu Asp Ala Val Lys Asp Lys Pro Arg
    130                 135                 140
Val Ala Glu Ile Ile Ala Leu Gln Thr Gln Leu Phe Ala Ser Arg Arg
145                 150                 155                 160
```

Gln Ala Leu Gln Ser Glu Ile Asp Gly Tyr Lys Gln Ser Met Asp Gly
165 170 175
Ile Arg Phe Gln Leu Lys Gly Leu Gln Asp Ser Arg Gly Asn Lys Gln
180 185 190
Ile Gln Leu Ser Ser Leu Arg Glu Gln Met Asn Ser Met Lys Gln Leu
195 200 205
Ala Ala Asp Gly Tyr Leu Pro Arg Asn Arg Tyr Leu Glu Val Gln Arg
210 215 220
Gln Phe Ala Glu Val Asn Ser Ser Ile Asp Glu Thr Val Gly Arg Ile
225 230 235 240
Gly Gln Leu Gln Lys Gln Leu Leu Glu Ser Gln Gln Arg Ile Asp Gln
245 250 255
Arg Phe Ala Asp Tyr Gln Arg Glu Val Arg Thr Gln Leu Ala Gln Thr
260 265 270
Gln Met Asp Ala Ser Glu Phe Arg Asn Lys Leu Gln Met Ala Asp Phe
275 280 285
Asp Leu Gly Asn Thr Ala Ile Thr Ser Pro Val Asp Gly Thr Val Val
290 295 300
Gly Leu Asn Ile Phe Thr Gln Gly Gly Val Val Gly Ala Gly Asp His
305 310 315 320
Leu Met Asp Val Val Pro Ser Gln Ala Thr Leu Val Val Asp Ser Arg
325 330 335
Leu Lys Val Asp Leu Phe Asp Lys Val Tyr Asn Gly Leu Pro Val Asp
340 345 350
Leu Met Phe Thr Ala Phe Asn Gln Asn Lys Thr Pro Lys Ile Pro Gly
355 360 365
Thr Val Thr Leu Val Ser Ala Asp Arg Leu Val Asp Lys Ala Asn Gly
370 375 380
Glu Pro Tyr Tyr Gln Met Gln Val Thr Val Ser Pro Glu Gly Met Lys
385 390 395 400
Met Leu Ser Gly Glu Asp Ile Lys Pro Gly Met Pro Val Glu Val Phe
405 410 415
Val Lys Thr Gly Ser Arg Ser Leu Leu Ser Tyr Leu Phe Lys Pro Ile
420 425 430

Leu Asp Arg Ala His Thr Ser Leu Thr Glu Glu
435 440

序列身份号：4
长度：437氨基酸
类型：氨基酸
局部解剖：线形
分子类型：肽
序列

Leu Ile His Ser Lys Arg Gln Ala Ala Gly Leu Val Ile Gly Thr Leu
1 5 10 15
Leu Phe Ala Met Ser Ala Pro Val Tyr Ser Ile Gly Ile Leu Asp Ala
20 25 30
Tyr Ser Leu Ala Leu Glu Lys Asp Pro Thr Phe Arg Ala Ala Ile Lys
35 40 45
Glu Lys Glu Ala Gly Asp Glu Asn Glu Asn Ile Gly Arg Ala Gly Leu
50 55 60
Leu Pro Lys Val Ser Leu Asn Tyr Gln Asn Ser Pro Arg Asn Trp Gln
65 70 75 80
Thr Gln Lys Tyr Pro Gln Ser Asp Phe Phe Gly Asn Val Ser Glu Val
85 90 95
Thr Arg Arg Gln Gln Tyr Arg Ser Tyr Ser Ser Ser Ile Thr Leu Thr
100 105 110
Gln Pro Leu Phe Asp Tyr Glu Ala Tyr Ala Arg Tyr Lys Ala Gly Val
115 120 125
Ala Gln Thr Met Met Ser Asp Glu Thr Tyr Arg Gly Lys Leu Leu Asp
130 135 140
Leu Ala Val Arg Val Ile Asn Ala Tyr Val Glu Val Ala Tyr Ser Lys
145 150 155 160
Asp Gln Ile Ala Leu Ala Glu Ala Gln Lys Ala Ala Tyr Lys Glu Gln
165 170 175
Leu Thr Leu Asn Asp Arg Leu Met Ser Ala Gly Glu Gly Thr Ile Thr
180 185 190
Asp Val Ser Glu Thr Gln Ala Arg Tyr Ser Leu Ala Glu Ala Gln Val
195 200 205

```
Ile Glu Ala Arg Asp Ala Leu Asp Ala Ala Gln Arg Gln Leu Glu Val
    210                 215                 220
Ile Ile Gly Met Pro Leu Asn Gln Leu Asp Glu Leu Gln Val Leu Arg
225                 230                 235                 240
Pro Gly Lys Phe Lys Val Ala Pro Leu Ile Pro Ser Lys Phe Glu Glu
                245                 250                 255
Trp Gln Lys Ile Ala Leu Glu Asn Asn Pro Val Leu Ala Ala Ser Arg
            260                 265                 270
His Gly Val Asp Ala Ala Lys Tyr Asp Val Glu Arg Lys Arg Ala Gly
        275                 280                 285
Phe Met Pro Gln Val Gln Leu Tyr Ala Ser His Ser Glu Asn Asp Ala
    290                 295                 300
Ser Ser Asp Asn Thr Val Asn Gln Lys Tyr Arg Thr Asp Ser Ile Gly
305                 310                 315                 320
Val Gln Val Ser Met Pro Ile Tyr Ser Gly Gly Gly Val Ser Ala Ser
                325                 330                 335
Thr Arg Gln Ala Ala Ala Arg Tyr Gly Gln Ala Met Tyr Glu Met Asp
            340                 345                 350
Ala Gln Thr Gly Thr Thr Leu Asn Asp Leu Arg Lys Gln Tyr Asn Leu
        355                 360                 365
Cys Ile Ser Ser Ser Ala Lys Val Ala Ala Tyr Glu Leu Ala Val Gln
    370                 375                 380
Ser Ala Thr Thr Gln Val Thr Ala Thr Arg Gln Ser Val leu Ala Gly
385                 390                 395                 400
Gln Arg Val Asn Val Asp Val Leu Asn Ala Glu Gln Gln Leu Tyr Ser
                405                 410                 415
Ala Gln Ala Ile Leu Ala Ser Ala Lys Tyr Thr Tyr Ile Lys Ser Trp
            420                 425                 430
Asp His Pro Ile Glu
        435
```

# 说 明 书 附 图

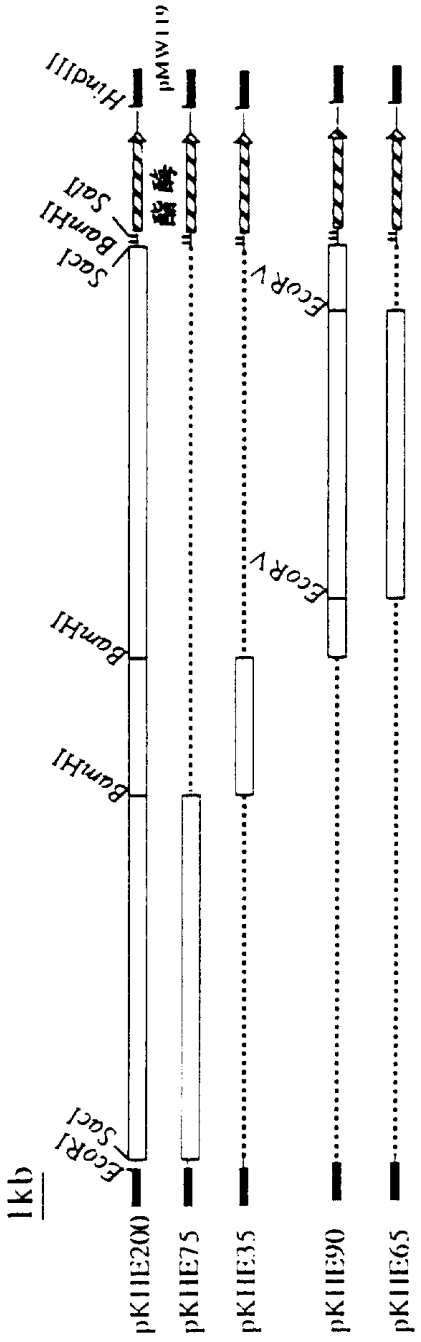

pMW119
HindIII
SacI
BamHI
SalI
EcoRV
EcoRV
BamHI
BamHI
EcoRI
SacI
1kb
pK11E200
pK11E75
pK11E35
pK11E90
pK11E65

图1