[19]中华人民共和国国家知识产权局

[51]Int.Cl[7]
C12N 15/56
C12N 9/28 C12Q 1/40
C11D 3/386

[12]发明专利申请公开说明书

[21]申请号 98801431.9

[43]公开日 2000年3月1日

[11]公开号 CN 1246153A

[22]申请日 1998.8.14 [21]申请号 98801431.9
[30]优先权
[32]1997.8.19 [33]US [31]08/914,679
[86]国际申请 PCT/US98/16906 1998.8.14
[87]国际公布 WO99/09183 英 1999.2.25
[85]进入国家阶段日期 1999.5.27
[71]申请人 金克克国际有限公司
地址 美国纽约州
[72]发明人 A·戴 B·斯旺森
[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 李 瑛

权利要求书2页 说明书34页 附图页数9页

[54]发明名称 包括相应于地衣芽孢杆菌A210、H405和/或、T412的残基的修饰的突变α-淀粉酶

[57]摘要

本发明公开了在与地衣芽孢杆菌A210、H405及T412位点相应有一个或多个残基发生突变的α-淀粉酶。公开的α-淀粉酶显示有改变或改良的稳定性和/或活性状态。

ISSN1008-4274

权　利　要　求　书

1. 突变体α-淀粉酶，它是通过相应于地衣芽孢杆菌α-淀粉酶A210、H405和/或T412的前体α-淀粉酶残基的删除、替换或添加产生的。

2. 根据权利要求1所述的突变体α-淀粉酶，其中所述的突变包括在地衣芽孢杆菌α-淀粉酶A210、H405和/或T412的两个或以上的位点进行的删除、替换或添加。

3. 根据权利要求2所述的α-淀粉酶，其中所述的α-淀粉酶包括相应于地衣芽孢杆菌α-淀粉酶A210T/H405D/T412A的残基的替换。

4. 根据权利要求1所述的α-淀粉酶，其中所述的α-淀粉酶来源是细菌或真菌。

5. 根据权利要求1所述的α-淀粉酶，其中所述的α-淀粉酶来源是芽孢杆菌。

6. 根据权利要求5所述的α-淀粉酶，其中所述的α-淀粉酶来源是地衣芽孢杆菌、嗜热脂肪芽孢杆菌或解淀粉芽孢杆菌。

7. 根据权利要求1所述的α-淀粉酶，其中所述的α-淀粉酶进一步包括相应于地衣芽孢杆菌α-淀粉酶M15、A33、A52、S85、N96、V129、H133、S148N、S187、N188、A209、A269和/或A379的残基的删除或替换。

8. 根据权利要求1所述的α-淀粉酶，其中的替换进一步包括相应于地衣芽孢杆菌M15T、W138Y和/或M179T的残基的替换或删除。

9. 编码根据权利要求1所述的α-淀粉酶的DNA。

10. 包含权利要求9的DNA的表达载体。

11. 使用权利要求10的表达载体转化的宿主细胞。

12. 根据权利要求1、3或7所述的具有增强的低pH性能和/或提高的温度稳定性的α-淀粉酶。

13. 包含根据权利要求1所述的α-淀粉酶的去污剂组合物。

14. 根据权利要求13所述的去污剂组合物，其中所述的去污剂可

用于清洁脏衣物和/或脏碗盘。

15. 一种液化淀粉的方法，该方法包括使淀粉浆与根据权利要求 1 所述的 α-淀粉酶接触。

说　　明　　书

包括相应于地衣芽孢杆菌 A210、H405 和/或 T412 的残基的修饰的突变α-淀粉酶

发明领域

本发明涉及其中引入了突变，在某些特定状况下提供额外的稳定性的α-淀粉酶。尤其期望突变体将有改变的性能特征，如改变的稳定性和/或改变的活性状态。

发明背景

α-淀粉酶(α-1，4-葡聚糖-4-葡聚糖水解酶，EC3.2.1.1)大部分随机水解淀粉内的内部α-1，4-糖苷键，产生分子量比较小的麦芽糖糊精。α-淀粉酶有相当大的商业价值，通常用于淀粉加工的初始阶段(液化)；酒精生产；去污剂基质中的清洁剂及纺织业中的退浆过程。α-淀粉酶由多种微生物制造，包括芽孢杆菌及曲霉。大部分商业淀粉酶的细菌来源通常是地衣芽孢杆菌、解淀粉芽孢杆菌、枯草芽孢杆菌或嗜热脂肪芽孢杆菌等。近年来，优选的商用酶通常来自地衣芽孢杆菌，因为它们在高温中稳定，并在商业操作条件下性能良好。

通常淀粉至果糖的加工过程有四个步骤：液化颗粒淀粉、糖化液态淀粉为葡萄糖、纯化及异构化为果糖。淀粉液化过程的目的是转变浓缩的淀粉聚合颗粒悬浮液为可溶性、低粘度的短链糊精溶液。这一步骤对于方便地使用标准设备和有效地转变为葡萄糖或其它糖非常重要。为了液化颗粒淀粉，必须将颗粒淀粉的温度增加至摄氏 72 度以上，使颗粒凝胶化。加热过程立即分裂不能溶解的淀粉颗粒，产生水溶性的淀粉液。溶化的淀粉液然后由α-淀粉酶(EC3.2.1.1)液化。

常用的酶液化过程需要添加氢氧化钙、氢氧化钠或碳酸钠以便调整颗粒淀粉浆的 pH 至 6.0-6.5 之间，这是地衣芽孢杆菌α-淀粉酶的

最适 pH。添加氢氧化钙的优点是同时提供钙离子，使α-淀粉酶稳定不至失活。添加α-淀粉酶后，悬浮液被泵过蒸汽喷嘴，使其温度立即提升至摄氏 80-115 度。淀粉立即凝胶化，由于α-淀粉酶的存在，淀粉通过(1-4)糖苷键的随机水解解聚成为易于泵送的液体。

液化过程的第二种方法是在淀粉悬浮液内添加α-淀粉酶，悬浮液保持在摄氏 80-100 度，使淀粉颗粒部分水解；部分水解的淀粉悬浮液被泵过温度超过约摄氏 105 度的喷嘴，使剩余的颗粒完全凝胶化。凝胶化淀粉冷却后，可作第二次添加α-淀粉酶以便进一步水解淀粉。

液化过程的第三种方法称为干磨过程。在干磨过程中，整谷粒被研磨后加水。胚芽可选择性地通过漂浮分离或类似技术除去。所得的混合物包括淀粉、纤维、蛋白质及谷粒的其它成分将使用α-淀粉酶进行液化。通常，在使用干磨过程时，酶液化将在较低的温度下进行。淀粉转变为可溶性糊精过程中使用低温液化的效率不如高温液化。

通常，在凝胶化后，淀粉液在添加α-淀粉酶后保持高温直至葡萄糖当量达到 10-20，这通常需要 1-3 小时。葡萄糖当量(DE)是量度总还原糖浓度的工业标准，以 D-葡萄糖的干重量为计算基础。非水解的颗粒淀粉的葡萄糖当量几乎是零，而 D-葡萄糖的葡萄糖当量为 100。

含α-淀粉酶的淀粉液可以保持的最高温度视酶的微生物来源及α-淀粉酶分子的分子结构而定。使用野生型枯草芽孢杆菌或解淀粉芽孢杆菌菌株所生产的α-淀粉酶的操作温度通常不超过约摄氏 90 度，因为在该温度以上将产生过度快速的热失活，而使用野生型地衣芽孢杆菌菌株生产的α-淀粉酶则可以在高达约摄氏 110 度的温度下操作。已知淀粉和钙离子的存在可稳定α-淀粉酶不至失活。然而，α-淀粉酶是在 pH 值 6 以上的环境下使用，以防快速失活。在低温下，来自地衣芽孢杆菌的α-淀粉酶可在 pH 值低于 5 的环境下对淀粉底物显示出水解活性。不过，当α-淀粉酶在普通喷射温度，例如摄氏 102 度至 109 度之间进行淀粉水解时，pH 值必须保持在至少 5.7 以上以防过度快速失活。pH 的要求使加工可能性减少，因为 pH 值 6.0 以上将产生不良的副产品，如 maltulose。因此，事实上，液化过程中的 pH

通常保持在 5.9 及 6.0 之间，以便得到令人满意的水解淀粉产量。

与液化的 pH 有关的另一个问题是必须将来自湿磨步骤的玉米淀粉悬浮液的 pH 从约 4 增加至 5.9-6.0。此 pH 调整需要添加昂贵的酸性中和化学品，及在最后的淀粉产品中进行附加的离子交换精制步骤以便去除化学品。此外，液化后的步骤通常是使用葡糖淀粉酶将液化淀粉糖化为葡萄糖，此步骤需要 4-4.5 的 pH；因此 pH 必须从 5.9-6.0 下调至 4-4.5，因而需要添加额外的化学品及精制步骤。

液化后，加工的淀粉由葡糖淀粉酶糖化为葡萄糖。目前加工过程的一个问题是由于淀粉不完全液化，例如淀粉酶所做的直链淀粉水解效率低，使糖化混合物内含残留淀粉。残留淀粉不容易被葡糖淀粉酶所水解。这将使产量减少，并干扰糖浆的顺流过滤。

此外，很多 α-淀粉酶都需要添加钙离子使其稳定，这进一步增加液化成本。

美国专利号 5，322，778 添加亚硫酸氢盐或同类盐、抗坏血酸或同类盐，异抗坏血酸等抗氧化剂，或丁基化羟茴香醚，2，6-二叔丁基对甲酚或 α-生育酚等酚类抗氧化剂至液化淤浆以便在 pH4.0 至 6.0 之间达到液化。根据本专利，添加的亚硫酸氢钠浓度必须在 5mM 以上。

美国专利号 5，180，669 添加超过缓冲所需的碳酸盐离子至研磨淀粉淤浆以便在 pH5.0 至 6.0 之间达到液化。添加碳酸盐离子会增加 pH，因此淤浆通常通过添加如盐酸或硫酸等无机酸作为氢离子源被中和。

PCT 公开号 WO95/35382 描述地衣芽孢杆菌 α-淀粉酶在 104、128、187 和/或 188 位点有改变的 α-淀粉酶突变体有改善的氧化稳定性。

PCT 公开号 WO96/23873 描述有不同突变的 α-淀粉酶突变体。

PCT 公开号 WO94/02597 描述 α-淀粉酶突变体内的一或多个蛋氨酸由半胱氨酸或蛋氨酸以外的任何氨基酸取代时，其氧化稳定性有所改善。

PCT 公开号 WO94/18314 描述 α-淀粉酶突变体内的一或多个蛋氨

酸、色氨酸、半胱氨酸、组氨酸或酪氨酸残基由不可氧化的氨基酸取代时，其氧化稳定性有所改善。

PCT 公开号 WO91/00353 描述与使用野生型地衣芽孢杆菌α-淀粉酶进行液化相关的性能特征和问题可经由对α-淀粉酶进行基因工程以包括特异性替换 Ala-111-Thr、His-133-Tyr 和/或 Thr-149-Ile 来处理。

使用重组 DNA 技术探讨哪些残基对淀粉酶的催化作用有重要性和/或探讨修饰不同淀粉酶及葡糖酶活性位点内的某些氨基酸的效果的研究已由多位科学家进行(Vihinen 等人，生物化学杂志，第 107 卷，第 267-272 页(1990)；Holm 等人，蛋白质工程，第 3 卷，第 181-191 页(1990)；Takase 等人，生物化学与生物物理学学报，第 1120 卷，第 281-288 页(1992)；Matsui 等人，FEBS 通讯，第 310 卷，第 216-218 页(1992)；Matsui 等人，生物化学，第 33 卷，第 451-458 页(1992)；Sogaard 等人，生物化学杂志，第 268 卷，第 22480-22484 页(1993)；Sogaard 等人，碳水化合物聚合物，第 21 卷，第 137-146 页(1993)；Svensson，植物分子生物学，第 25 卷，第 141-157 页(1994)；Svensson 等人，生物技术杂志，第 29 卷，第 1-37 页(1993))。此外，科学家也研究了哪些残基对热稳定性重要(Suzuki 等人，生物化学杂志，第 264 卷，第 18933-18938 页(1989)；Watanabe 等人，欧洲生物化学杂志，第 226 卷，第 277-283 页(1994))；其中一组科学家使用此等方法在地衣芽孢杆菌淀粉酶不同组氨酸残基上引入突变，根据的理论是与其它类似芽孢杆菌淀粉酶比较，地衣芽孢杆菌淀粉酶比较有耐热力、有过量的组氨酸，因此取代组氨酸可能影响酶的耐热力。此研究结果发现在组氨酸残基的+133 位点及丙氨酸残基的+209 位点有稳定性的突变(Declerck 等人，生物化学杂志，第 265 卷，第 15481-15488 页(1990)；FR2 665 178-A1；Joyet 等人，生物技术，第 10 卷，第 1579-1583 页(1992))。

虽然现有技术有所进展，仍然有必要研发一种在商业液化过程更有效的α-淀粉酶，但容许在比目前可行的 pH 低的环境下操作。此外，

有必要研发有改良性能的淀粉酶，使它们在使用去污剂的条件下更为有效。因为稳定性问题使商用淀粉酶在很多条件下都不能被接受，例如与去污剂有关的高碱度及氧化剂(漂白剂)水平、或它们操作环境的温度，因此有必要研发一种在这些条件下有更高性能的改良淀粉酶。

发明概要

本发明的目的是提供一种有改变性能的α-淀粉酶。

本发明的另一目的是提供一种在高温下有改善的稳定性的α-淀粉酶。

因此，本发明提供一种在与地衣芽孢杆菌α-淀粉酶 A210、H405 和/或 T412 位点相应的残基上添加、取代或删除以引入突变的α-淀粉酶。本发明的优选实施方案是从细菌或真菌来源产生α-淀粉酶，其包含相应于地衣芽孢杆菌的替换残基。最理想的情况是α-淀粉酶从芽孢杆菌产生，而突变则与地衣芽孢杆菌 A210T、H405D 和/或 T412A 位点相应。

此外，本发明包括编码这种突变α-淀粉酶的核酸、含此核酸的载体、以此载体所转化的宿主细胞及使用这些宿主细胞表达突变α-淀粉酶的方法。

本发明还包括使用本发明的突变α-淀粉酶在淀粉加工过程中液化淀粉成葡萄糖或其它淀粉衍生物、作为去污剂中的添加剂如洗衣或洗碗剂、烘焙辅助剂及用于纺织品退浆。

附图简单说明

图 1 显示地衣芽孢杆菌(NCIB8061)α-淀粉酶基因的 DNA 序列(SEQ ID NO: 1)，及根据 Gray 等人，细菌学杂志，第 166 卷，第 635-643 页(1986)的说明推定的转译产物的氨基酸序列(SEQ ID NO: 2)。

图 2 显示地衣芽孢杆菌成熟的α-淀粉酶的氨基酸序列(SEQ ID NO: 3)。

图 3 显示三种芽孢杆菌α-淀粉酶一级结构的序列对比。地衣芽孢

杆菌α-淀粉酶(Am-Lich)(SEQ ID NO:4)在 Gray 等人，细菌学杂志，第 166 卷，第 635-643 页(1986)中有描述；解淀粉芽孢杆菌α-淀粉酶(Am-Amylo)(SEQ ID NO:5)在 Takkinen 等人，生物化学杂志，第 258 卷，第 1007-1013 页(1983)中有描述；嗜热脂肪芽孢杆菌α-淀粉酶(Am-Stearo)(SEQ ID NO:6)在 Ihara 等人，生物化学杂志，第 98 卷，第 95-103 页(1985)中有描述。

图 4 显示质粒 pHP13，其中 Cm^R 指的是氯霉素抗性，Em^R 指的是红霉素抗性，ReppTA1060 指的是质粒 pTA1060 的复制起点。

图 5 显示质粒 pBLapr，其中 BLAA 指的是地衣芽孢杆菌α-淀粉酶基因；aprE 指的是 aprE 基因的启动子及信号肽编码区域；AmpR 指的是 BR322 的氨苄青霉素抗性基因；而 CAT 指的是 pC194 氯霉素抗性基因。

图 6 显示携带地衣芽孢杆菌α-淀粉酶基因的 pHP.BL 质粒。

详细说明

“α-淀粉酶”指的是切割或水解α(1-4)糖苷键(例如淀粉、支链淀粉或直链淀粉聚合物中的)的酶活性。本发明所用的α-淀粉酶包括自然产生的α-淀粉酶及重组的α-淀粉酶。本发明优选的α-淀粉酶来自地衣芽孢杆菌、解淀粉芽孢杆菌或嗜热脂肪芽孢杆菌及来自曲霉(即米曲霉及黑曲霉)的真菌α-淀粉酶。

“重组α-淀粉酶”指的是编码自然产生的α-淀粉酶的 DNA 序列经修饰产生突变 DNA 序列的α-淀粉酶，与自然产生的α-淀粉酶比较，此突变 DNA 序列编码α-淀粉酶序列内一个或多个氨基酸的替换、插入或删除。

“表达载体”指的是包含可操作地连接到一个可在适当宿主内表达所述 DNA 的适当控制序列上的 DNA 序列的 DNA 构建体。此控制序列可以包括转录启动子、控制转录的选择性操纵基因序列、编码适合 mRNA 核糖体结合位点的序列及控制转录与转译终止的序列。优选的启动子是枯草芽孢杆菌 aprE 启动子。载体可以是质粒、噬菌体微粒或潜

在的基因组插入物。一旦转化到适当的宿主中，载体便可独立于宿主基因组复制及运作，或在某些情况下，可整合至宿主基因组内。在本说明书中，质粒与载体有时交替使用，因为质粒是目前最常用形式的载体。不过，本发明将包括其它形式的提供同等功能、在本领域中已知，或以后已知的表达载体。

“宿主菌株”或“宿主细胞”指的是包含编码本发明的α-淀粉酶的 DNA 的表达载体的适当宿主。本发明适用的宿主细胞通常为原核或真核宿主，包括任何可表达本发明的α-淀粉酶的可转化微生物。具体地说，与产生α-淀粉酶同种或同属的宿主菌株(如芽孢杆菌菌株)都是适合的宿主。优选使用的是α-淀粉酶阴性芽孢杆菌菌株(删除基因)和/或删除α-淀粉酶及蛋白酶的芽孢杆菌菌株(Δ amyE、Δ apr、Δ npr)。宿主细胞的转化或转染是以使用重组 DNA 技术构建的载体进行的。经过这样转化的宿主细胞可复制编码α-淀粉酶及其变异体(突变体)的载体或表达所需的α-淀粉酶。

“液化”指的是淀粉转变为较短的链及粘性比较低的糊精的过程。通常，此过程包括在添加α-淀粉酶的同时或之后将淀粉凝胶化。

根据本发明，提供在 A210、H405 和/或 T412 引入替代、添加或删除的突变α-淀粉酶。本发明中所述删除、添加或替代氨基酸指的是α-淀粉酶前体氨基酸序列的任何修饰，但优选指的是使用基因工程使编码前体α-淀粉酶的核酸产生突变，以便编码在表达的蛋白质中的删除、替换或添加的残基。前体α-淀粉酶包括自然产生的α-淀粉酶及重组α-淀粉酶。修饰编码前体α-淀粉酶氨基酸序列的前体 DNA 序列可使用本发明及共同拥有的美国专利第 4，760，025 号及 5，185，258 号(并入本发明以供参考)所述的方法。

此外，提供编码包含本发明所提供的突变α-淀粉酶的氨基酸序列的核酸分子(DNA)、掺加了此 DNA 的表达系统，包括载体及噬菌体、以此 DNA 转化的宿主细胞、与编码氨基酸序列的 DNA 分子相应的 DNA 反义链。同样，本发明包括产生突变α-淀粉酶的方法，该方法是在已经转化到宿主细胞中的表达系统中表达掺入的 DNA。本发明的突变α-

淀粉酶可用于淀粉液化过程、作为洗衣剂的成分、自动洗碗机的清洁剂、硬表面的清洁产品、包括烘焙的食品加工、包括退浆的纺织品加工或任何其它需要使用α-淀粉酶的场合。

前体α-淀粉酶由任何可产生α-淀粉酶的来源产生。适当的α-淀粉酶来源为原核或真核生物体，包括真菌、细菌、植物或动物。优选的前体α-淀粉酶由芽孢杆菌产生；更优选的是地衣芽孢杆菌、解淀粉芽孢杆菌或嗜热脂肪芽孢杆菌；最优选的前体α-淀粉酶来自地衣芽孢杆菌。

到目前为止几乎全部已测序的内淀粉酶都发现同源性，范围包括植物、哺乳动物及细菌(Nakajima 等人，应用微生物学及生物技术，第 23 卷，第 355-360 页(1986)；Rogers 等人，生物化学、生物物理研究通讯，第 128 卷，第 470-476 页(1985)；Janecek 等人，欧洲生物化学杂志，第 224 卷，第 519-524 页(1994))。某些芽孢杆菌淀粉酶在四个区域有特别高的同源性，如图 3 所示，在其下划线的部分为高同源性的区域。此外，序列对比也被用来制作不同芽孢杆菌内淀粉酶的关系图(Feng 等人，分子进化杂志，第 35 卷，第 351-360 页(1987))。Holm 等人，蛋白质工程，第 3 卷，第 3 期，第 181-191 页(1990)确定嗜热脂肪芽孢杆菌与地衣芽孢杆菌淀粉酶间的相对序列同源性约为 66%、地衣芽孢杆菌与解淀粉芽孢杆菌淀粉酶间的相对序列同源性约为 81%。虽然序列同源性有其重要性，一般认为在比较淀粉酶或其它酶时，结构同源性也相当重要。例如已有人提出真菌淀粉酶及细菌淀粉酶有结构同源性，因此本发明包括了真菌淀粉酶。

此外，本发明鉴定了与地衣芽孢杆菌α-淀粉酶 A210、H405 和/或 T412 位点相应残基上的添加、删除或替换。因此，特定的残基如 A210 指的是氨基酸位点号码(即+210)，这是指定给图 1 所示的成熟地衣芽孢杆菌α-淀粉酶序列的参考号。不过，本发明并不局限于某一成熟地衣芽孢杆菌α-淀粉酶的突变，而是包括含有与所鉴定的地衣芽孢杆菌α-淀粉酶残基位点相同的氨基酸残基的前体α-淀粉酶。如果前体α-淀粉酶的残基与地衣芽孢杆菌α-淀粉酶的特定残基或该残基的

部分同源(即在一级或三级结构有相同的位点)或类似(即在化学或结构上有相同或相似的结合、反应和相互作用的功能),则它与地衣芽孢杆菌α-淀粉酶残基相同。

为了确定一级结构的同源性,将一个前体α-淀粉酶的氨基酸序列与地衣芽孢杆菌α-淀粉酶的的一级序列作直接比较,尤其是使用一组在所有已知序列的α-淀粉酶内不变的残基(例如见图 3)。此外,还可以使用对猪胰α-淀粉酶所报道的三级结构分析晶体结构以确定相同的残基(Buisson 等人,EMBO 杂志,第 6 卷,第 3909-3916 页(1987); Qian 等人,生物化学,第 33 卷,第 6284-6294 页(1994); Larson 等人,分子生物学杂志,第 235 卷,第 1560-1584 页(1994)); 来自米曲霉的 Taka-淀粉酶 A(Matsuura 等人,生物化学杂志(东京),第 95 卷,第 697-702 页(1984)); 及来自黑曲霉的酸性α-淀粉酶(Boel 等人,生物化学,第 29 卷,第 6244-6249 页(1990),前两者的结构比较相近及大麦α-淀粉酶(Vallee 等人,分子生物学杂志,第 236 卷,第 368-371 页(1994); Kadziola 等人,分子生物学杂志,第 239 卷,第 104-121 页(1994))。有几项初步研究发表了有关α-淀粉酶的二级结构,即(Suzuki 等人,生物化学杂志,第 108 卷,第 379-381 页(1990); Lee 等人,生物化学、生物物理学进展,第 291 卷,第 255-257 页(1991); Chang 等人,分子生物学杂志,第 229 卷,第 235-238 页(1993); Mizuno 等人,分子生物学杂志,第 234 卷,第 1282-1283 页(1993)),而晶体地衣芽孢杆菌α-淀粉酶至少有一种结构已发表(Machius 等人,分子生物学杂志,第 246 卷,第 545-549 页(1995))。不过,几位科学家预测了葡聚糖酶间的共同超二级结构(MacGregor 等人,生物化学杂志,第 259 卷,第 145-152 页(1989))及在α-淀粉酶与其它淀粉代谢酶内(Jaspersen,蛋白质化学杂志,第 12 卷,第 791-805 页(1993); MacGregor,Starke,第 45 卷,第 232-237 页(1993)); 及与α-淀粉酶有相似的超二级结构酶之间的序列相似性(Janecek 等人,FEBS 通讯,第 316 卷,第 23-26 页(1993); Janecek 等人,蛋白质化学杂志,第 12 卷,第 509-514 页(1993))。嗜热脂肪

芽孢杆菌酶的结构是根据 Taka-淀粉酶 A 的结构模型构造的(Holm 等人，蛋白质工程，第 3 卷，第 181-191 页(1990))。图 3 所示的四个高度保守区域含有很多被认为是活性位点一部分的残基(Matsuura 等人，生物化学杂志(东京)，第 95 卷，第 697-702 页(1984)；Buisson 等人，EMBO 杂志，第 6 卷，第 3909-3916 页(1987)；Vihinen 等人，生物化学杂志，第 107 卷，第 267-272 页(1990))包括在地衣芽孢杆菌编号系统下的 His+105；Arg+229；Asp+231；His+235；Glu+261 及 Asp+328。

本发明的α-淀粉酶显现改变的性能特征，在α-淀粉酶常用的各种用途中非常有用，因为它提供所要求及没有预期到的结果。例如，本发明的α-淀粉酶在低 pH 下显现改变的性能特征包括改善的热稳定性、改善的 pH 稳定性和/或改善的氧化稳定性等对于在低 pH 条件下液化淀粉非常有用。加强热稳定性有助延长添加了α-淀粉酶的产品之货架寿命。加强氧化稳定性或改善性能对清洁产品尤其重要，并有助于延长α-淀粉酶在含漂白剂、过硼酸盐、过碳酸盐或过酸等清洁产品内的货架寿命。与此相反，降低热稳定性和氧化稳定性可能有助于要求快速及有效淬灭淀粉分解活性的工业生产过程。

本发明的α-淀粉酶表现改善的低 pH 稳定性，对淀粉加工，特别是淀粉液化尤为重要。理想的商业液化过程的条件通常包括低 pH、高温及潜在氧化条件，需要α-淀粉酶有改善的低 pH 性能、改善的热稳定性及改善的氧化稳定性。本发明的α-淀粉酶在液化过程中尤其有用，因为它在低于约 6.0，优选的是约 5.5，最优选的是低于约 5.0 的 pH 下表现改善的性能。此外，本发明α-淀粉酶在约摄氏 80-120 度，优选的是在约摄氏 100-110 度的温度下热稳定性增强，及在有氧化剂的条件下稳定性增强，因而特别有用。

本领域的技术人员所知对液化过程有用的其它成分包括例如抗氧化剂、钙、离子、盐或其它酶如内切糖苷酶、纤维素酶、蛋白酶、脂酶或其它淀粉酶等可根据实际反应条件而添加。例如，混合本发明的α-淀粉酶及其它来源的α-淀粉酶可提供在某些特定液化条件下有用

的独特作用。尤其是，预期混合本发明的α-淀粉酶及来自嗜热脂肪芽孢杆菌的α-淀粉酶将因其互补作用模式而在 pH 值 5.5 以下加强液化。

在液化过程中，淀粉，尤其是来自湿磨或干磨过程的颗粒淀粉淤浆将按照已知的液化技术用本发明的α-淀粉酶进行处理。通常，淀粉降解过程的第一步骤是以相当高的温度(约摄氏 80 度至 110 度之间)使淀粉淤浆凝胶化。淀粉淤浆凝胶化后使用α-淀粉酶使之液化。

在本发明的另一实施方案中，液态、凝胶或颗粒形式的包含本发明的α-淀粉酶的去污剂组合物是有用的。本发明的α-淀粉酶有增强的热稳定性，可增加货架寿命，或有增强的氧化稳定性，因此对通常存在于去污剂内的漂白剂或过酸化合物有加强的耐受性，因此对这种去污剂组合物特别有益。因此，本发明的α-淀粉酶可有利地添加到已知的 pH 在 6.5 至 12.0 之间的粉状、液态或凝胶去污剂中。含本发明的α-淀粉酶的去污剂组合物还可以包括其它酶，如内切糖苷酶、纤维素酶、蛋白酶、脂酶或其它淀粉酶，尤其是来自嗜热脂肪芽孢杆菌的α-淀粉酶及在本领域中公知的其它成分。

除了本发明提供的替换、添加或删除残基外，本发明优选的实施方案进一步包含本领域已知可增加稳定性或活性的任何一种或多种替代。例如删除或替换蛋氨酸残基或色氨酸残基，例如 WO94/18314(其公开内容引入本文作为参考)所述的 M15、M197 或 W138; PCT 公开号 WO91/00353 所述在 H133Y 上的替换; 或 DeClerck 等人，生物化学杂志，第 265 卷，第 15481-15488 页(1990)所述在 A209 上的替换; 或在 PCT 公开号 WO95/10603、WO96/23873 及 WO96/23874 所述的任何替换。尤其优选的实施方案是本发明的α-淀粉酶可进一步包含与地衣芽孢杆菌的 M15、A33、A52、S85、N96、V129、H133、S148、S187、N188、A209、A269 和/或 A379 相应的一或多个残基的删除或替换。

本发明的实施方案包含本发明的α-淀粉酶与蛋白酶混合，优选地包括美国 Re. 34,606 所述(在此引用供参考)的具有氧化稳定性的蛋白酶及商业酶如 DURAZYM (Novo Nordisk)及 PURAFECT® OxP

(Genencor International, Inc.)。美国 Re. 34,606 描述了有关制造这样的蛋白酶突变体(具有氧化稳定性的蛋白酶)，尤其是在与解淀粉芽孢杆菌 M222 相同位点上有蛋氨酸替换残基的这样的突变体。

本发明的另一实施方案包含编码本发明的α-淀粉酶的 DNA 及包含此 DNA 的表达载体。DNA 序列的表达是将它们可操作地连接到在适当的表达载体内的表达控制序列上，然后使用已知的技术以该表达载体转化适当的宿主。在表达本发明的 DNA 序列时，可使用多种宿主/表达载体的组合。有用的表达载体，例如包括染色体、非染色体及合成 DNA 序列片段，如在这方面有用的各种已知的质粒和噬菌体。此外，多种表达控制序列的任何一种通常都可在这些载体中使用。例如，申请人已经发现芽孢杆菌转化体的优选的表达控制序列是来自枯草芽孢杆菌的 aprE 信号肽。

在表达本发明的 DNA 序列方面，多种宿主细胞也有用。这些宿主可以包括常见的真核及原核宿主，如大肠杆菌、假单胞菌、芽孢杆菌、链霉菌，各种真菌、酵母菌及动物细胞。优选的是宿主在细胞外表达本发明的α-淀粉酶以便于纯化及下游加工。本发明的突变α-淀粉酶之表达及纯化可通过本领域所认可的方法进行。

与野生型芽孢杆菌α-淀粉酶比较时，根据本发明产生的改善的α-淀粉酶预期可提供重要的优点。例如，一个优点是在低 pH 及高温(这些都是一般淀粉液化方法的典型条件)下活性增强。其它优点包括增加的高 pH 及氧化稳定性，这使它们适用于去污剂；更完全地水解淀粉分子，减少在加工蒸汽中的残留淀粉；在没有钙离子的条件下有增强的稳定性；此外，与野生型地衣芽孢杆菌α-淀粉酶比较，添加同等蛋白质剂量的本发明的α-淀粉酶可提供更优异的性能，因为在逆境下，本发明的α-淀粉酶的比活性及稳定性有所增强。

以下实施例仅供参考，并非限定权利要求的范围。本文使用的缩写，特别是三字母或一字母的氨基酸代号描述于 Dale, J.W., 细菌分子遗传学, John Wiley & Sons, (1989)附录 B。

实施例

实施例 1

构建质粒 pHP. BL

图 1 所示的α-淀粉酶基因是从地衣芽孢杆菌 NCIB8061(Gray 等人，细菌学杂志，第 166 卷，第 635-643 页(1986))克隆的。将编码信号序列的最后三个残基、整个成熟蛋白质及终止子区域的 1.72kb Pstl-Sstl 片段亚克隆进 M13mp18。使用下列形式的合成低聚核苷酸盒将一个合成终止子加进 Bcll 与 Sstl 位点之间:

5'-GATCAAAACATAAAAAACCGGCCTTGGCCCCGCCGGTTTTTTATTATTTTTGAGCT-3' (SEQ ID NO:7)
3'-TTTTGTATTTTTTGGCCGGAACCGGGGCGGCCAAAAAATAATAAAAAC-5' (SEQ ID NO:8)

其设计是包含解淀粉芽孢杆菌枯草芽孢杆菌蛋白酶转录终止子(Wells 等人，核酸研究，第 11 卷，第 7911-7925 页(1983))。

被构建的 pBLapr 质粒携带着地衣芽孢杆菌α-淀粉酶的基因。如图 5 所示，pBLapr 含一个 6.1kb 质粒，包括来自 pBR322 的抗氨苄青霉素基因及来自 pC194 的抗氯霉素基因、aprE 启动子及编码地衣芽孢杆菌α-淀粉酶(“BLAA”)的基因。aprE 启动子是从编码启动子的一个 660bp Hindlll-Pstl 片段及枯草芽孢杆菌碱性蛋白酶的信号序列所构建。Pstl 位点被去除，一个 Sfil 位点被加到 aprEl BLAA 接合点的附近。BLAA 基因含上述 1720bp Pstl-Sstl 片段。在这里描述的工作中，pBLapr 是使用与成熟淀粉酶基因编码序列开端 5′ 端相邻的 Sfil 位点所构建。具体地说，在 EcoRI-Sstil 片段上将构建 pBLapr 的 5′ 端从 pBLapr 亚克隆进 M13BM20 (Boehringer Mannheim)以便获得下列诱变低聚核苷酸的编码链模板:

5'- CCC ATT AAG ATT GGC CGC CTG GGC CGA CAT GTT GCT GG - 3' (SEQ ID NO:9)

此引物导入了一个 Sfil 位点(在其下划线显示)，由于此独特限制位点的存在可筛选正确的形式。将 EcoRI-Sstil 片段亚克隆回 pBLapr 载体提供一个含 Sfil 位点的质粒。

质粒 pHP13(Haima 等人，Mol. Gen. Genet.，第 209 卷，第 335-342

页(1987))(图 4)由限制酶 EcoRI 及 Hindlll 所消化，所得的载体由聚丙烯酰胺凝胶纯化，然后洗脱。质粒 pBLapr 由 Hindlll、Asp718 所消化，并与 Asp718、EcoRI 作独立保温培养，然后以凝胶纯化。两条染色带，Hindlll-Asp718 (1203bp)及 Asp718-EcoRI (1253bp)经凝胶纯化后，从凝胶中洗脱，并以三路连接的方法连接至载体以产生用于表达α-淀粉酶的 pHP.BL 质粒(图 6)。

实施例 2

构建编码含 A210T/H405A/T412D 的 α-淀粉酶的质粒

使用美国专利申请序号 08/194,664 (PCT 公开号 WO94/18314)所述的方法构建一个在氨基酸 15 的蛋氨酸由苏氨酸取代的 pBLapr 质粒。引入突变的方法如下：使用编码 A210T/H405D/T412A 的替代的下列诱变引物，同时使用非诱变引物引入所需的突变至质粒的线形多串联重复体。

H405D (L)
(411) CCA GCC GAC AAT GTC ATG GTC GTC GAA ATA ATC (401) (SEQ ID NO:10)
A210T (R)
(206) CCT GAT GTC GCA ACA GAA ATT AAG AGA TGG (215) (SEQ ID NO:11)
T412A (L)
(416) GTC GCC TTC CCT TGC CCA GCC GAC AAT GTC (407) (SEQ ID NO:12)

在所需突变的适当诱变引物(上列)开始、在非诱变引物末端结束的片段由 PCR 产生。此片段由凝胶纯化、并根据下列方法产生含编码所需突变的长、线形串联重复的质粒：

载体(pBLapr)由限制消化(Sall)方法线形化，并使用 Qiagen 试剂盒纯化。此多聚反应在 100μl 反应混合物中通常含 5.4mM Tris 缓冲液 pH8.0、1xXL 缓冲液(Perkin Elmer, Branchburg, NJ)、0.2mM dNTPs、1.1mMMg $(OAc)_2$、3ng/μl 进入片段、0.15ng/μl 线形化载体、4U rTth DNA 聚合酶、XL(Perkin Elmer)。PCR 反应通常是在下

列条件的热循环器内进行: 20 周期(摄氏 94 度 15 秒、摄氏 68 度 5 分钟)及 15 周期(摄氏 94 度 15 秒、摄氏 68 度 10 分钟)。

使用标准技术将所得的多聚体直接转化至枯草芽孢杆菌感受态细胞。使用标准技术从转化体中分离质粒 DNA。

突变由双脱氧测序来确认(Sanger 等人, 美国国家科学院院报, 第 74 卷, 第 5463-5467 页(1977))。

实施例 3

转化质粒进枯草芽孢杆菌, 表达及纯化α-淀粉酶突变体

经过上述质粒转化后, α-淀粉酶可在枯草芽孢杆菌内表达。pHP13 是一种可以在大肠杆菌及枯草芽孢杆菌内复制的质粒。含不同变异体的质粒是使用大肠杆菌菌株 MM294 所构建, 质粒根据 Anagnostopoulos 等人, 细菌学杂志, 第 81 卷, 第 741-746 页(1961)的描述分离然后转化进枯草芽孢杆菌内。芽孢杆菌菌株的两个蛋白酶(△apr、△npr)(见例如 Ferrari 等人, 美国专利号 5, 264, 366)及淀粉酶(△amyE)(见例如 Stahl 等人, 细菌学杂志, 第 158 卷, 第 411-418 页(1984)已被删除。转化后, sacU(Hy)突变(Henner 等人, 细菌学杂志, 第 170 卷, 第 296-300 页(1988)由 PBS-1 介导的转导作用(Hoch, 细菌学杂志, 第 154 卷, 第 1513-1515 页(1983)引入。

分泌的淀粉酶由下列方法从培养的枯草芽孢杆菌中回收: 氯化钠被加到培养上清液内至 20mM, 使用 1M 的 Tris 缓冲液(pH7.2)调整 pH 至约 7.0。上清液然后被加热 15 分钟至摄氏 70 度, 沉淀物使用离心法除去。硫酸铵被加进上清液至 1.3M, 然后再加 20ml 苯基琼脂糖凝胶快流 6(高置换)树脂(Pharmacia)。搅拌后, 过滤分离树脂, 然后以 1M 硫酸铵、20mM pH7.0 乙酸铵及 5mM 氯化钙清洗。结合的淀粉酶被洗脱进 20mM pH7.0 乙酸铵及 5mM 氯化钙, 然后加入硫酸铵至 70%饱和以便沉淀。沉淀物以离心法分离, 然后再使用最小体积的 20mM pH7.0 乙酸铵及 5mM 氯化钙溶解并使用同一缓冲液进行透析。

使用可溶性底物测定法，假定野生型比活性，确定浓度。

实施例 4

确定α-淀粉酶活性的测定

可溶性底物测定：速率测定是基于由Megazyme (Aust.) Pty. Ltd.提供的终点测定试剂盒发展而来的。一瓶底物(对硝基苯基麦芽七糖，BPNPG7)被溶解在 10ml 无菌水中，然后以 1：4 的比例在测定缓冲液(50mM pH6.7 顺丁烯二酸盐缓冲液、5mM 氯化钙，0.002%Tween20)中稀释。测定进行的方法是在摄氏 25 度的小池内将 10μl 淀粉酶加入 790μl 底物。水解速率是根据在延迟 75 秒后，在 410nm 的吸光度变化率来衡量的。在每分钟 0.2 吸收单位的吸光度变化率之前测定为线形。

α-淀粉酶蛋白质浓度是使用标准的 Bio-Rad 测定(Bio-Rad Laboratories)方法来量度的，该方法基于 Bradford，生物化学分析，第 72 卷，第 248 页(1976)使用牛血清白蛋白标准的方法。

实施例 5

制备及测试其它突变

α-淀粉酶以了解其热稳定性

制备了在 A210T/H405D/T412A 有替代的地衣芽孢杆菌α-淀粉酶突变体。突变体的热失活率根据下列步骤量度。淀粉酶原液广泛地透析进 20mM 乙酸铵、4mM pH6.5 氯化钙。每一样品的存储温度为摄氏 4 度。为了量度稳定性，使用 50mM 乙酸铵、5mM 氯化钙 0.02% Tween 20pH4.8 将原液稀释 50 倍以上到 30 至 50μg/ml 的终浓度。六个 100μl 等分试样被放进 eppendorf 管内置于温度在摄氏 83 度的水浴或热块中。eppendorf 管以 30 秒至 5 分钟的固定测定间隔被取出置于冰上以停止失活。残余活性使用实施例 4 说明的可溶性底物测定。活性的自然对数对保温时间作图，从直线斜率得到失活的速率常数。所得结果列于表 1。

表 1

	相对半寿期	
淀粉酶	Exp. #1	Exp. #2
野生型	1.00	1.00
野生型	1.01	XX
A210T/H405D/T412A	1.06	1.05

如表 1 所示，根据本发明引进突变的突变体酶在测定条件下的稳定性有显著的改善。

93.05.27

序列表

(1)一般资料:

(i) 申请者: Anthony G. Day
Barbara A. Swanson

(ii)发明题目: 包括相应于地衣芽孢杆菌 A210、H405 和/或 T412 的残基的修饰的突变α-淀粉酶

(iii)序列数目: 12

(iv)通信地址:

(A)收信人: Genencor International, Inc.
(B)街道: 925 Page Mill Road
(C)城市: Palo Alto
(D)州: CA
(E)国家: USA
(F)邮政编码: 94304-1013

(v)计算机可读形式:

(A)媒介类型: 磁盘
(B)计算机: IBM 兼容
(C)操作系统: DOS
(D)软件: 供 Windows2.0 版本使用的 FastSEQ

(vi)目前申请资料:

(A)申请号: 待指定
(B)提交日: 待指定

(viii)律师/代理人资料:

(A)姓名: Stone, Christopher

(B)注册号: 35, 696

(C)参考/摘要号: GC387

(ix)电讯资料:

(A)电话: (650)846-7555

(B)电传: (650)845-6504

(2)序列识别号: 1 的资料:

(i)序列特征:

(A)长度: 1968 碱基对

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 线型

(ii)分子类型: DNA(基因组)

(xi)序列描述: 序列识别号: 1:

AGCTTGAAGA AGTGAAGAAG CAGAGAGGCT ATTGAATAAA TGAGTAGAAA GCGCCATATC 60

GGCGCTTTTC TTTTGGAAGA AAATATAGGG AAAATGGTAC TTGTTAAAAA TTCGGAATAT 120

TTATACAACA TCATATGTTT CACATTGAAA GGGGAGGAGA ATCATGAAAC AACAAAAACG 180

GCTTTACGCC CGATTGCTGA CGCTGTTATT TGCGCTCATC TTCTTGCTGC CTCATTCTGC 240

AGCAGCGGCG GCAAATCTTA ATGGGACGCT GATGCAGTAT TTTGAATGGT ACATGCCCAA 300

TGACGGCCAA CATTGGAAGC GTTTGCAAAA CGACTCGGCA TATTTGGCTG AACACGGTAT 360

TACTGCCGTC TGGATTCCCC CGGCATATAA GGGAACGAGC CAAGCGGATG TGGGCTACGG 420

TGCTTACGAC CTTTATGATT TAGGGGAGTT TCATCAAAAA GGGACGGTTC GGACAAAGTA 480

CGGCACAAAA GGAGAGCTGC AATCTGCGAT CAAAAGTCTT CATTCCCGCG ACATTAACGT 540

TTACGGGGAT GTGGTCATCA ACCACAAAGG CGGCGCTGAT GCGACCGAAG ATGTAACCGC 600

GGTTGAAGTC GATCCCGCTG ACCGCAACCG CGTAATTTCA GGAGAACACC TAATTAAAGC 660

CTGGACACAT TTTCATTTTC CGGGGCGCGG CAGCACATAC AGCGATTTTA AATGGCATTG 720

GTACCATTTT GACGGAACCG ATTGGGACGA GTCCCGAAAG CTGAACCGCA TCTATAAGTT 780

TCAAGGAAAG GCTTGGGATT GGGAAGTTTC CAATGAAAAC GGCAACTATG ATTATTTGAT 840

GTATGCCGAC ATCGATTATG ACCATCCTGA TGTCGCAGCA GAAATTAAGA GATGGGGCAC 900

TTGGTATGCC AATGAACTGC AATTGGACGG TTTCCGTCTT GATGCTGTCA AACACATTAA 960

ATTTTCTTTT TTGCGGGATT GGGTTAATCA TGTCAGGGAA AAAACGGGGA AGGAAATGTT 1020

TACGGTAGCT GAATATTGGC AGAATGACTT GGGCGCGCTG GAAAACTATT TGAACAAAAC 1080

AAATTTTAAT CATTCAGTGT TTGACGTGCC GCTTCATTAT CAGTTCCATG CTGCATCGAC 1140

ACAGGGAGGC GGCTATGATA TGAGGAAATT GCTGAACGGT ACGGTCGTTT CCAAGCATCC 1200

GTTGAAATCG GTTACATTTG TCGATAACCA TGATACACAG CCGGGGCAAT CGCTTGAGTC 1260

GACTGTCCAA ACATGGTTTA AGCCGCTTGC TTACGCTTTT ATTCTCACAA GGGAATCTGG 1320

ATACCCTCAG GTTTTCTACG GGGATATGTA CGGGACGAAA GGAGACTCCC AGCGCGAAAT 1380

TCCTGCCTTG AAACACAAAA TTGAACCGAT CTTAAAAGCG AGAAAACAGT ATGCGTACGG 1440

AGCACAGCAT GATTATTTCG ACCACCATGA CATTGTCGGC TGGACAAGGG AAGGCGACAG 1500

CTCGGTTGCA AATTCAGGTT TGGCGGCATT AATAACAGAC GGACCCGGTG GGGCAAAGCG 1560

AATGTATGTC GGCCGGCAAA ACGCCGGTGA GACATGGCAT GACATTACCG GAAACCGTTC 1620

GGAGCCGGTT GTCATCAATT CGGAAGGCTG GGGAGAGTTT CACGTAAACG GCGGGTCGGT 1680

TTCAATTTAT GTTCAAAGAT AGAAGAGCAG AGAGGACGGA TTTCCTGAAG GAAATCCGTT 1740

TTTTTATTTT GCCCGTCTTA TAAATTTCTT TGATTACATT TTATAATTAA TTTTAACAAA 1800

GTGTCATCAG CCCTCAGGAA GGACTTGCTG ACAGTTTGAA TCGCATAGGT AAGGCGGGGA 1860

TGAAATGGCA ACGTTATCTG ATGTAGCAAA GAAAGCAAAT GTGTCGAAAA TGACGGTATC 1920

GCGGGTGATC AATCATCCTG AGACTGTGAC GGATGAATTG AAAAAGCT 1968

(2)序列识别号: 2 的资料:

(i)序列特征:

(A)长度: 511 氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线型

(ii)分子类型: 蛋白质

(xi)序列描述: 序列识别号: 2:

```
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
1               5                  10                  15

Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala Ala Ala Asn Leu
            20                  25                  30

Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn Asp Gly
        35                  40                  45

His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His Gly
    50                  55                  60

Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Ala
65                  70                  75                  80

Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe His
                85                  90                  95
```

```
Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu Gln
            100                 105                 110

Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly Asp
        115                 120                 125

Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp Val Thr
    130                 135                 140

Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly Glu
145                 150                 155                 160

His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg Gly Ser
                165                 170                 175

Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Thr Asp
            180                 185                 190

Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln Gly Lys
        195                 200                 205

Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr Leu
    210                 215                 220

Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val Ala Ala Glu Ile
225                 230                 235                 240

Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp Gly Phe
                245                 250                 255

Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp Trp
            260                 265                 270

Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val Ala
        275                 280                 285

Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn Lys
    290                 295                 300

Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln Phe
305                 310                 315                 320

His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu Leu
                325                 330                 335

Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr Phe Val
            340                 345                 350

Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val Gln
        355                 360                 365
```

Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu Ser
370 375 380

Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly Asp
385 390 395 400

Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile Leu
405 410 415

Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe Asp
420 425 430

His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val Ala
435 440 445

Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala Lys
450 455 460

Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp Ile
465 470 475 480

Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp Gly
485 490 495

Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg
500 505 510

(2)序列识别号: 3的资料:

(i)序列特征:

(A)长度: 483氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 未知

(ii)分子类型: 蛋白质

(xi)序列描述: 序列识别号: 3:

```
Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro
1               5                  10                  15

Asn Asp Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu
            20                  25                  30

Ala Glu His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly
        35                  40                  45

Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu
    50                  55                  60

Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys
65                  70                  75                  80

Gly Glu Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn
                85                  90                  95
```

Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr
100 105 110

Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val
115 120 125

Ile Ser Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro
130 135 140

Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe
145 150 155 160

Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
165 170 175

Phe Gln Gly Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn
180 185 190

Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val
195 200 205

Ala Ala Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln
210 215 220

Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe
225 230 235 240

Leu Arg Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met
245 250 255

Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn
260 265 270

Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu
275 280 285

His Tyr Gln Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met
290 295 300

Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser
305 310 315 320

Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu
325 330 335

Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu
340 345 350

Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly
355 360 365

29.05.27

```
Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile
    370                 375                 380

Glu Pro Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His
385                 390                 395                 400

Asp Tyr Phe Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp
                405                 410                 415

Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro
            420                 425                 430

Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr
        435                 440                 445

Trp His Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser
    450                 455                 460

Glu Gly Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr
465                 470                 475                 480

Val Gln Arg
```

(2)序列识别号: 4的资料:

(i)序列特征:

(A)长度: 511氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 未知

(ii)分子类型: 蛋白质

(xi)序列描述: 序列识别号: 4:

```
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
1               5                  10                  15

Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala Ala Ala Asn Leu
            20                  25                  30

Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn Asp Gly
        35                  40                  45

His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His Gly
    50                  55                  60

Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Ala
65                  70                  75                  80

Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe His
                85                  90                  95

Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu Gln
            100                 105                 110

Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly Asp
        115                 120                 125

Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp Val Thr
    130                 135                 140

Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly Glu
145                 150                 155                 160

His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg Gly Ser
            165                 170                 175

Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Thr Asp
        180                 185                 190

Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln Gly Lys
    195                 200                 205

Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr Leu
210                 215                 220                 225
```

Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val Ala Ala Glu Ile
230 235 240

Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp Gly Phe
245 250 255

Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp Trp
260 265 270
Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val Ala
275 280 285

Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn Lys
290 295 300 305

Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln Phe
310 315 320

His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu Leu
325 330 335

Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr Phe Val
340 345 350

Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val Gln
355 360 365

Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu Ser
370 375 380 385

Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly Asp
390 395 400

Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile Leu
405 410 415

Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe Asp
420 425 430

His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val Ala
435 440 445

Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala Lys
450 455 460 465

Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp Ile
470 475 480

Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp Gly
485 490 495

Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg
500 505 510

(2)序列识别号: 5的资料:

(i)序列特征:

(A)长度: 520氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 未知

(ii)分子类型: 蛋白质

(xi)序列描述: 序列识别号: 5:

```
Met Arg Gly Arg Gly Asn Met Ile Gln Lys Arg Lys Arg Thr Val Ser
1               5                  10                  15

Phe Arg Leu Val Leu Met Cys Thr Leu Leu Phe Val Ser Leu Pro Ile
            20                  25                  30

Thr Lys Thr Ser Ala Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp
        35                  40                  45

Tyr Thr Pro Asn Asp Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala
    50                  55                  60

Glu His Leu Ser Asp Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala
65                  70                  75                  80

Tyr Lys Gly Leu Ser Gln Ser Asp Asn Gly Tyr Gly Pro Tyr Asp Leu
                85                  90                  95

Tyr Asp Leu Gly Glu Phe Gln Gln Lys Gly Thr Val Arg Thr Lys Tyr
            100                 105                 110

Gly Thr Lys Ser Glu Leu Gln Asp Ala Ile Gly Ser Leu His Ser Arg
        115                 120                 125

Asn Val Gln Val Tyr Gly Asp Val Val Leu Asn His Lys Ala Gly Ala
    130                 135                 140

Asp Ala Thr Glu Asp Val Thr Ala Val Glu Val Asn Pro Ala Asn Arg
145                 150                 155                 160

Asn Gln Glu Thr Ser Glu Glu Tyr Gln Ile Lys Ala Trp Thr Asp Phe
                165                 170                 175

Arg Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp His Trp
            180                 185                 190

Tyr His Phe Asp Gly Ala Asp Trp Asp Glu Ser Arg Lys Ile Ser Arg
        195                 200                 205

Ile Phe Lys Phe Arg Gly Glu Gly Lys Ala Trp Asp Trp Glu Val Ser
    210                 215                 220
```

Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Tyr
225 230 235 240

Asp His Pro Asp Val Val Ala Glu Thr Lys Lys Trp Gly Ile Trp Tyr
245 250 255

Ala Asn Glu Leu Ser Leu Asp Gly Phe Arg Ile Asp Ala Ala Lys His
260 265 270

Ile Lys Phe Ser Phe Leu Arg Asp Trp Val Gln Ala Val Arg Gln Ala
275 280 285

Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asn Ala
290 295 300

Gly Lys Leu Glu Asn Tyr Leu Asn Lys Thr Ser Phe Asn Gln Ser Val
305 310 315 320

Phe Asp Val Pro Leu His Phe Asn Leu Gln Ala Ala Ser Ser Gln Gly
325 330 335

Gly Gly Tyr Asp Met Arg Arg Leu Leu Asp Gly Thr Val Val Ser Arg
340 345 350

His Pro Glu Lys Ala Val Thr Phe Val Glu Asn His Asp Thr Gln Pro
355 360 365

Gly Gln Ser Leu Glu Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala
370 375 380

Tyr Ala Phe Ile Leu Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr
385 390 395 400

Gly Asp Met Tyr Gly Thr Lys Gly Thr Ser Pro Lys Glu Ile Pro Ser
405 410 415

Leu Lys Asp Asn Ile Glu Pro Ile Leu Lys Ala Arg Lys Glu Tyr Ala
420 425 430

Tyr Gly Pro Gln His Asp Tyr Ile Asp His Pro Asp Val Ile Gly Trp
435 440 445

Thr Arg Glu Gly Asp Ser Ser Ala Ala Lys Ser Gly Leu Ala Ala Leu
450 455 460

Ile Thr Asp Gly Pro Gly Gly Ser Lys Arg Met Tyr Ala Gly Leu Lys
465 470 475 480

Asn Ala Gly Glu Thr Trp Tyr Asp Ile Thr Gly Asn Arg Ser Asp Thr
485 490 495

Val Lys Ile Gly Ser Asp Gly Trp Gly Glu Phe His Val Asn Asp Gly
500 505 510

Ser Val Ser Ile Tyr Val Gln Lys
515 520

(2)序列识别号: 6 的资料:

(i)序列特征:

(A)长度: 548 氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 未知

(ii)分子类型: 蛋白质

(xi)序列描述: 序列识别号: 6:

```
Val Leu Thr Phe His Arg Ile Ile Arg Lys Gly Trp Met Phe Leu Leu
1               5                  10                  15

Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Arg His Ala
            20                  25                  30

Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp
        35                  40                  45

Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala
    50                  55                  60

Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Ser Leu Pro Pro Ala
65                  70                  75                  80

Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu
                85                  90                  95

Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr
            100                 105                 110

Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala
        115                 120                 125

Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala
    130                 135                 140

Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg
145                 150                 155                 160

Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe
                165                 170                 175

Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp
            180                 185                 190

Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg
        195                 200                 205

Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp
    210                 215                 220
```

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met
225 230 235 240

Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr
245 250 255

Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Gly Leu Lys His
260 265 270

Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln
275 280 285

Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile
290 295 300

Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asn Gly Thr Met Ser Leu
305 310 315 320

Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly
325 330 335

Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp
340 345 350

Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Asn Pro
355 360 365

Ala Lys Arg Cys Ser His Gly Arg Pro Trp Phe Lys Pro Leu Ala Tyr
370 375 380

Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly
385 390 395 400

Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys
405 410 415

Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln
420 425 430

His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly
435 440 445

Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly
450 455 460

Ala Gly Arg Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys
465 470 475 480

Val Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn
485 490 495

Ser Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val
500 505 510

Trp Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile Thr
515 520 525

Thr Arg Pro Trp Thr Gly Glu Phe Val Arg Trp His Glu Pro Arg Leu
530 535 540

Val Ala Trp Pro
545

(2)序列识别号: 7 的资料:

(i)序列特征:

(A)长度: 56 碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线型

(ii)分子类型: cDNA

(xi)序列描述: 序列识别号: 7:

GATCAAAACA TAAAAAACCG GCCTTGGCCC CGCCGGTTTT TTATTATTTT TGAGCT 56

(2)序列识别号: 8 的资料:

(i)序列特征:

(A)长度: 48 碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线型

(ii)分子类型: cDNA

(xi)序列描述: 序列识别号: 8:

CAAAAATAAT AAAAAACCGG CGGGGCCAAG GCCGGTTTTT TATGTTTT 48

(2)序列识别号: 9 的资料:

(i)序列特征:

(A)长度: 38 碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线型

(ii)分子类型: cDNA

(xi)序列描述: 序列识别号: 9:

CCCATTAAGA TTGGCCGCCT GGGCCGACAT GTTGCTGG 38

(2)序列识别号: 10 的资料:

(i)序列特征:

(A)长度: 33 碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线型

(xi)序列描述: 序列识别号: 10:

CCAGCCGACA ATGTCATGGT CGTCGAAATA ATC 33

(2)序列识别号: 11 的资料:

(i)序列特征:

(A)长度: 30 碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线型

(xi)序列描述: 序列识别号: 11:

CCTGATGTCG CAACAGAAAT TAAGAGATGG 30

(2)序列识别号: 12 的资料:

(i)序列特征:

(A)长度: 30 碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线型

(xi)序列描述: 序列识别号: 12:

GTCGCCTTCC CTTGCCCAGC CGACAATGTC 30

说 明 书 附 图

```
        10                  30                  50
AGCTTGAAGAAGTGAAGAAGCAGAGAGGCTATTGAATAAATGAGTAGAAAGCGCCATATC

        70                  90                  110
GGCGCTTTTCTTTTGGAAGAAAATATAGGGAAAATGGTACTTGTTAAAAATTCGGAATAT

       130                 150                 170
TTATACAACATCATATGTTTCACATTGAAAGGGGAGGAGAATCATGAAACAACAAAAACG
                                                M  K  Q  Q  K  R

       190                 210                 230
GCTTTACGCCCGATTGCTGACGCTGTTATTTGCGCTCATCTTCTTGCTGCCTCATTCTGC
 L  Y  A  R  L  L  T  L  L  F  A  L  I  F  L  L  P  H  S  A

       250                 270                 290
AGCAGCGGCGGCAAATCTTAATGGGACGCTGATGCAGTATTTTGAATGGTACATGCCCAA
 A  A  A  A  N  L  N  G  T  L  M  Q  Y  F  E  W  Y  M  P  N

       310                 330                 350
TGACGGCCAACATTGGAAGCGTTTGCAAAACGACTCGGCATATTTGGCTGAACACGGTAT
 D  G  Q  H  W  K  R  L  Q  N  D  S  A  Y  L  A  E  H  G  I

       370                 390                 410
TACTGCCGTCTGGATTCCCCCGGCATATAAGGGAACGAGCCAAGCGGATGTGGGCTACGG
 T  A  V  W  I  P  P  A  Y  K  G  T  S  Q  A  D  V  G  Y  G

       430                 450                 470
TGCTTACGACCTTTATGATTTAGGGGAGTTTCATCAAAAAGGGACGGTTCGGACAAAGTA
 A  Y  D  L  Y  D  L  G  E  F  H  Q  K  G  T  V  R  T  K  Y

       490                 510                 530
CGGCACAAAAGGAGAGCTGCAATCTGCGATCAAAAGTCTTCATTCCCGCGACATTAACGT
 G  T  K  G  E  L  Q  S  A  I  K  S  L  H  S  R  D  I  N  V

       550                 570                 590
TTACGGGGATGTGGTCATCAACCACAAAGGCGGCGCTGATGCGACCGAAGATGTAACCGC
 Y  G  D  V  V  I  N  H  K  G  G  A  D  A  T  E  D  V  T  A

       610                 630                 650
GGTTGAAGTCGATCCCGCTGACCGCAACCGCGTAATTTCAGGAGAACACCTAATTAAAGC
 V  E  V  D  P  A  D  R  N  R  V  I  S  G  E  H  L  I  K  A

       670                 690                 710
CTGGACACATTTTCATTTTCCGGGGCGCGGCAGCACATACAGCGATTTTAAATGGCATTG
 W  T  H  F  H  F  P  G  R  G  S  T  Y  S  D  F  K  W  H  W

       730                 750                 770
GTACCATTTTGACGGAACCGATTGGGACGAGTCCCGAAAGCTGAACCGCATCTATAAGTT
 Y  H  F  D  G  T  D  W  D  E  S  R  K  L  N  R  I  Y  K  F

       790                 810                 830
TCAAGGAAAGGCTTGGGATTGGGAAGTTTCCAATGAAAACGGCAACTATGATTATTTGAT
 Q  G  K  A  W  D  W  E  V  S  N  E  N  G  N  Y  D  Y  L  M
```

图 1A

```
       850                       870                       890
GTATGCCGACATCGATTATGACCATCCTGATGTCGCAGCAGAAATTAAGAGATGGGGCAC
 Y  A  D  I  D  Y  D  H  P  D  V  A  A  E  I  K  R  W  G  T

       910                       930                       950
TTGGTATGCCAATGAACTGCAATTGGACGGTTTCCGTCTTGATGCTGTCAAACACATTAA
 W  Y  A  N  E  L  Q  L  D  G  F  R  L  D  A  V  K  H  I  K

       970                       990                      1010
ATTTTCTTTTTTGCGGGATTGGGTTAATCATGTCAGGGAAAAAACGGGGAAGGAAATGTT
 F  S  F  L  R  D  W  V  N  H  V  R  E  K  T  G  K  E  M  F

      1030                      1050                      1070
TACGGTAGCTGAATATTGGCAGAATGACTTGGGCGCGCTGGAAAACTATTTGAACAAAAC
 T  V  A  E  Y  W  Q  N  D  L  G  A  L  E  N  Y  L  N  K  T

      1090                      1110                      1130
AAATTTTAATCATTCAGTGTTTGACGTGCCGCTTCATTATCAGTTCCATGCTGCATCGAC
 N  F  N  H  S  V  F  D  V  P  L  H  Y  Q  F  H  A  A  S  T

      1150                      1170                      1190
ACAGGGAGGCGGCTATGATATGAGGAAATTGCTGAACGGTACGGTCGTTTCCAAGCATCC
 Q  G  G  G  Y  D  M  R  K  L  L  N  G  T  V  V  S  K  H  P

      1210                      1230                      1250
GTTGAAATCGGTTACATTTGTCGATAACCATGATACACAGCCGGGGCAATCGCTTGAGTC
 L  K  S  V  T  F  V  D  N  H  D  T  Q  P  G  Q  S  L  E  S

      1270                      1290                      1310
GACTGTCCAAACATGGTTTAAGCCGCTTGCTTACGCTTTTATTCTCACAAGGGAATCTGG
 T  V  Q  T  W  F  K  P  L  A  Y  A  F  I  L  T  R  E  S  G

      1330                      1350                      1370
ATACCCTCAGGTTTTCTACGGGGATATGTACGGGACGAAAGGAGACTCCCAGCGCGAAAT
 Y  P  Q  V  F  Y  G  D  M  Y  G  T  K  G  D  S  Q  R  E  I

      1390                      1410                      1430
TCCTGCCTTGAAACACAAAATTGAACCGATCTTAAAAGCGAGAAAACAGTATGCGTACGG
 P  A  L  K  H  K  I  E  P  I  L  K  A  R  K  Q  Y  A  Y  G

      1450                      1470                      1490
AGCACAGCATGATTATTTCGACCACCATGACATTGTCGGCTGGACAAGGGAAGGCGACAG
 A  Q  H  D  Y  F  D  H  H  D  I  V  G  W  T  R  E  G  D  S

      1510                      1530                      1550
CTCGGTTGCAAATTCAGGTTTGGCGGCATTAATAACAGACGGACCCGGTGGGGCAAAGCG
 S  V  A  N  S  G  L  A  A  L  I  T  D  G  P  G  G  A  K  R

      1570                      1590                      1610
AATGTATGTCGGCCGGCAAAACGCCGGTGAGACATGGCATGACATTACCGGAAACCGTTC
 M  Y  V  G  R  Q  N  A  G  E  T  W  H  D  I  T  G  N  R  S

      1630                      1650                      1670
GGAGCCGGTTGTCATCAATTCGGAAGGCTGGGGAGAGTTTCACGTAAACGGCGGGTCGGT
 E  P  V  V  I  N  S  E  G  W  G  E  F  H  V  N  G  G  S  V
```

图 1B

```
      1690                  1710                  1730
TTCAATTTATGTTCAAAGATAGAAGAGCAGAGAGGACGGATTTCCTGAAGGAAATCCGTT
 S   I   Y   V   Q   R   *

      1750                  1770                  1790
TTTTTATTTTGCCCGTCTTATAAATTTCTTTGATTACATTTTATAATTAATTTTAACAAA

      1810                  1830                  1850
GTGTCATCAGCCCTCAGGAAGGACTTGCTGACAGTTTGAATCGCATAGGTAAGGCGGGGA

      1870                  1890                  1910
TGAAATGGCAACGTTATCTGATGTAGCAAAGAAAGCAAATGTGTCGAAAATGACGGTATC

      1930                  1950
GCGGGTGATCAATCATCCTGAGACTGTGACGGATGAATTGAAAAAGCT
```

图 1C

```
         10                  30                  50
ANLNGTLMQYFEWYMPNDGQHWKRLQNDSAYLAEHGITAVWIPPAYKGTSQADVGYGAYD

         70                  90                  110
LYDLGEFHQKGTVRTKYGTKGELQSAIKSLHSRDINVYGDVVINHKGGADATEDVTAVEV

         130                 150                 170
DPADRNRVISGEHLIKAWTHFHFPGRGSTYSDFKWHWYHFDGTDWDESRKLNRIYKFQGK

         190                 210                 230
AWDWEVSNENGNYDYLMYADIDYDHPDVAAEIKRWGTWYANELQLDGFRLDAVKHIKFSF

         250                 270                 290
LRDWVNHVREKTGKEMFTVAEYWQNDLGALENYLNKTNFNHSVFDVPLHYQFHAASTQGG

         310                 330                 350
GYDMRKLLNGTVVSKHPLKSVTFVDNHDTQPGQSLESTVQTWFKPLAYAFILTRESGYPQ

         370                 390                 410
VFYGDMYGTKGDSQREIPALKHKIEPILKARKQYAYGAQHDYFDHHDIVGWTREGDSSVA

         430                 450                 470
NSGLAALITDGPGGAKRMYVGRQNAGETWHDITGNRSEPVVINSEGWGEFHVNGGSVSIY

VQR
```

图 2

Am-Lich = 地衣芽孢杆菌　　Am-Amylo = 解淀粉芽孢杆菌　　Am-Stearo = 嗜热脂肪芽孢杆菌

```
                                                                                                  1
                                                                                                                  19
                                                                                                                  60
                1
Am-Lich         .....MKQQ KRLYARLLTL    LFALIFLLPH    .....SAAA AANLNGTLMQ    YFEWYMPNDG
Am-Amylo        MRGRGNMIQK RKRTVSFRLV    LMCTLLFVSL    .....PITK TSAVNGTLMQ    YFEWYTPNDG
Am-Stearo       .....VLTF HRIIRKGWMF     LLAFLLTASL    FCPTGRHAKA AAPFNGTMMQ    YFEWYLPDDG

                                                                                                  79
                                                                                                  120
                61
Am-Lich         QHWKRLQNDS AYLAEHGITA    VWIPPAYKGT    SQADVGYGAY DLYDLGEFHQ    KGTVRTKYGT
Am-Amylo        QHWKRLQNDA EHLSDIGITA    VWIPPAYKGL    SQSDNGYGPY DLYDLGEFQQ    KGTVRTKYGT
Am-Stearo       TLWTKVANEA NNLSSLGITA    LSLPPAYKGT    SRSDVGYGVY DLYDLGEFNQ    KGTVRTKYGT

                                                                                                  139
                                                                                                  180
                121
Am-Lich         KGELQSAIKS LHSRDINVYG    DVVINHKGGA    DATEDVTAVE VDPADRNRVI    SGEHLIKAWT
Am-Amylo        KSELQDAIGS LHSRNVQVYG    DVVLNHKAGA    DATEDVTAVE VNPANRNQET    SEEYQIKAWT
Am-Stearo       KAQYLQAIQA AHAAGMQVYA    DVVFDHKGGA    DGTEWVDAVE VNPSDRNQEI    SGTYQIQAWT

                                                                                                  197
                                                                                                  240
                181
Am-Lich         HFHFPGRGST YSDFKWHWYH    FDGTDWDESR    KLNRIYKF....  QGKAWDWEVS NENGNYDYLM
Am-Amylo        DFRFPGRGNT YSDFKWHWYH    FDGADWDESR    KISRIFKFRG    EGKAWDWEVS SENGNYDYLM
Am-Stearo       KFDFPGRGNT YSSFKWRWYH    FDGVDWDESR    KLSRIYKFRG    IGKAWDWEVD TENGNYDYLM

                                                                                                  257
                                                                                                  300
                241
Am-Lich         YADIDYDHPD VAAEIKRWGT    WYANELQLDG    FRLDAVKHIK    FSFLRDWVNH VREKTGKEMF
Am-Amylo        YADVDYDHPD VVAETKKWGI    WYANELSLDG    FRIDAAKHIK    FSFLRDWVQA VRQATGKEMF
Am-Stearo       YADLDMDHPE VVTELKNWGK    WYVNTTNIDG    FRLDGLKHIK    FSFFPDWLSY VRSQTGKPLF

                                                                                                  317
                                                                                                  360
                301
Am-Lich         TVAEYWQNDL GALENYLNKT    NFNHSVFDVP    LHYQFHAAST    QGGGYDMRKL LNGTVVSKHP
Am-Amylo        TVAEYWQNNA GKLENYLNKT    SFNQSVFDVP    LHFNLQAASS    QGGGYDMRRL LDGTVVSRHP
Am-Stearo       TVGEYWSYDI NKLHNYITKT    NGTMSLFDAP    LHNKFYTASK    SGGAFDMRTL MTNTLMKDQP
```

图 3A

```
                                                                                                  377
                                                                                                  420
            361
Am-Lich     LKSVTFVDNH DTQPGQSLES TVQTWFKPLA YAFILTRESG   YPQVFYGDMY GTKGDSQREI
Am-Amylo    EKAVTFVENH DTQPGQSLES TVQTWFKPLA YAFILTRESG   YPQVFYGDMY GTKGTSPKEI
Am-Stearo   TLAVTFVDNH DTNPAKR..CS HGRPWFKPLA YAFILTRQEG   YPCVFYGDYY GI.....PQYNI

                                                                                                  437
                                                                                                  480
            421
Am-Lich     PALKHKIEPI LKARKQYAYG AQHDYFDHHD IVGWTREGDS SVANSGLAAL ITDGPGGAKR
Am-Amylo    PSLKDNIEPI LKARKEYAYG PQHDYIDHPD VIGWTREGDS SAAKSGLAAL ITDGPGGSKR
Am-Stearo   PSLKSKIDPL LIARRDYAYG TQHDYLDHSD IIGWTREGVT EKPGSGLAAL ITDGAGRSKW

                                                                                 483
                                                                                                  540
            481
Am-Lich     MYVGRQNAGE TWHDITGNRS EPVVINSEGW GEFHVNGGSV SIYVQR..... .....
Am-Amylo    MYAGLKNAGE TWYDITGNRS DTVKIGSDGW GEFHVNDGSV SIYVQK..... .....
Am-Stearo   MYVGKQHAGK VFYDLTGNRS DTVTINSDGW GEFKVNGGSV SVWVPRKTTV STIARPITTR

            541                 559
Am-Lich     ..................... .....
Am-Amylo    ..................... .....
Am-Stearo   PWTGEFVRWH EPRLVAWP*
```

图 3B

99-05-27

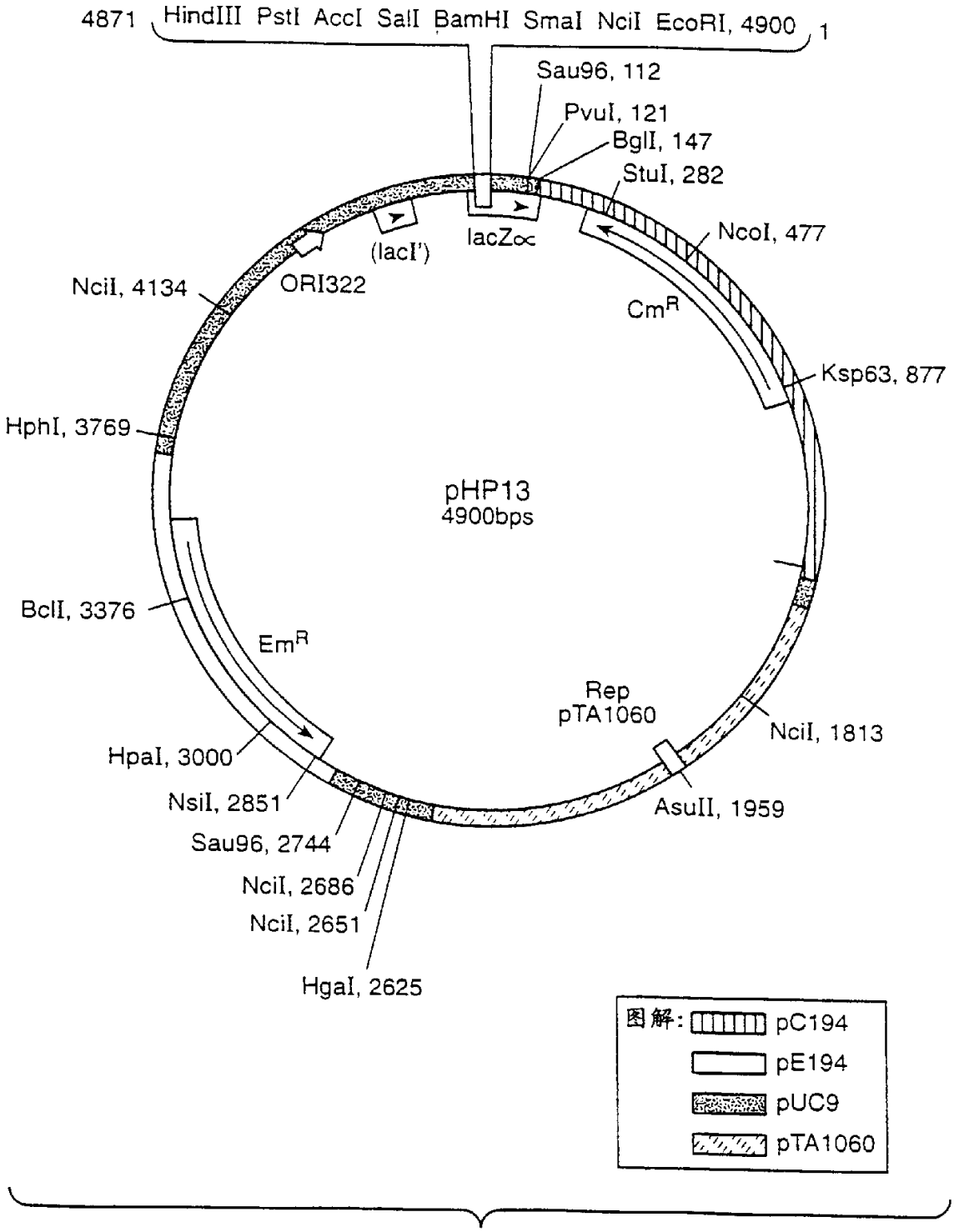

4871 HindIII PstI AccI SalI BamHI SmaI NciI EcoRI, 4900 1
Sau96, 112
PvuI, 121
BglI, 147
StuI, 282
NcoI, 477
lacZ∝
(lacI')
ORI322
NciI, 4134
Cm^R
Ksp63, 877
HphI, 3769
pHP13
4900bps
BclI, 3376
Em^R
Rep
pTA1060
NciI, 1813
HpaI, 3000
NsiI, 2851
AsuII, 1959
Sau96, 2744
NciI, 2686
NciI, 2651
HgaI, 2625
图解:
pC194
pE194
pUC9
pTA1060

图 4

99.05.27

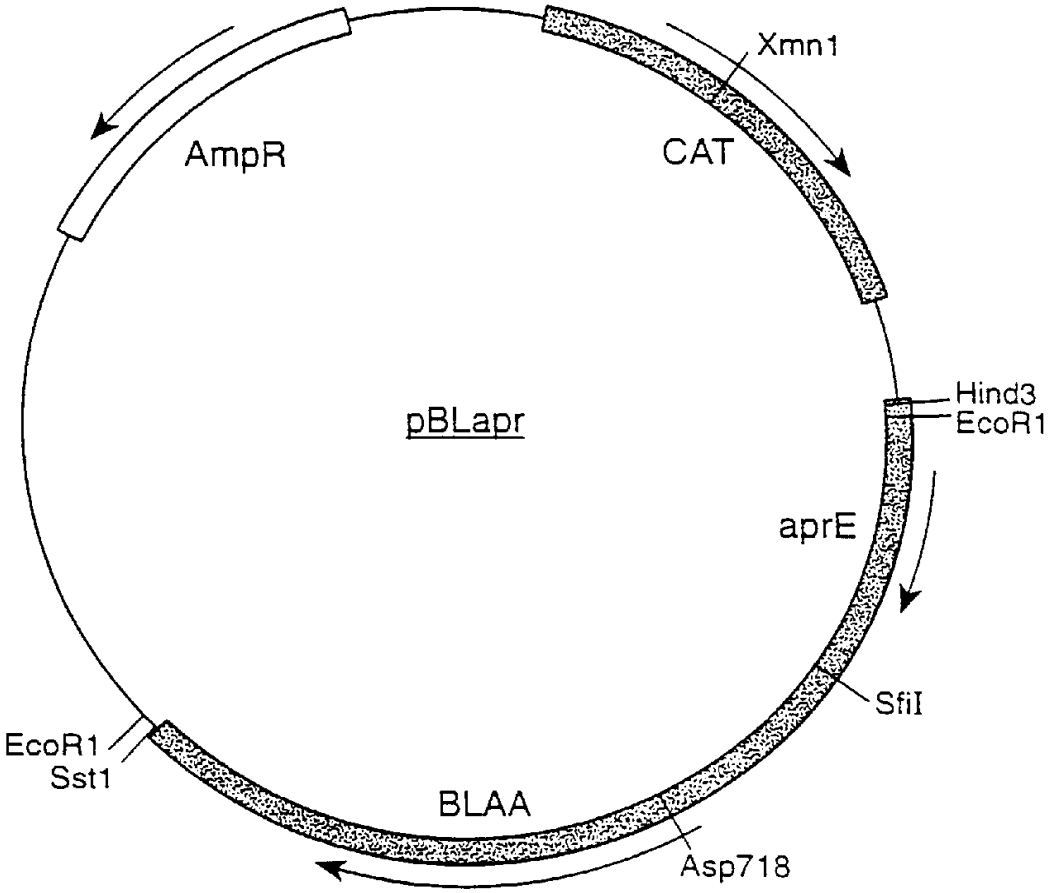
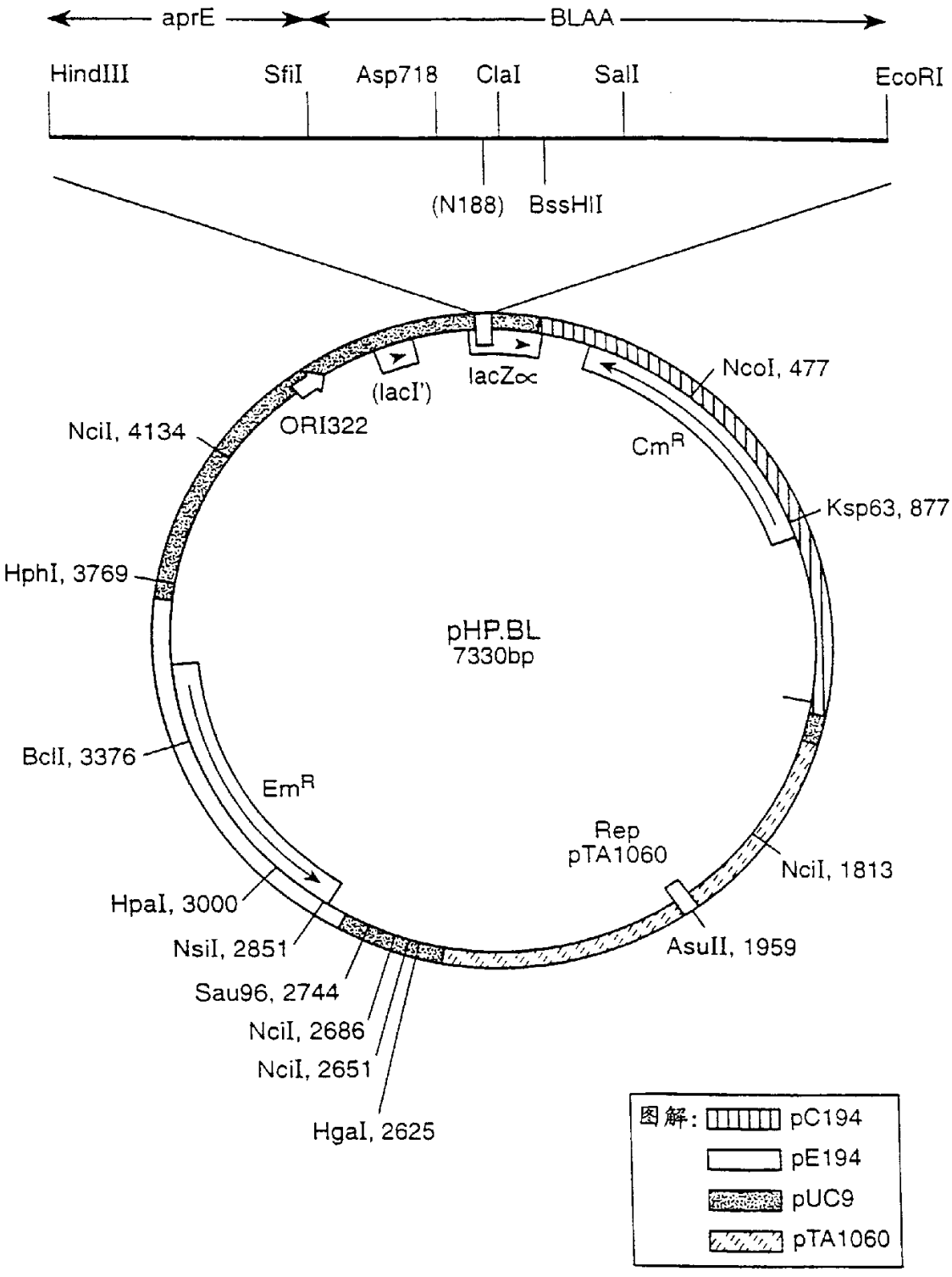

Xmn1
AmpR
CAT
Hind3
EcoR1
pBLapr
aprE
SfiI
EcoR1
Sst1
BLAA
Asp718

图 5



pHP.BL = 含有来自 pBLapr 的 2460bp HindIII-EcoRI 插入片段的 pHP13

图 6