



[12] 发明专利申请公开说明书

[21] 申请号 02816345.1

[43] 公开日 2005 年 7 月 6 日

[11] 公开号 CN 1636066A

[22] 申请日 2002.8.20 [21] 申请号 02816345.1

[30] 优先权

[32] 2001.8.22 [33] DE [31] 10141132.4

[32] 2001.10.11 [33] DE [31] 10150359.8

[86] 国际申请 PCT/EP2002/009272 2002.8.20

[87] 国际公布 WO2003/018826 德 2003.3.6

[85] 进入国家阶段日期 2004.2.20

[71] 申请人 巴斯福股份公司

地址 德国路德维希港

[72] 发明人 A·金克尔 J-H·马腾斯

D·耶恩 H·巴尔格 M·瓦伦

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 隗永良

权利要求书 2 页 说明书 18 页 序列表 15 页
附图 10 页

[54] 发明名称 制备维生素 B12 的方法

[57] 摘要

本发明涉及用巨大芽孢杆菌制备维生素 B12 的方法。

1. 使用包含巨大芽孢杆菌的培养物制备维生素 B12 的方法，其特征在于在厌氧条件下进行发酵。

2. 如权利要求 1 所述的方法，其特征在于巨大芽孢杆菌首先在需氧条件下发酵，然后在厌氧条件下发酵。

3. 如权利要求 1 或 2 所述的方法，其特征在于从需氧发酵到厌氧发酵的转变发生在需氧发酵细胞的指数生长期。

4. 如权利要求 1 至 3 中任意一项所述的方法，其特征在于从需氧发酵到厌氧发酵的转变发生在需氧生长细胞的指数生长期的中期或末期，优选地在末期。

5. 如权利要求 1 至 4 中任意一项所述的方法，其特征在于向培养基中至少添加钴。

6. 如权利要求 1 至 5 中任意一项所述的方法，其特征在于发酵其中 hemAXCDBL 操纵子或其部分显示表达增强的巨大芽孢杆菌菌株。

7. 如权利要求 1 至 6 中任意一项所述的方法，其特征在于发酵其细胞中 hemAXCDBL 操纵子以增加的拷贝数存在的巨大芽孢杆菌菌株。

8. 编码如 SEQ ID No.2-6 所示的参与尿卟啉原 III 生物合成的酶或它们的同工型的分离的核苷酸序列或它们的等位基因，所述分离的核苷酸序列被组织在 hemAXCDBL 操纵子中，所述操纵子具有示于 SEQ ID No.1 中的核苷酸序列。

9. 基因结构，其包含如权利要求 8 所述的分离的核苷酸序列或其部分并包含与其可操作连接的且具有调节功能的核苷酸序列。

10. 载体，其包含如权利要求 8 所述的分离的核苷酸序列或其部分或包含如权利要求 9 所述的基因结构，并包含用于选择、在宿主细胞内复制和/或整合入宿主细胞基因组的额外的核苷酸序列。

11. 如权利要求 1-7 中任意一项所述用于维生素 B12 生产的方法中的经转化巨大芽孢杆菌菌株，其特征在于它表现出如权利要求 8 所述核苷酸

序列或其部分的表达增强和/或拷贝数增加。

12. 如权利要求 11 所述的经转化巨大芽孢杆菌菌株, 其特征在于它包含了以复制型存在的如权利要求 9 所述的基因结构或如权利要求 10 所述的载体。

13. 如权利要求 8 所述的分离的核苷酸序列或其部分或如权利要求 9 所述的基因结构或如权利要求 10 所述的载体的用途, 用于产生如权利要求 11 或 12 所述的经转化巨大芽孢杆菌菌株。

14. 如权利要求 11 或 12 所述的经转化巨大芽孢杆菌菌株的用途, 用于产生维生素 B12。

制备维生素 B12 的方法

本发明涉及用巨大芽孢杆菌(*Bacillus megaterium*)制备维生素 B12 的方法。

早在本世纪三十年代, George Minot 和 William Murphy 通过维生素 B12 在人体中的作用而间接发现了维生素 B12 (Stryer, L., 1988, 生物化学 (Biochemie), 第四版, 528 -531 页, Spektrum Akademischer Verlag GmbH, Heidelberg, 柏林, 纽约)。在 1948 年, 维生素 B12 首次得到纯化和分离, 而直到八年后的 1956 年, Dorothy Hodgkin 才阐明了维生素 B12 复杂的三维晶体结构(Hodgkin, D.C.等人, 1956, 维生素 B12 的结构 (Structure of Vitamin B₁₂), 自然 (Nature), 176, 325-328 和 自然, 178, 64-70)。维生素 B12 生物合成的天然最终产物是 5'-脱氧腺苷钴胺素 (辅酶 B12) 和甲基钴胺素 (MeCbl), 然而维生素 B12 被定义为氰钴胺素, 这是工业上制备和处理的主要形式。在本发明中, 除非专门指出, 维生素 B12 总是指所有这三种类似分子。

100 多年前(1884)De Bary 首次对巨大芽孢杆菌这一物种进行了描述。尽管通常将巨大芽孢杆菌归类于土壤细菌, 但在多种其它生境如海水、沉积物、水稻、干肉、牛奶或蜂蜜中也可检测到巨大芽孢杆菌。巨大芽孢杆菌通常与假单胞菌属 (*pseudomonads*) 和放线菌属 (*actinomyces*) 相关联。象它的近亲枯草芽孢杆菌 (*Bacillus subtilis*) 一样, 巨大芽孢杆菌是革兰氏阳性细菌, 并且可以通过, 特别是通过其相对明显的个体大小 $2 \times 5 \mu\text{m}$, 据此而得名巨大芽孢杆菌, 约 38% 的 G+C 含量和很强的芽孢形成能力辨别出来。在生长培养基中即便是微量的锰也足以使巨大芽孢杆菌进行完全的芽孢形成, 这种能力只有一些嗜热细菌 (*thermophilic bacilli*) 的芽孢形成效率能比得上。正是由于它的大小和非常有效的芽孢形成和发芽能力,

针对巨大芽孢杆菌中这些过程的分子基础开展了不同的研究，以至于目前已有 150 多个涉及芽孢形成和发芽的巨大芽孢杆菌基因得以阐述。关于巨大芽孢杆菌的生理学方面的研究 (Priest, F.G.等人, 1988, 芽孢杆菌属的一种数值分类法 (A Numerical Classification of the Genus *Bacillus*), 普通微生物学杂志 (J. Gen. Microbiol.), 134, 1847-1822) 将这一物种归类于专性需氧的芽孢形成菌, 该类细菌为尿素酶阳性和伏-波试验阴性, 且不能够还原硝酸盐。巨大芽孢杆菌最显著的特征之一是它利用大量碳源的能力。因此它可以利用大量的糖并且发现存在于如玉米糖浆、肉类加工业的废弃物甚至石油化学废弃物中。鉴于其能够代谢极广泛碳源的能力, 巨大芽孢杆菌可以不受限地等同于假单胞菌属 (Vary, P.S., 1994, 微生物学 (Microbiology), 40, 1001-1013, 研究巨大芽孢杆菌的全盛期 (Prime time for *Bacillus megaterium*))。

在各种各样的酶、维生素等的工业生产中广泛应用巨大芽孢杆菌的优点是多样的。这些优点包括, 其中首要且确定的优点是转化进巨大芽孢杆菌的质粒的环境证明是很稳定的。这被认为是与迄今已建立起来的例如用聚乙二醇处理来转化这一物种的可行性有直接关系。直到几年前, 转化的可行性仍然是利用巨大芽孢杆菌作为生产菌株的一个主要障碍。认为相对充分研发的遗传学的优势与转化的可行性相关联, 在芽孢杆菌属中这一点只有枯草芽孢杆菌超过了它。第二, 巨大芽孢杆菌没有碱性蛋白酶, 所以在产生异源蛋白质时几乎不发生降解作用。另外我们还知道巨大芽孢杆菌可高效分泌一些具有商业价值的产物, 如利用它生产 α -和 β -淀粉酶。而且, 巨大芽孢杆菌的大小使它能够在达到致死的过密种群密度之前积累大量的生物量。在用巨大芽孢杆菌进行工业生产中一个更有利且非常重要的情况是该物种能够从废弃物和低质量物质中制备高价值和非常高质量的产品。巨大芽孢杆菌的这种能够代谢极广泛底物的能力也反映在它可以作为土壤解毒菌的用途上, 它甚至可以分解各种氰化物、除草剂和残留性农药。最后, 巨大芽孢杆菌完全是非病原性的且不产生毒素, 这是非常重要的, 尤其是在生产食品和化妆品方面。正是由于以上这些优点, 巨大芽孢杆菌

已经广泛应用于大量工业生产，如生产 α -和 β -淀粉酶、青霉素酰胺酶、有毒废弃物的处理或需氧维生素 B12 的生产（总结见 Vary, P.S., 1994, 微生物学, 40, 1001-1013, 研究巨大芽孢杆菌的全盛期（Prime time for *Bacillus megaterium*））。

由于在用生物技术生产各种具有工业利益的产品时，巨大芽孢杆菌具有大量优点，因此巨大芽孢杆菌的使用就具有了巨大的经济利益。然而，在需氧微生物的大规模工业发酵中经常会出现一些问题，特别是对细菌培养物进行有效氧供给所引起的问题，这与产量可观的损失相关联。

本发明的一个目的是用巨大芽孢杆菌优化维生素 B12 的制备。

我们已经发现通过使用一种含有巨大芽孢杆菌的培养物制备维生素 B12 的方法可以达到这个目的，其中发酵作用在厌氧条件下进行。

为了本发明的目的，原则上可使用所有适于作为维生素 B12 生产菌的普通巨大芽孢杆菌菌株。为了本发明的目的维生素 B12 的生产菌是指巨大芽孢杆菌菌株或同源微生物，它们已经用经典方法和/或分子遗传学方法进行了改造，使它们的代谢流向在生物合成维生素 B12 或其衍生物的方向上得以增加（代谢工程）。例如，在这些生产菌中将一个或多个位于代谢途径关键位置上的基因和/或相应的酶进行调节上的修饰或甚至去调节，其中所述的基因和/或相应的酶是至关重要的并受到相应的复杂的调节作用（瓶颈）。在这方面本发明包含了所有以前已知的维生素 B12 生产菌，优选地为芽孢杆菌属或同源生物。根据本发明这些具优势的菌株包括，特别是包括巨大芽孢杆菌的 DSMZ 32 菌株、DSMZ 509 菌株和 DSMZ 2894 菌株。

根据本发明，已经表明巨大芽孢杆菌在厌氧条件下生存是可能的，而且在这些条件下维生素 B12 的产量高于需氧条件下的产量。通过比较巨大芽孢杆菌在需氧和厌氧生长条件下的维生素 B12 的产量，清楚地表明在所有被研究的菌株中厌氧条件下维生素 B12 的产量比需氧条件下的产量至少提高 3-4 倍。通过对生长条件、培养基成分和所使用的细菌菌株进行系统的优化实现维生素 B12 产量的进一步提高。

根据本发明，用巨大芽孢杆菌增加维生素 B12 的制备是可能的，如通

过向培养基中至少加入钴。应用本发明的方法，加入如甜菜碱、甲硫氨酸、谷氨酸、二甲基苯并咪唑或胆碱或它们的组合对于维生素 B12 的生产有有利的作用。上述化合物单独或组合后再与钴配合使用可能也是有利的。

本发明相应地也涉及到一种方法，这种方法的不同在于它在培养基中至少加入了钴。也就是说，例如钴可以被单独地加入或至少与甜菜碱、甲硫氨酸、谷氨酸、二甲基苯并咪唑或胆碱或与最后提到的化合物的组合配合加入。在本发明方法的一个改变形式中，通过在每升培养基中加入约 200-750 μM ，优选地加入 250-500 μM 的钴能够提高维生素 B-12 的含量。

本发明的进一步改变形式包括一种上述类型的方法，其中巨大芽孢杆菌最初进行需氧发酵然后再进行厌氧发酵。在一个有利的改变形式中，从需氧发酵到厌氧发酵的转变发生在需氧发酵细胞的指数生长期。在这方面，从需氧到厌氧的转变发生在需氧生长细胞指数生长期的中期或末期，优选地是在末期，是有利的。

在这方面本发明优选的方法是一旦需氧培养物达到最大的光密度就将发酵作用从需氧转变成厌氧，但光密度至少要达到约 2-3。

为了本发明的目的，在本发明方法的单阶段方法和两阶段方法中，厌氧条件都是指将需氧培养后的细菌转移到厌氧瓶并在其中发酵时的种种条件。转移进厌氧瓶内的时间，特别是在两阶段方法中，发生在需氧培养的细菌刚刚达到指数生长期。这表明，转移到厌氧瓶后，细菌只消耗瓶中存在的氧而不再额外供氧。这些条件也可以说成是半厌氧。相应的步骤是常规的实验室操作并且为技术人员所熟知。

起初将细菌在发酵罐里进行需氧发酵然后将氧的供应逐渐减少并最终建立起半厌氧条件也是一种流行的相似方法。

在本发明的一个特殊改变形式中，例如通过向培养基中加入还原剂建立严格的厌氧条件也是可能的。

对于本发明在厌氧条件下（无论是半厌氧或严格的厌氧条件）的发酵，细菌一般不是绝对需要需氧培养（预培养）。这意味着细菌也可在厌氧条件

下培养,随后进一步在半厌氧或严格厌氧条件下发酵。也可以想象到保存的菌株可以直接用于接种并用来在厌氧条件下制备维生素 B12。

在本发明的一个改变形式中,需氧条件下的发酵作用要加入大约 250 μM 的钴;而厌氧条件下的发酵作用加入大约 500 μM 的钴是有利的。

本发明还包括遗传学修饰的细菌菌株,这种菌株可通过经典的诱变或靶向分子生物学技术和恰当的选择方法获得。靶向遗传学操作的令人感兴趣的起始点特别是导致维生素 B12 分支的生物合成途径的点,通过它可使代谢流向有意地引导到维生素 B12 最大产量一方。涉及代谢流向调节作用的基因的靶向修饰也包括对结构基因上游和下游调节区的研发和改变,比如对启动子、增强子、终止子、核糖体结合位点等的优化和/或互换。本发明也包括提高 DNA、mRNA 或由它们编码的蛋白质的稳定性,例如通过降低或阻止核酸酶或蛋白酶的降解作用。

本发明涉及编码参与维生素 B12 生物合成的酶的相应的核苷酸序列。本发明也特别涉及到分离的编码参与尿卟啉原 III 生物合成的酶的核苷酸序列或它的等位基因,其中所述核苷酸序列被组织到巨大芽孢杆菌 hemAXCDBL 操纵子中,这些核苷酸序列显示于 SEQ ID No.1 中。根据本发明,也包括由巨大芽孢杆菌 hemAXCDBL 操纵子编码的且其氨基酸序显示于 SEQ ID No.2-6 的酶和它们的同工酶或修饰型。

同工型的酶是指一级结构不同,但具有相同或相当的底物特异性和作用特异性的酶。

根据本发明,修饰型的酶是指这样一些酶,即它们的序列有所改变,如多肽的 N 末端和/或 C 末端或在氨基酸保守区域发生改变,但酶的功能没有受到破坏。这些改变能够通过氨基酸互换的形式实现,而氨基酸互换可通过本来就已知的方法产生。

本发明的一个特定实施方案中包含本发明多肽的变体,与相应的原始蛋白质相比这些多肽的活性已经被减弱或增强,例如通过氨基酸互换的方法产生变体。该方法同样应用于细胞内的本发明酶的稳定性,例如这些酶对于蛋白酶降解的敏感性会有所增强或减弱。

根据本发明，分离的核酸或分离的核酸片段是指 RNA 或 DNA 的聚合物，它们可以是单链的或双链的，且可任选地含有天然的、化学合成的、修饰的或人造的核苷酸。在这方面，术语 DNA 聚合物也包括 DNA、cDNA 或它们的混合物。

根据本发明，等位基因是指功能相当的核苷酸序列，即本质上具有相同作用的核苷酸序列。功能相当的核苷酸序列是这样一些序列，尽管它们的核苷酸序列不同，但却具有如通过遗传密码的简并性仍具有所需的功能。因此，功能等同物包含此处所述序列的天然存在变体、和人工的核苷酸序列如通过化学合成得到的并且合适时适用于宿主生物密码子使用的核苷酸序列。此外，功能相当的核苷酸序列还包含具有改变了的核苷酸序列的那些序列，其中所述改变了的核苷酸序列赋予酶诸如对抑制剂的不敏感性或抗性。

功能等同物也特别是指最初分离序列的天然或人工突变，这些突变仍显示出所需的功能。突变包括一个以上核苷酸残基的替换、添加、缺失、转位或插入。它们包括所谓的有义突变，它导致在蛋白质水平上出现诸如保守氨基酸的互换，但蛋白质的活性基本不发生改变，因此在功能上是无作用的。突变还包括那些在蛋白质水平影响其 N 末端或 C 末端但对其功能只产生微不足道影响的核苷酸序列的变化。这些变化事实上可对蛋白质的结构起稳定作用。

本发明还包含那些如通过核苷酸序列修饰得到的导致相应衍生物产生的核苷酸序列。此类修饰的目的可以是，如对其中存在的编码序列的进一步定位，或如其它限制性酶切位点的插入。功能等同物也是与原始的基因或基因片段相比功能已经减弱或加强的变体。

此外，本发明还涉及到能够赋予如上所述的预期特性的人工 DNA 序列。如通过蛋白质的反翻译可发现这样的人工 DNA 序列，其中所述蛋白质的反翻译是通过计算机辅助程序（分子模建）或体外选择建立的。根据宿主生物特异的密码子使用对多肽序列进行反翻译所得到的 DNA 编码序列是特别合适的。掌握了分子遗传学方法的技术人员通过对欲用于转化的

生物的其它以前已知的基因进行计算机分析可以容易地发现特异的密码子使用。

根据本发明，同源序列是指那些与本发明核苷酸序列互补的序列和/或与后者杂交的序列。根据本发明，杂交序列这一术语包括能够与先前提到的核苷酸序列在本身已知的严格条件下发生特异相互作用（结合）的基本相似的 DNA 或 RNA 核苷酸序列。这些还包括一些短的核苷酸序列，如长度为 10-30 个核苷酸，优选地是 12-15 个核苷酸。根据本发明，这还包括尤其是所谓的引物或探针。

本发明也包括编码区（结构基因）的前面序列区域（5'-或上游）和/或后面序列区域（3'-或下游）。这些序列区域特别包括那些具有调节功能的序列区域。它们可以影响转录、RNA 的稳定性或 RNA 加工和翻译。调节序列的例子特别是启动子、增强子、操纵基因、终止子或翻译增强子。

本发明进一步包括这样的基因结构，该结构包含上述类型的分离的核苷酸序列或其中部分序列和可操作地与其连接的具有调节功能的序列。

可操作的连接是指诸如启动子、编码序列、终止子和适当时进一步的调节元件的按序排列，这种排列使得在编码序列的表达过程中这些调节元件中的每一个都可以发挥其适当的功能。这些调节性核苷酸序列可以是天然来源的，或是通过化学合成获得的。合适的启动子原则上是指能在适宜的宿主生物中控制基因表达的任何启动子。根据本发明，启动子也可能是化学诱导的启动子，该启动子能够控制受其调控的基因在宿主细胞特定时期表达。 β -半乳糖苷酶或阿拉伯糖系统在此作为实例被提到。

通过常规的重组和克隆技术将合适的启动子和至少一个本发明的核苷酸序列融合构建了基因结构，这些重组和克隆技术是由诸如 Sambrook, J. 等人在 1989, 分子克隆：实验室手册，第二版，冷泉港实验室出版社，冷泉港，纽约一书中描述。

连接物或接头可以附加到片段中用于将 DNA 片段连接起来。

本发明还包括载体，所述载体包含上述类型的分离的核苷酸序列或其部分核苷酸序列或包含上述类型的基因结构，还包含一些额外的核苷酸序

列, 这些额外的核苷酸序列是用于选择、用于在宿主细胞内复制和/或整合到宿主细胞基因组。关于这些额外核苷酸序列的很多例子在文献中有所叙述, 在此不再进一步叙述。

所提及的基因在巨大芽孢杆菌中转化和过表达的适宜体系为诸如 pWH1510 和 pWH1520 质粒以及无质粒的过表达巨大芽孢杆菌菌株 WH320, 这些是由 Rygus T.等人描述的(1991, 在巨大芽孢杆菌中异源基因的诱导型高效表达 (Inducible High-Level Expression of heterologous genes in *Bacillus megaterium*), 应用微生物学和生物技术 (Appl. Microbiol. Biotechnol.), 35, 5: 5594-599)。质粒也是商业可得的 (Q biogens 和 MoBiTec)。然而, 所提到的体系不用于限制本发明。

本发明还涉及到一种方法, 该方法的不同在于它应用 hemAXCDBL 操纵子或其部分呈现增强表达的巨大芽孢杆菌菌株进行发酵。

本发明的一个改变形式中包含了这样一种方法, 在该方法中使用了与没有进行遗传学改变的芽孢杆菌相比其 hemAXCDBL 操纵子或其部分在细胞中以增加的拷贝数存在的遗传学改变的巨大芽孢杆菌菌株进行发酵。拷贝数可在 2 到几百范围内变动。

通过增加适宜基因的拷贝数可以提高基因的表达(过表达)。以一种能够使表达呈现速率增加的方式适当地改变位于结构基因上游的启动子区和/或调节区和/或核糖体结合位点也是进一步可能的。整合到结构基因上游的表达盒也能够以同样的方式发挥作用。此外, 在维生素 B12 的生产中, 通过使用诱导型启动子来增加表达也是可能的。

通过延长 mRNA 寿命的方法同样可以提高表达。基因或基因构建体以不同拷贝数存在于质粒上和/或整合和/或扩增至染色体上。

通过阻止酶蛋白质的降解作用提高或增强酶自身的活性也是进一步可能的。

通过改变培养基的组成和培养物的处理方法也是获得相应基因过表达的另一种可供选择的可能方法。

本发明进一步涉及到用于上述类型的维生素 B12 产生方法中的已经转

化了的巨大芽孢杆菌菌株,这种菌株呈现 hemAXCDBL 操纵子或其部分的核苷酸序列表达增强和/或拷贝数得以增加。

根据本发明,还包括一种转化了的巨大芽孢杆菌菌株,该菌株包含以复制型存在的上述类型的基因结构或载体。

本发明进一步涉及到分离的 hemAXCDBL 操纵子或其部分的核苷酸序列的用途,或进一步涉及到上述类型的基因结构或载体的用途,用于产生上述类型的经转化巨大芽孢杆菌菌株。本发明还包括上述类型的转化了的巨大芽孢杆菌菌株的用途,用于生产维生素 B12。

以下的实施例用于阐明本发明。然而,它们对本发明没有限制作用。

1. 化学试剂和分子生物学试剂

DNA 分离试剂盒	Qiagen
Fast-Link 连接试剂盒	Biozym
分子生物学酶类	Amersham-Pharmacia, NEN-LifeScience
生长培养基	Difco

2. 细菌菌株和质粒

本研究中使用的所有细菌菌株和质粒都列于表 1 和表 2 中。

3. 缓冲液和溶液

3. 1 基本培养基

大肠杆菌 (E.coli) 基本培养基

K ₂ HPO ₄	60.3 mM
KH ₂ PO ₄	33.1 mM
(NH ₄) ₂ SO ₄	7.6 mM
柠檬酸钠	1.7 mM
MgSO ₄	1.0 mM
D-葡萄糖	10.1 mM
硫胺素	3.0 μM
酪蛋白氨基酸	0.0 25%(w/v)

固体培养基中加入 15 g/l 琼脂。

Mopso 基本培养基

Mopso (pH 7.0)	50.0 mM
N-三(羟甲基)甲基甘氨酸 (Tricine) (pH 7.0)	5.0 mM
MgCl₂	520.0 μM
K₂SO₄	276.0 μM
FeSO₄	50.0 μM
CaCl₂	1.0 mM
MnCl₂	100.0 μM
NaCl	50.0 mM
KCl	10.0 mM
K₂HPO₄	1.3 mM
(NH₄)₆Mo₇O₂₄	30.0 pM
H₃BO₃	4.0 nM
CoCl₂	300.0 pM
CuSO₄	100.0 pM
ZnSO₄	100.0 pM
D-葡萄糖	20.2 mM
NH₄Cl	37.4 mM

滴定试剂为 KOH 溶液。

鼠伤寒沙门氏杆菌 (S. typhimurium) 基本培养基

NaCl	8.6 mM
Na₂HPO₄	33.7 mM
KH₂PO₄	22.0 mM
NH₄Cl	18.7mM
D-葡萄糖	20.2 mM
MgSO₄	2.0 mM

CaCl₂ 0.1 mM

固体培养基中加入 15 g/l 琼脂。

3. 2 用于巨大芽孢杆菌的原生质体转化的溶液

SMMP 缓冲液

3 号抗生素培养基 (Difco)	17.5 g/l
蔗糖	500.0 mM
马来酸钠 (pH 6.5)	20.0 mM
MgCl ₂	20.0 mM

滴定试剂为 NaOH 溶液。

PEG-P 溶液

PEG 6000	40.0 %(w/v)
蔗糖	500.0 mM
马来酸钠 (pH 6.5)	20.0 mM
MgCl ₂	20.0 mM

滴定试剂为 NaOH 溶液。

cR5 上层琼脂

蔗糖	300.0 mM
Mops (pH 7.3)	31.1 mM
NaOH	15.0mM
L-脯氨酸	52.1 mM
D-葡萄糖	50.5mM
K ₂ SO ₄	1.3 mM
MgCl ₂ × 6 H ₂ O	45.3 mM
KH ₂ PO ₄	313.0 μ M
CaCl ₂	13.8 mM
琼脂	4.0 g/l
酪蛋白氨基酸	0.2 g/l
酵母提取物	10.0 g/l

滴定试剂为 NaOH 溶液。

4. 培养基和培养基添加剂

4.1 培养基

除非特别声明，均使用 Sambrook J.等人（1989，分子克隆；实验室手册，第2版，冷泉港实验室出版社，冷泉港，纽约）所述的 Luria-Bertani 肉汤(LB)完全培养基。

4.2 添加剂

添加剂，诸如碳源、氨基酸或抗生素，可以加入到培养基中一起高压灭菌，或者用水配成浓缩贮液灭菌，适当时通过过滤除菌。将这些物质加入到已高压灭菌并冷却到 50℃ 以下的培养基中。若含有对光敏感的物质如四环素，要注意在暗处培养。通常使用的终浓度如下：

氨苄青霉素	296 μM
四环素	21 μM
ALA	298 μM
血红素	153 μM
X-Gal	98 μM
甲硫氨酸	335 μM
半胱氨酸	285 μM
硝酸钠	10 mM
亚硝酸钠	10 mM
富马酸二钠	10 mM
葡萄糖	10 mM
氯化铵	37 mM
木糖	33 mM
溶菌酶	10 μg/ml
酪蛋白氨基酸	0.025 % (w/v)

5. 微生物学技术

5.1 灭菌

除非特别说明，所有的培养基和缓冲液均在 120℃，1 巴压力下，蒸汽灭菌 20 分钟。对热敏感物质需通过过滤除菌，玻璃器具在 180℃ 高热灭菌至少 3 小时。

5.2 一般生长条件

需氧细菌培养物在挡板摇瓶中 37℃ 培养，最小转速为 180 转/分钟。培养时间根据预期所要达到得细菌培养物的光密度而有所不同。

5.3 巨大芽孢杆菌生长的条件

为了尽可能给需氧培养物通气，巨大芽孢杆菌应培养于转速为 250 转/分钟的挡板摇瓶中，除非特别说明，温度均为 30℃。厌氧培养物在 30℃ 以 100ml 的体积在小厌氧瓶中发酵，转速为 100 转/分钟。在这两种情况下，都应注意使用恒定质量的培养基，对过夜培养物以 1:100 的比例进行接种，并使用同过夜培养一致的条件。

5.4 细胞密度测定

通过测定 578nm 波长下的光密度来测定细菌培养物的细胞密度，假定 1 OD₅₇₈ 相当于 1×10^9 个细胞。

5.5 比较生长研究

开展了不同的巨大芽孢杆菌菌株的需氧和厌氧生长行为的比较研究，并在图 1、图 2 和图 3 中加以叙述。所使用的菌株是巨大芽孢杆菌菌株 DSMZ32(野生型)或“生产菌株” DSMZ509 和 DSMZ2894，它们适于在需氧条件下进行维生素 B12 的生产。然而，本发明不局限于仅使用这些菌株。也可以想到适于制备维生素 B12 的其它菌株，包括用经典的突变或特殊的分子生物学技术和恰当的筛选方法产生的遗传学修饰的细菌菌株。

在对巨大芽孢杆菌能够厌氧生长的程度进行研究时，将一些备选的电子受体硝酸盐、亚硝酸盐和富马酸盐加入到培养基中以代替需氧电子受体氧(图 4-6)。平行的研究是针对巨大芽孢杆菌是否也能够没有添加这些电子受体的培养基中进行发酵而开展的。

从本发明的结果，在这方面很清楚地看到电子受体亚硝酸盐对细菌的生长具有毒性效应。富马酸盐也抑制生长。所加入的这些潜在电子受体

没有一个能够刺激超出发酵生长程度的厌氧生长。

5.6 维生素 B12 产量的比较研究

对巨大芽孢杆菌在厌氧和需氧条件下维生素 B12 产量也进行了分析。本发明所用的实施例是在厌氧条件下于含葡萄糖 (LB 完全) 培养基中生长的巨大芽孢杆菌菌株 DSMZ32、DSMZ509 和 DSMZ2894, 并且在指数生长期的末期对维生素 B12 的含量进行了测定, 表示为 pmol/OD₅₇₈。作为平行, 对需氧生长的指数生长期中期维生素 B12 的产量也进行了研究。结果示于图 7。

5.7 5-氨基乙酰丙酸 (ALA) 对维生素 B12 产量的影响

合成维生素 B12 中的四吡咯生物合成途径的中心调节点是 5-氨基乙酰丙酸 (ALA) 的形成, 因此本发明研究了 ALA 对巨大芽孢杆菌中维生素 B-12 产量的影响。加入或不加入外源的添加物 ALA 时巨大芽孢杆菌的需氧维生素 B-12 产量描绘于图 8。在此应用的培养基中 ALA 浓度大约为 50 μ g/ml。

根据本发明, ALA 的添加对细菌在需氧和厌氧发酵时维生素 B-12 的产量可能没有影响。据此, 维生素 B-12 生物合成的调节似乎表现为不单独受 ALA 形成的限制。

5.8 维生素 B12 的定量分析

测定 OD₅₇₈ 后, 收获巨大芽孢杆菌需氧生长的指数生长期的中期培养物和厌氧生长的指数生长期末期的培养物, 5000 转/分钟, 离心 15 分钟 (Centrifuge 5403, Eppendorf) 收获培养物。用 40 ml 盐水洗涤, 然后 5000 转/分钟, 离心 15 分钟 (Centrifuge 5403, Eppendorf)。所得到的细胞沉淀物被最终冷冻干燥。鼠伤寒沙门氏杆菌 metE cysG 双突变体 (Raux, E. 等人, 1996, 细菌学杂志 (J.Bacteriol.), 178: 753-767) 接种在含有甲硫氨酸和半胱氨酸的基本培养基中 37℃ 培养过夜, 然后从平板上刮下来并用 40ml 等渗的盐水洗涤。离心后, 细胞沉淀物重悬于等渗的盐水。将洗涤后的细菌培养物小心地与 400 ml 含有半胱氨酸的基本培养基琼脂在 47-48℃ 混合。已经重悬于无菌去离子水并经沸水浴 15 分钟后的 10 μ l 巨大芽孢

杆菌样本涂于冷的平板上, 37℃ 孵育 18 小时。长出的沙门氏菌的菌落直径与所用的巨大芽孢杆菌样本的维生素 B12 含量成比例。通过加入 0.01、0.1、1、10 和 40pmol 的维生素 B12 做出校准图, 通过与此相比较可以推测出所研究的样本中维生素 B12 的含量。这种标准方法可对生物材料中小量维生素 B12 进行快速和可重复性的监测。

6. 分子生物学技术

如 DNA 制备、DNA 限制性酶切、连接、PCR、测序、功能互补、蛋白质表达等普通技术是常规实验室操作的一部分并且由 Sambrook J. 等人 (1989, 分子克隆; 实验室手册, 第 2 版, 冷泉港实验室出版社, 冷泉港, 纽约) 描述。

6.1 巨大芽孢杆菌的原生质体转化

原生质体制备:

将 1ml 巨大芽孢杆菌过夜培养物接种于 50ml LB 培养基中, 置 37℃ 进行培养。当 OD₅₇₈ 为 1 时, 4℃, 15 000 转/分钟离心 15 分钟 (RC 5B Plus, Sorvall), 并将细胞重悬于 5 ml SMMP 缓冲液中。在 SMMP 缓冲液中加入溶菌酶后, 将悬浮物 37℃ 培养 60 分钟, 并在显微镜下监测原生质体的形成。室温 3000 转/分钟离心 (Centrifuge 5403, Eppendorf) 收获细胞, 然后将细胞沉淀物小心重悬于 5ml SMMP 缓冲液中, 重复离心和洗涤步骤一次。当加入 10% (v /v) 甘油后, 可以将原生质体悬浮物分装成小份冻于 -80℃。

转化:

将 500 μl 的原生质体悬浮物与 0.5-1 μg 在 SMMP 缓冲液中的 DNA 进行混合, 并加入 1.5 ml PEG-P 溶液。室温孵育 2 分钟后, 加入 5 ml SMMP 缓冲液, 小心混匀后将悬浮物室温 3000 转/分钟离心 10 分钟 (Centrifuge 5403, Eppendorf)。之后, 立即去除上清液并且将几乎看不见的沉淀物重悬于 500 μl SMMP 缓冲液中。悬浮物在 37℃ 轻轻振荡孵育 90 分钟。然后将 50-200 μl 转化了的细胞与 2.5 ml cR5 上层琼脂混合, 铺到含有适于选择的抗生素的 LB-琼脂平板上。37℃ 培养两天后可见到转化了的菌落。

6.2 巨大芽孢杆菌 hemAXCDBL 操纵子的鉴定和测序

所选择的克隆 hemAXCDBL 操纵子的策略是血红素营养缺陷型大肠杆菌突变体的功能互补。为了达到该目的，通过常规方法制备了巨大芽孢杆菌基因文库。通过大肠杆菌 hemB 突变体 RP523 与已经构建到质粒中的巨大芽孢杆菌基因组文库进行功能互补分析，来分离那些又能够完成四吡咯生物合成即表型上表现为血红素原养野生型的菌落是可能的。制备质粒 DNA 并将存在于质粒上的巨大芽孢杆菌 DNA 进行测序后，可鉴定到编码我们所寻找的 hemAXCDBL 操纵子的大小为 3855bp 的插入片段。相应的核苷酸序列列于 SEQ ID No.1，由此推导出的氨基酸序列列于 SEQ ID No.2-6。

用功能互补法得到的 DNA 片段的上下游序列可以用 Sigma Genosis 公司的 Vectorette TM 系统得到。Stratagene 公司的 Turbo Pfu DNA 聚合酶，由于具有校对功能，因此错误率极其低。

6.3 转化和表达系统

用于在巨大芽孢杆菌中转化并过表达基因的合适质粒是 pWH1510 和 pWH1520，且无质粒的过表达菌株是巨大芽孢杆菌 WH320 (Rygus T. 等人, 1991, 在巨大芽孢杆菌中异源基因的可诱导型高效表达 (Inducible High-Level Expression of heterologous genes in *Bacillus megaterium*), 应用微生物学和生物技术 (Appl. Microbiol And. Biotechnol.), 35, 5: 594-599)。

对照质粒 pWH1510 在打断的 xylA 阅读框中包含 spoVG-lacZ 融合结构。在本例中，spoVG-lacZ 融合结构是指枯草芽孢杆菌芽孢形成蛋白的非常强的核糖体结合序列 (spoVG) 与大肠杆菌编码 β 半乳糖苷酶的基因 (lacZ) 融合。因此这种质粒非常适于研究对于巨大芽孢杆菌的转化效率和过表达条件。

质粒 pWH1520 是真正的能够起作用的克隆和表达载体。这两种载体具有四环素抗性和氨苄青霉素抗性，并含有在大肠杆菌和芽孢杆菌亚种 (*Bacillus* spp) 中复制至关重要的元件。因此，对于建立在大肠杆菌基础

上的用于操作质粒 pBR322 衍生物的所有技术也同样适用于它们。这两个载体均包含巨大芽孢杆菌 xyl 操纵子的 xylA 和 xylR 基因和它们的调节序列 (Rygus 等人, 1991, 分子克隆, 结构; 用于转录巨大芽孢杆菌中编码木糖应用型调节子的启动子和调节元件 (Molecular Cloning, Structure; Promoters and Regulatory Elements for Transcription of the *Bacillus megaterium* Encoded Regulon for Xylose Utilization)), 微生物学文献 (Arch. Microbiol.), 155: 535-542)。xylA 编码木糖异构酶而 xylR 编码一种调节蛋白质, 这种调节蛋白质对 xylA 进行强转录调控。当缺乏木糖时, xylA 受到抑制。当加入木糖时, 通过 xylA 的去抑制作用可以产生约 200 倍的诱导。处于 xylA 阅读框的多接头使得将基因与 xylA 融合在一起成为可能, 同样, 所述基因也处于 XylR 的强转录控制之下。而且, 由于位于多接头上游的 xylA 阅读框是完整的, 所以选择是形成转录融合还是形成翻译融合也是可能的。

图表说明

所使用的细菌菌株和质粒列于表 1 和表 2 中。

表 1: 所使用的细菌菌株

表 2: 所使用的质粒

用下列图进一步解释本发明。

图 1 显示 30℃ 下巨大芽孢杆菌 DSMZ32 (野生型) 在需氧和厌氧条件下生长情况的比较。测定了在加入 10mM 硝酸盐 (空心菱形)、10mM 亚硝酸盐 (空心三角形) 和 10mM 富马酸盐 (十字叉) 条件下的厌氧生长情况。测定了在不含添加物的 LB 培养基中的发酵 (空心圆形) 和需氧 (实心菱形) 生长情况。在所示时间抽取样本并测定 578nm 波长下的光密度值。

图 2 显示 30℃ 下巨大芽孢杆菌 DSMZ509 在需氧和厌氧条件下生长情况的比较。测定了在加入 10mM 硝酸盐 (空心菱形)、10mM 亚硝酸盐 (空心三角形) 和 10mM 富马酸盐 (十字叉) 条件下的厌氧生长情况。测定了在不含添加物的 LB 培养基中的发酵 (空心圆形) 和需氧 (实心菱形) 生

长情况。在所示时间抽取样本并测定 578nm 波长下的光密度值。

图 3 显示 30℃ 下巨大芽孢杆菌 DSMZ2894 在需氧和厌氧条件下生长情况的比较。测定了在加入 10mM 硝酸盐（空心菱形）、10mM 亚硝酸盐（空心三角形）和 10mM 富马酸盐（十字叉）条件下的厌氧生长情况。测定了在不含添加物的 LB 培养基中的发酵（空心圆形）和需氧（实心菱形）生长情况。在所示时间抽取样本并测定 578nm 波长下的光密度值。

图 4 显示 30℃ 下巨大芽孢杆菌 DSMZ32（野生型）在加入 10mM 硝酸盐（菱形）、10mM 亚硝酸盐（三角形）和 10mM 富马酸盐（十字叉）条件下的厌氧生长情况。在不含添加物的 LB 培养基中的发酵（圆形）生长情况。在所示时间抽取样本并测定 578nm 波长下的光密度值。

图 5 显示 30℃ 下巨大芽孢杆菌 DSMZ509 在加入 10mM 硝酸盐（菱形）、10mM 亚硝酸盐（三角形）和 10mM 富马酸盐（十字叉）条件下的厌氧生长情况。在不含添加物的 LB 培养基中的发酵（圆形）生长情况。在所示时间抽取样本并测定 578nm 波长下的光密度值。

图 6 显示 30℃ 下巨大芽孢杆菌 DSMZ2894 在加入 10mM 硝酸盐（菱形）、10mM 亚硝酸盐（三角形）和 10mM 富马酸盐（十字叉）条件下的厌氧生长情况。在不含添加物的 LB 培养基中的发酵（圆形）生长情况。在所示时间抽取样本并测定 578nm 波长下的光密度值。

图 7 显示在需氧和厌氧生长条件下，巨大芽孢杆菌维生素 B12 的产量。对需氧生长的（1）和厌氧生长的（2）野生型巨大芽孢杆菌菌株 DSMZ32、需氧生长的（3）和厌氧生长的（4）巨大芽孢杆菌 DSMZ509、需氧生长的（5）和厌氧生长的（6）巨大芽孢杆菌 DSMZ2894 进行了生物量中维生素 B12 含量测定，以 pmol/OD₅₇₈ 为单位表示。

图 8 显示添加和不添加 50 μg/ml 外源 ALA 的条件下巨大芽孢杆菌的需氧维生素 B12 的产量。对不添加（1）和添加（2）ALA 的野生型巨大芽孢杆菌 DSMZ32、不添加（3）和添加（4）ALA 的巨大芽孢杆菌 DSMZ509、不添加（5）和添加（6）ALA 的巨大芽孢杆菌 DSMZ2894 进行了每生物量中维生素 B12 含量的测定，以 pmol/OD₅₇₈ 为单位表示。

<110> 巴斯福股份公司

<120> 制备维生素 B12 的方法

<130> Sample File Reference

<140> j

<141> 2001-08-01

<160> 7

<170> PatentIn 版本 2.1

<210> 1

<211> 7193

<212> DNA

<213> 巨大芽孢杆菌(Bacillus megaterium)

<220>

<221> CDS

<222> (439)..(1779)

<223> hemA

<220>

<221> CDS

<222> (1794)..(2627)

<223> hemX

<220>

<221> CDS

<222> (2641)..(3570)

<223> hemC

<220>

<221> CDS

<222> (3573)..(4346)

<223> hemD

<220>

<221> CDS

<222> (4364)..(5344)

<223> hemB

<220>

<221> CDS

<222> (5361)..(6650)

<223> hemL

<400> 1

caggccacgt gtataccttt gctgcaatgc tcatatgtag ttcattcagt aaaaaaaga 60

ttccaaatcc aactagtaag acaccaaaaa acattgattg ctttttcata gtaacctcca 120

ggtgtttcaa aacagaaaaac gtgataaaac gctgcttttt tgatacaatt tattacggaa 180

taacagttaa tatcttaaca catgttttca aagactgttc acatTTTTTT ttcgaatgac 240

ttatagaaca tgttataata gttagagata aggcgaagct tacacctgtg taatgtacgc 300

```

cgcttttttac gaatagtaga atgaagaacg ccggatttta cgggcagtta tctcactgta 360
gctcttttggg aatttagggg ttactgccccg tcaaattcgg ataaataagc acatttcaat 420

tataggggggt atgtacaa atg cat att ata gca gtc ggc tta aac ttt cga 471
                Met His Ile Ile Ala Val Gly Leu Asn Phe Arg
                  1                5                10

aca gcc cct gtt gaa ata aga gaa aag ctt agt ttt aat gag caa gaa 519
Thr Ala Pro Val Glu Ile Arg Glu Lys Leu Ser Phe Asn Glu Gln Glu
                  15                20                25

cta gca tct gcc atg aaa acg tta agc ggc caa aaa agt att ttg gaa 567
Leu Ala Ser Ala Met Lys Thr Leu Ser Gly Gln Lys Ser Ile Leu Glu
                  30                35                40

aac atc att gtt tca aca tgc aat cgt acg gaa att tat gct gtt gtt 615
Asn Ile Ile Val Ser Thr Cys Asn Arg Thr Glu Ile Tyr Ala Val Val
                  45                50                55

gat caa tta cac acg ggc cgt tat tat gta aaa gca ttt tta gct gaa 663
Asp Gln Leu His Thr Gly Arg Tyr Tyr Val Lys Ala Phe Leu Ala Glu
                  60                65                70                75

tgg ttc gga ata gat aaa gaa gag ttt tcg cct tat tta acc atc tat 711
Trp Phe Gly Ile Asp Lys Glu Glu Phe Ser Pro Tyr Leu Thr Ile Tyr
                  80                85                90

gaa aat gat gga gcc atc gaa cat tta tac cga gta gca tgt gga tta 759
Glu Asn Asp Gly Ala Ile Glu His Leu Tyr Arg Val Ala Cys Gly Leu
                  95                100                105

gat tca atg gtg att ggt gaa aca caa att ctt gga caa gtg cgt tca 807
Asp Ser Met Val Ile Gly Glu Thr Gln Ile Leu Gly Gln Val Arg Ser
                  110                115                120

agc ttt ttg ctt gct caa gaa gaa gaa aca att gga acg gtt ttt aat 855
Ser Phe Leu Leu Ala Gln Glu Glu Glu Thr Ile Gly Thr Val Phe Asn
                  125                130                135

cag ctg ttt aaa caa gct gtg aca tta gcc aaa aaa gct cat cat gaa 903
Gln Leu Phe Lys Gln Ala Val Thr Leu Ala Lys Lys Ala His His Glu
                  140                145                150                155

aca gaa att ggt gca aat gca gtt tct gtc agc tat gca gcc gta gag 951
Thr Glu Ile Gly Ala Asn Ala Val Ser Val Ser Tyr Ala Ala Val Glu
                  160                165                170

ctt gct aaa aag att ttt gga gac tta tcg tcc aaa cat gtg ttg att 999
Leu Ala Lys Lys Ile Phe Gly Asp Leu Ser Ser Lys His Val Leu Ile
                  175                180                185

ctt ggt gcg gga aaa atg ggt cag ctt gcc gta cag aat tta tat ggc 1047
Leu Gly Ala Gly Lys Met Gly Gln Leu Ala Val Gln Asn Leu Tyr Gly
                  190                195                200

agc gga gct aaa aaa gta acg gta gtg aat cga aca ttt gaa aag gca 1095
Ser Gly Ala Lys Lys Val Thr Val Val Asn Arg Thr Phe Glu Lys Ala
                  205                210                215

caa gag cta gca aac cgt ttt tct gga gaa gcc aaa cca ttt gct gat 1143
Gln Glu Leu Ala Asn Arg Phe Ser Gly Glu Ala Lys Pro Phe Ala Asp
                  220                225                230                235

```

ctt caa cat gct tta agc gaa gca gat atc tta att agt tca aca ggg	1191
Leu Gln His Ala Leu Ser Glu Ala Asp Ile Leu Ile Ser Ser Thr Gly	
240 245 250	
gct aat gac tat gta gta acg aaa caa atg atg agt gaa gcc gag cgt	1239
Ala Asn Asp Tyr Val Val Thr Lys Gln Met Met Ser Glu Ala Glu Arg	
255 260 265	
acg cgt aaa ggt cgt cca ttg ttt atg gta gac atc gcc gtt cca cgt	1287
Thr Arg Lys Gly Arg Pro Leu Phe Met Val Asp Ile Ala Val Pro Arg	
270 275 280	
gac tta gat cca gaa cta gat gag cta gag acg gtc ttt tta tat gat	1335
Asp Leu Asp Pro Glu Leu Asp Glu Leu Glu Thr Val Phe Leu Tyr Asp	
285 290 295	
atc gac gat tta aat ggg att gta gaa tcc aat ttg caa gag cgt caa	1383
Ile Asp Asp Leu Asn Gly Ile Val Glu Ser Asn Leu Gln Glu Arg Gln	
300 305 310 315	
aaa gct gcg gat gaa att gaa att atg ctt gaa gct gaa att gtt gct	1431
Lys Ala Ala Asp Glu Ile Glu Ile Met Leu Glu Ala Glu Ile Val Ala	
320 325 330	
ttc aag tct tgg ctt ggt aca tta ggc gtt gtg cct gtt att tct gca	1479
Phe Lys Ser Trp Leu Gly Thr Leu Gly Val Val Pro Val Ile Ser Ala	
335 340 345	
ctt cgt caa aaa gca ctt acc att caa gct gaa acg atg aaa agc att	1527
Leu Arg Gln Lys Ala Leu Thr Ile Gln Ala Glu Thr Met Lys Ser Ile	
350 355 360	
gat cgt aag ctc cca gat tta tct gag cgt gag cgc aaa gtt tta aat	1575
Asp Arg Lys Leu Pro Asp Leu Ser Glu Arg Glu Arg Lys Val Leu Asn	
365 370 375	
aaa cat aca aaa agt ata att aat caa ctt ctt cgc gat cca att cca	1623
Lys His Thr Lys Ser Ile Ile Asn Gln Leu Leu Arg Asp Pro Ile Pro	
380 385 390 395	
cat gcg aaa gaa ttg gct gga gag aaa cat gct gaa gag tct tta gag	1671
His Ala Lys Glu Leu Ala Gly Glu Lys His Ala Glu Glu Ser Leu Glu	
400 405 410	
ctg ttc atg aag att ttt aat att gaa cag gaa gtt gct ttg caa aaa	1719
Leu Phe Met Lys Ile Phe Asn Ile Glu Gln Glu Val Ala Leu Gln Lys	
415 420 425	
gaa aaa gaa gag cag ctt cat aca tcg att acg tct cat tca gta gct	1767
Glu Lys Glu Glu Gln Leu His Thr Ser Ile Thr Ser His Ser Val Ala	
430 435 440	
tat cag tct taa gtgaggattg aata atg gtt gat atg gaa tta acc aga	1817
Tyr Gln Ser Met Val Asp Met Glu Leu Thr Arg	
445 450 455	
atc tat gaa ttg att gtg att tta tac gcg gtg agc gtt tta tta tat	1865
Ile Tyr Glu Leu Ile Val Ile Leu Tyr Ala Val Ser Val Leu Leu Tyr	
460 465 470	
ttt att gat ttt gta caa aat aac cgg aag gca aac tcg ata gcc ttc	1913
Phe Ile Asp Phe Val Gln Asn Asn Arg Lys Ala Asn Ser Ile Ala Phe	

475					480					485						
tgg	tta	ctt	tct	att	gtg	tgg	gtt	ttg	caa	aca	gcc	ttt	tta	att	tta	1961
Trp	Leu	Leu	Ser	Ile	Val	Trp	Val	Leu	Gln	Thr	Ala	Phe	Leu	Ile	Leu	
		490					495					500				
cga	atg	ttt	gag	acg	gga	agg	ttt	cct	att	ctt	acc	gtt	tct	gaa	ggg	2009
Arg	Met	Phe	Glu	Thr	Gly	Arg	Phe	Pro	Ile	Leu	Thr	Val	Ser	Glu	Gly	
		505				510					515					
ctg	tat	ttt	tac	gcg	tgg	gtg	tta	att	act	ctt	tca	ctt	gtg	att	aat	2057
Leu	Tyr	Phe	Tyr	Ala	Trp	Val	Leu	Ile	Thr	Leu	Ser	Leu	Val	Ile	Asn	
		520				525					530				535	
cgc	ctg	ttc	cgc	ggt	gat	ttt	atg	gtg	ttt	ttt	aca	aat	gtg	att	ggg	2105
Arg	Leu	Phe	Arg	Val	Asp	Phe	Met	Val	Phe	Phe	Thr	Asn	Val	Ile	Gly	
				540					545					550		
ttc	tta	att	atg	gct	att	cat	acg	ttt	gct	cca	ctt	gaa	aca	tac	act	2153
Phe	Leu	Ile	Met	Ala	Ile	His	Thr	Phe	Ala	Pro	Leu	Glu	Thr	Tyr	Thr	
			555					560						565		
gct	gtc	cag	tcg	gag	cgc	ctt	gtg	tca	gaa	tta	tta	atc	atc	cat	att	2201
Ala	Val	Gln	Ser	Glu	Arg	Leu	Val	Ser	Glu	Leu	Leu	Ile	Ile	His	Ile	
		570					575					580				
acc	atg	gct	att	tta	tct	tac	ggc	gct	ttt	tca	gtc	tct	ttt	ggt	ttt	2249
Thr	Met	Ala	Ile	Leu	Ser	Tyr	Gly	Ala	Phe	Ser	Val	Ser	Phe	Val	Phe	
		585				590					595					
tcc	gct	ctg	tat	ctt	ggt	cag	tat	aat	tta	ttg	aaa	aag	aag	aaa	tgg	2297
Ser	Ala	Leu	Tyr	Leu	Val	Gln	Tyr	Asn	Leu	Leu	Lys	Lys	Lys	Lys	Trp	
		600				605				610					615	
ggc	aag	cgc	ttg	cta	aga	att	gaa	gat	tta	aca	aag	cta	gat	tac	atg	2345
Gly	Lys	Arg	Leu	Leu	Arg	Ile	Glu	Asp	Leu	Thr	Lys	Leu	Asp	Tyr	Met	
				620					625					630		
tcc	tat	gta	tta	att	tta	att	ggg	gta	ccg	ctg	tta	ttg	ctt	tcg	tta	2393
Ser	Tyr	Val	Leu	Ile	Leu	Ile	Gly	Val	Pro	Leu	Leu	Leu	Leu	Ser	Leu	
			635					640						645		
atc	tta	gga	atc	ata	tgg	gct	tat	atc	aaa	tac	gta	gat	ttt	cat	tgg	2441
Ile	Leu	Gly	Ile	Ile	Trp	Ala	Tyr	Ile	Lys	Tyr	Val	Asp	Phe	His	Trp	
		650				655					660					
tac	gat	aca	aaa	gtg	tta	ggg	tca	ttt	tta	atg	ttt	gtt	gca	tac	agc	2489
Tyr	Asp	Thr	Lys	Val	Leu	Gly	Ser	Phe	Leu	Met	Phe	Val	Ala	Tyr	Ser	
		665				670					675					
gct	tac	ttg	tat	gct	aga	ctg	cgt	aga	ggg	att	caa	gga	aag	gca	atc	2537
Ala	Tyr	Leu	Tyr	Ala	Arg	Leu	Arg	Arg	Gly	Ile	Gln	Gly	Lys	Ala	Ile	
		680				685				690					695	
gcg	atg	tgg	aat	att	gga	ttg	ttt	tta	gtc	tta	ctt	att	aat	ttc	ttt	2585
Ala	Met	Trp	Asn	Ile	Gly	Leu	Phe	Leu	Val	Leu	Leu	Ile	Asn	Phe	Phe	
				700					705					710		
tta	ttt	ggc	agt	tta	tcg	aac	ttt	cac	ttt	tgg	ggg	tcg	taa			2627
Leu	Phe	Gly	Ser	Leu	Ser	Asn	Phe	His	Phe	Trp	Gly	Ser				
			715					720				725				
ttaggaggac	aac	atg	cga	aaa	att	att	gtc	ggc	tct	cgt	cga	agc	aaa			2676

Met Arg Lys Ile Ile Val Gly Ser Arg Arg Ser Lys															
730								735							
ctt gcg tta aca cag acc aag tgg gtc att gaa cag tta aaa aaa caa	2724														
Leu Ala Leu Thr Gln Thr Lys Trp Val Ile Glu Gln Leu Lys Lys Gln															
740	745							750							
ggt ctt cct ttt gag ttt gaa att aaa gaa atg gtt aca aaa ggg gac	2772														
Gly Leu Pro Phe Glu Phe Glu Ile Lys Glu Met Val Thr Lys Gly Asp															
755	760							765							
caa att tta aat gta acg ctg tcc aaa gta ggc ggt aaa ggt ctt ttt	2820														
Gln Ile Leu Asn Val Thr Leu Ser Lys Val Gly Gly Lys Gly Leu Phe															
770	775							780	785						
gta aaa gaa att gaa caa gcc atg tta gat aaa gaa att gat atg gcc	2868														
Val Lys Glu Ile Glu Gln Ala Met Leu Asp Lys Glu Ile Asp Met Ala															
790	795							800							
gtt cat agt atg aaa gat atg ccg gct gtt ttg cca gaa gga tta acg	2916														
Val His Ser Met Lys Asp Met Pro Ala Val Leu Pro Glu Gly Leu Thr															
805	810							815							
att gga tgt att cct tta cga gaa gat cat cgt gat gcg ctt att tca	2964														
Ile Gly Cys Ile Pro Leu Arg Glu Asp His Arg Asp Ala Leu Ile Ser															
820	825							830							
aaa aat ggt gag cga ttt gag gag cta cca agc ggt gct gtg att ggg	3012														
Lys Asn Gly Glu Arg Phe Glu Glu Leu Pro Ser Gly Ala Val Ile Gly															
835	840							845							
aca agc agc ttg cgt cga ggg gca cag ctt cta tct atg cgt tca gat	3060														
Thr Ser Ser Leu Arg Arg Gly Ala Gln Leu Leu Ser Met Arg Ser Asp															
850	855							860	865						
atc gaa att aaa tgg att cga gga aat atc gac aca cgc cta gag aaa	3108														
Ile Glu Ile Lys Trp Ile Arg Gly Asn Ile Asp Thr Arg Leu Glu Lys															
870	875							880							
tta aaa aat gaa gat tac gat gcg att atc ttg gca gct gca ggg tta	3156														
Leu Lys Asn Glu Asp Tyr Asp Ala Ile Ile Leu Ala Ala Ala Gly Leu															
885	890							895							
tct cgt atg ggc tgg tcg aaa gat acg gtc act cag tac cta gag ccg	3204														
Ser Arg Met Gly Trp Ser Lys Asp Thr Val Thr Gln Tyr Leu Glu Pro															
900	905							910							
gaa ata agc gtt ccc gct gta gga caa gga gct tta gct atc gaa tgc	3252														
Glu Ile Ser Val Pro Ala Val Gly Gln Gly Ala Leu Ala Ile Glu Cys															
915	920							925							
cga gaa aat gac cat gaa tta tta tcg ctt ctt caa gct cta aat cat	3300														
Arg Glu Asn Asp His Glu Leu Leu Ser Leu Leu Gln Ala Leu Asn His															
930	935							940	945						
gat gaa acg gct cgt gct gtt aga gct gaa cgc gta ttt tta aaa gag	3348														
Asp Glu Thr Ala Arg Ala Val Arg Ala Glu Arg Val Phe Leu Lys Glu															
950	955							960							
atg gaa ggc gga tgc caa gtg ccg att gcg ggt tat ggg cgt ata ttg	3396														
Met Glu Gly Gly Cys Gln Val Pro Ile Ala Gly Tyr Gly Arg Ile Leu															
965	970							975							

gac ggc gga aac att gag cta acg tcg cta gtc gcg tct cct gat ggc	3444
Asp Gly Gly Asn Ile Glu Leu Thr Ser Leu Val Ala Ser Pro Asp Gly	
980 985 990	
aaa acg att tac aaa gag cac att acg gga aaa gac ccc att gct atc	3492
Lys Thr Ile Tyr Lys Glu His Ile Thr Gly Lys Asp Pro Ile Ala Ile	
995 1000 1005	
ggt tca gaa gct gct gag cgc tta aca tct caa gga gct aag ctt tta	3540
Gly Ser Glu Ala Ala Glu Arg Leu Thr Ser Gln Gly Ala Lys Leu Leu	
1010 1015 1020 1025	
att gac cgt gta aaa gag gag ctg gac aaa ta atg agt cat gag gct	3587
Ile Asp Arg Val Lys Glu Glu Leu Asp Lys Met Ser His Glu Ala	
1030 1035 1040	
cct ctt ttt ttg aaa aaa gtt ttg att aca aga gaa gcg aag cag gct	3635
Pro Leu Phe Leu Lys Lys Val Leu Ile Thr Arg Glu Ala Lys Gln Ala	
1045 1050 1055	
cat tca ttt gca gca cag att aaa aag ctt gga ggc ata cct tta agc	3683
His Ser Phe Ala Ala Gln Ile Lys Lys Leu Gly Gly Ile Pro Leu Ser	
1060 1065 1070	
tta cca gtc ctc acg ttc gca cgt tct aac aat gaa aaa gaa gtg gaa	3731
Leu Pro Val Leu Thr Phe Ala Arg Ser Asn Asn Glu Lys Glu Val Glu	
1075 1080 1085	
gat gtc ctg ctg tcg att gac tgt ttt gat tgg att gtt ttt aca agt	3779
Asp Val Leu Leu Ser Ile Asp Cys Phe Asp Trp Ile Val Phe Thr Ser	
1090 1095 1100	
caa aac ggc att tct ttt ttc ttt gaa tgg cta aaa gag tta aac atc	3827
Gln Asn Gly Ile Ser Phe Phe Phe Glu Trp Leu Lys Glu Leu Asn Ile	
1105 1110 1115 1120	
agt tgg agt gct tta aaa aga tta aaa gtg gct gcg gtt gga gaa aaa	3875
Ser Trp Ser Ala Leu Lys Arg Leu Lys Val Ala Ala Val Gly Glu Lys	
1125 1130 1135	
acc gca aaa ttg cta gaa aaa cac gat gtt aac gtg caa ctt gtt cca	3923
Thr Ala Lys Leu Leu Glu Lys His Asp Val Asn Val Gln Leu Val Pro	
1140 1145 1150	
aaa gaa tac gtt gcg gaa tca ctt ctt gag tcg tta aaa gca aat gtg	3971
Lys Glu Tyr Val Ala Glu Ser Leu Leu Glu Ser Leu Lys Ala Asn Val	
1155 1160 1165	
tca cat cac gag cga gtt ctc ctt gta cga gga caa tta gcc cgc tct	4019
Ser His His Glu Arg Val Leu Leu Val Arg Gly Gln Leu Ala Arg Ser	
1170 1175 1180	
ggt tta aaa gat ggg tta aaa gag tgt gag gta acg gat tta gtt gta	4067
Val Leu Lys Asp Gly Leu Lys Glu Cys Glu Val Thr Asp Leu Val Val	
1185 1190 1195 1200	
tat cag acc att cga aat caa gac tct gaa cgc ttg gtt aag caa cat	4115
Tyr Gln Thr Ile Arg Asn Gln Asp Ser Glu Arg Leu Val Lys Gln His	
1205 1210 1215	
gta gaa aca gga att gat gcc att act ttt aca agt tct tct aca gtg	4163
Val Glu Thr Gly Ile Asp Ala Ile Thr Phe Thr Ser Ser Thr Val	
1220 1225 1230	

aga ttt ttt gtc gag tca gtt aaa gat att ccg aat tgg ctt gag ctt	4211
Arg Phe Phe Val Glu Ser Val Lys Asp Ile Pro Asn Trp Leu Glu Leu	
1235 1240 1245	
ata aat aga acc aaa gtt gtt tgc atc ggt cct att acg agt cag aca	4259
Ile Asn Arg Thr Lys Val Val Cys Ile Gly Pro Ile Thr Ser Gln Thr	
1250 1255 1260	
gcg agc gaa tat gga att aag cat ctt gtt cct aac aca tac aca ata	4307
Ala Ser Glu Tyr Gly Ile Lys His Leu Val Pro Asn Thr Tyr Thr Ile	
1265 1270 1275 1280	
aat ggt atg ata gat atg ttg gtt agt tgt ttt gaa taa taaaggagga	4356
Asn Gly Met Ile Asp Met Leu Val Ser Cys Phe Glu	
1285 1290	
ttttaac atg aaa gat tta caa ttt aca cgt cat cgt cgt ctt cgt aat	4405
Met Lys Asp Leu Gln Phe Thr Arg His Arg Arg Leu Arg Asn	
1295 1300 1305	
tct gta aat atg cgt gcg ctt gta cgt gaa aca cat tta cat cct gaa	4453
Ser Val Asn Met Arg Ala Leu Val Arg Glu Thr His Leu His Pro Glu	
1310 1315 1320	
gac ttt ata tat cct att ttt att gta gaa ggt gaa caa aaa aga aat	4501
Asp Phe Ile Tyr Pro Ile Phe Ile Val Glu Gly Glu Gln Lys Arg Asn	
1325 1330 1335	
gct gta aaa tca atg cca ggt gtg gat caa att tct tta gac tac ttg	4549
Ala Val Lys Ser Met Pro Gly Val Asp Gln Ile Ser Leu Asp Tyr Leu	
1340 1345 1350 1355	
aat gat gaa atc caa gag tta gtc gat tta gga atc aag tca gct atg	4597
Asn Asp Glu Ile Gln Glu Leu Val Asp Leu Gly Ile Lys Ser Ala Met	
1360 1365 1370	
gta ttt ggt gtt ccc gct gaa aaa gat gcg gtt ggc tca cag gca tac	4645
Val Phe Gly Val Pro Ala Glu Lys Asp Ala Val Gly Ser Gln Ala Tyr	
1375 1380 1385	
cat gac cac gga att gtc caa aaa ggt att cgt caa att aaa gaa aat	4693
His Asp His Gly Ile Val Gln Lys Gly Ile Arg Gln Ile Lys Glu Asn	
1390 1395 1400	
ttt cca gac ttt gtt gtt atc gca gat aca tgc ttg tgt caa tat aca	4741
Phe Pro Asp Phe Val Val Ile Ala Asp Thr Cys Leu Cys Gln Tyr Thr	
1405 1410 1415	
gat cat ggc cac tgt ggt gta att gaa gac gga aaa atc tta aac gat	4789
Asp His Gly His Cys Gly Val Ile Glu Asp Gly Lys Ile Leu Asn Asp	
1420 1425 1430 1435	
cct agc ctt gac tta tta gcg cgt aca gcg gtt agt caa gct aaa gca	4837
Pro Ser Leu Asp Leu Leu Ala Arg Thr Ala Val Ser Gln Ala Lys Ala	
1440 1445 1450	
gga gcg gat att att gcg cct tca aac atg atg gac ggt ttc gtg gca	4885
Gly Ala Asp Ile Ile Ala Pro Ser Asn Met Met Asp Gly Phe Val Ala	
1455 1460 1465	
gct att cgc cac ggt cta gac gaa gct gga ttt gaa gat gtg cct gtg	4933
Ala Ile Arg His Gly Leu Asp Glu Ala Gly Phe Glu Asp Val Pro Val	

1470	1475	1480	
atg tct tat gct gtt aaa tat gca tct gct ttt tac gga cca ttc cgc Met Ser Tyr Ala Val Lys Tyr Ala Ser Ala Phe Tyr Gly Pro Phe Arg 1485 1490 1495			4981
gat gca gcg cat tct tct cct caa ttt gga gac cgt aaa aca tat caa Asp Ala Ala His Ser Ser Pro Gln Phe Gly Asp Arg Lys Thr Tyr Gln 1500 1505 1510 1515			5029
atg gac cca gcc aat cgc tta gaa gca ctt cga gaa gcg gaa tct gat Met Asp Pro Ala Asn Arg Leu Glu Ala Leu Arg Glu Ala Glu Ser Asp 1520 1525 1530			5077
gtc gaa gaa gga gca gat ttc tta att gta aaa cca gct cta tca tat Val Glu Glu Gly Ala Asp Phe Leu Ile Val Lys Pro Ala Leu Ser Tyr 1535 1540 1545			5125
tta gat att atg cgc gat gta aaa aat acg ttt aat tta ccg gta gtg Leu Asp Ile Met Arg Asp Val Lys Asn Thr Phe Asn Leu Pro Val Val 1550 1555 1560			5173
gct tat aat gtg agc ggt gaa tat tca atg att aaa gcg gca gct caa Ala Tyr Asn Val Ser Gly Glu Tyr Ser Met Ile Lys Ala Ala Ala Gln 1565 1570 1575			5221
aac ggc tgg gta aat gaa aaa gag att gtt ctt gaa aaa tta att agc Asn Gly Trp Val Asn Glu Lys Glu Ile Val Leu Glu Lys Leu Ile Ser 1580 1585 1590 1595			5269
atg aag cgt gca gga gcg gat tta att gta acg tat cat gct aaa gac Met Lys Arg Ala Gly Ala Asp Leu Ile Val Thr Tyr His Ala Lys Asp 1600 1605 1610			5317
gca gca cgc tgg tta tct gac aag taa ttaaggaggc ttacac atg aga agt Ala Ala Arg Trp Leu Ser Asp Lys Met Arg Ser 1615 1620			5369
tat gaa aag tca aaa gca gct ttt tta gaa gcg cag caa tta atg ccg Tyr Glu Lys Ser Lys Ala Ala Phe Leu Glu Ala Gln Gln Leu Met Pro 1625 1630 1635			5417
ggt gga gta aac agt cct gtt cgt gca ttt aaa tca gta gat atg gat Gly Gly Val Asn Ser Pro Val Arg Ala Phe Lys Ser Val Asp Met Asp 1640 1645 1650 1655			5465
cct att ttt atg gaa cga gga aaa ggc gct acg att tac gat att gac Pro Ile Phe Met Glu Arg Gly Lys Gly Ala Thr Ile Tyr Asp Ile Asp 1660 1665 1670			5513
ggc aac aaa tat att gac tac gtc tta tca tgg gga cct tta att cac Gly Asn Lys Tyr Ile Asp Tyr Val Leu Ser Trp Gly Pro Leu Ile His 1675 1680 1685			5561
gga cat tcc aac gat caa gta gtt gaa gca gtc aaa aaa gta acg gaa Gly His Ser Asn Asp Gln Val Val Glu Ala Val Lys Lys Val Thr Glu 1690 1695 1700			5609
tct ggt aca agc ttt ggt gca cca aca tta att gaa aat gag ctt gct Ser Gly Thr Ser Phe Gly Ala Pro Thr Leu Ile Glu Asn Glu Leu Ala 1705 1710 1715			5657
aag gtt gtg att gag cgc gtt cct tca atc gaa atc gtc cgt atg gtg			5705

Lys Val Val Ile Glu Arg Val Pro Ser Ile Glu Ile Val Arg Met Val	
1720 1725 1730 1735	
agt tca ggt acg gaa gca acg atg agc gca ctg cgt tta gca cgt ggc	5753
Ser Ser Gly Thr Glu Ala Thr Met Ser Ala Leu Arg Leu Ala Arg Gly	
1740 1745 1750	
tat acg ggt cgc aac aaa att tta aaa ttt gaa ggc tgt tac cac gga	5801
Tyr Thr Gly Arg Asn Lys Ile Leu Lys Phe Glu Gly Cys Tyr His Gly	
1755 1760 1765	
cac ggt gac tct ctt cta att aaa gct gga tca ggc gtt gca acg tta	5849
His Gly Asp Ser Leu Leu Ile Lys Ala Gly Ser Gly Val Ala Thr Leu	
1770 1775 1780	
ggc cta cca gat agt ccg gga gtg cct gaa gga atc gct aaa aac acg	5897
Gly Leu Pro Asp Ser Pro Gly Val Pro Glu Gly Ile Ala Lys Asn Thr	
1785 1790 1795	
att aca gtg ccg tac aat gat cta gaa agc gta caa tat gct ttt aaa	5945
Ile Thr Val Pro Tyr Asn Asp Leu Glu Ser Val Gln Tyr Ala Phe Lys	
1800 1805 1810 1815	
gaa ttc gga gat gac att gcc gga gta atc gtt gaa ccg gta gca gga	5993
Glu Phe Gly Asp Asp Ile Ala Gly Val Ile Val Glu Pro Val Ala Gly	
1820 1825 1830	
aat atg ggt gtg gta ccc cct cag cca ggt ttc tta caa ggc ctt cgt	6041
Asn Met Gly Val Val Pro Pro Gln Pro Gly Phe Leu Gln Gly Leu Arg	
1835 1840 1845	
gat atc aca tca gaa tat ggc tca ctg ctc att ttt gat gaa gta atg	6089
Asp Ile Thr Ser Glu Tyr Gly Ser Leu Leu Ile Phe Asp Glu Val Met	
1850 1855 1860	
aca ggc ttc cgc gta gcg tac aac tgt gca cag gga tac tac aat att	6137
Thr Gly Phe Arg Val Ala Tyr Asn Cys Ala Gln Gly Tyr Tyr Asn Ile	
1865 1870 1875	
act cct gac tta act tgc tta gga aaa gta atc ggt ggt ggt tta cca	6185
Thr Pro Asp Leu Thr Cys Leu Gly Lys Val Ile Gly Gly Gly Leu Pro	
1880 1885 1890 1895	
gta ggt gct tac ggt ggt aaa gct gaa att atg gag cga gtt gca cca	6233
Val Gly Ala Tyr Gly Gly Lys Ala Glu Ile Met Glu Arg Val Ala Pro	
1900 1905 1910	
agc ggc ccg att tat caa gca ggt aca tta tca ggt aat ccg ctt gcg	6281
Ser Gly Pro Ile Tyr Gln Ala Gly Thr Leu Ser Gly Asn Pro Leu Ala	
1915 1920 1925	
atg aca gct gga tac gag acg tta acg cag ctg aca cct caa tcc tac	6329
Met Thr Ala Gly Tyr Glu Thr Leu Thr Gln Leu Thr Pro Gln Ser Tyr	
1930 1935 1940	
gat gag ttc att cgt aaa gca gat cgc tta gaa gaa gga ttg atg gca	6377
Asp Glu Phe Ile Arg Lys Ala Asp Arg Leu Glu Glu Gly Leu Met Ala	
1945 1950 1955	
gca gct gaa aaa tat gat gtg ccg ttt acc gta aac cga gct ggt tca	6425
Ala Ala Glu Lys Tyr Asp Val Pro Phe Thr Val Asn Arg Ala Gly Ser	
1960 1965 1970 1975	

```

atg att ggt ttc ttc ttt aca aac gaa cag gta ata aat tac gaa aca 6473
Met Ile Gly Phe Phe Phe Thr Asn Glu Gln Val Ile Asn Tyr Glu Thr
              1980                      1985                      1990

gcg aaa aca tcg aat tta gat tat ttt gca gct tac tat cgt gaa atg 6521
Ala Lys Thr Ser Asn Leu Asp Tyr Phe Ala Ala Tyr Tyr Arg Glu Met
              1995                      2000                      2005

gct aat gaa ggc gta ttt tta cct cct tct caa ttc gaa gga atg ttt 6569
Ala Asn Glu Gly Val Phe Leu Pro Pro Ser Gln Phe Glu Gly Met Phe
              2010                      2015                      2020

tta tca act tct cat aca gat gca gac att gaa cat aca att gca gca 6617
Leu Ser Thr Ser His Thr Asp Ala Asp Ile Glu His Thr Ile Ala Ala
              2025                      2030                      2035

gct gaa aaa gct ttt tcg aag tta cgt gga taa acagctaaat aagattgact 6670
Ala Glu Lys Ala Phe Ser Lys Leu Arg Gly
2040                      2045                      2050

tttatatcaa ggcttacgga ggattccgta agcttttttta gtttttaaaga catatttttaa 6730

tttatttggt catacatctt taatgaccgc gctattacct ggcgggtggt ggaaggggga 6790

gaacaatttg tcgcaggata atttatcatg tctgcgtttt tccgtagaag aatcagtatg 6850

gtttcaaaaa ggacaggaag tctcacaact tttgtctatc tccttagagc ctcaagtatc 6910

tattcaagaa tatgaccaat acgtgtctat tcgaggagca ttacaattaa caggtgaata 6970

cgaaacgtac gagggagagc actcgttacg cgaatataca aatgagaatc aagtgcaatc 7030

tatctctgtc cgtgaagacg gaacggctga attatcacat caatttcctg ttgatattac 7090

cattcctaaa aatcgatttc aaagtattga agacgtatat gtatctattg atttatttga 7150

ttatgaattg ccaggcacia atcagctgca gctaattggct gac 7193

```

<210> 2

<211> 446

<212> PRT

<213> 巨大芽孢杆菌

<400> 2

```

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

```

```

      130              135              140
Ala Val Thr Leu Ala Lys Lys Ala His His Glu Thr Glu Ile Gly Ala
145              150              155              160
Asn Ala Val Ser Val Ser Tyr Ala Ala Val Glu Leu Ala Lys Lys Ile
      165              170              175
Phe Gly Asp Leu Ser Ser Lys His Val Leu Ile Leu Gly Ala Gly Lys
      180              185              190
Met Gly Gln Leu Ala Val Gln Asn Leu Tyr Gly Ser Gly Ala Lys Lys
      195              200              205
Val Thr Val Val Asn Arg Thr Phe Glu Lys Ala Gln Glu Leu Ala Asn
      210              215              220
Arg Phe Ser Gly Glu Ala Lys Pro Phe Ala Asp Leu Gln His Ala Leu
225              230              235              240
Ser Glu Ala Asp Ile Leu Ile Ser Ser Thr Gly Ala Asn Asp Tyr Val
      245              250              255
Val Thr Lys Gln Met Met Ser Glu Ala Glu Arg Thr Arg Lys Gly Arg
      260              265              270
Pro Leu Phe Met Val Asp Ile Ala Val Pro Arg Asp Leu Asp Pro Glu
      275              280              285
Leu Asp Glu Leu Glu Thr Val Phe Leu Tyr Asp Ile Asp Asp Leu Asn
      290              295              300
Gly Ile Val Glu Ser Asn Leu Gln Glu Arg Gln Lys Ala Ala Asp Glu
305              310              315              320
Ile Glu Ile Met Leu Glu Ala Glu Ile Val Ala Phe Lys Ser Trp Leu
      325              330              335
Gly Thr Leu Gly Val Val Pro Val Ile Ser Ala Leu Arg Gln Lys Ala
      340              345              350
Leu Thr Ile Gln Ala Glu Thr Met Lys Ser Ile Asp Arg Lys Leu Pro
      355              360              365
Asp Leu Ser Glu Arg Glu Arg Lys Val Leu Asn Lys His Thr Lys Ser
      370              375              380
Ile Ile Asn Gln Leu Leu Arg Asp Pro Ile Pro His Ala Lys Glu Leu
385              390              395              400
Ala Gly Glu Lys His Ala Glu Glu Ser Leu Glu Leu Phe Met Lys Ile
      405              410              415
Phe Asn Ile Glu Gln Glu Val Ala Leu Gln Lys Glu Lys Glu Glu Gln
      420              425              430
Leu His Thr Ser Ile Thr Ser His Ser Val Ala Tyr Gln Ser
      435              440              445

```

<210> 3

<211> 277

<212> PRT

<213> 巨大芽孢杆菌

<400> 3

```

Met Val Asp Met Glu Leu Thr Arg Ile Tyr Glu Leu Ile Val Ile Leu
 1              5              10              15
Tyr Ala Val Ser Val Leu Leu Tyr Phe Ile Asp Phe Val Gln Asn Asn
      20              25              30
Arg Lys Ala Asn Ser Ile Ala Phe Trp Leu Leu Ser Ile Val Trp Val
      35              40              45
Leu Gln Thr Ala Phe Leu Ile Leu Arg Met Phe Glu Thr Gly Arg Phe
      50              55              60
Pro Ile Leu Thr Val Ser Glu Gly Leu Tyr Phe Tyr Ala Trp Val Leu
      65              70              75              80
Ile Thr Leu Ser Leu Val Ile Asn Arg Leu Phe Arg Val Asp Phe Met
      85              90              95
Val Phe Phe Thr Asn Val Ile Gly Phe Leu Ile Met Ala Ile His Thr
      100              105              110
Phe Ala Pro Leu Glu Thr Tyr Thr Ala Val Gln Ser Glu Arg Leu Val

```

```

      115      120      125
Ser Glu Leu Leu Ile Ile His Ile Thr Met Ala Ile Leu Ser Tyr Gly
      130      135      140
Ala Phe Ser Val Ser Phe Val Phe Ser Ala Leu Tyr Leu Val Gln Tyr
145      150      155      160
Asn Leu Leu Lys Lys Lys Lys Trp Gly Lys Arg Leu Leu Arg Ile Glu
      165      170      175
Asp Leu Thr Lys Leu Asp Tyr Met Ser Tyr Val Leu Ile Leu Ile Gly
      180      185      190
Val Pro Leu Leu Leu Leu Ser Leu Ile Leu Gly Ile Ile Trp Ala Tyr
      195      200      205
Ile Lys Tyr Val Asp Phe His Trp Tyr Asp Thr Lys Val Leu Gly Ser
      210      215      220
Phe Leu Met Phe Val Ala Tyr Ser Ala Tyr Leu Tyr Ala Arg Leu Arg
225      230      235      240
Arg Gly Ile Gln Gly Lys Ala Ile Ala Met Trp Asn Ile Gly Leu Phe
      245      250      255
Leu Val Leu Leu Ile Asn Phe Phe Leu Phe Gly Ser Leu Ser Asn Phe
      260      265      270
His Phe Trp Gly Ser
      275

```

<210> 4

<211> 310

<212> PRT

<213> 巨大芽孢杆菌

<400> 4

```

Met Arg Lys Ile Ile Val Gly Ser Arg Arg Ser Lys Leu Ala Leu Thr
1      5      10      15
Gln Thr Lys Trp Val Ile Glu Gln Leu Lys Lys Gln Gly Leu Pro Phe
      20      25      30
Glu Phe Glu Ile Lys Glu Met Val Thr Lys Gly Asp Gln Ile Leu Asn
      35      40      45
Val Thr Leu Ser Lys Val Gly Gly Lys Gly Leu Phe Val Lys Glu Ile
      50      55      60
Glu Gln Ala Met Leu Asp Lys Glu Ile Asp Met Ala Val His Ser Met
      65      70      75      80
Lys Asp Met Pro Ala Val Leu Pro Glu Gly Leu Thr Ile Gly Cys Ile
      85      90      95
Pro Leu Arg Glu Asp His Arg Asp Ala Leu Ile Ser Lys Asn Gly Glu
      100      105      110
Arg Phe Glu Glu Leu Pro Ser Gly Ala Val Ile Gly Thr Ser Ser Leu
      115      120      125
Arg Arg Gly Ala Gln Leu Leu Ser Met Arg Ser Asp Ile Glu Ile Lys
      130      135      140
Trp Ile Arg Gly Asn Ile Asp Thr Arg Leu Glu Lys Leu Lys Asn Glu
145      150      155      160
Asp Tyr Asp Ala Ile Ile Leu Ala Ala Ala Gly Leu Ser Arg Met Gly
      165      170      175
Trp Ser Lys Asp Thr Val Thr Gln Tyr Leu Glu Pro Glu Ile Ser Val
      180      185      190
Pro Ala Val Gly Gln Gly Ala Leu Ala Ile Glu Cys Arg Glu Asn Asp
      195      200      205
His Glu Leu Leu Ser Leu Leu Gln Ala Leu Asn His Asp Glu Thr Ala
      210      215      220
Arg Ala Val Arg Ala Glu Arg Val Phe Leu Lys Glu Met Glu Gly Gly
225      230      235      240
Cys Gln Val Pro Ile Ala Gly Tyr Gly Arg Ile Leu Asp Gly Gly Asn
      245      250      255
Ile Glu Leu Thr Ser Leu Val Ala Ser Pro Asp Gly Lys Thr Ile Tyr

```



```

Lys Ser Met Pro Gly Val Asp Gln Ile Ser Leu Asp Tyr Leu Asn Asp
  50                      55                      60
Glu Ile Gln Glu Leu Val Asp Leu Gly Ile Lys Ser Ala Met Val Phe
  65                      70                      75                      80
Gly Val Pro Ala Glu Lys Asp Ala Val Gly Ser Gln Ala Tyr His Asp
                      85                      90                      95
His Gly Ile Val Gln Lys Gly Ile Arg Gln Ile Lys Glu Asn Phe Pro
                      100                    105                    110
Asp Phe Val Val Ile Ala Asp Thr Cys Leu Cys Gln Tyr Thr Asp His
                      115                    120                    125
Gly His Cys Gly Val Ile Glu Asp Gly Lys Ile Leu Asn Asp Pro Ser
                      130                    135                    140
Leu Asp Leu Leu Ala Arg Thr Ala Val Ser Gln Ala Lys Ala Gly Ala
145                      150                    155                    160
Asp Ile Ile Ala Pro Ser Asn Met Met Asp Gly Phe Val Ala Ala Ile
                      165                    170                    175
Arg His Gly Leu Asp Glu Ala Gly Phe Glu Asp Val Pro Val Met Ser
                      180                    185                    190
Tyr Ala Val Lys Tyr Ala Ser Ala Phe Tyr Gly Pro Phe Arg Asp Ala
                      195                    200                    205
Ala His Ser Ser Pro Gln Phe Gly Asp Arg Lys Thr Tyr Gln Met Asp
                      210                    215                    220
Pro Ala Asn Arg Leu Glu Ala Leu Arg Glu Ala Glu Ser Asp Val Glu
225                      230                    235                    240
Glu Gly Ala Asp Phe Leu Ile Val Lys Pro Ala Leu Ser Tyr Leu Asp
                      245                    250                    255
Ile Met Arg Asp Val Lys Asn Thr Phe Asn Leu Pro Val Val Ala Tyr
                      260                    265                    270
Asn Val Ser Gly Glu Tyr Ser Met Ile Lys Ala Ala Ala Gln Asn Gly
                      275                    280                    285
Trp Val Asn Glu Lys Glu Ile Val Leu Glu Lys Leu Ile Ser Met Lys
                      290                    295                    300
Arg Ala Gly Ala Asp Leu Ile Val Thr Tyr His Ala Lys Asp Ala Ala
305                      310                    315                    320
Arg Trp Leu Ser Asp Lys
                      325

```

<210> 7

<211> 429

<212> PRT

<213> 巨大芽孢杆菌

<400> 7

```

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

```

```

Tyr His Gly His Gly Asp Ser Leu Leu Ile Lys Ala Gly Ser Gly Val
145                      150                      155                      160
Ala Thr Leu Gly Leu Pro Asp Ser Pro Gly Val Pro Glu Gly Ile Ala
                      165                      170                      175
Lys Asn Thr Ile Thr Val Pro Tyr Asn Asp Leu Glu Ser Val Gln Tyr
                      180                      185                      190
Ala Phe Lys Glu Phe Gly Asp Asp Ile Ala Gly Val Ile Val Glu Pro
                      195                      200                      205
Val Ala Gly Asn Met Gly Val Val Pro Pro Gln Pro Gly Phe Leu Gln
210                      215                      220
Gly Leu Arg Asp Ile Thr Ser Glu Tyr Gly Ser Leu Leu Ile Phe Asp
225                      230                      235                      240
Glu Val Met Thr Gly Phe Arg Val Ala Tyr Asn Cys Ala Gln Gly Tyr
                      245                      250                      255
Tyr Asn Ile Thr Pro Asp Leu Thr Cys Leu Gly Lys Val Ile Gly Gly
260                      265                      270
Gly Leu Pro Val Gly Ala Tyr Gly Gly Lys Ala Glu Ile Met Glu Arg
275                      280                      285
Val Ala Pro Ser Gly Pro Ile Tyr Gln Ala Gly Thr Leu Ser Gly Asn
290                      295                      300
Pro Leu Ala Met Thr Ala Gly Tyr Glu Thr Leu Thr Gln Leu Thr Pro
305                      310                      315                      320
Gln Ser Tyr Asp Glu Phe Ile Arg Lys Ala Asp Arg Leu Glu Glu Gly
                      325                      330                      335
Leu Met Ala Ala Ala Glu Lys Tyr Asp Val Pro Phe Thr Val Asn Arg
                      340                      345                      350
Ala Gly Ser Met Ile Gly Phe Phe Phe Thr Asn Glu Gln Val Ile Asn
                      355                      360                      365
Tyr Glu Thr Ala Lys Thr Ser Asn Leu Asp Tyr Phe Ala Ala Tyr Tyr
370                      375                      380
Arg Glu Met Ala Asn Glu Gly Val Phe Leu Pro Pro Ser Gln Phe Glu
385                      390                      395                      400
Gly Met Phe Leu Ser Thr Ser His Thr Asp Ala Asp Ile Glu His Thr
                      405                      410                      415
Ile Ala Ala Ala Glu Lys Ala Phe Ser Lys Leu Arg Gly
                      420                      425

```

表1

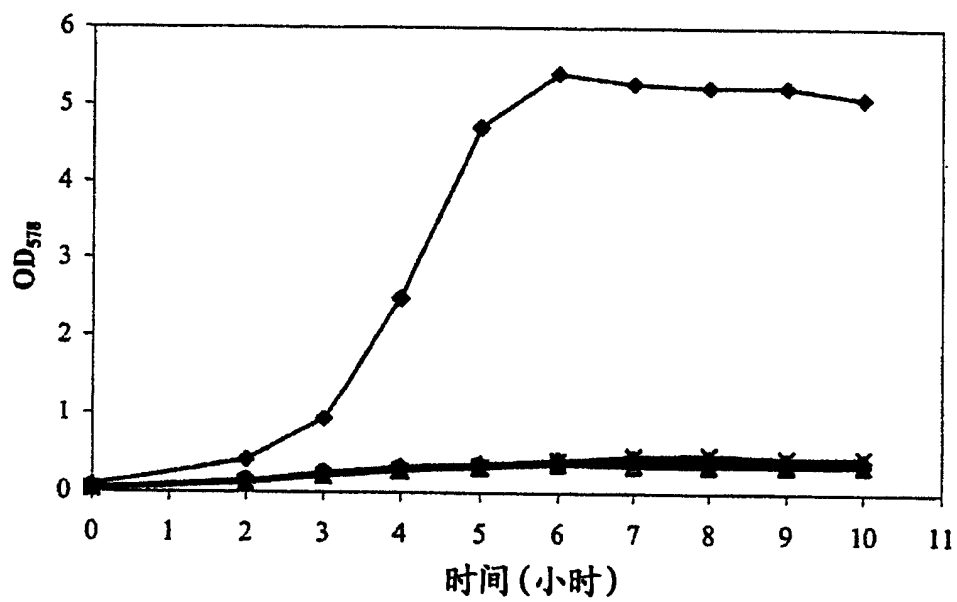
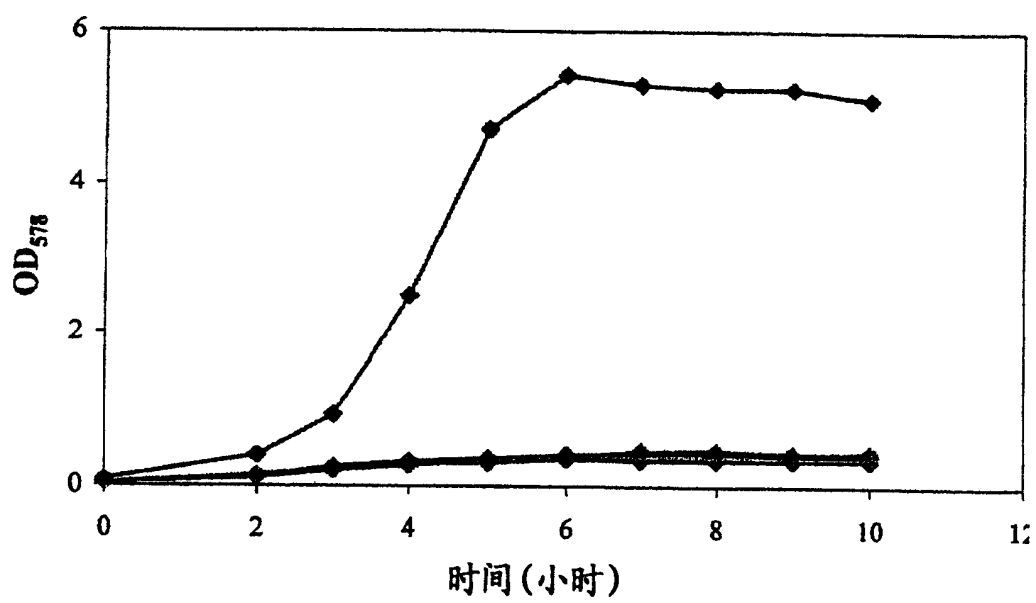
菌株	描述	参考文献/来源
大肠杆菌		
DH5 α	<i>F</i> λ ^{supE44} Δ (argF-lac)U169 ϕ 89dlacZ Δ M15hsdR17recA1endA1gyrA96thi-1relA1	Sambrook 等人, 1989
RP523	<i>F</i> λ ^{supE44} thr-1leuB6thi-1lacY1tonA21hemB	Li 等人, 1987
巨大芽孢杆菌		
DSMZ32	野生型	DSMZ *
DSMZ509	维生素 B ₁₂ 生产菌	DSMZ *
DSMZ2894	核黄素和维生素 B ₁₂ 生产菌	DSMZ *
WH320	DSMZ319 lac	Rygus 等人, 1991
鼠伤寒沙门氏杆菌		
AR 3612	Leu ⁺ cysG metE, Sm ^r	Raux 等人, 1996

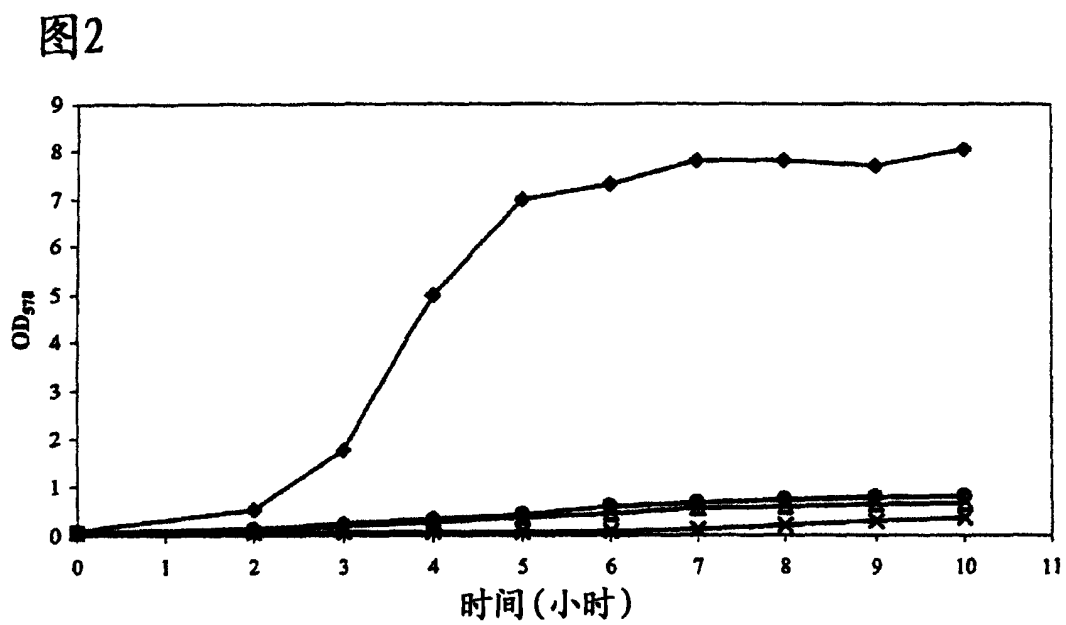
* DSMZ: 德国微生物菌种保藏中心 (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Braunschweig

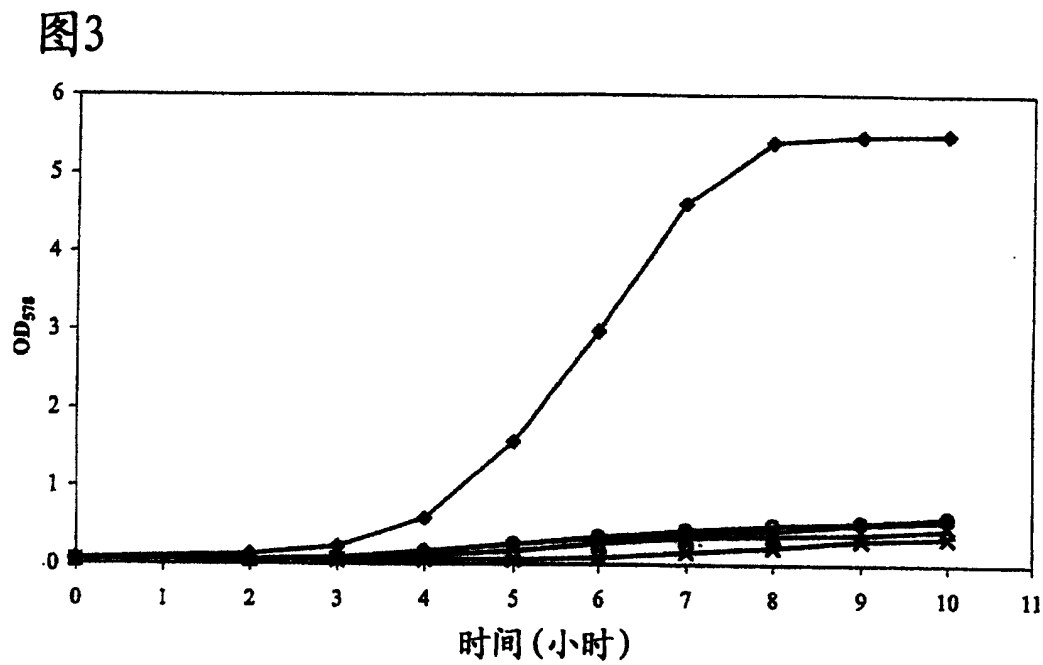
表2

质粒	描述	参考文献/来源
pBlueskript II SK+	具有非常高拷贝数的噬菌粒 克隆和表达载体, Ap ^r	
pWH1510	<i>xylA-lacZ</i> , Ap ^r , Tc ^r	Rygus 等人, 1991
pWH1520	用于芽孢杆菌亚种的克隆和 表达载体, Ap ^r , Tc ^r	Rygus 等人, 1991

图1







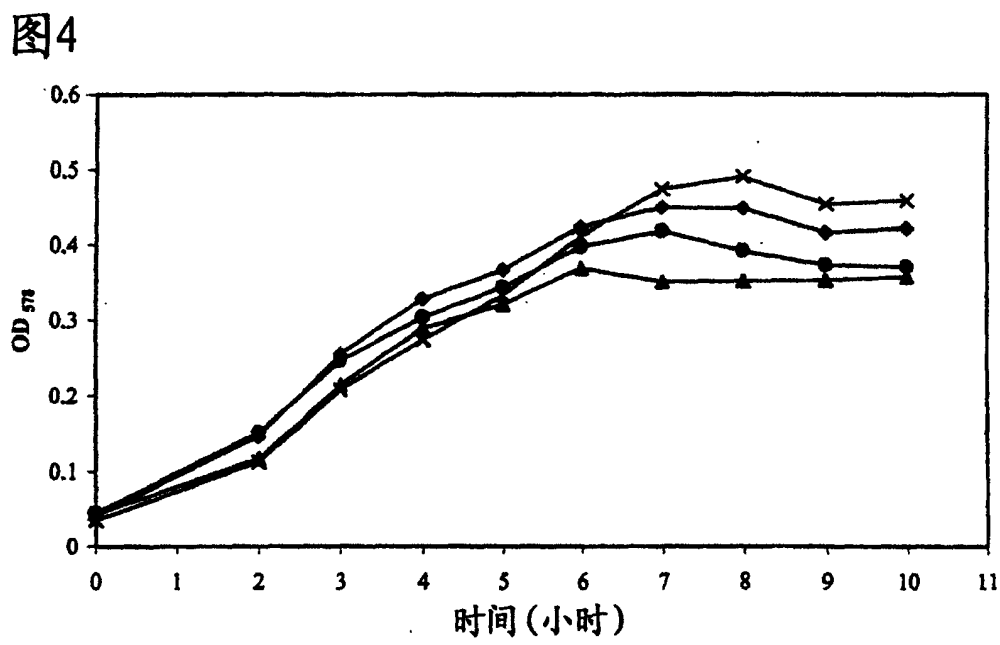


图5

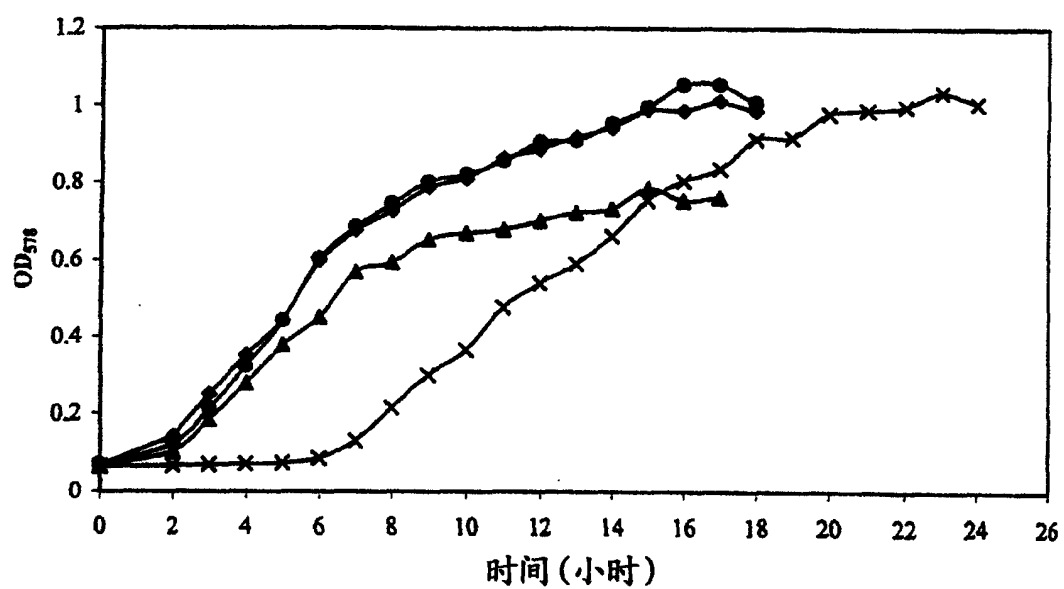


图6

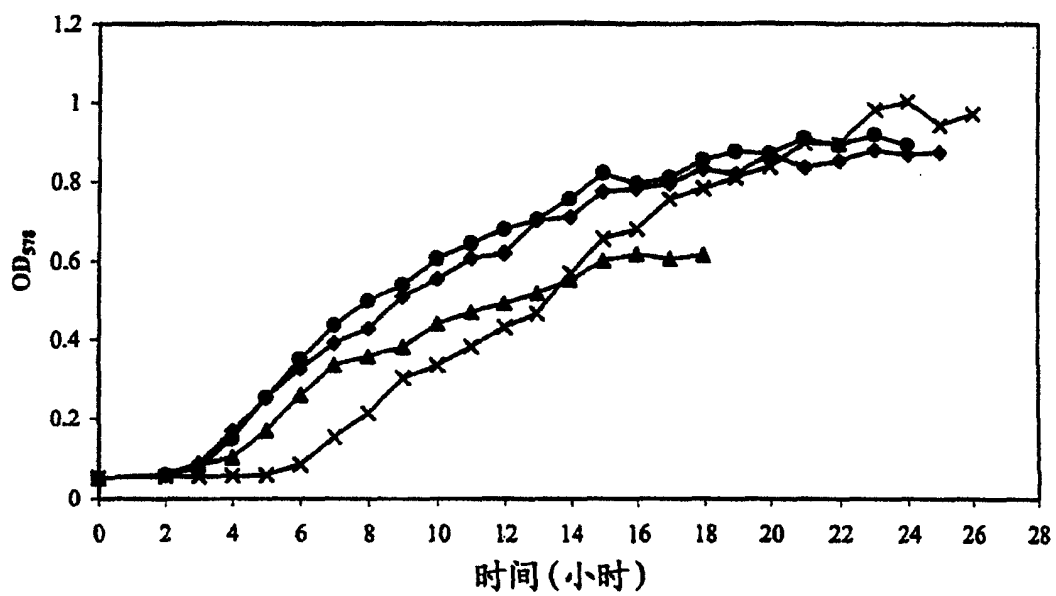


图7

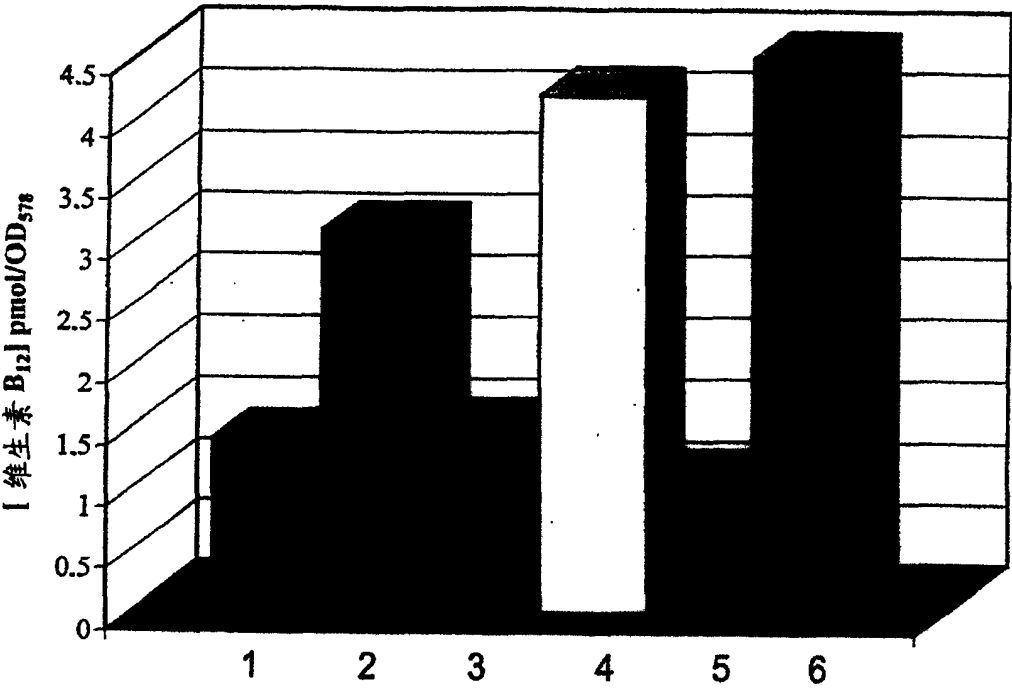


图8

