



(12) 发明专利

(10) 授权公告号 CN 108841851 B

(45) 授权公告日 2021.12.24

(21) 申请号 201810789194.6  
 (22) 申请日 2018.07.18  
 (65) 同一申请的已公布的文献号  
 申请公布号 CN 108841851 A  
 (43) 申请公布日 2018.11.20  
 (73) 专利权人 中国科学院微生物研究所  
 地址 100101 北京市朝阳区北辰西路1号院  
 3号  
 (72) 发明人 王丽敏 王玉 付丽红 于波  
 (74) 专利代理机构 北京纪凯知识产权代理有限公司 11245  
 代理人 关畅 张立娜  
 (51) Int. Cl.  
 C12N 15/75 (2006.01)  
 C12N 9/80 (2006.01)  
 C12N 1/21 (2006.01)  
 (56) 对比文件  
 CN 101506233 A, 2009.08.12  
 CN 101506233 A, 2009.08.12  
 CN 1427078 A, 2003.07.02  
 CN 107586764 A, 2018.01.16  
 CN 102174456 A, 2011.09.07  
 Yi-Sin Lin等. Cloning and expression

of the transglutaminase gene from  
 Streptovercillium ladakanum in  
 Streptomyces lividans.《Process  
 Biochemistry》.2004,第39卷591-598.  
 Dongdong Mu等.Heterologous signal  
 peptides-directing secretion of  
 Streptomyces mobaraensis transglutaminase  
 by Bacillus subtilis.《Applied  
 Microbiology and Biotechnology》.2018,第  
 102卷第5533-5543页.  
 GenBank.Streptomyces mobaraensis  
 transglutaminase (tgB1) gene, complete  
 cds  
 and unknown genes.《GenBank:  
 AY241675.1》.2003,  
 Ulf Brockmeier等.Systematic Screening  
 of All Signal Peptides from Bacillus  
 subtilis: A Powerful Strategy in  
 Optimizing Heterologous Protein Secretion  
 in Gram-positive Bacteria.《J. Mol.  
 Biol.》.2006,第362卷393-402,doi:10.1016/  
 j.jmb.2006.07.034.

审查员 陈中伟

权利要求书2页 说明书12页  
 序列表11页 附图2页

(54) 发明名称  
 一种利用食源安全性宿主表达谷氨酰胺转氨酶的方法

(57) 摘要

本发明公开了一种利用食源安全性宿主表达谷氨酰胺转氨酶的方法。本发明提供了构建产谷氨酰胺转氨酶的工程菌的方法,包括:对枯草芽孢杆菌进行改造,使其表达达卡轮枝链霉菌来源的谷氨酰胺转氨酶,改造后的所述枯草芽孢杆菌即为工程菌。本发明与现有的以枯草芽孢杆菌为宿主的TG酶表达具有下述区别:①TG酶活力高,最终酶活达到130U/mg,满足工业生产需要;②基因整合到基因组上,性能稳定;③TG酶稳定

性好,放置15天TG酶仍具有80%的活力;④非钙离子依赖性,钙离子稳定性好,存在25mM钙离子的条件下,TG酶具有更高的活性(113%),应用范围更为广泛。本发明对于谷氨酰胺转氨酶的工业化生产具有重要意义。

CN 108841851 B

1. 一种构建产谷氨酰胺转氨酶的工程菌的方法,包括如下步骤:对枯草芽孢杆菌进行改造,使所述枯草芽孢杆菌表达达卡轮枝链霉菌(*S. ladakanum*)来源的谷氨酰胺转氨酶,改造后的所述枯草芽孢杆菌即为工程菌;

所述方法包括如下步骤:将达卡轮枝链霉菌(*S. ladakanum*)来源的谷氨酰胺转氨酶的编码基因导入所述枯草芽孢杆菌,得到产氨酰胺转氨酶的重组菌,即为所述工程菌;

所述达卡轮枝链霉菌(*S. ladakanum*)来源的谷氨酰胺转氨酶的编码基因是通过DNA片段的形式导入所述枯草芽孢杆菌中的;所述DNA片段自5'端到3'端依次含有编码信号肽的核酸序列、编码前导肽和谷氨酰胺转氨酶的活性区域的核酸序列;其中,所述前导肽和谷氨酰胺转氨酶的活性区域为来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体;

来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽的氨基酸序列如SEQ ID No.12所示;来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的活性区域的氨基酸序列如SEQ ID No.13所示;来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶从N-端信号肽到C-端活性区域的氨基酸序列如SEQ ID No.27所示;

来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点处的氨基酸变化与SEQ ID No.27相比,为如下任一所示:将第74-75位的PS突变为AA并将第80-82位的DSD突变为AAA、将第74-75位的PS突变为AA。

2. 根据权利要求1所述的方法,其特征在于:所述信号肽包括N-端正电荷区,疏水核心区 and C-端亲水信号肽酶识别区域。

3. 根据权利要求2所述的方法,其特征在于:所述信号肽为如下任一: SacB信号肽、AmyE信号肽、AmyQ信号肽、Epr信号肽、LytD信号肽、NucB信号肽、PenP信号肽、WapA信号肽、WprA信号肽、YhcR信号肽、YncM信号肽。

4. 根据权利要求3所述的方法,其特征在于:所述SacB信号肽的氨基酸序列如SEQ ID No.1所示;所述AmyE信号肽的氨基酸序列如SEQ ID No.2所示;所述AmyQ信号肽的氨基酸序列如SEQ ID No.3所示;所述Epr信号肽的氨基酸序列如SEQ ID No.4所示;所述LytD信号肽的氨基酸序列如SEQ ID No.5所示;所述NucB信号肽的氨基酸序列如SEQ ID No.6所示;所述PenP信号肽的氨基酸序列如SEQ ID No.7所示;所述WapA信号肽的氨基酸序列如SEQ ID No.8所示;所述WprA信号肽的氨基酸序列如SEQ ID No.9所示;所述YhcR信号肽的氨基酸序列如SEQ ID No.10所示;所述YncM信号肽的氨基酸序列如SEQ ID No.11所示。

5. 根据权利要求3或4所述的方法,其特征在于:编码所述SacB信号肽的核酸序列如SEQ ID No.14所示;编码所述AmyE信号肽的核酸序列如SEQ ID No.15所示;编码所述AmyQ信号肽的核酸序列如SEQ ID No.16所示;编码所述Epr信号肽的核酸序列如SEQ ID No.17所示;编码所述LytD信号肽的核酸序列如SEQ ID No.18所示;编码所述NucB信号肽的核酸序列如SEQ ID No.19所示;编码所述PenP信号肽的核酸序列如SEQ ID No.20所示;编码所述WapA信号肽的核酸序列如SEQ ID No.21所示;编码所述WprA信号肽的核酸序列如SEQ ID No.22所示;编码所述YhcR信号肽的核酸序列如SEQ ID No.23所示;编码所述YncM信号肽的核酸序列如SEQ ID No.24所示。

6. 根据权利要求1所述的方法,其特征在于:编码来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点处的核苷酸变化

与SEQ ID No.28相比,为如下任一所示:将第220-225位的ccgtcc突变为gcagca并将第238-246位的gactccgac突变为gcagcagca、将第220-225位的ccgtcc突变为gcagca。

7. 根据权利要求1所述的方法,其特征在于:所述DNA片段是通过重组载体的形式导入所述枯草芽孢杆菌中的。

8. 根据权利要求7所述的方法,其特征在于:所述重组载体为将所述编码前导肽和谷氨酰胺转氨酶的活性区域的核酸序列插入到pWB980载体的酶切位点XmaI和XbaI之间后得到的重组质粒。

9. 根据权利要求1所述的方法,其特征在于:所述DNA片段被整合到所述枯草芽孢杆菌的基因组的lacA位点处。

10. 根据权利要求1所述的方法,其特征在于:所述枯草芽孢杆菌为枯草芽孢杆菌菌株WB600。

11. 利用权利要求1-10中任一所述方法制备得到的产谷氨酰胺转氨酶的工程菌。

12. 蛋白质,自N端到C端依次为权利要求1-5任一中所述的信号肽、权利要求1-6任一中所述的来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体。

13. 编码权利要求12所述蛋白质的核酸分子。

14. 含有权利要求13所述核酸分子的表达盒、重组载体或重组菌。

15. 权利要求11所述工程菌或权利要求12所述蛋白质或权利要求13所述核酸分子或权利要求14所述表达盒、重组载体或重组菌在制备谷氨酰胺转氨酶中的应用。

16. 一种制备谷氨酰胺转氨酶的方法,包括如下步骤:培养权利要求11中所述的产谷氨酰胺转氨酶的工程菌,从培养产物中获得谷氨酰胺转氨酶。

17. 利用权利要求16所述方法制备得到的谷氨酰胺转氨酶。

## 一种利用食源安全性宿主表达谷氨酰胺转氨酶的方法

### 技术领域

[0001] 本发明涉及生物技术领域,具体涉及一种利用食源安全性宿主表达谷氨酰胺转氨酶的方法。

### 背景技术

[0002] 谷氨酰胺转氨酶(Transglutaminase, EC 2.3.2.13, 简称为TG酶)能够使蛋白质发生交联作用,可以改善肉制品的外观、口感、风味、营养特性,可以提升酸奶品质,提高植物蛋白的弹性和持水能力。TG酶在食品加工、包装等领域具有极其广泛的应用前景,此外, TG酶还可用于制备薄膜,替代用高分子合成材料制造的食品包装薄膜(赵华梅,王曼霞,牟志春,徐琴,蔡发(2008). 食品工业中的谷氨酰胺转氨酶及其酶制剂. 食品科技, 9:137-141; Gabe J, Zohar B, Moran J, Ayelet F(2017) Constitutive expression of active microbial transglutaminase in *Escherichia coli* and comparative characterization to a known variant. *BMC Biotechnol* 17:23-33)。1998年美国食品药品监督管理局(FDA)认定TG酶为公认安全的添加剂(GRAS),我国已将TG酶作为新增补品种,列入食品添加剂行列。

[0003] TG酶商业化的生产方法主要有微生物发酵法和天然提取法。天然提取法是从凝血因子XⅢ,或者豚鼠肝脏中提取,不仅成本昂贵,而且天然提取的产品会产生色素沉积,影响产品的外观,在食品工业中应用面临很多困难(Lin Y, Chao ML, Liu C, Chu W(2004) Cloning and expression of the transglutaminase gene from *Streptoverticillium ladakanum* in *Streptomyces lividans*. *Process Biochem* 39:591-598)。微生物来源TG酶属于胞外酶,近几年正在逐步取代动物来源的TG酶,成为商业生产上的主要方法。微生物来源的TG酶主要来自链霉菌属和芽孢杆菌属,而目前仅链霉菌属(*S. mobaraensis*)来源的TG酶实现了商业化(Liu S, Zhang D, Wang M, Cui W, Chen K, Liu Y, Du G, Chen Jian, Zhou Z(2011) The pro-region of *Streptomyces hygroscopicus* transglutaminase affects its secretion by *Escherichia coli*. *FEMS Microbiol Lett* 324:98-105)。虽然链霉菌来源的TG酶是钙离子非依赖性,具有更为广阔的应用空间,但是由于链霉菌生长比缓慢,需要比较长的培养期,增加了TG酶的生产成本(中国发明专利CN96105596.0)。而且,由于链霉菌本身含有降解TG酶的蛋白酶,因此,随着时间的延长,链霉菌来源TG酶活力逐渐降低。为了克服上述缺陷,国内外研究者开展了大量基因工程菌株构建工作,将不同来源的TG酶在不同宿主中表达,包括 *Corynebacterium* spp., *Escherichia coli*, *Candida boidinii* 以及 *Streptomyces* spp., 但均未实现商业化(CN 1230529C; CN 104818291A; CN 103555739A; WO 2016170447)。

[0004] 枯草芽孢杆菌是美国食品药品监督管理局和中国农业部等部门批准为食品级安全菌株。已有大量在枯草芽孢杆菌中表达的报道,如蛋白酶、淀粉酶。中国专利CN201110034957.4报道了以枯草芽孢杆菌为宿主表达TG酶,酶活仅为2U/ml,无法满足工业生产需要。中国发明专利CN96105596.0报道了从枯草芽孢杆菌AJ12866和枯草芽孢杆菌

AJ1307中纯化得到TG酶,虽然该酶是钙离子非依赖性,但是5mM的钙离子就抑制了20%的酶活。该专利尝试利用枯草芽孢杆菌为宿主构建重组酶,但是TG酶以包涵体形式存在。虽然枯草芽孢杆菌克隆体系已建立的较为完善,但是以食源安全性的枯草芽孢杆菌为宿主表达TG酶的专利鲜有报道。TG酶是非常有发展潜力的酶制剂,寻找更多途径实现其大规模工业化生产是未来的必然趋势。以枯草芽孢杆菌为宿主异源表达TG酶,不仅有助于缩短发酵周期,降低成本,而且可以用于食品行业,具有非常重要的意义。

## 发明内容

[0005] 本发明的目的是提供一种利用食源安全性宿主表达谷氨酰胺转氨酶的方法。

[0006] 第一方面,本发明要求保护一种构建产谷氨酰胺转氨酶的工程菌的方法。

[0007] 本发明所提供的构建产谷氨酰胺转氨酶的工程菌的方法,包括如下步骤:对枯草芽孢杆菌进行改造,使所述枯草芽孢杆菌表达达卡轮枝链霉菌(*S. ladakanum*)来源的谷氨酰胺转氨酶,改造后的所述枯草芽孢杆菌即为工程菌。

[0008] 进一步地,所述方法可包括如下步骤:将达卡轮枝链霉菌(*S. ladakanum*)来源的谷氨酰胺转氨酶的编码基因导入所述枯草芽孢杆菌得到产谷氨酰胺转氨酶的重组菌,即为所述工程菌。

[0009] 在所述方法中,所述达卡轮枝链霉菌(*S. ladakanum*)来源的谷氨酰胺转氨酶的编码基因可通过DNA片段的形式导入所述枯草芽孢杆菌中;所述DNA片段自5'端到3'端依次含有编码信号肽的核酸序列、编码前导肽和谷氨酰胺转氨酶的活性区域的核酸序列;其中,所述前导肽和谷氨酰胺转氨酶的活性区域为来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域或其突变体。

[0010] 其中,所述信号肽包括N-端正电荷区,疏水核心区 and C-端亲水信号肽酶识别区域。

[0011] 进一步地,所述信号肽具体可为如下任一: SacB信号肽、AmyE信号肽、AmyQ信号肽、Epr信号肽、LytD信号肽、NucB信号肽、PenP信号肽、WapA信号肽、WprA信号肽、YhcR信号肽、YncM信号肽。

[0012] 更进一步地,所述SacB信号肽的氨基酸序列具体如SEQ ID No.1所示;所述AmyE信号肽的氨基酸序列具体如SEQ ID No.2所示;所述AmyQ信号肽的氨基酸序列具体如SEQ ID No.3所示;所述Epr信号肽的氨基酸序列具体如SEQ ID No.4所示;所述LytD信号肽的氨基酸序列具体如SEQ ID No.5所示;所述NucB信号肽的氨基酸序列具体如SEQ ID No.6所示;所述PenP信号肽的氨基酸序列具体如SEQ ID No.7所示;所述WapA信号肽的氨基酸序列具体如SEQ ID No.8所示;所述WprA信号肽的氨基酸序列具体如SEQ ID No.9所示;所述YhcR信号肽的氨基酸序列具体如SEQ ID No.10所示;所述YncM信号肽的氨基酸序列具体如SEQ ID No.11所示。

[0013] 其中,来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽的氨基酸序列可如SEQ ID No.12所示。来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的活性区域的氨基酸序列可如SEQ ID No.13所示。

[0014] 相应的,来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点可位于SEQ ID No.27(来源于所述达卡轮枝链霉菌(*S. ladakanum*)的谷氨酰胺转氨酶从N-端信号肽到C-端活性区域的氨基酸序列)的如下位

置中的部分或全部：第34-36位氨基酸、第74-75位氨基酸、第80-82位氨基酸、第83位氨基酸。

[0015] 更进一步地，来源于所述达卡轮枝链霉菌 (*S. ladakanum*) 的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点处的氨基酸变化与SEQ ID No.27相比，为如下任一所示：将第74-75位的PS突变为AA并将第80-82位的DAD突变为AAA、将第74-75位的PS突变为AA、将第80-82位的DAD突变为AAA、缺失第74-75位的PS、缺失第80-82位的DAD、将第34-36位的SGS突变为AAA、将第83位的E突变为D。

[0016] 对应于基因水平，编码所述SacB信号肽的核酸序列可如SEQ ID No.14所示；编码所述AmyE信号肽的核酸序列如SEQ ID No.15所示；编码所述AmyQ信号肽的核酸序列可如SEQ ID No.16所示；编码所述Epr信号肽的核酸序列如SEQ ID No.17所示；编码所述LytD信号肽的核酸序列可如SEQ ID No.18所示；编码所述NucB信号肽的核酸序列如SEQ ID No.19所示；编码所述PenP信号肽的核酸序列可如SEQ ID No.20所示；编码所述WapA信号肽的核酸序列如SEQ ID No.21所示；编码所述WprA信号肽的核酸序列可如SEQ ID No.22所示；编码所述YhcR信号肽的核酸序列如SEQ ID No.23所示；编码所述YncM信号肽的核酸序列可如SEQ ID No.24所示。

[0017] 编码来源于所述达卡轮枝链霉菌 (*S. ladakanum*) 的谷氨酰胺转氨酶的前导肽的核酸序列可如SEQ ID No.25所示。

[0018] 编码来源于所述达卡轮枝链霉菌 (*S. ladakanum*) 的谷氨酰胺转氨酶的活性区域的核酸序列可如SEQ ID No.26所示。

[0019] 编码来源于所述达卡轮枝链霉菌 (*S. ladakanum*) 的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点处的核苷酸变化与SEQ ID No.28 (编码来源于所述达卡轮枝链霉菌 (*S. ladakanum*) 的谷氨酰胺转氨酶从N-端信号肽到C-端活性区域的核苷酸序列) 相比，为如下任一所示：将第220-225位的ccgtcc突变为gcagca并将第238-246位的gactccgac突变为gcagcagca (对应“将第74-75位的PS突变为AA并将第80-82位的DAD突变为AAA”)、将第220-225位的ccgtcc突变为gcagca (对应“将第74-75位的PS突变为AA”)、将第238-246位的gactccgac突变为gcagcagca (对应“将第80-82位的DAD突变为AAA”)、缺失第220-225位的ccgtcc (对应“缺失第74-75位的PS”)、缺失第238-246位的gactccgac (对应“缺失第80-82位的DAD”)、将第100-108位的agtggcagc突变为gcagcagca (对应“将第34-36位的SGS突变为AAA”)、将第247-249位的gag突变为gac (对应“将第83位的E突变为D”)。

[0020] 在本发明的具体实施方式中，所述重组载体具体为将所述编码前导肽和谷氨酰胺转氨酶的活性区域的核酸序列插入到pWB980载体的酶切位点XmaI和XbaI之间后得到的重组质粒。

[0021] 在所述方法中，为了保证最终所得工程菌性能稳定，将所述DNA片段整合到了所述枯草芽孢杆菌的基因组的lacA位点处。

[0022] 进一步地，所述DNA片段具体被整合到所述枯草芽孢杆菌的基因组中的SEQ ID No.29和SEQ ID No.31之间。

[0023] 在本发明的具体实施方式中，所述枯草芽孢杆菌具体为枯草芽孢杆菌菌株WB600。

[0024] 第二方面，本发明要求保护如下任一生物材料：

[0025] (A1) 利用前文所述方法制备得到的产谷氨酰胺转氨酶的工程菌。

[0026] (A2) 蛋白质,自N端到C端依次为前文所述的信号肽、前文所述的前导肽和前文所述的谷氨酰胺转氨酶的活性区域。

[0027] 特别的,所述蛋白质自N端到C端依次为SEQ ID No.3所示的AmyQ信号肽、来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体。其中,来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的前导肽的氨基酸序列如SEQ ID No.12所示;来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的活性区域的氨基酸序列如SEQ ID No.13所示;来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点处的氨基酸变化与SEQ ID No.27(来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶从N-端信号肽到C-端活性区域的氨基酸序列)相比,为将第74-75位的PS突变为AA并将第80-82位的DAD突变为AAA。

[0028] (A3) 编码(A2)所示蛋白质的核酸分子。

[0029] 特别的,所述核酸分子自5'端到3'端依次为SEQ ID No.16所示的编码AmyQ信号肽的核酸序列、编码来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体的核酸序列。其中,编码来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的前导肽的核酸序列如SEQ ID No.25所示;编码来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的活性区域的核酸序列如SEQ ID No.26所示;编码来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶的前导肽和活性区域的突变体的突变位点处的核苷酸变化与SEQ ID No.28(编码来源于所述达卡轮枝链霉菌(*S.ladakanum*)的谷氨酰胺转氨酶从N-端信号肽到C-端活性区域的核苷酸序列)相比,为将第220-225位的ccgtcc突变为gcagca并将第238-246位的gactccgac突变为gcagcagca(对应“将第74-75位的PS突变为AA并将第80-82位的DAD突变为AAA”)。该核酸分子对应实施例部分表1中的突变体“PS 74-75AA/DSD 80-82AAA”。

[0030] (A4) 含有(A3)所示核酸分子的表达盒、重组载体或重组菌。

[0031] 第三方面,本发明要求保护前文第二方面中所述的生物材料在制备谷氨酰胺转氨酶中的应用。

[0032] 第四方面,本发明要求保护一种制备谷氨酰胺转氨酶的方法。

[0033] 本发明所提供的制备谷氨酰胺转氨酶的方法,可包括如下步骤:培养前文所述的产谷氨酰胺转氨酶的工程菌,从培养产物中获得谷氨酰胺转氨酶。

[0034] 在所述方法中,进行所述培养时,采用的培养基为LB培养基,条件为:37℃,200rpm培养48小时。

[0035] 在所述方法中,可按照包括如下步骤的方法从所述培养产物中获得谷氨酰胺转氨酶:培养结束后,将整个培养体系进行离心(6000rpm,4℃,20分钟),上清液过膜(截留分子量为30KDa)浓缩干燥,制成干粉,即为谷氨酰胺转氨酶的粗酶粉。

[0036] 第五方面,本发明要求保护利用前文第四方面中所述的方法制备得到的谷氨酰胺转氨酶。

[0037] 采用该方法制备得到的所述谷氨酰胺转氨酶具有如下特性:最适反应pH为6-7,最适反应温度为50℃,放置15天后所述谷氨酰胺转氨酶仍具有80%的活力,存在25mM钙离子的条件下,所述谷氨酰胺转氨酶的活性为无钙离子的条件下的113%。

[0038] 本发明与现有的以枯草芽孢杆菌为宿主的TG酶表达具有下述区别:①TG酶活力

高,最终酶活达到约130U/mg,满足工业生产需要;②基因整合到基因组上,性能稳定;③TG酶稳定性好,放置15天TG酶仍具有80%的活力;④非钙离子依赖性,钙离子稳定性好,存在25mM钙离子的条件下,TG酶具有更高的活性(113%),应用范围更为广泛。本发明对于谷氨酰胺转氨酶的工业化生产具有重要意义。

### 附图说明

[0039] 图1为重组表达载体pWB980-pro-tg和pWB980-pro-sacb-tg的质粒图谱。A为pWB980-pro-tg;B为pWB980-pro-sacb-tg。

[0040] 图2为各重组芽孢杆菌的SDS-PAGE验证结果。1:蛋白Marker;2:PWB980-WB600;3:pro-sacb-tg-PWB980-WB600。箭头处为目标蛋白。

[0041] 图3为菌株*B. subtilis* TG所产TG酶的最适反应pH和最适反应温度测定结果。A为最适反应pH;B为最适反应温度。

[0042] 图4为菌株*B. subtilis* TG所产TG酶的钙离子依赖性测定结果。

### 具体实施方式

[0043] 下述实施例中所使用的实验方法如无特殊说明,均为常规方法。

[0044] 下述实施例中所用的材料、试剂等,如无特殊说明,均可从商业途径得到。

[0045] 达卡轮枝链霉菌(*S. ladakanum*):记载于“Cloning and expression of the transglutaminase gene from *Streptovercillium ladakanum* in *Streptomyces lividans*,《Process Biochemistry》,2004,39(5):591-598”一文,公众可从申请人处获得,仅可用于重复本发明实验使用。

[0046] pWB980载体:TakaRa Biotechnology (Dalian) Co.,Ltd. (Dalian,China) 该载体的酶切位点XmaI上游有sacB信号肽编码序列,sacB信号肽的氨基酸序列如SEQ ID No.1所示,其对应的核苷酸序列如SEQ ID No.14所示。

[0047] 枯草芽孢杆菌菌株WB600(简称*B. subtilis* WB600):为上海今迈生物科技有限公司(中国,上海)产品。

[0048] 分散酶:北京索莱宝科技有限公司产品,货号为Z8030。活力: $\geq 50$ units/mg;活力定义:1g固体酶粉,在30℃,pH7.5条件下,1分钟水解酪素产生1 $\mu$ g酪氨酸为1个酶活力单位,以u/g表示;稳定pH:5.5~8.5;最适pH:6.8~7.0;最适温度:45~50℃。

[0049] 实施例1、产谷氨酰胺转氨酶的重组枯草芽孢杆菌的构建及鉴定

[0050] 一、达卡轮枝链霉菌(*S. ladakanum*)来源TG酶基因的克隆

[0051] 达卡轮枝链霉菌(*S. ladakanum*)来源TG酶由信号肽,前导肽和活性区域三部分组成。链霉菌TG酶是以无活性的酶原进行合成和分泌,其N端前导肽被胞外活化蛋白酶切割后,才会使TG酶表现出催化活性。由于N端区域决定着TG酶的活性表达,因此,本发明采取两种表达策略,分别为信号肽替换和前导区重排。

[0052] 针对第一种策略,用正向引物 pro-TG-XmaI-F:5'-cggtaccccggggccaccggcagtggcagtg-3';和反向引物 pro-TG-R-XbaI:5'-ctctagaggatcgagcggccagccctgtgtcacct-3',以达卡轮枝链霉菌(*S. ladakanum*)基因组为模板,扩增得到了含有TG酶自身前导肽编码序列和TG酶活性区域编码序列的DNA片段pro-TG。所述DNA片段



pro-TG的序列为“5'-cggtagccggg+SEQ ID No.25+SEQ ID No.26的第1-993位+ctcgatcctctagag-3'”。其中,SEQ ID No.25为达卡轮枝链霉菌(*S.ladakanum*)来源的TG酶的自身前导肽编码核酸序列,编码SEQ ID No.12所示的*S.ladakanum*来源的TG酶自身前导肽;SEQ ID No.26为达卡轮枝链霉菌(*S.ladakanum*)来源的TG酶活性区域编码核酸序列,编码SEQ ID No.13所示的达卡轮枝链霉菌(*S.ladakanum*)来源的TG酶活性区域。

[0053] 针对第二种策略,需要分别克隆TG酶的前导肽和活性区域,以及SacB序列。用PRO-F-XmaI:5'-cctcgagctcggtagccggggccaccggcagtggtggcagcgcg-3'和PRO-R:5'-cagccctcctggtaccgct atcacttttagggggcccggaaggaccgaccg-3'这对引物从达卡轮枝链霉菌(*S.ladakanum*)基因组扩增得到了谷氨酰胺转氨酶的前导肽的核酸序列pro(5'-cctcgagctcggtagccggg+SEQ ID No.25+agtgatagcggtagcaggaggctg-3')。用TG-F:5'-aggaggcgcaactcaagcttttgccgatcctctagaggactccgacgagc-3'和TG-R-NheI:5'-aagctagcttgcctgcagg tggaggccagccctgtgtcaccttgtc-3'这对引物从达卡轮枝链霉菌(*S.ladakanum*)基因组扩增得到了TG酶活性区域的基因序列TG(5'-aggaggcgcaactcaagcttttgccgatcctctagag+SEQ ID No.26的第1-992位+tcgacctgcaggcatgcaagctagctt-3')。用sacB-F:5'-cggtcctccttccggggccccctaaagtgatagcggtagcaggaggctg-3'和sacB-R:5'-gctcgtcggagtcctctagaggatcgcaaaagcttgagttgcctcct-3'为引物,以pWB980载体为模板,克隆得到sacB信号肽序列(5'-cggtcctccttccggggccccctaaagtgatagcggtagcaggaggctggaagaagcagaccgctaacacagtacataaaaaaggagacatgaacg+SEQ ID No.14+gaccccttag aggactccgacgagc-3')。

[0054] 二、表达载体的构建及在枯草芽孢杆菌中的表达

[0055] 1、pWB980-pro-tg的构建

[0056] 利用sacB信号肽替换链霉菌来源TG酶原有信号肽,以pWB980为表达载体,使用XmaI和XbaI分别酶切pWB980载体和步骤一中第一种策略所得的pro-TG。然后将酶切后的pro-TG片段和pWB980载体骨架大片段16℃过夜连接,化转*B.subtilis* WB600。PCR验证获得重组载体pWB980-pro-tg(图1中A)。

[0057] 重组表达载体pWB980-pro-tg的结构描述为:将pWB980载体的酶切位点XmaI和XbaI之间的小片段替换为DNA片段pro-TG后得到的重组质粒(XmaI上游有SacB信号肽的核酸序列)。

[0058] 将向*B.subtilis* WB600中转入重组表达载体pWB980-pro-tg后所得的重组枯草芽孢杆菌命名为pro-tg-PWB980-WB600。

[0059] 2、pWB980-pro-sacB-tg的构建

[0060] 将pro、sacB和TG这三个片段(步骤一中第二种策略所得)通过引物PRO-F-XmaI和TG-R-NheI(具体序列同上)overlap合在一起,得到长片段pro-sacB-TG。然后连在pMD19T上,转化*E.coli* TOP10,PCR验证选择正确的转化子pro-sacB-TG-19T。使用XmaI和NheI将长片段pro-sacB-TG从pro-sacB-TG-19T载体上切下来,并同时XmaI和NheI酶切pWB980载体。将得到的片段pro-sacB-TG和pWB980载体骨架大片段16℃过夜连接,转化*B.subtilis* WB600,PCR验证获得重组载体pWB980-pro-sacB-tg(图1中B)。

[0061] 重组表达载体pWB980-pro-sacB-tg的结构描述为:将pWB980载体的酶切位点XmaI和NheI之间的小片段替换为DNA片段pro-sacB-TG后得到的重组质粒。pWB980-pro-sacB-tg

中,DNA片段pro-sacB-TG连接到质粒pWB980sacB信号肽后面。

[0062] 将向*B.subtilis* WB600中转入重组表达载体pWB980-pro-sacB-tg后所得的重组枯草芽孢杆菌命名为pro-sacB-tg-PWB980-WB600。

### [0063] 三、酶活检测

#### [0064] 1、SDS-PAGE蛋白胶验证

[0065] 取步骤二获得的各重组芽孢杆菌的培养上清1mL,加入130微升100%TCA冰上放置30min,离心收集沉淀,丙酮洗两次,挥发后沉淀溶解于20微升蒸馏水。加入蛋白Loading沸水浴5min。以带有pWB980空质粒的枯草芽孢杆菌(命名为PWB980-WB600)做对照,重组蛋白的大小和纯度检测使用SDS-PAGE电泳。SDS-PAGE结果如图2。由图可见,在重组芽孢杆菌pro-tg-pWB980-WB600的培养上清中检测到目的单一的蛋白条带,条带大小约为44KDa,说明TG酶被成功表达。

#### [0066] 2、酶活测定方法

[0067] 用比色法测定酶活:通过测定CBZ-Gln-Gly与羟胺反应生成氧肟酸的量来计算。反应液A为0.2M Tris-HCl (pH=6.0)、0.03M CBZ-Gln-Gly、0.01M还原型谷胱甘肽、0.1M盐酸羟胺。显色液B为3M盐酸、12% (w/v) (12g/100ml) 三氯乙酸、5% (w/v) (5g/100ml) 六水三氯化铁(溶解于0.1M盐酸中)三种试剂等体积混合。

[0068] 将步骤二获得的各重组芽孢杆菌置于LB培养基中培养,37°C,200rpm培养48小时。培养结束后,按照如下方法测定重组芽孢杆菌胞外酶活:取500μl发酵上清,加入10μl浓度为2.4mg/ml的分散酶,37°C反应20min。取100μl分散酶处理的上清,加入200μl的反应液A,37°C反应60分钟,加入200μl的显色液B,在525nm下检测吸光值。根据吸光值计算氧肟酸含量。对照为100μl分散酶处理过的发酵上清液加入200μl显色液B终止反应,37°C处理60min之后,加入200μl反应液A。1单位TG酶活定义为:37°C每分钟催化底物,形成1μmol氧肟酸所需的酶量(U/mL)。采用Bradford方法计算蛋白浓度(mg/mL),利用单位酶活除以蛋白浓度得到比酶活(U/mg)。

[0069] 结果:48小时重组菌pro-tg-pWB980-WB600的胞外TG酶的比酶活达到31U/mg。重组菌pro-sacB-tg-PWB980-WB600未检测到胞外TG酶活。

### [0070] 3、突变体构建及酶活检测

[0071] 根据步骤2的结果,利用重组表达载体pWB980-pro-tg和*B.subtilis* WB600构建重组菌pro-tg-pWB980-WB600的突变体。

[0072] 利用定点突变技术,对来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域(SEQ ID No.27)做如下几种突变(突变位点主要集中在前导肽及前导肽与活性区的交界部分):

[0073] (1) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第83位的E突变为D(对应的突变体记为E83D)。正向引物5' -ttccgggccccgactccgacgaccgggtgactcct-3',反向引物5' -aggagtcacccgggtcgtcggagtcggggccccgaa-3'。

[0074] (2) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第80-82位的DSD缺失(对应的突变体记为DSD缺失)。正向引物5' -gtcctccttccgggccccgagcgggtgactcctcccg-3',反向引物5' -gcgggaggagtcaccc

gctcgggggccccggaaggacggac-3'。

[0075] (3) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第80-82位的DSD突变成AAA(对应的突变体记为DSD80-82AAA)。正向引物5'-gtccttccgggccccgcagcagcagagcgggtgactcctccgc-3', 反向引物5'-gcgggaggagtcacccg ctctgctgctgcgggggccccggaaggac-3'。

[0076] (4) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第74-75位的PS缺失(对应的突变体记为PS缺失)。正向引物5'-gccgcttcgagcgcggtttccgggccccgactcc-3', 反向引物5'-ggagtcgggggccccgaaa ccggcgctcgaagcggc-3'。

[0077] (5) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第74-75位的PS突变为AA(对应的突变体记为PS 74-75AA)。正向引物5'-gcttcgagcgcggtgagcagcattccgggccccgac-3', 反向引物5'-gtcgggggccccggaatgctgcacccg cgctcgaagc-3'。

[0078] (6) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第34-36位的SGS突变为AAA(对应的突变体记为SGS34-36AAA)。正向引物5'-gccaccggcagtggcgcagcagcagggcaccggggaagag-3', 反向引物5'-cttccccgggtgcctgctgctgc gccactgccggtggc-3'。

[0079] (7) 将SEQ ID No.27所示的来源于*S.ladakanum*的TG酶从N-端信号肽到C-端活性区域的氨基酸序列的第74-75位的PS突变为AA, 同时将第80-82位的DSD突变成AAA(对应的突变体记为PS 74-75AA/DSD 80-82AAA)。在(1)的基础上(第80-82位的DSD突变成AAA), 正向引物5'-tcgagcgc cggtgagcagcattccgggcccc-3', 反向引物5'-gggggccccggaatgctgcacccggcgctcga-3'进行第二个位点(第74-75位的PS突变为AA)的突变。

[0080] 以重组表达载体pWB980-pro-tg为模板, 以(1)-(7)中所示的各正向引物和相应的反应引物进行PCR扩增, 得到构建各突变体所对应的重组突变表达载体。将各重组突变表达载体分别导入*B.subtilis* WB600中, 得到对应于上述(1)-(7)的重组菌pro-tg-PWB980-WB600的各种突变体。

[0081] 将获得重组菌pro-tg-PWB980-WB600的各种突变体以及重组菌pro-tg-PWB980-WB600(野生型)分别置于LB培养基中培养, 37°C, 200rpm培养48小时。培养结束后, 按照上述步骤2中的方法测定各重组菌所得TG酶的比酶活(U/mg)。

[0082] 结果如表1所示, 由表可知: 与原始菌相比, 除了N-端第83位突变导致酶活降低, 其余位点突变均得到较高的酶活和比酶活, 其中N-端74-75位和80-82位双点突变酶活和比酶活最高, 酶活达到26U/ml, 比酶活达到128U/mg。

[0083] 表1重组菌pro-tg-pWB980-WB600突变体的突变信息以及比酶活测定结果

重组菌	前导肽突变信息	比酶活 (U/mg)
野生型	蛋白: SEQ ID No.27	31
	基因: SEQ ID No.28	
E83D	蛋白: E83D	25
	基因: gag 247-249 gac	
DSD 缺失	蛋白: $\Delta$ 80-82	41
	基因: $\Delta$ 238-246	
DSD 80-82 AAA	蛋白: DSD 80-8 2AAA	54
	基因: gactccgac 238-246 gcagcagca	
PS 缺失	蛋白: $\Delta$ 74-75	60
	基因: $\Delta$ 220-225	
PS 74-75 AA	蛋白: PS 74-75 AA	86
	基因: ccgtcc 220-225 gcagca	
SGS 34-36 AAA	蛋白: SGS 34-36 AAA	34
	基因: agtggcagc100-108 gcagcagca	
PS 74-75 AA / DSD 80-82 AAA	蛋白: PS 74-75 AA / DSD 80-82 AAA	128
[0085]	基因: ccgtcc 220-225 gcagca / gactccgac 238-246 gcagcagca	

[0086] 注:蛋白取代/缺失的编号是从SEQ ID No.27所示氨基酸序列的N端为起始的;基因取代/缺失的编号是从SEQ ID No.27所示核苷酸序列的5'端为起始的。表中,对于氨基酸取代,使用下述命名法:原始氨基酸(野生型),位置(即在SEQ ID No.27中的位置),取代氨基酸。如,在SEQ ID No.27的第83位用天冬氨酸(D)取代原有的谷氨酸(E)命名为“E83D”;在SEQ ID No.27的第80-82位用三个丙氨酸(AAA)取代原有的“天冬氨酸-丝氨酸-天冬氨酸”(DSD)命名为“DSD80-82AAA”。对于碱基取代,使用下述命名法:原始碱基(野生型),位置(即在SEQ ID No.28中的位置),取代碱基。如,在SEQ ID No.28的第247-249位用gac取代原有的gag命名为“gag 247-249gac”。对于氨基酸和碱基的缺失,均以符号 $\Delta$ 表示,如“蛋白: $\Delta$ 80-82”表示缺失SEQ ID No.27的第80-82位氨基酸,“基因: $\Delta$ 238-246”表示缺失SEQ ID No.28的第238-246位核苷酸”。对于氨基酸和碱基的突变而言,均以“/”表示多重突变。

#### [0087] 四、基因组整合与酶学性质分析

##### [0088] 1、基因组整合

[0089] 由于步骤三表1中的突变体PS74-75AA/DSD80-82AAA具有较高的酶活,因此将该重组菌中的TG突变体基因整合到B.subtilis WB600基因lacA位点,获得菌株B.subtilis TG。

[0090] 具体操作如下:

[0091] 使用引物LacA-up-F和LacA-up-R扩增WB600基因组LacA位点的5'端序列,扩增所得序列为:5'-gccagtcactatggcgtgctgctagc+SEQ ID No.29+atcgattttgctgaggtggcagagg-3'。其中,SEQ ID No.29为来源于B.subtilis WB600基因

lacA位点的5'端的核苷酸序列。

[0092] LacA-up-F:5'-gccagtcactatggcgtgctgctagc gatcagaccctcatagaaatcacgct-3' ;

[0093] LacA-up-R:5'-cctctgccacctcagcaaaatcgat atccatgcaaatgattgcccacggc-3'。

[0094] 使用引物P43-F和TG-R扩增表达载体PS74-75AA/DSD80-82AAA-PWB980(表达突变体PS74-75AA/DSD80-82AAA的载体)的TG表达框,扩增所得序列为:5'-gccgtgggcaatcatttgcattgat+SEQ ID No.30+acagcagattcatgcccgggcttt-3'。其中,SEQ ID No.30为来源于S.ladakanum的双点突变(PS74-75AA/DSD80-82AAA)TG酶的核苷酸序列。

[0095] P43-F:5'-gccgtgggcaatcatttgcattgat atcgatttctgctgaggtggcagagg-3' ;

[0096] TG-R:5'-aaagcggcgggcatgaatctgc tgtggtaatcgagcggccagccctgtgt-3'

[0097] 使用引物LacA-Down-F和引物LacA-Down-R扩增WB600基因组的LacA的3'端序列,扩增所得序列为:5'-acacagggctggccgctcgattaacc+SEQ ID No.31+gagctcctaacttataggggtaaacacttaa-3'。其中,SEQ ID No.31为来源于B.subtilis WB600基因lacA位点的3'端的核苷酸序列。

[0098] LacA-Down-F:5'-acacagggctggccgctcgattaacc acagcagattcatgcccgggcttt-3' ;

[0099] LacA-Down-R:5'-ttaagtgttaccctataagttaggagctccagtataggcggcgggtcatattag-3'

[0100] 将上述三片段overlap融合在一起。利用RMGE基因编辑方法(A Novel Tool for Microbial Genome Editing Using the Restriction-Modification System.ACS Synth Biol.2018Jan 19;7(1):98-106.doi:10.1021/acssynbio.7b00254.Epub 2017Oct 12.)方法整合到B.subtilis WB600基因组上,并经测序验证正确后的菌株即为B.subtilis TG。

[0101] 将菌株B.subtilis TG在LB培养基中,37℃,200rpm,培养48小时。培养结束后,收集整个培养体系,离心(6000rpm,4℃,20分钟),上清液过膜浓缩干燥(截留分子量为30KDa),制成干粉,即为TG酶的粗酶粉。

[0102] 2、最适反应pH

[0103] 将酶干粉(即步骤1获得的TG酶的粗酶粉)用水溶解,制成浓度为2.5g/L的酶液(以下测酶学性质所用浓度均是一样),取500μl酶液,加入10μl浓度为2.4mg/ml的分散酶,37℃反应20min。加入200μl的反应液A(配方同上)。取100μl分散酶处理的上清,在pH为3,4,5(0.2mol/mL磷酸氢二钠-柠檬酸缓冲液),6,7,8,9(0.2mol/mL Tris-盐酸缓冲液)中,于37℃反应1个小时,加入200μl的显色液B(配方同上),在525nm条件下检测酶活(U/mg),以最高点酶活为100%,其余与之相比得相对酶活。结果见图3中A,最适反应pH为6-7。

[0104] 3、最适反应温度

[0105] 将500μl酶液,加入10μl浓度为2.4mg/ml的分散酶,37℃反应20min。加入200μl的反应液A(配方同上)。取100μl分散酶处理的上清,在0.2mol/mL Tris-盐酸缓冲液(pH 6)中,分别在20℃,30℃,40℃,50℃,60℃,70℃的条件下反应1个小时。加入200μl的显色液B(配方同上),在525nm条件下检测酶活(U/mg)。设其中的最高点酶活为100%,其余与之相比得相对酶活。结果见图3中B,最适反应温度为50℃。

[0106] 4、金属离子依赖性

[0107] 向分散酶处理后的酶液(反应条件同步骤3)加入终浓度为5mM的金属离子( $\text{Na}^+$ 、 $\text{K}^+$ 、 $\text{Ca}^{2+}$ 、 $\text{Mg}^{2+}$ 、 $\text{Cu}^{2+}$ 、 $\text{Zn}^{2+}$ 、 $\text{Mn}^{2+}$ ),反应条件同步骤3(温度为37℃)。测定添加不同金属离子后的酶活,以不加金属离子的酶活为100%,其他与之相比得相对酶活。结果如表2所示:铜离子和锌离子对*B.subtilis* TG来源TG酶具有较强的抑制作用,而钾离子,钠离子,钙离子等对*B.subtilis* TG来源TG酶没有抑制作用。

[0108] 表2 *B.subtilis* TG来源TG酶对金属离子的依赖性测定结果

[0109]

序号	金属离子	相对酶活/%
1	对照	100.0±2
2	$\text{Na}^+$	109.2±0.5
3	$\text{K}^+$	90.8±2.4
4	$\text{Ca}^{2+}$	101.4±2.2
5	$\text{Mg}^{2+}$	102.4±0.1
6	$\text{Cu}^{2+}$	59.7±0.9
7	$\text{Zn}^{2+}$	9.8±0.5
8	$\text{Mn}^{2+}$	91.7±1.7

[0110] 5、钙离子依赖性

[0111] 向分散酶处理后的酶液(反应条件同步骤3)加入不同浓度的 $\text{Ca}^{2+}$ (5mM,10mM,15mM,20mM,25mM,为在反应体系中的终浓度),反应条件同步骤3(温度为37℃)。以不加 $\text{Ca}^{2+}$ 的酶液为对照,其酶活设为100%,其余与之相比得相对酶活。结果如图4所示,可见,*B.subtilis* TG来源TG酶是 $\text{Ca}^{2+}$ 非依赖性,但是 $\text{Ca}^{2+}$ 存在不仅没有抑制TG酶活,而且酶活有所增加。存在25mM钙离子的条件下,TG酶具有更高的活性(113%)。因此,与其它来源TG酶相比,*B.subtilis* TG来源TG酶具有更加广阔的利用空间。

[0112] 五、信号肽优化

[0113] 信号肽影响着蛋白的分泌,从而影响胞外TG酶含量。从枯草芽孢杆菌信号肽数据库中(SignalP 3.0),选择不同强度的信号肽,替换酶活最高突变体(PS74-75AA/DSD80-82AAA)中的sacB信号肽,候选的10个信号肽如表3所示。利用上下游引物(上游引物中带有BamH I酶切位点,下游引物带有Xma I酶切位点),以枯草芽孢杆菌WB600为模板,PCR得到信号肽片段,将重组信号肽表达载体导入*B.subtilis* WB600中,得到对应的重组菌。测定各菌株的比酶活,测定方法同上文。比酶活结果显示,不同信号肽对胞外酶活影响较小,其中与sacB相比,AmyQ为信号肽比酶活有略微提高。

[0114] 表3 10个信号肽以及对应重组菌的比酶活测定结果

[0115]

信号肽名称	序列	上下游引物(5'-3')	比酶活(U/mg)
AmyE	氨基酸: SEQ ID No.2 核苷酸: SEQ ID No.15	AmyE-F: gaacatctggatccatgtttgcaaacgattcaaaacc AmyE-R: tggccccgggtaccgagctcgaggcactcgagccgcccgg	120
AmyQ	氨基酸: SEQ ID No.3 核苷酸: SEQ ID No.16	AmyQ-F: catctggatccatgattcaaaaacgaaagcggacagtttcgt AmyQ-R: tggccccgggtaccgagctcgaggctgatgttttgaatcgcaaa	130
Epr	氨基酸: SEQ ID No.4 核苷酸: SEQ ID No.17	Epr-F: atgaacatctggatccatgaaaaacatgtcttgcaaac Epr-R: tggccccgggtaccgagctcgagcgcgatgagcggagggccta	125

[0116]

LytD	氨基酸: SEQ ID No.5 核苷酸: SEQ ID No.18	lytD-F: catctggatccatgaaaaagagactaatcgcacct lytD-R: ccccggtaccgagctcgaggctgggcagaaccagacat	122
NucB	氨基酸: SEQ ID No.6 核苷酸: SEQ ID No.19	NucB-F: tgaacatctggatccatgaaaaatggatggcaggcct NucB-R: ggccccgggtaccgagctcgacgaagatgcgccttggatctg	119
PenP	氨基酸: SEQ ID No.7 核苷酸: SEQ ID No.20	PenP-F: acatctggatccatgaagttgaaaactaaagcgtcaataa PenP-R: tggccccgggtaccgagctcgaggctcggcatgtgttgagt	110
WapA	氨基酸: SEQ ID No.8 核苷酸: SEQ ID No.21	WapA-F: acatctggatccatgaaaaaagaaagaggcga WapA-R: gccccgggtaccgagctcgaggctagtagatcggtggca	115
WprA	氨基酸: SEQ ID No.9 核苷酸: SEQ ID No.22	WprA-F: atgaacatctggatccatgaaacgcagaaaattcagct WprA-R: ggccccgggtaccgagctcgacgctgcagcttggttccgg	118
YhcR	氨基酸: SEQ ID No.10 核苷酸: SEQ ID No.23	YhcR-F: atgaacatctggatccatgctgtctgtcgaaatgataa YhcR-R: ggccccgggtaccgagctcgaaagctcgaacgtgtacattac	120
YncM	氨基酸: SEQ ID No.11 核苷酸: SEQ ID No.24	YncM-F: catctggatccatggcgaaccactatcaaaaagggga YncM-R: ggccccgggtaccgagctcgaggctctgcccggtgtaa	126

[0001] <110> 中国科学院微生物研究所  
 [0002] <120> 一种利用食源安全性宿主表达谷氨酰胺转氨酶的方法  
 [0003] <130> GNCLN181470  
 [0004] <160> 31  
 [0005] <170> PatentIn version 3.5  
 [0006] <210> 1  
 [0007] <211> 29  
 [0008] <212> PRT  
 [0009] <213> Artificial sequence  
 [0010] <400> 1  
 [0011] Met Asn Ile Lys Lys Phe Ala Lys Gln Ala Thr Val Leu Thr Phe Thr  
 [0012] 1 5 10 15  
 [0013] Thr Ala Leu Leu Ala Gly Gly Ala Thr Gln Ala Phe Ala  
 [0014] 20 25  
 [0015] <210> 2  
 [0016] <211> 33  
 [0017] <212> PRT  
 [0018] <213> Artificial sequence  
 [0019] <400> 2  
 [0020] Met Phe Ala Lys Arg Phe Lys Thr Ser Leu Leu Pro Leu Phe Ala Gly  
 [0021] 1 5 10 15  
 [0022] Phe Leu Leu Leu Phe His Leu Val Leu Ala Gly Pro Ala Ala Ala Ser  
 [0023] 20 25 30  
 [0024] Ala  
 [0025] <210> 3  
 [0026] <211> 31  
 [0027] <212> PRT  
 [0028] <213> Artificial sequence  
 [0029] <400> 3  
 [0030] Met Ile Gln Lys Arg Lys Arg Thr Val Ser Phe Arg Leu Val Leu Met  
 [0031] 1 5 10 15  
 [0032] Cys Thr Leu Leu Phe Val Ser Leu Pro Ile Thr Lys Thr Ser Ala  
 [0033] 20 25 30  
 [0034] <210> 4  
 [0035] <211> 27  
 [0036] <212> PRT  
 [0037] <213> Artificial sequence  
 [0038] <400> 4  
 [0039] Met Lys Asn Met Ser Cys Lys Leu Val Val Ser Val Thr Leu Phe Phe  
 [0040] 1 5 10 15  
 [0041] Ser Phe Leu Thr Ile Gly Pro Leu Ala His Ala



[0042]	20	25	
[0043]	<210> 5		
[0044]	<211> 27		
[0045]	<212> PRT		
[0046]	<213> Artificial sequence		
[0047]	<400> 5		
[0048]	Met Lys Lys Arg Leu Ile Ala Pro Met Leu Leu Ser Ala Ala Ser Leu		
[0049]	1	5	10 15
[0050]	Ala Phe Phe Ala Met Ser Gly Ser Ala Gln Ala		
[0051]	20	25	
[0052]	<210> 6		
[0053]	<211> 29		
[0054]	<212> PRT		
[0055]	<213> Artificial sequence		
[0056]	<400> 6		
[0057]	Met Lys Lys Trp Met Ala Gly Leu Phe Leu Ala Ala Ala Val Leu Leu		
[0058]	1	5	10 15
[0059]	Cys Leu Met Val Pro Gln Gln Ile Gln Gly Ala Ser Ser		
[0060]	20	25	
[0061]	<210> 7		
[0062]	<211> 36		
[0063]	<212> PRT		
[0064]	<213> Artificial sequence		
[0065]	<400> 7		
[0066]	Met Lys Leu Lys Thr Lys Ala Ser Ile Lys Phe Gly Ile Cys Val Gly		
[0067]	1	5	10 15
[0068]	Leu Leu Cys Leu Ser Ile Thr Gly Phe Thr Pro Phe Phe Asn Ser Thr		
[0069]	20	25	30
[0070]	His Ala Glu Ala		
[0071]	35		
[0072]	<210> 8		
[0073]	<211> 32		
[0074]	<212> PRT		
[0075]	<213> Artificial sequence		
[0076]	<400> 8		
[0077]	Met Lys Lys Arg Lys Arg Arg Asn Phe Lys Arg Phe Ile Ala Ala Phe		
[0078]	1	5	10 15
[0079]	Leu Val Leu Ala Leu Met Ile Ser Leu Val Pro Ala Asp Val Leu Ala		
[0080]	20	25	30
[0081]	<210> 9		
[0082]	<211> 31		
[0083]	<212> PRT		

[0084] <213> Artificial sequence  
 [0085] <400> 9  
 [0086] Met Lys Arg Arg Lys Phe Ser Ser Val Val Ala Ala Val Leu Ile Phe  
 [0087] 1 5 10 15  
 [0088] Ala Leu Ile Phe Ser Leu Phe Ser Pro Gly Thr Lys Ala Ala Ala  
 [0089] 20 25 30  
 [0090] <210> 10  
 [0091] <211> 46  
 [0092] <212> PRT  
 [0093] <213> Artificial sequence  
 [0094] <400> 10  
 [0095] Met Leu Ser Val Glu Met Ile Ser Arg Gln Asn Arg Cys His Tyr Val  
 [0096] 1 5 10 15  
 [0097] Tyr Lys Gly Gly Asn Met Met Arg Arg Ile Leu His Ile Val Leu Ile  
 [0098] 20 25 30  
 [0099] Thr Ala Leu Met Phe Leu Asn Val Met Tyr Thr Phe Glu Ala  
 [0100] 35 40 45  
 [0101] <210> 11  
 [0102] <211> 42  
 [0103] <212> PRT  
 [0104] <213> Artificial sequence  
 [0105] <400> 11  
 [0106] Met Ala Lys Pro Leu Ser Lys Gly Gly Ile Leu Val Lys Lys Val Leu  
 [0107] 1 5 10 15  
 [0108] Ile Ala Gly Ala Val Gly Thr Ala Val Leu Phe Gly Thr Leu Ser Ser  
 [0109] 20 25 30  
 [0110] Gly Ile Pro Gly Leu Pro Ala Ala Asp Ala  
 [0111] 35 40  
 [0112] <210> 12  
 [0113] <211> 51  
 [0114] <212> PRT  
 [0115] <213> S. ladakanum  
 [0116] <400> 12  
 [0117] Ala Thr Gly Ser Gly Ser Gly Ser Gly Thr Gly Glu Glu Lys Arg Ser  
 [0118] 1 5 10 15  
 [0119] Tyr Ala Glu Thr His Arg Leu Thr Ala Asp Asp Val Asp Asp Ile Asn  
 [0120] 20 25 30  
 [0121] Ala Leu Asn Glu Ser Ala Pro Ala Ala Ser Ser Ala Gly Pro Ser Phe  
 [0122] 35 40 45  
 [0123] Arg Ala Pro  
 [0124] 50  
 [0125] <210> 13

[0126] <211> 331  
 [0127] <212> PRT  
 [0128] <213> S. ladakanum  
 [0129] <400> 13  
 [0130] Asp Ser Asp Glu Arg Val Thr Pro Pro Ala Glu Pro Leu Asp Arg Met  
 [0131] 1 5 10 15  
 [0132] Pro Asp Pro Tyr Arg Pro Ser Tyr Gly Arg Ala Glu Thr Ile Val Asn  
 [0133] 20 25 30  
 [0134] Asn Tyr Ile Arg Lys Trp Gln Gln Val Tyr Ser His Arg Asp Gly Arg  
 [0135] 35 40 45  
 [0136] Lys Gln Gln Met Thr Glu Glu Gln Arg Glu Trp Leu Ser Tyr Gly Cys  
 [0137] 50 55 60  
 [0138] Val Gly Val Thr Trp Val Asn Ser Gly Gln Tyr Pro Thr Asn Arg Leu  
 [0139] 65 70 75 80  
 [0140] Ala Phe Ala Phe Phe Asp Glu Asp Lys Tyr Lys Asn Glu Leu Lys Asn  
 [0141] 85 90 95  
 [0142] Gly Arg Pro Arg Ser Gly Glu Thr Arg Ala Glu Phe Glu Gly Arg Val  
 [0143] 100 105 110  
 [0144] Ala Lys Asp Ser Phe Asp Glu Ala Lys Gly Phe Gln Arg Ala Arg Asp  
 [0145] 115 120 125  
 [0146] Val Ala Ser Val Met Asn Lys Ala Leu Glu Asn Ala His Asp Glu Gly  
 [0147] 130 135 140  
 [0148] Ala Tyr Leu Asp Asn Leu Lys Lys Glu Leu Ala Asn Gly Asn Asp Ala  
 [0149] 145 150 155 160  
 [0150] Leu Arg Asn Glu Asp Ala Arg Ser Pro Phe Tyr Ser Ala Leu Arg Asn  
 [0151] 165 170 175  
 [0152] Thr Pro Ser Phe Lys Asp Arg Asn Gly Gly Asn His Asp Pro Ser Lys  
 [0153] 180 185 190  
 [0154] Met Lys Ala Val Ile Tyr Ser Lys His Phe Trp Ser Gly Gln Asp Arg  
 [0155] 195 200 205  
 [0156] Ser Gly Ser Ser Asp Lys Arg Lys Tyr Gly Asp Pro Glu Ala Phe Arg  
 [0157] 210 215 220  
 [0158] Pro Asp Arg Gly Thr Gly Leu Val Asp Met Ser Arg Asp Arg Asn Ile  
 [0159] 225 230 235 240  
 [0160] Pro Arg Ser Pro Thr Ser Pro Gly Glu Ser Phe Val Asn Phe Asp Tyr  
 [0161] 245 250 255  
 [0162] Gly Trp Phe Gly Ala Gln Thr Glu Ala Asp Ala Asp Lys Thr Val Trp  
 [0163] 260 265 270  
 [0164] Thr His Gly Asn His Tyr His Ala Pro Asn Gly Ser Leu Gly Ala Met  
 [0165] 275 280 285  
 [0166] His Val Tyr Glu Ser Lys Phe Arg Asn Trp Ser Asp Gly Tyr Ser Asp  
 [0167] 290 295 300

[0168] Phe Asp Arg Gly Ala Tyr Val Val Thr Phe Val Pro Lys Ser Trp Asn  
[0169] 305 310 315 320  
[0170] Thr Ala Pro Asp Lys Val Thr Gln Gly Trp Pro  
[0171] 325 330  
[0172] <210> 14  
[0173] <211> 87  
[0174] <212> DNA  
[0175] <213> Artificial sequence  
[0176] <400> 14  
[0177] atgaacatca aaaagtttgc aaaacaagca acagtattaa cctttactac cgcactgctg 60  
[0178] gcaggaggcg caactcaagc ttttgcc 87  
[0179] <210> 15  
[0180] <211> 99  
[0181] <212> DNA  
[0182] <213> Artificial sequence  
[0183] <400> 15  
[0184] atgtttgcaa aacgattcaa aacctcttta ctgccgttat tcgctggatt tttattgctg 60  
[0185] tttcatttgg ttctggcagg accggcgct gcgagtgc 99  
[0186] <210> 16  
[0187] <211> 93  
[0188] <212> DNA  
[0189] <213> Artificial sequence  
[0190] <400> 16  
[0191] atgattcaaa aacgaaagcg gacagtttcg ttcagacttg tgcttatgtg cacgctgtta 60  
[0192] tttgtcagtt tgccgattac aaaaacatca gcc 93  
[0193] <210> 17  
[0194] <211> 81  
[0195] <212> DNA  
[0196] <213> Artificial sequence  
[0197] <400> 17  
[0198] atgaaaaca tgtcttgc aaactgttga tcagtcactc tgtttttcag ttttctacc 60  
[0199] atagccctc tcgctcatgc g 81  
[0200] <210> 18  
[0201] <211> 81  
[0202] <212> DNA  
[0203] <213> Artificial sequence  
[0204] <400> 18  
[0205] atgaaaaga gactaatgc acctatgett ctatccgccg cgtcccttgc cttttttgcc 60  
[0206] atgtctggtt ctgcccaggc c 81  
[0207] <210> 19  
[0208] <211> 87  
[0209] <212> DNA

- [0210] <213> Artificial sequence  
[0211] <400> 19  
[0212] atgaaaaaat ggatggcagg cctgtttctt gctgcagcag ttcttctttg tttaatggtt 60  
[0213] ccgcaacaga tccaaggcgc atcttcg 87  
[0214] <210> 20  
[0215] <211> 108  
[0216] <212> DNA  
[0217] <213> Artificial sequence  
[0218] <400> 20  
[0219] atgaagtga aactaaagc gtcaataaaa ttcggaatat gtgttgggct tttatgttta 60  
[0220] agcattactg gtttcacacc tttttcaac tcaacacatg ccgaagcc 108  
[0221] <210> 21  
[0222] <211> 96  
[0223] <212> DNA  
[0224] <213> Artificial sequence  
[0225] <400> 21  
[0226] atgaaaaaaa gaaagaggcg aaactttaa aggttcattg cagcattttt agtgttggct 60  
[0227] ttaatgattt cattagtgcc agccgatgta ctagcc 96  
[0228] <210> 22  
[0229] <211> 93  
[0230] <212> DNA  
[0231] <213> Artificial sequence  
[0232] <400> 22  
[0233] atgaaacgca gaaattcag ctcggttggt gcggcagtc ttatTTTTGC actgattttc 60  
[0234] agcctTTTTT ctccgggaac caaagetgca gcg 93  
[0235] <210> 23  
[0236] <211> 138  
[0237] <212> DNA  
[0238] <213> Artificial sequence  
[0239] <400> 23  
[0240] atgctgtctg tcgaaatgat aagcagacaa aatcggtgtc attatgtga taaggaggga 60  
[0241] aatatgatga ggcgtattct gcatattgtg ttgatcacgg cattaatgtt cttaaagtga 120  
[0242] atgtacacgt tcgaagct 138  
[0243] <210> 24  
[0244] <211> 126  
[0245] <212> DNA  
[0246] <213> Artificial sequence  
[0247] <400> 24  
[0248] atggcgaaac cactatcaaa agggggaatt ttggtgaaaa aagtattgat tgcaggtgca 60  
[0249] gtagaacag cagttctttt cggaaccctt tcatcaggta taccaggttt acccgcgca 120  
[0250] gacgc 126  
[0251] <210> 25

[0252] <211> 153  
 [0253] <212> DNA  
 [0254] <213> S. ladakanum  
 [0255] <400> 25  
 [0256] gccaccggca gtggcagtgg cagcggcacc ggggaagaga agaggtccta cgccgaaacg 60  
 [0257] caccgcctga cggcggatga cgctgacgac atcaacgcgc tgaacgaaag cgctccggcc 120  
 [0258] gcttcgagcg ccggtccgtc ctccggggcc ccc 153  
 [0259] <210> 26  
 [0260] <211> 996  
 [0261] <212> DNA  
 [0262] <213> S. ladakanum  
 [0263] <400> 26  
 [0264] gactccgacg agcgggtgac tcctcccgcc gagccgctcg accggatgcc cgacccttac 60  
 [0265] cggccctcgt acggcagggc cgagacgac gtcaacaact acatacgcaa gtggcagcag 120  
 [0266] gtctacagcc accgcgacgg caggaaacag cagatgaccg aggaacagcg ggagtggctg 180  
 [0267] tcctacggtt gcgtcgggtg cacctgggtc aactcgggcc agtatccgac gaacaggtctg 240  
 [0268] gctttcgcgt tcttcgacga ggacaagtac aagaacgagc tgaagaacgg caggccccgg 300  
 [0269] tccggcgaaa cgcggcgga gttcgagggg cgcgtcgcca aggacagctt cgacgaggcg 360  
 [0270] aaggggttcc agcggcgcg tgactggcg tccgtcatga acaaggcctt ggagaacgcc 420  
 [0271] cagcagcagg gggcgtacct cgacaacctc aagaaggagc tggcgaacgg caacgacgcc 480  
 [0272] ctgcggaacg aggatgcccg ctcgcccttc tactcggcgc tgcggaacac gccgtccttc 540  
 [0273] aaggaccgca acggcgcaa tcacgaccg tccaagatga aggccgtcat ctactcgaag 600  
 [0274] cacttctgga gcggccagga ccggtcgggc tcctccgaca agaggaagta cggcgaccgg 660  
 [0275] gaggccttcc gccccgaccg cggcaccggc ctggtcgaca tgtcgaggga caggaacatt 720  
 [0276] ccgcgcagcc ccaccagccc cggcgagagt ttcgtcaatt tcgactacgg ctggttcgga 780  
 [0277] gcgcagacgg aagcggacgc cgacaagacc gtatggacc accgcaacca ctaccacgcg 840  
 [0278] cccaatggca gcctgggtgc catgcacgtg tacgagagca agttccgcaa ctggtccgac 900  
 [0279] ggttactcgg acttcgaccg cggagcctac gtggtcacgt tcgtcccaa gagctggaac 960  
 [0280] accgcccccg acaaggtgac acagggetgg ccgtga 996  
 [0281] <210> 27  
 [0282] <211> 410  
 [0283] <212> DNA  
 [0284] <213> S. ladakanum  
 [0285] <400> 27  
 [0286] Met Ser Gln Arg Gly Arg Thr Leu Val Phe Ala Ala Leu Gly Ala Val  
 [0287] 1 5 10 15  
 [0288] Met Cys Thr Thr Ala Leu Met Pro Ser Ala Gly Ala Ala Thr Gly Ser  
 [0289] 20 25 30  
 [0290] Gly Ser Gly Ser Gly Thr Gly Glu Glu Lys Arg Ser Tyr Ala Glu Thr  
 [0291] 35 40 45  
 [0292] His Arg Leu Thr Ala Asp Asp Val Asp Asp Ile Asn Ala Leu Asn Glu  
 [0293] 50 55 60

[0294] Ser Ala Pro Ala Ala Ser Ser Ala Gly Pro Ser Phe Arg Ala Pro Asp  
[0295] 65 70 75 80  
[0296] Ser Asp Glu Arg Val Thr Pro Pro Ala Glu Pro Leu Asp Arg Met Pro  
[0297] 85 90 95  
[0298] Asp Pro Tyr Arg Pro Ser Tyr Gly Arg Ala Glu Thr Ile Val Asn Asn  
[0299] 100 105 110  
[0300] Tyr Ile Arg Lys Trp Gln Gln Val Tyr Ser His Arg Asp Gly Arg Lys  
[0301] 115 120 125  
[0302] Gln Gln Met Thr Glu Glu Gln Arg Glu Trp Leu Ser Tyr Gly Cys Val  
[0303] 130 135 140  
[0304] Gly Val Thr Trp Val Asn Ser Gly Gln Tyr Pro Thr Asn Arg Leu Ala  
[0305] 145 150 155 160  
[0306] Phe Ala Phe Phe Asp Glu Asp Lys Tyr Lys Asn Glu Leu Lys Asn Gly  
[0307] 165 170 175  
[0308] Arg Pro Arg Ser Gly Glu Thr Arg Ala Glu Phe Glu Gly Arg Val Ala  
[0309] 180 185 190  
[0310] Lys Asp Ser Phe Asp Glu Ala Lys Gly Phe Gln Arg Ala Arg Asp Val  
[0311] 195 200 205  
[0312] Ala Ser Val Met Asn Lys Ala Leu Glu Asn Ala His Asp Glu Gly Ala  
[0313] 210 215 220  
[0314] Tyr Leu Asp Asn Leu Lys Lys Glu Leu Ala Asn Gly Asn Asp Ala Leu  
[0315] 225 230 235 240  
[0316] Arg Asn Glu Asp Ala Arg Ser Pro Phe Tyr Ser Ala Leu Arg Asn Thr  
[0317] 245 250 255  
[0318] Pro Ser Phe Lys Asp Arg Asn Gly Gly Asn His Asp Pro Ser Lys Met  
[0319] 260 265 270  
[0320] Lys Ala Val Ile Tyr Ser Lys His Phe Trp Ser Gly Gln Asp Arg Ser  
[0321] 275 280 285  
[0322] Gly Ser Ser Asp Lys Arg Lys Tyr Gly Asp Pro Glu Ala Phe Arg Pro  
[0323] 290 295 300  
[0324] Asp Arg Gly Thr Gly Leu Val Asp Met Ser Arg Asp Arg Asn Ile Pro  
[0325] 305 310 315 320  
[0326] Arg Ser Pro Thr Ser Pro Gly Glu Ser Phe Val Asn Phe Asp Tyr Gly  
[0327] 325 330 335  
[0328] Trp Phe Gly Ala Gln Thr Glu Ala Asp Ala Asp Lys Thr Val Trp Thr  
[0329] 340 345 350  
[0330] His Gly Asn His Tyr His Ala Pro Asn Gly Ser Leu Gly Ala Met His  
[0331] 355 360 365  
[0332] Val Tyr Glu Ser Lys Phe Arg Asn Trp Ser Asp Gly Tyr Ser Asp Phe  
[0333] 370 375 380  
[0334] Asp Arg Gly Ala Tyr Val Val Thr Phe Val Pro Lys Ser Trp Asn Thr  
[0335] 385 390 395 400

[0336] Ala Pro Asp Lys Val Thr Gln Gly Trp Pro  
 [0337] 405 410  
 [0338] <210> 28  
 [0339] <211> 1233  
 [0340] <212> DNA  
 [0341] <213> *S. ladakanum*  
 [0342] <400> 28  
 [0343] atgtccaac gcgggagaac tctcgtcttc gccgctctcg gtgcggtcat gtgcaccacc 60  
 [0344] gcgttaatgc cgtccgcagg cgcgccacc ggcagtgga gtggcagcgg caccgggaa 120  
 [0345] gagaagaggt cctacccga aacgcaccgc ctgacggcgg atgacgtcga cgacatcaac 180  
 [0346] gcgctgaac aaagcgtcc ggccgcttc agcgcggtc cgtccttcgg ggccccgac 240  
 [0347] tccgacgagc gggtagctcc tcccgcgag ccgctcgacc ggatgcccga cccgtaccgg 300  
 [0348] ccctcgtacg gcagggccga gacgatcgtc aacaactaca tacgcaagt gcagcagtc 360  
 [0349] tacagccacc gcgacggcag gaaacagcag atgaccgagg aacagcggga gtggctgtcc 420  
 [0350] tacggttcg tcggtgtcac ctgggtcaac tcgggccagt atccgacgaa caggctggct 480  
 [0351] ttcgcgttct tcgacagga caagtacaag aacgagctga agaacggcag gcccccgtcc 540  
 [0352] ggcgaaacgc gggcggagtt cgagggcgc gtcgccaagg acagcttcga cgaggcgaag 600  
 [0353] gggttccagc gggcgcgtga cgtggcgtcc gtcatgaaca aggccctgga gaacgccac 660  
 [0354] gacgaggggg cgtacctga caacctcaag aaggagctgg cgaacggcaa cgacgccctg 720  
 [0355] cggaacgagg atgcccgtc gcccttctac tcggcgtgc ggaacacgc gtccttcaag 780  
 [0356] gaccgcaacg gcgcaatca cgaccgtcc aagatgaagg ccgtcatcta ctggaagcac 840  
 [0357] ttctggagcg gccaggaccg gtcgggctcc tccgacaaga ggaagtacgg cgaccggag 900  
 [0358] gccttccgcc ccgaccgagg caccggctg gtcgacatgt cgaggacag gaacattccg 960  
 [0359] cgcagcccca ccagccccg cgagagttc gtcaatttc actacggctg gttcggagcg 1020  
 [0360] cagacggaag cggacgccga caagaccgta tggaccacg gcaaccaata ccacgcgcc 1080  
 [0361] aatggcagcc tgggtgcat gcacgtgtac gagagcaagt tccgcaactg gtccgacggt 1140  
 [0362] tactcggact tcgaccgagg agcctacgtg gtcacgttc tcccgaagag ctggaacacc 1200  
 [0363] gccccgaca aggtgacaca gggctggccg tga 1233  
 [0364] <210> 29  
 [0365] <211> 799  
 [0366] <212> DNA  
 [0367] <213> *B. subtilis*  
 [0368] <400> 29  
 [0369] gatcagacc tcatagaaat cacgctgaaa ttgatcctcc aaacgcgcgc cgataaaata 60  
 [0370] cgcttgccc tgctgatact catggctgtg gaccgctggc gtgcgcgcat aaaaatcttc 120  
 [0371] ttgatacacc gttccactg aagctgtctt tacatcaatc acggttgcat aatccttcat 180  
 [0372] ttcatatatt tggtcgggt agctgacagc gtttcgatcc ttcgataca ggggtgtccg 240  
 [0373] ttcaagagge tcaactcaa atatagctt aagatccgga tgccatccgc ctgtgtatgt 300  
 [0374] taagtcatgc tcattcaca ccccgctgat atacgtcatg actaaggtgc cgccgtcage 360  
 [0375] cgtaaacgct ttaaaccgg aaacgggtgc ctcgctgatt aaatacagca tcgggacgat 420  
 [0376] cagcagttta tatggtgaaa agtcttgttc tttcgtgatg acgtcgacag ggatatcgtg 480  
 [0377] ttcccagaat gtgcggtaat gctgctgaag cgtttgcgga taacgttttg tcgccttcgc 540



[0378] aaaccctga gcatcctcga gcgccaatg attttccag tcatataaaa tcgcggttg 600  
[0379] agccggcctc ttcgttccga caacttcgga cagecgttc aatgtctcgc ctaccttggc 660  
[0380] cacttcttga aagacgcgt tcttcgggct attgtcatga tccacaaccg ctccgtgtaa 720  
[0381] ttttctgat gacccccgtg atttgcgta ttgaaatag agaacgctgt ccgagccgtg 780  
[0382] ggcaatcatt tgcattgat 799  
[0383] <210> 30  
[0384] <211> 1734  
[0385] <212> DNA  
[0386] <213> Artificial sequence  
[0387] <400> 30  
[0388] atcgattttg ctgagggtgc agagggcagg tttttttgt tctttttct cgtaaaaaaa 60  
[0389] agaaaggtct taaagtttt atggttttg tcggcactga attcgagctc agcattattg 120  
[0390] agtggatgat tatattcctt ttgatagtg gtatgtttc gcttgaactt taaatacag 180  
[0391] ccattgaaca tacggtgat ttaataactg acaaacatca ccctcttgc aaagcggcca 240  
[0392] aggacgctgc cgccgggct gtttgcgtt ttgccgtgat ttcgtgtatc attggtttac 300  
[0393] ttatttttt gccaaagctg taatggetga aaattcttac atttattta catttttaga 360  
[0394] aatgggcgtg aaaaaagcg cgcgattatg taaaataaa agtgatagcg gtaccaggag 420  
[0395] ggctggaaga agcagaccgc taacacagta cataaaaaag gagacatgaa cgatgaacat 480  
[0396] caaaaagttt gcaaaaacaag caacagtatt aacctttact accgcactgc tggcaggagg 540  
[0397] cgcaactcaa gcttttgcct cgagctcggc acccggggcc accggcagtg gcagtggcag 600  
[0398] cggcaccggg gaagagaaga ggtcctacgc cgaaacgcac cgcctgacgg cggatgacgt 660  
[0399] cgacgacatc aacgcgctga acgaaagcgc tccggccgct tcgagcgcgg gtccgtcctt 720  
[0400] ccgggcccc gcagcagcag agcgggtgac tcctcccgc gagccgctc accggatgcc 780  
[0401] cgacccttac cggcctcgt acggcaggc cgagacgac gtcaacaact acatacgea 840  
[0402] gtggcagcag gtctacagcc accgcagcg caggaaacag cagatgaccg aggaacagcg 900  
[0403] ggagtggctg tcttacggt gcgtcgggt cacctgggtc aactcgggccc agtatccgac 960  
[0404] gaacaggctg gctttcgtt tcttcgacga ggacaagtac aagaacgagc tgaagaacgg 1020  
[0405] caggccccgg tccggcgaaa cgcgggcgga gttcgagggg cgcgtcgcca aggacagctt 1080  
[0406] cgacgaggcg aaggggttcc agcgggcgcg tgacgtggcg tccgtcatga acaaggcctt 1140  
[0407] ggagaacgcc cacgacgagg gggcgtacct cgacaacctc aagaaggagc tggcgaacgg 1200  
[0408] caacgacgcc ctgcggaacg aggatgccg ctcgccctc tactcggcgc tgcggaacac 1260  
[0409] gccgtcctc aaggaccgca acggcggcaa tcacgaccg tccaagatga aggccgtcat 1320  
[0410] ctactcgaag cacttctgga gcggccagga ccggtcgggc tctccgaca agaggaagta 1380  
[0411] cggcgaccg gaggccttc gccccgaccg cggcaccgac ctggtcgaca tgcgaggga 1440  
[0412] caggaacatt ccgagcagcc ccaccagccc cggcgagagt ttcgtcaatt tcgactacgg 1500  
[0413] ctggttcgga gcgagcagc aagcggacgc cgacaagacc gtatggacc accgcaacca 1560  
[0414] ctaccacgcg ccaatggca gcctgggtgc catgcacgtg tacgagagca agttccgca 1620  
[0415] ctggtccgac ggttactcgg acttcgaccg cggagcctac gtggtcactg tcgtcccaaa 1680  
[0416] gagctggaac accgccccg acaaggtgac acagggtg cgcctcgatt aacc 1734  
[0417] <210> 31  
[0418] <211> 800  
[0419] <212> DNA

[0420] <213> B. subtilis  
[0421] <400> 31  
[0422] acagcagatt catgcccggg cgctttgct tgttgacgtt atgccaattg accgcgcttg 60  
[0423] gcgtacactc cattaataag aagggtgct gcttcaagct tcggtacaga tcattgataa 120  
[0424] agccgacctt catgccc aaa tcagctgtgc tttcccagtc attgtgccag acaggataag 180  
[0425] cgtcccagct gatggcatcg acatgctttg caaatttctg gtagtcgagg ccctgatacg 240  
[0426] ggatcaaatc cgggtgtgtca gccataaaat tcgttgtgat agggatatca ggcgtcaatt 300  
[0427] ctttcagcgg aatgatttca ttttcataaa acgaaatcgt ttgatcgggtg acgaaccggc 360  
[0428] gccaatctaa attcaggcca tgcaagccat tttcaccgat cggcgaaggg ctttcaattt 420  
[0429] gtgaccagtc attgaacgta tggctccaaa aagggtcca ccacgcatgg ttcaatgtct 480  
[0430] tgaggctgtt gtcatatttc gatttcagcc actcccggaa agcatgctgg cataaatcac 540  
[0431] agtggcaatc tccccgat tcgtttgaaa tgtgccacat taacagcgcg gggatgatgtc 600  
[0432] cgtatcgttc tgctaataag cggttgatgt gccgtgtttt ttctcggtag acttttagatg 660  
[0433] tgaggcagtg gttgtgcctt ccgccgtgca gctgtttgac gcgggaggca ttgacgcgca 720  
[0434] aaacttccgg ataggtttgc gacagccagg ccggacgggc tccgctcggc gttgctaata 780  
[0435] tgaccggcc gcctatactg 800

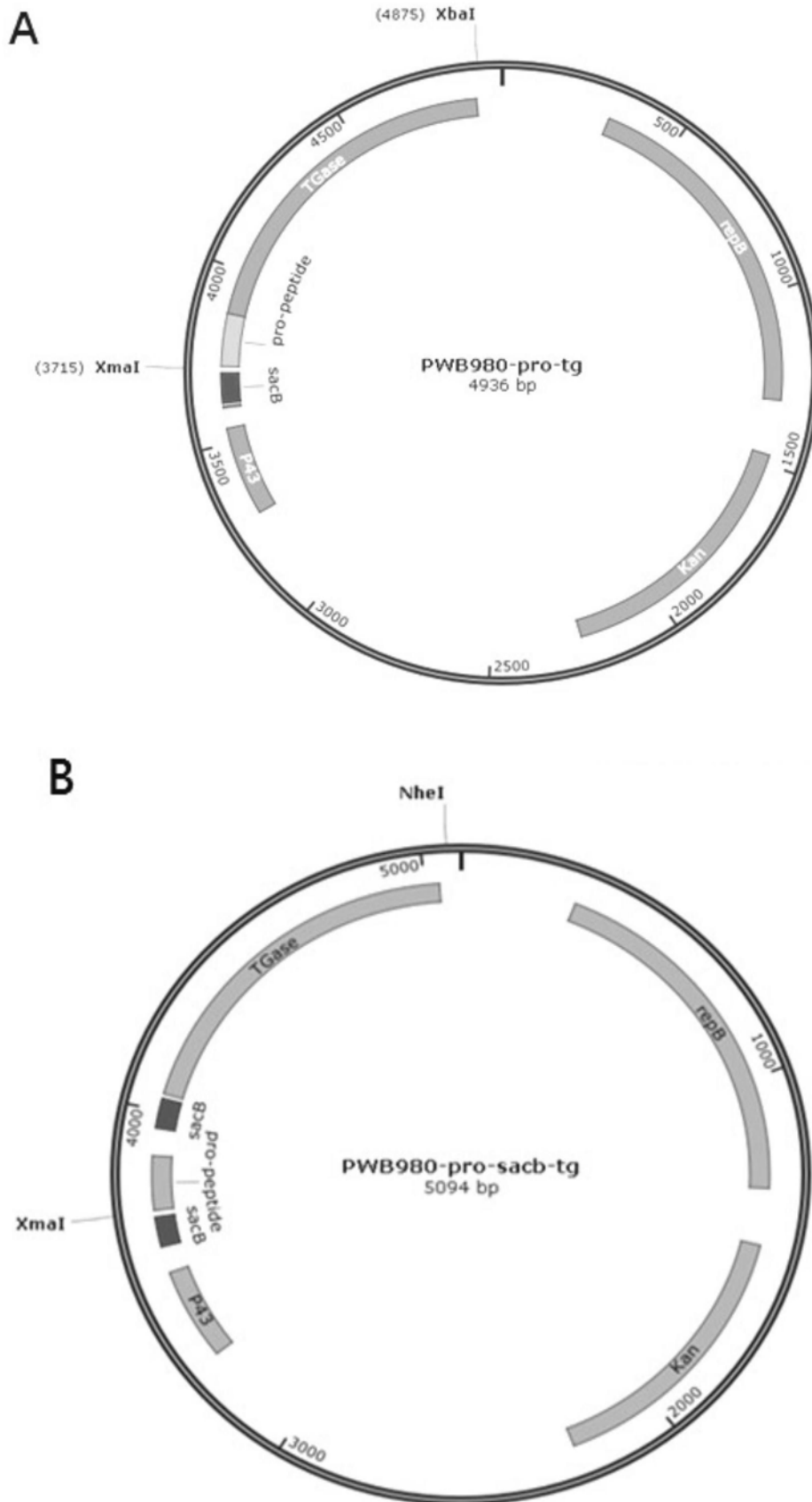


图1

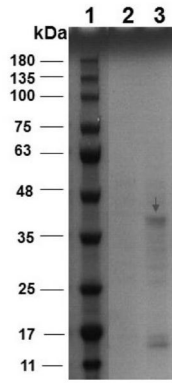


图2

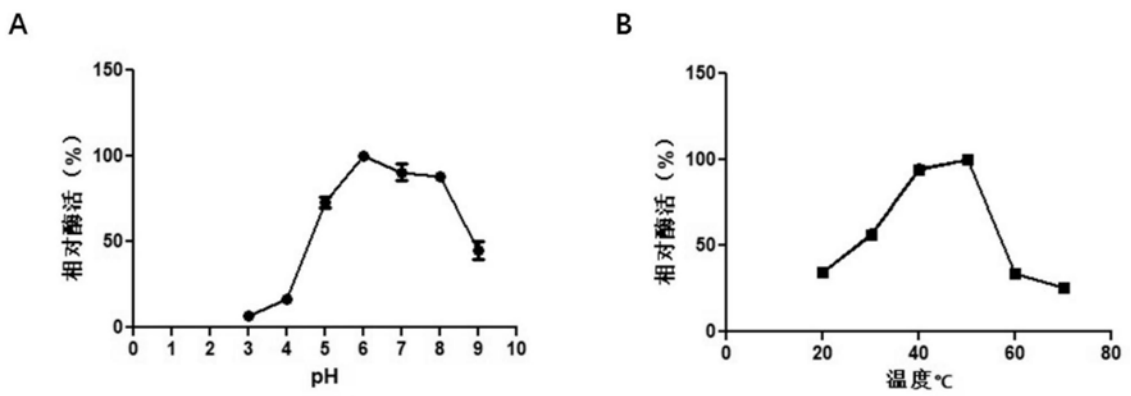


图3

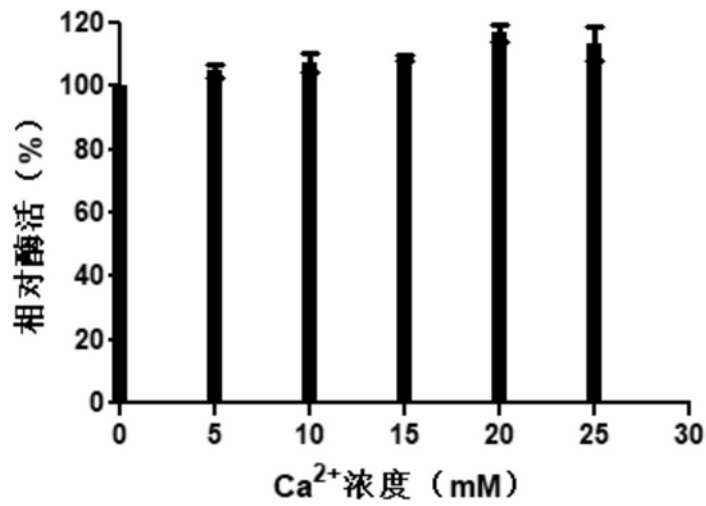


图4