



(12) 发明专利



(10) 授权公告号 CN 107109352 B

(45) 授权公告日 2022.03.04

(21) 申请号 201580070585.7

C12N 9/38 (2006.01)

(22) 申请日 2015.11.06

A23C 9/13 (2006.01)

(65) 同一申请的已公布的文献号

A23C 9/152 (2006.01)

申请公布号 CN 107109352 A

C12R 1/125 (2006.01)

C12R 1/01 (2006.01)

(43) 申请公布日 2017.08.29

(30) 优先权数据

(56) 对比文件

1419894.9 2014.11.07 GB

WO 2013182686 A1, 2013.12.12

1419897.2 2014.11.07 GB

WO 2009071539 A1, 2009.06.11

1419900.4 2014.11.07 GB

KARINA PAKUSAEVA等. Carbohydrate metabolism in Bifidobacteria.《GENES& NUTRITION》.2011,第6卷(第3期),

1515645.8 2015.09.03 GB

KARINA PAKUSAEVA等. Carbohydrate metabolism in Bifidobacteria.《GENES& NUTRITION》.2011,第6卷(第3期),

(85) PCT国际申请进入国家阶段日

2017.06.23

(86) PCT国际申请的申请数据

PCT/EP2015/075950 2015.11.06

BOLOTIN A.等. THE COMPLETE GENOME SEQUENCE OF THE LACTIC ACID BACTERIUM LACTOCOCCUS LACTIS SSP. LACTIS IL1403.《GENOME RESEARCH》.2001,第11卷(第5期),

(87) PCT国际申请的公布数据

W02016/071504 EN 2016.05.12

M.N. HUNG等. Purification and characterization of a recombinant β -galactosidase with transgalactosylation activity from BBifidobacterium infantis HL96.《APPLIED MICROBIOLOGY AND BIOTECHNOLOGY》.2002,第58卷(第4期),

(73) 专利权人 杜邦营养生物科学有限公司

地址 丹麦哥本哈根

(72) 发明人 M·K·拉尔森 J·F·克拉米尔

J·拉巴戈 T·艾泽勒

K·汉森谢尔

(74) 专利代理机构 北京市中咨律师事务所

11247

审查员 习洋

代理人 柴云峰 黄革生

(51) Int.Cl.

C12N 1/20 (2006.01)

权利要求书2页 说明书61页

序列表44页 附图3页

(54) 发明名称

甘露聚糖酶、纤维素酶和果胶酶缺陷型的表达 β -半乳糖苷酶和/或转半乳糖基化活性的重组宿主细胞

(57) 摘要

披露了一种宿主细胞,该宿主细胞能够表达具有 β -半乳糖苷酶活性和/或转半乳糖基化活性的多肽,并且该宿主细胞是被修饰为纤维素酶、甘露聚糖酶和果胶酶缺陷型。

1. 一种能够表达编码具有 β -半乳糖苷酶活性和/或转半乳糖基化活性的多肽的异源核酸的重组宿主细胞,其中所述宿主细胞相对于亲本宿主细胞被修饰为具有敲除失活的纤维素酶、甘露聚糖酶、果胶酶和淀粉酶,且其中所述宿主细胞是枯草芽孢杆菌细胞;

其中所述宿主细胞包含含有编码该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽的核酸的表达载体,或者所述宿主细胞用编码该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽的核酸转化,

其中该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽选自由以下各项组成的组:

a. 来自两歧双歧杆菌且与SEQ ID NO:1具有至少99%序列同一性的氨基酸序列的多肽,

b. 来自两歧双歧杆菌且与SEQ ID NO:2具有至少99%序列同一性的氨基酸序列的多肽,

c. 来自两歧双歧杆菌且与SEQ ID NO:3具有至少99%序列同一性的氨基酸序列的多肽,

d. SEQ ID NO:1、2、3、4或5的多肽;

且其中所述多肽包含SEQ ID NO:7的催化核心序列。

2. 根据权利要求1所述的宿主细胞,其中通过遗传操作修饰该宿主细胞。

3. 根据权利要求1所述的宿主细胞,其中该纤维素酶、甘露聚糖酶和果胶酶多肽以及该淀粉酶多肽就酶活性而言在功能上是无活性的。

4. 根据权利要求1所述的宿主细胞,其中通过遗传操作使得该具有纤维素酶、甘露聚糖酶和果胶酶活性的多肽以及该具有淀粉酶活性的多肽无活性。

5. 根据权利要求2或4所述的宿主细胞,其中该遗传操作是一步基因破坏法、标志物插入、定点诱变、缺失、RNA干扰或反义RNA。

6. 根据权利要求1所述的宿主细胞,其中该宿主细胞包含一种表达载体,该表达载体包含编码该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽的核酸。

7. 根据权利要求1所述的宿主细胞,其中用一种核酸转化该宿主细胞,该核酸编码该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽。

8. 根据权利要求1所述的宿主细胞,其中所述失活的甘露聚糖酶基因是gmuG甘露聚糖酶。

9. 根据权利要求1所述的宿主细胞,其中所述失活的纤维素酶基因是bgLC内切葡聚糖酶。

10. 根据权利要求1所述的宿主细胞,其中所述失活的果胶酶基因是pe1果胶酸裂解酶。

11. 根据权利要求1所述的宿主细胞,其中所述宿主细胞具有失活的淀粉酶基因,其中所述基因是amyE α -淀粉酶。

12. 一种生产具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽、或包含所述多肽的组合物的方法,该方法包括在培养基中,在适合表达该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽的条件下,培养如权利要求1至11中任一项所述的宿主细胞,并且任选地从该培养基或该宿主细胞中回收该具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽。

13. 根据权利要求12的方法,其中从所述培养基或所述宿主细胞获得所述具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽或获得包含所述多肽的组合物,其中所述组合物包

含具有所述 β -半乳糖苷酶和/或转半乳糖基化活性的多肽并且具有降低量的所述纤维素酶、甘露聚糖酶和果胶酶的酶副活性。

14. 根据权利要求12或权利要求13所述的方法, 其中所述具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽、或包含所述多肽的组合物没有纤维素酶、甘露聚糖酶和果胶酶活性。

15. 一种生产奶产品的方法, 该方法包括: 通过权利要求12或13的方法提供所述具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽或包含所述多肽的组合物; 向含有乳糖的奶产品中添加所述的具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽或所述的组合物。

16. 一种生产包含GOS的奶产品的方法, 该方法包括: 通过权利要求12或13的方法提供所述具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽或包含所述多肽的组合物; 向含有乳糖的奶产品中添加所述的具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽或所述的组合物。

17. 根据权利要求15或权利要求16所述的方法, 其中与使用由表达纤维素酶、甘露聚糖酶和果胶酶的宿主细胞制备的、具有 β -半乳糖苷酶和/或转半乳糖基化活性的多肽或包含所述多肽的组合物相比, 所述方法用于防止奶产品的粘度和/或质感降低。

甘露聚糖酶、纤维素酶和果胶酶缺陷型的表达 β -半乳糖苷酶和/或转半乳糖基化活性的重组宿主细胞

技术领域

[0001] 本发明涉及一种其中不希望的酶失活的宿主细胞,并且涉及该宿主细胞在酶组合物的制备中的用途,该酶组合物用于制造具有降低的不想要的特性的食品。

背景技术

[0002] 普遍使用酶改进食品的化学性质。另外,在牛乳和其他动物衍生底物的加工中,酶的使用显著增加了最终产品的价值。

[0003] 半乳寡糖(GOS)是在人和动物中不易消化的碳水化合物,其包括通过糖苷键连接的两个或更多个(通常多达九个)半乳糖分子。GOS还可以包括一个或多个葡萄糖分子。GOS的有益作用之一是它们作为益生元化合物通过选择性地刺激有益的结肠微生物(如细菌)以向消费者给出生理益处的能力。已经形成的健康作用导致对作为各种类型食物的食物成分的GOS的兴趣日益增加。

[0004] 酶 β -半乳糖苷酶(EC 3.2.1.23)通常将乳糖水解成单糖D-葡萄糖和D-半乳糖。在 β -半乳糖苷酶的正常酶反应中,酶水解乳糖并瞬时结合半乳糖单糖在半乳糖-酶复合物中,该复合物将半乳糖转移到水的羟基基团上,导致D-半乳糖和D-葡萄糖的释放。然而,在高乳糖浓度下,在称为转半乳糖基化的过程中,一些 β -半乳糖苷酶能够将半乳糖转移到D-半乳糖或D-葡萄糖的羟基中,由此产生半乳寡糖。同样,在高乳糖浓度下,一些 β -半乳糖苷酶能够将半乳糖转移到乳糖或更高级寡糖的羟基基团上。

[0005] 双歧杆菌是奶业中最常用类型的细菌培养物之一,用于发酵各种奶产品。另外,摄入含双歧杆菌的产品具有促进健康的作用。这种作用不仅通过降低肠内容物的pH值而且通过双歧杆菌在个体中重新填充肠道菌群的能力来实现,这些个体的肠道菌群受到例如抗生素的摄入的干扰。双歧杆菌还具有战胜潜在有害肠道微生物的潜力。

[0006] 已知半乳寡糖可增强双歧杆菌的生长。这种作用可能通过双歧杆菌利用半乳寡糖作为碳源的独特能力来实现。还考虑了半乳寡糖的膳食补充剂具有许多长期的疾病保护作用。例如,已经显示出半乳寡糖摄入对大鼠结肠直肠癌发展具有高度保护作用。因此,有兴趣开发用于生产在工业中使用以改进膳食补充剂和奶产品的半乳寡糖的廉价且有效的方法。

[0007] 已经将来自两歧双歧杆菌DSM20215的具有约580个氨基酸的截短的细胞外乳糖酶(BIF3-d3)描述为在包含溶解在水中的乳糖的溶液中的转糖半乳糖酶(乔根森(Jørgensen)等人,(2001),应用微生物学和生物技术(Appl.Microbiol.Biotechnol.), 57:647-652)。W0 01/90317还描述了作为转半乳糖基化酶的截短变体(OLGA347),并且在W0 2012/010597中,显示OLGA347将半乳糖部分转移到D-岩藻糖、N-乙酰基-半乳糖胺和木糖。

[0008] 在W0 2009/071539中,与BIF3-d3相比,不同的截短片段被描述为在乳中测试时导致GOS的高效水解和非常低的产量。

[0009] W0 2013/182686描述了一种多肽,该多肽具有反式半乳糖基化与水解活性的有用

比率,并且因此即使在低乳糖水平如乳基产品中与乳糖一起孵育时也是GOS的有效生产者。
[0010] 然而,仍然需要提供改进的方法,用于在奶产品中现场从乳糖生产处于GOS形式的膳食纤维。本发明解决了这一需要。

发明内容

[0011] 酶的副活性通常用于描述不想要的酶,其活性对方法或产物产生负面影响。它们主要存在于作为例如生产宿主的副产物的酶产物中,并且可能阻碍这些酶的商业应用。奶应用对少量的纤维素酶、果胶酶、淀粉酶和甘露聚糖酶特别敏感,因为许多奶产品用水状胶体例如CMC、GUAR、淀粉和果胶进行配制/稳定(参见表1)。

[0012] 表1

[0013]	酶	可降解的水状胶体/稳定剂
	纤维素酶	CMC
	甘露聚糖酶	GUAR、LBG
	果胶酶(果胶酸裂解酶)	果胶
	淀粉酶	淀粉

[0014] 我们现在已经能够开发包含 β -半乳糖苷酶活性,任选地具有转半乳糖基化活性的产物或由其产生的GOS,其中纤维素、甘露聚糖酶、果胶酶和淀粉酶的水平得以降低。作为增稠剂、凝胶剂和稳定剂,水状胶体对许多食品和饮料产品作出了重要贡献。本发明允许提供这样的食品和饮料产品,其中不仅保留由水状胶体的使用产生的所期望的特性,而且其中另外地,食品和饮料产品能够保持处于GOS形式的膳食纤维。

[0015] 水状胶体或树胶是不同类的长链聚合物,其特征在于当分散在水中时其形成粘性分散体和/或凝胶的特性。这些材料首先在树木或灌木丛的渗出物、植物或海藻的提取物、种子或谷物的粉状物质、发酵过程中的粘液和许多其他天然产物中发现。大量羟基基团的出现明显增加了它们结合水分子的亲和力,使它们成为亲水性化合物。此外,它们产生一种分散体,该分散体是在真溶液和悬浮液之间的中间体,并表现出胶体的特性。考虑到这两种特性,它们被适当地称为‘亲水胶体’或‘水状胶体’。

[0016] 水状胶体在食品中具有广泛的功能特性,这些功能特性包括:增稠、胶凝、乳化、稳定、涂层等。当以如下水平使用时,水状胶体对食品特性具有深远的影响,这些水平范围从在加热处理奶产品中角叉菜胶的百万分之几至胶冻质糖食中的阿拉伯树胶、淀粉或明胶的高水平。水状胶体在食品中充分利用的主要原因是它们能够改变食品系统的流变学。这包括食品系统的两个基本特性,即流动行为(粘度)和机械固体特性(质感)。食品系统的质感和/或粘度的改变有助于改变其感官特性,因此将水状胶体用作重要的食品添加剂来执行特定目的。很显然,若干种水状胶体属于世界上许多国家中允许的食品添加剂类别。各种食物配制品如汤、肉汁、沙拉酱、调味汁和浇头使用水状胶体作为添加剂以达到优选的粘度和口感。它们也被用于许多食品像冰淇淋、果酱、果冻、凝胶甜点、蛋糕和糖果中,以产生所期望的质感。

[0017] 除了功能属性之外,未来的接受以及可能的积极的认可可能来自于认识到纤维对身体的自然功能和健康有许多生理益处。

[0018] 由于其水结合特性,水状胶体对食品的质感和口感有重大影响-通常为质感创新

创造机会。这些产品中的若干种还与蛋白质相互作用,是用于蛋白质稳定和保护的有用特性。本发明提供了基于使用水状胶体的溶液,其中不仅保留使用水状胶体上述所期望的特性,而且其中另外地,食品和饮料产品能够包含处于GOS形式的膳食纤维。

[0019] 在所培养的产品中,水状胶体提供细腻的质感和光泽的外观。水状胶体可以在降低的乳固体配制品中优化成本,并且在整个保质期中保持质感。水状胶体还改进所培养的产品本体,尤其是处于较高消耗温度下。本发明提供了基于使用水状胶体的溶液,其中不仅保留使用水状胶体上述所期望的特性,而且其中另外地,食品和饮料产品能够包含处于GOS形式的膳食纤维。

[0020] 在奶产品和低pH蛋白饮料中,水状胶体可以稳定乳和大豆蛋白质,防止沉降和乳清脱落,并可以实现各种质感。此外,更多的水状胶体可以用减少的乳固体、糖和/或脂肪替代配制品中的质感,并且从而给食品制造提供更好的低脂肪产品。本发明提供了基于使用水状胶体的溶液,其中不仅保留使用水状胶体上述所期望的特性,而且其中另外地,食品和饮料产品能够包含处于GOS形式的膳食纤维。

[0021] 本发明通过提供包含具有转半乳糖基化活性的多肽的酶组合物来解决这些问题,但是该酶组合物没有或基本上没有归因于以下酶的活性:纤维素、甘露聚糖酶、果胶酶和任选地淀粉酶。

[0022] 根据本发明的第1方面,提供一种能够表达具有 β -半乳糖苷酶活性的多肽的宿主细胞,并且将该所述宿主细胞修饰为纤维素酶、甘露聚糖酶和果胶酶缺陷型。

[0023] 在一个优选的实施例中,本发明中使用的具有 β -半乳糖苷酶活性的多肽具有转半乳糖基化活性。

[0024] 根据本发明的第2方面,提供一种能够表达具有转半乳糖基化活性的多肽的宿主细胞,并且将该所述宿主细胞修饰为纤维素酶、甘露聚糖酶和果胶酶缺陷型。

[0025] 优选地,该宿主细胞还被修饰为淀粉酶缺陷型。

[0026] 在一个实施例中,通过常规诱变技术修饰该宿主细胞。

[0027] 在另一个实施例中,通过常规遗传操作技术修饰该宿主细胞。

[0028] 根据本发明的第3方面,提供一种能够表达具有 β -半乳糖苷酶活性的多肽的宿主细胞,并且其中具有纤维素酶、甘露聚糖酶和果胶酶活性的多肽基本上是无活性的。

[0029] 根据本发明的第4方面,提供一种能够表达具有转半乳糖基化活性的多肽的宿主细胞,并且其中具有纤维素酶、甘露聚糖酶和果胶酶活性的多肽基本上是无活性的。

[0030] 另外,优选地,在宿主细胞中,具有淀粉酶活性的多肽基本上是无活性的。

[0031] 在第三和第四方面的一个实施例中,关于酶活性,该基本上无活性的纤维素酶、甘露聚糖酶和果胶酶多肽以及任选地该淀粉酶多肽在功能上是无活性的。

[0032] 在一个实施例中,通过常规诱变技术使得该具有纤维素酶、甘露聚糖酶和果胶酶活性的多肽以及任选地该具有淀粉酶活性的多肽基本上无活性。

[0033] 在另一个实施例中,通过常规遗传操作技术使得该具有纤维素酶、甘露聚糖酶和果胶酶活性的多肽以及任选地该具有淀粉酶活性的多肽基本上无活性。

[0034] 可以使用的常规诱变技术是化学或物理诱变。

[0035] 可以使用的常规遗传操作技术是一步基因破坏法、标志物插入、定点诱变、缺失、RNA干扰或反义RNA。

[0036] 优选地,该宿主细胞是一种细菌。

[0037] 该宿主细胞可以是乳酸菌。

[0038] 优选地,该宿主细胞是枯草芽孢杆菌。

[0039] 优选地,本发明中使用的具有转半乳糖基化活性的多肽选自由以下各项组成的组:

[0040] a.包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列的多肽,其中所述多肽由至多980个氨基酸残基组成,

[0041] b.包含与SEQ ID NO:2具有至少97%序列同一性的氨基酸序列的多肽,其中所述多肽由至多975个氨基酸残基组成,

[0042] c.包含与SEQ ID NO:3具有至少96.5%序列同一性的氨基酸序列的多肽,其中所述多肽由至多1300个氨基酸残基组成,

[0043] d.由多核苷酸编码的多肽,该多核苷酸在至少低严格条件下与以下各项杂交:i)包含在SEQ ID NO:9、10、11、12或13中的编码SEQ ID NO:1、2、3、4或5的多肽的核酸序列;或ii) i)的互补链,

[0044] e.由包含如下各项的多核苷酸编码的多肽:与编码SEQ ID NO:1、2、3、4或5的多肽的核苷酸序列具有至少70%同一性的核苷酸序列,或包含在SEQ ID NO:9、10、11、12或13中的编码成熟多肽的核苷酸序列,和

[0045] f.包括SEQ ID NO:1、2、3、4或5的一个或多个氨基酸残基的缺失、插入和/或保守取代的多肽。

[0046] 在一个实施例中,该宿主细胞包含一种表达载体,该表达载体包含编码具有 β -半乳糖苷酶活性的多肽的核酸。

[0047] 在一个实施例中,用编码具有 β -半乳糖苷酶活性的多肽的核酸转化该宿主细胞。

[0048] 在一个实施例中,该宿主细胞包含一种表达载体,该表达载体包含编码具有转半乳糖基化活性的多肽的核酸。

[0049] 在一个实施例中,用编码具有转半乳糖基化活性的多肽的核酸转化该宿主细胞。

[0050] 根据本发明的第5方面,提供了一种用于提供多肽组合物的方法,该多肽组合物包含具有 β -半乳糖苷酶活性并且具有降低量的所不希望的纤维素、甘露聚糖酶和果胶酶酶副活性的多肽,该方法包括:提供能够表达该具有 β -半乳糖苷酶活性的多肽、具有纤维素酶活性的多肽、具有甘露聚糖酶活性的多肽和具有果胶酶活性的多肽的宿主细胞;并且使所述纤维素、甘露聚糖酶和果胶酶失活。

[0051] 根据本发明的第6方面,提供了一种用于提供多肽组合物的方法,该多肽组合物包含具有转半乳糖基化活性并且具有降低量的所不希望的纤维素、甘露聚糖酶和果胶酶酶副活性的多肽,该方法包括:提供能够表达该具有转半乳糖基化活性的多肽、具有纤维素酶活性的多肽、具有甘露聚糖酶活性的多肽和具有果胶酶活性的多肽的宿主细胞;并且使所述纤维素、甘露聚糖酶和果胶酶失活。

[0052] 优选地,该方法还包括使具有淀粉酶活性的多肽失活。

[0053] 在本方法的一个实施例中,通过常规诱变技术使得该具有纤维素酶、甘露聚糖酶和果胶酶活性的多肽以及任选地该具有淀粉酶活性的多肽基本上无活性。

[0054] 在本方法的另一个实施例中,通过遗传操作使得该具有纤维素酶、甘露聚糖酶和

果胶酶活性的多肽以及任选地该具有淀粉酶活性的多肽基本上无活性。

[0055] 优选地,在该方法中,该宿主细胞是一种细菌。

[0056] 在该方法的一个实施例中,该宿主细胞是乳酸菌。

[0057] 优选地,在该方法中,该宿主细胞是枯草芽孢杆菌。

[0058] 优选地,在该方法中,具有转半乳糖基化活性的多肽选自由以下各项组成的组:

[0059] a. 包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列的多肽,其中所述多肽由至多980个氨基酸残基组成,

[0060] b. 包含与SEQ ID NO:2具有至少97%序列同一性的氨基酸序列的多肽,其中所述多肽由至多975个氨基酸残基组成,

[0061] c. 包含与SEQ ID NO:3具有至少96.5%序列同一性的氨基酸序列的多肽,其中所述多肽由至多1300个氨基酸残基组成,

[0062] d. 由多核苷酸编码的多肽,该多核苷酸在至少低严格条件下与以下各项杂交:i) 包含在SEQ ID NO:9、10、11、12或13中的编码SEQ ID NO:1、2、3、4或5的多肽的核酸序列;或 ii) i) 的互补链,

[0063] e. 由包含如下各项的多核苷酸编码的多肽:与编码SEQ ID NO:1、2、3、4或5的多肽的核苷酸序列具有至少70%同一性的核苷酸序列,或包含在SEQ ID NO:9、10、11、12或13中的编码成熟多肽的核苷酸序列,和

[0064] f. 包括SEQ ID NO:1、2、3、4或5的一个或多个氨基酸残基的缺失、插入和/或保守取代的多肽。

[0065] 在本方法的一个实施例中,该宿主细胞包含一种表达载体,该表达载体包含编码具有 β -半乳糖苷酶活性的多肽的核酸。

[0066] 在本方法的另一个实施例中,用编码具有 β -半乳糖苷酶活性的多肽的核酸转化该宿主细胞。

[0067] 在本方法的一个实施例中,该宿主细胞包含一种表达载体,该表达载体包含编码具有转半乳糖基化活性的多肽的核酸。

[0068] 在本方法的另一个实施例中,用编码具有转半乳糖基化活性的多肽的核酸转化该宿主细胞。

[0069] 根据本发明的第7方面,提供了生产具有 β -半乳糖苷酶活性的多肽的方法,该方法包括在培养基中,在适合表达具有 β -半乳糖苷酶活性的多肽的条件下培养本发明的宿主细胞,并且任选地从该培养基或该宿主细胞回收具有 β -半乳糖苷酶活性的多肽。

[0070] 根据本发明的第8方面,提供了生产具有转半乳糖基化活性的多肽的方法,该方法包括在培养基中,在适合表达具有转半乳糖基化活性的多肽的条件下培养本发明的宿主细胞,并且任选地从该培养基或该宿主细胞回收具有转半乳糖基化活性的多肽。

[0071] 根据本发明的第9方面,提供了使用本发明宿主细胞或本发明方法产生的具有 β -半乳糖苷酶活性的多肽。

[0072] 优选地,根据本发明的具有 β -半乳糖苷酶活性的多肽没有或基本上没有纤维素、甘露聚糖酶和果胶酶活性,以及任选地淀粉酶活性。

[0073] 在一个实施例中,根据本发明的具有 β -半乳糖苷酶活性的多肽没有或基本上没有纤维素、甘露聚糖酶和果胶酶活性,这样使得以粘度相对降低表示的粘度降低是至少0.85

或至少0.9、或更少,该粘度降低由与没有添加酶或水的水状胶体溶液相比添加具有 β -半乳糖苷酶活性的多肽的含水状胶体溶液的粘度计算得出。

[0074] 根据本发明的第10方面,提供了包括使用本发明宿主细胞或本发明方法产生的具有 β -半乳糖苷酶活性的多肽的多肽组合物。

[0075] 根据本发明的第11方面,提供了使用本发明宿主细胞或本发明方法产生的具有转半乳糖基化活性的多肽。

[0076] 优选地,根据本发明的具有转半乳糖基化活性的多肽没有或基本上没有纤维素、甘露聚糖酶和果胶酶活性,以及任选地淀粉酶活性。

[0077] 在一个实施例中,根据本发明的具有转半乳糖基化活性的多肽没有或基本上没有纤维素、甘露聚糖酶和果胶酶活性,这样使得以粘度相对降低表示的粘度降低是至少0.85或至少0.9、或更少,该粘度降低由与没有添加酶或水的水状胶体溶液相比添加具有转半乳糖基化活性的多肽的含水状胶体溶液的粘度计算得出。

[0078] 根据本发明的第12方面,提供了包括使用本发明宿主细胞或本发明方法产生的具有转半乳糖基化活性的多肽的多肽组合物。

[0079] 优选地,本发明的多肽组合物没有或基本上没有纤维素、甘露聚糖酶和果胶酶活性,以及任选地淀粉酶活性。

[0080] 在一个实施例中,多肽组合物没有或基本上没有纤维素、甘露聚糖酶和果胶酶活性,这样使得以粘度相对降低表示的粘度降低是至少0.85或至少0.9、或更少,该粘度降低由与没有添加酶或水的水状胶体溶液相比添加本发明的多肽组合物的含水状胶体溶液的粘度计算得出。

[0081] 根据本发明的第13方面,提供了包括本发明具有 β -半乳糖苷酶活性的多肽或本发明的多肽组合物的奶产品。

[0082] 根据本发明的第14方面,提供了生产奶产品的方法,该方法包括向含有乳糖的奶产品中添加本发明具有 β -半乳糖苷酶活性的多肽或本发明的多肽组合物。

[0083] 根据本发明的第15方面,提供了生产包含GOS的奶产品的方法,该方法包括向含有乳糖的奶产品中添加本发明具有 β -半乳糖苷酶活性的多肽或本发明的多肽组合物。

[0084] 根据本发明的第16方面,提供了根据本发明的具有转半乳糖基化活性的多肽或根据本发明的多肽组合物用于制备奶产品的用途。

[0085] 根据本发明的第17方面,提供了根据本发明的具有转半乳糖基化活性的多肽或根据本发明的多肽组合物用于制备包含GOS的奶产品的用途。

[0086] 根据本发明的第18方面,提供了包含本发明的具有转半乳糖基化活性的多肽或本发明的多肽组合物的奶产品。

[0087] 根据本发明的第19方面,提供了生产奶产品的方法,该方法包括向含有乳糖的奶产品中添加本发明的具有转半乳糖基化活性的多肽或本发明的多肽组合物。

[0088] 根据本发明的第20方面,提供了生产包含GOS的奶产品的方法,该方法包括向含有乳糖的奶产品中添加本发明的具有转半乳糖基化活性的多肽或本发明的多肽组合物。

[0089] 根据本发明的第21方面,提供了根据本发明的具有转半乳糖基化活性的多肽或根据本发明的多肽组合物用于制备奶产品的用途。

[0090] 根据本发明的第22方面,提供了根据本发明的具有转半乳糖基化活性的多肽或根

据本发明的多肽组合物用于制备包含GOS的奶产品的用途。

[0091] 在一个实施例中,本发明的用途是防止奶产品的粘度和/或质感降低,该用途是与具有 β -半乳糖苷酶和/或转半乳糖基化活性的、由表达纤维素酶、甘露聚糖酶和果胶酶的宿主细胞制备的多肽的用途相比而言的。

[0092] 本发明具体可用于在包含乳糖的组合物如奶产品中现场生产GOS膳食纤维。

[0093] 序列表

[0094] SEQ ID NO:1 (在此又称为(BIF_917))是SEQ ID NO:22的887个氨基酸的截短的片段。

[0095] SEQ ID NO:2 (在此又称为(BIF_995))是SEQ ID NO:22的965个氨基酸的截短的片段。

[0096] SEQ ID NO:3 (在此又称为(BIF_1068))是SEQ ID NO:22的1038个氨基酸的截短的片段。

[0097] SEQ ID NO:4 (在此又称为(BIF_1172))是SEQ ID NO:22的1142个氨基酸的截短的片段。

[0098] SEQ ID NO:5 (在此又称为(BIF_1241))是SEQ ID NO:22的1211个氨基酸的截短的片段。

[0099] SEQ ID NO:6 (在此又称为(BIF_1326))是SEQ ID NO:22的1296个氨基酸的截短的片段。

[0100] SEQ ID NO:7是两歧双歧杆菌糖苷水解酶催化核心

[0101] SEQ ID NO:8是编码来自两歧双歧杆菌DSM20215的胞外乳糖酶的核苷酸序列

[0102] SEQ ID NO:9是编码BIF_917的核苷酸序列

[0103] SEQ ID NO:10是编码BIF_995的核苷酸序列

[0104] SEQ ID NO:11是编码BIF_1068的核苷酸序列

[0105] SEQ ID NO:12是编码BIF_1172的核苷酸序列

[0106] SEQ ID NO:13是编码BIF_1241的核苷酸序列

[0107] SEQ ID NO:14是编码BIF_1326的核苷酸序列

[0108] SEQ ID NO:15是用于生成上述BIF变体的正向引物

[0109] SEQ ID NO:16是BIF917的反向引物

[0110] SEQ ID NO:17是BIF995的反向引物

[0111] SEQ ID NO:18是BIF1068的反向引物

[0112] SEQ ID NO:19是BIF1241的反向引物

[0113] SEQ ID NO:20是BIF1326的反向引物

[0114] SEQ ID NO:21是BIF1478的反向引物

[0115] SEQ ID NO:22是来自两歧双歧杆菌DSM20215的胞外乳糖酶。

[0116] SEQ ID NO:23是来自两歧双歧杆菌DSM20215的胞外乳糖酶的信号序列。

[0117] 发明详细披露

[0118] 定义

[0119] 根据此详细描述,适用下列简称和定义。应当注意,如在此所用,单数形式“一种(a)”、“一种(an)”以及“该”包括复数个指示物,除非上下文中另外明确指明。因而,例如,提

及“一个/种多肽”时包括多个/种此类多肽,并且提及“所述/该配制品”时包括提及一个或多个配制品和本领域技术已知的其等效物,等等。

[0120] 除非另外定义,否则在此使用的所有技术和科学术语具有与本领域普通技术人员通常所理解的相同的含义。下文中提供了下述术语。

[0121] 除其他事项之外,“转半乳糖基酶”意指能够将半乳糖转移到D-半乳糖或D-葡萄糖的羟基基团的酶,由此产生半乳寡糖。在一方面,通过酶在乳糖上的反应来鉴定转半乳糖基酶,其中产生的半乳糖的量小于在任何给定时间产生的葡萄糖的量。

[0122] 在上下文中,术语“转半乳糖基化活性”意指将半乳糖部分转移至除水以外的分子。活性可以测定为[葡萄糖]-反应期间的任何给定时间产生的[半乳糖],或通过直接定量在反应期间的任何给定时间产生的GOS来测量。该测量可以以若干种方式进行,如通过实例中所示的HPLC方法进行。当比较转半乳糖基化活性的测量值时,它们已经在给定的初始乳糖浓度下进行,例如像,3%、4%、5%、6%、7%、8%、9%或10% (w/w)。

[0123] 在上下文中,术语“β-半乳糖苷酶活性”意指酶将β-半乳糖苷例如像乳糖水解成单糖、葡萄糖和半乳糖的能力。

[0124] 在计算转半乳糖基化活性:β-半乳糖苷酶活性的情况下,β-半乳糖苷酶活性被测量为在反应期间的任何给定时间产生[半乳糖]。该测量可以以若干种方式进行,如通过实例中所示的HPLC方法进行。

[0125] 在上下文中,使用邻硝基苯酚-β-D-吡喃半乳糖苷(ONPG),如下计算术语“转半乳糖基化活性的比率”:将比率计算为具有存在的受体的Abs420除以没有存在的受体的Abs420之间的比率乘以100。指数100的或以下的变体是纯水解变体,而上述水平描绘了相对的转半乳糖基化活性。

[0126] 转半乳糖基化活性的比率 = $(\text{Abs}420^{+\text{纤维二糖}} / \text{Abs}420^{-\text{纤维二糖}}) * 100\%$, 其中 $\text{Abs}420^{+\text{纤维二糖}}$ 是使用以下描述的方法3在420nm处读取的吸光度,在反应中包括纤维二糖,并且 $\text{Abs}420^{-\text{纤维二糖}}$ 是使用以下描述的方法3在420nm处读取的吸光度,而在反应中没有纤维二糖。以上方程仅适用于吸光度在0.5和1.0之间的稀释。

[0127] 在一方面,在15min反应、30min反应、60min反应、90min反应、120min反应或180min反应后测量本发明涉及的任何酶的活性。因此,在一方面,作为实例,在添加酶后15分钟,如添加酶后30分钟,如添加酶后60分钟,如添加酶后90分钟,如添加酶后120分钟或如添加酶后180分钟,测量相对转半乳糖基化活性。

[0128] 在上下文中,术语“转半乳糖基化活性:β-半乳糖苷酶活性的比率”意指([葡萄糖]-[半乳糖])/[半乳糖])。

[0129] 在上下文中,术语[葡萄糖]意指如通过HPLC测量的以重量%计的葡萄糖浓度。

[0130] 在上下文中,术语[半乳糖]意指如通过HPLC测量的以重量%计的半乳糖浓度。

[0131] 在上下文中,术语“已经将乳糖转半乳糖基化”意指半乳糖分子已经共价连接至乳糖分子,例如像共价连接至乳糖分子中的任何游离羟基基团,或者如通过内部转半乳糖基化产生,例如形成异乳糖。

[0132] 在上下文中,在“乳基测定”(酸奶应用模拟物)中测试了在半乳寡糖(GOS)生产中在此披露的多肽的性能评估。使用由98,60% (w/v) 新鲜巴氏杀菌低脂乳(爱氏晨曦迷你乳(Arla Mini-mælk)) 和1.4% (w/v) Nutrilac YQ-5075乳清成分(爱氏晨曦)组成的酸奶混合

物在96孔MTP板中进行体积为100 μ l的分批实验。为了完全使Nutrilac YQ-5075与水化合,将混合物搅拌20h,并且然后添加20mM磷酸钠(pH 6.5),以确保pH为6.5。将这种酸奶基简单地使用或与各种补充剂如另外的乳糖、海藻糖、麦芽糖、木糖或盐一起使用。将90 μ l酸奶与10 μ l纯化的酶或粗制酵素混合,用胶带密封并在43 $^{\circ}$ C下孵育3小时。通过100 μ l 10%Na₂CO₃终止反应。将样品储存于-20 $^{\circ}$ C下。通过HPLC定量半乳寡糖(GOS)、乳糖、葡萄糖和半乳糖。在Dionex ICS 3000上进行样品分析。IC参数如下:流动相:150mM NaOH,流速:等度,0,25ml/min,柱:Carbopac PA1,柱温:RT,注射体积:10 μ l,检测器:PAD,积分:手动,样品制备:在Milli-Q水中稀释100倍(0,1ml样品+9,9ml水),并通过0.45 μ m针筒式滤器过滤,定量:标准品的峰面积的百分比中的峰面积。将GOS糖浆(Vivanal GOS,菲仕兰坎皮纳公司(Friesland Campina))用作GOS定量的标准品。

[0133] 转半乳糖基化活性可以通过如WO 2013/182686中所述的HPLC定量或酶测定来测量。

[0134] 在上下文中,术语“将该多肽喷雾干燥”意指已经通过将溶液或悬浮液中的多肽在合适的温度下喷雾干燥并且持续适当的时间除去水而获得多肽。

[0135] 在上下文中,术语“该多肽在溶液中”涉及可溶于溶剂而不会沉淀出溶液的多肽。用于该目的的溶剂包括多肽可以存在的任何环境,如水性缓冲液或盐溶液、发酵液或表达宿主的细胞质。

[0136] 在上下文中,术语“稳定剂”意指用于稳定多肽的任何稳定剂,例如,多元醇,例如像甘油或丙二醇、糖或糖醇,乳酸,硼酸或硼酸衍生物(例如芳族硼酸酯)。在一方面,稳定剂不是多元醇,或该多元醇以0.1wt%或更低水平存在。

[0137] 术语“分离的”意指多肽至少基本上不含与该序列在自然界中天然缔合或者在自然界中存在的至少一种其他组分,和/或基本上不含纤维素、甘露聚糖、果胶酶或淀粉酶。在一方面,如在此所用“分离的多肽”是指如通过SDS-PAGE确定的是至少30%纯、至少40%纯、至少60%纯、至少80%纯、至少90%纯、和至少95%纯的多肽。

[0138] 因此,术语“基本上不含纤维素酶”在此意指按重量计包含至多10%,优选至多8%,更优选至多6%,更优选至多5%,更优选至多4%,至多3%,甚至更优选至多2%,最优选至多1%,并且甚至最优选至多0.5%的纤维素酶的制剂。因此,在此术语“基本上不含”可以被看作与术语“分离的多肽”和“分离形式的多肽”同义。

[0139] 因此,术语“基本上不含甘露聚糖酶”在此意指按重量计包含至多10%,优选至多8%,更优选至多6%,更优选至多5%,更优选至多4%,至多3%,甚至更优选至多2%,最优选至多1%,并且甚至最优选至多0.5%的甘露聚糖酶的制剂。因此,在此术语“基本上不含”可以被看作与术语“分离的多肽”和“分离形式的多肽”同义。

[0140] 因此,术语“基本上不含果胶酶”在此意指按重量计包含至多10%,优选至多8%,更优选至多6%,更优选至多5%,更优选至多4%,至多3%,甚至更优选至多2%,最优选至多1%,并且甚至最优选至多0.5%的果胶酶的制剂。因此,在此术语“基本上不含”可以被看作与术语“分离的多肽”和“分离形式的多肽”同义。

[0141] 因此,术语“基本上不含淀粉酶”在此意指按重量计包含至多10%,优选至多8%,更优选至多6%,更优选至多5%,更优选至多4%,至多3%,甚至更优选至多2%,最优选至多1%,并且甚至最优选至多0.5%的淀粉酶的制剂。因此,在此术语“基本上不含”可以被看

作与术语“分离的多肽”和“分离形式的多肽”同义。

[0142] 术语“基本上纯的多肽”在此意指包含按重量计至多10%，优选至多8%，更优选至多6%，更优选至多5%，更优选至多4%，至多3%，甚至更优选至多2%，最优选至多1%，并且甚至最优选至多0.5%的与其天然缔合的其他多肽材料的多肽制剂。因此，优选的是该基本上纯的多肽是按制剂中存在的所有的多肽材料的重量计至少92%纯，优选至少94%纯，更优选至少95%纯，更优选至少96%纯，更优选至少96%纯，更优选至少97%纯，更优选至少98%纯，甚至更优选至少99%纯，最优选至少99.5%纯，并且甚至最优选100%纯的。在此披露的多肽优选为基本上纯的形式。特别地，优选的是这些多肽是“基本上纯的形式”，即多肽制剂基本上不含与其天然缔合的其他多肽材料。例如，这可以通过熟知的重组方法或通过经典的纯化方法制备多肽来完成。在此，术语“基本上纯的多肽”与术语“分离的多肽”和“分离形式的多肽”同义。

[0143] 术语“纯化”或“纯”意指给定的组分以高水平状态存在-例如，至少约51%纯如至少51%纯，或至少约75%如至少75%纯，或至少约80%如至少80%纯，或至少约90%如至少90%纯，或至少约95%如至少95%纯，或至少约98%如至少98%纯。希望的是，该组分是组合物中存在的主要活性组分。

[0144] 关于本发明的术语“微生物”包括任何“微生物”，该任何“微生物”可以包括根据本发明所述的核苷酸序列或编码具有如在此定义的特定特性的多肽的核苷酸序列和/或由其获得的产物。在上下文中，“微生物”可以包括能够发酵乳基质的任何细菌或真菌。

[0145] 关于本发明的术语“宿主细胞”包括任何细胞，该任何细胞包括编码具有如在此定义的特定特性的多肽的核苷酸序列或如上所述的并用于生产具有如在此定义的特定特性的多肽的表达载体。在一方面，生产是重组生产。

[0146] 在本发明的上下文中，术语“乳”应理解为从任何哺乳动物如牛、绵羊、山羊、水牛或骆驼获得的乳状分泌物。

[0147] 在上下文中，术语“乳基底物”意指任何生的和/或加工的乳材料或衍生自乳成分的材料。有用的乳基底物包括但不限于包括乳糖的任何乳或乳样产品的溶液/悬浮液，如全脂或低脂乳、脱脂乳、酪乳、复原乳粉、炼乳、奶粉溶液、UHT乳、乳清、乳清渗透物、酸乳清或奶油。优选地，乳基底物是乳或脱脂乳粉的水溶液。乳基底物可以比生乳更浓。在一个实施例中，乳基底物的蛋白质与乳糖的比率为至少0.2，优选至少0.3，至少0.4，至少0.5，至少0.6，或最优选至少0.7。可以根据本领域已知的方法对乳基底物进行均化和/或巴氏杀菌。

[0148] 如在此所用“均质化”意指强烈混合以获得可溶性悬浮液或乳液。可以进行该均质化以便将乳脂分解成更小的尺寸，这样使得乳脂不再与该乳分离。这可以通过在高压下使该乳通过小孔来实现。

[0149] 如在此所用“巴氏杀菌”意指减少或消除活生物体如微生物在乳基底物中的存在。优选地，通过将规定温度保持规定的时间来实现巴氏杀菌。指定的温度通常通过加热来实现。可以选择温度和持续时间以便杀灭或使某些细菌，如有害细菌失活，和/或使该乳中的酶失活。可以遵循快速冷却步骤。本发明上下文中的“食品”或“食物组合物”可以是适合于动物或人类消费的任何可食用食品或饲料产品。

[0150] 在本发明的上下文中的“奶产品”可以是主要成分之一是乳基的任何食品。优选地，主要成分是乳基的。更优选地，主要成分是已经用具有转半乳糖基化活性的酶处理过的

乳基底物。

[0151] 在上下文中,“主要成分之一”意指具有干物质的成分,该干物质占奶产品总干物质的20%以上,优选30%以上或40%以上,而“主要成分”意指具有干物质的成分,该干物质占奶产品总干物质的50%以上,优选60%以上或70%以上。

[0152] 在上下文中,“发酵的奶产品”应被理解为任何奶产品,其中任何类型的发酵形成生产过程的一部分。发酵的奶产品的实例是产品像酸奶、酪乳、鲜奶油、奶渣和清爽干酪。发酵的奶产品的另一个实例是奶酪。发酵的奶产品可以通过本领域已知的任何方法来生产。

[0153] 术语“发酵”意指通过微生物如起子培养物的作用将碳水化合物转化为醇或酸。在一方面,发酵包括将乳糖转化为乳酸。

[0154] 在上下文中,“微生物”可以包括能够发酵乳基质的任何细菌或真菌。

[0155] 在上下文中,术语“Pfam结构域”意指基于多重序列比对和隐马尔可夫基序的存在,被鉴定为Pfam-A或Pfam-B的蛋白质序列内的区域(“Pfam蛋白家族数据库(The Pfam protein families database)” :R.D.芬恩(Finn),J.米斯特里(Mistry),J.塔特(Tate),P.柯吉尔(Coggill),A.赫格尔(Heger),J.E.普林顿(Pollington),O.L.加文(Gavin),P.古恩赛卡伦(Gunasekaran),G.赛利克(Ceric),K.福斯隆德(Forslund),L.霍尔姆(Holm),E.L.松哈默(Sonnhammer),S.R.艾迪(Eddy),A.贝特曼(Bateman),核酸研究(Nucleic Acids Research)(2010),数据库专辑38:D211-222)。作为Pfam结构域的实例,可以提及Glyco_hydro2N(PF02837)、Glyco_hydro(PF00703)、Glyco_hydro 2C(PF02836)和细菌Ig样结构域(组4)(PF07532)。

[0156] 如在此所用,术语“表达”是指基于基因的核酸序列产生多肽的过程。此过程包括转录和翻译。

[0157] 如在此所用,“转化的细胞”包括已经通过使用重组DNA技术转化的细胞,包括细菌和真菌细胞。转化通常通过将个或多个核苷酸序列插入细胞而发生。所插入的核苷酸序列可以是异源核苷酸序列,即,对于待转化的细胞不是天然的序列,如融合蛋白。

[0158] 如在此所用,“可操作地连接”意指所描述的组分处于允许它们以其预期方式起作用的关系。例如,将可操作地连接到编码序列的调控序列以如下这样方式连接:在与控制序列相容的条件下实现编码序列的表达。

[0159] 如在此所用,“载体”是指设计用于将核酸引入一种或多种细胞类型的多核苷酸序列。载体包括克隆载体、表达载体、穿梭载体、质粒、噬菌体颗粒、盒等。

[0160] 如在此所用,术语“表达”是指基于基因的核酸序列产生多肽的过程。此过程包括转录和翻译。表达可能涉及使用宿主生物来产生多肽。宿主生物体,又简称为宿主,可以包括原核生物和真核生物,并且在一些实施例中可以包括细菌和真菌物种。

[0161] 如在此所用,“表达载体”是指包含DNA编码序列(例如,基因序列)的DNA构建体,其可操作地连接到能够影响编码序列在宿主中表达的一种或多种合适的控制序列。这样的控制序列包括影响转录的启动子,控制这种转录的任意的操纵子序列,编码合适的mRNA核糖体结合位点的序列,以及控制转录和翻译终止的序列。载体可以是质粒、噬菌体颗粒,或简单地是潜在的基因组插入物。一旦转化入合适的宿主中,载体可以独立于宿主基因组复制和起作用,或者在某些情况下可以整合入基因组本身。质粒是最常用的表达载体形式。然而,该描述旨在包括发挥等效功能并且本领域已知或将成为已知的其他形式的表达载体。

[0162] “启动子”是指参与结合RNA聚合酶以启动基因转录的调控序列。启动子可以是诱导型启动子或组成型启动子。可以使用的诱导型启动子的非限制性实例是里氏木霉cbh1，其是诱导型启动子。

[0163] 术语“可操作地连接”是指并置，其中元件处于使得它们在功能上相关的布置。例如，如果启动子控制编码序列的转录，则启动子可操作地连接到编码序列。

[0164] “在转录控制下”是本领域中众所周知的术语，该术语表明多核苷酸序列的转录取决于其与有助于启动或促进转录的元件可操作地连接。

[0165] “在翻译控制下”是本领域中众所周知的术语，其表示在mRNA形成后发生的调节过程。

[0166] “基因”是指参与产生多肽并且包括编码区之前和之后的区域以及个体编码区段(外显子)之间的间插序列(内含子)的DNA区段。

[0167] 如在此所用，术语“宿主细胞”是指用于产生多肽的重组表达载体可以转染入其中以表达该多肽的细胞或细胞系。宿主细胞包括单个宿主细胞的后代，并且由于天然的、意外的或故意的突变，后代可能不一定(在形态学或总基因组DNA补体方面)与原始亲本细胞完全同一。宿主细胞包括用表达载体体内转染或转化的细胞。

[0168] 术语“重组”当用于提及一种细胞、核酸、蛋白或载体时是指该细胞、核酸、蛋白或载体已经通过导入异源核酸或蛋白或者改变天然的核酸或蛋白进行修饰，或者表明该细胞衍生自经这样修饰的细胞。因此，例如，这些重组细胞表达在细胞的天然形式(非重组)中未发现的基因或表达以其他方式异常表达的、低表达的或根本不表达的天然基因。

[0169] “信号序列”(又称为“前序列”、“信号肽”、“前导序列”或“前导肽”)是指与蛋白质的N端部分结合的氨基酸序列，该氨基酸序列促进了成熟形式的蛋白质从细胞的分泌(例如，SEQ ID NO:5)。信号序列将多肽靶向分泌途径，并且一旦其在内质网膜中被转位就从新生多肽切割。成熟形式的细胞外蛋白质(例如SEQ ID NO:1)缺乏在分泌过程中被切除的信号序列。

[0170] 术语“选择性标志物”或“可选择的标志物”是指能够在宿主细胞中表达的基因，其允许容易地选择包含引入的核酸或载体的那些宿主。可选择的标志物的实例包括但不限于在宿主细胞上赋予代谢优势，如营养优势的抗微生物物质(例如，潮霉素、博来霉素或氯霉素)和/或基因。

[0171] 术语“培养”是指在液体或固体培养基中在生长适合生长条件下使微生物细胞群生长。术语“培养基”是指在该过程中使用的介质

[0172] 在将核酸序列插入细胞的上下文中，术语“引入的”包括“转染”、“转化”或“转导”，并且是指将核酸序列并入真核或原核细胞中，其中可以将核酸序列并入细胞的基因组(例如，染色体、质粒、质体或线粒体DNA)中，转化成自主复制子或进行瞬时表达。

[0173] 如在此所用，术语“转化的”，“稳定转化的”和“转基因的”是指具有整合到其基因组中的或作为通过多代维持的附加型质粒的非天然(例如，异源的)核酸序列的细胞。

[0174] 如在此所用，术语“修改”和“改变”可互换使用并且意指变化或变更。在修饰或改变多肽的上下文中，这些术语可以意指直接或通过改变编码核酸来改变氨基酸序列，或如将酶糖基化来改变多肽的结构。

[0175] 除非另有说明，本发明的实践将使用在本领域技术范围内的分子生物学(包括重

组技术)、微生物学、细胞生物学和生物化学的常规技术。这些技术在文献中得到充分解释,这些文献例如,分子克隆:实验室手册(Molecular Cloning:A Laboratory Manual),第二版(萨姆布鲁克(Sambrook)等人,1989,分子克隆:实验室手册);寡核苷酸合成(Oligonucleotide Synthesis)(M.J.盖特(Gait)编辑,1984;分子生物学现代方法(Current Protocols in Molecular Biology)(F.M.奥苏贝尔(Ausubel)等人编辑,1994);PCR:聚合酶链式反应(PCR:The Polymerase Chain Reaction)(穆利斯(Mullis)等人编辑,1994);和基因转移和表达:实验室手册(Gene Transfer and Expression:A Laboratory Manual)(克里格勒(Kriegler),1990)。

[0176] 除非在此另外定义,否则在此所用的所有技术与科学术语均具有与本发明所属领域的普通技术人员通常理解的相同含义。辛格尔顿(Singleton)等人,“微生物与分子生物学杂志(Dictionary of Microbiology and Molecular Biology)”,第二版,约翰威利父子出版公司(John Wiley and Sons),纽约(1994)和黑尔(Hale)与马卡姆(Markham),哈珀柯林斯生物学词典(The Harper Collins Dictionary of Biology),哈珀永久出版社(Harper Perennial),纽约(1991)向技术人员提供本发明中使用的许多术语的通用词典。与在此所述那些类似或等效的任何方法和材料可用于本发明的实践或测试。

[0177] 除非另有说明,分别地,核酸以5'至3'取向从左向右书写;氨基酸序列以氨基至羧基取向从左向右书写。

[0178] 在此提供的数字范围包括定义范围的数字。

[0179] 在酶产物中,少量不需要的酶活性主要来源于生产宿主,并且称为酶副活性。本发明涉及减少不需要的酶副活性。乳制品应用中的酶活性应优选在适用于应用的Ph和温度下进行测量。在乳中,pH值从6.4至6.8不等,酸奶:pH约4,婴儿配方奶pH 5.9-7.3,马苏里拉奶酪pH 5.2-5.5,以及蛋黄酱pH 4。最佳地,可以通过每个预期应用中的应用测试来确定不希望的活性的水平。

[0180] 纤维素酶

[0181] 在一方面,本发明涉及不含或基本上不含纤维素酶活性的组合物。在另一方面,本发明提供了一种新颖的细菌,其中使编码纤维素酶的基因失活。这样的组合物包含具有转半乳糖基化活性的多肽。根据本发明的宿主细胞能够表达具有转半乳糖基化活性的多肽。这种转半乳糖基化活性应以如下这样水平存在:GOS能够在产品中由乳糖生产,即优选多肽存在于组合物中,或者宿主细胞能够以如下这种水平表达它:GOS形式的膳食纤维可以由添加有组合物或宿主细胞的奶产品中存在的乳糖产生。

[0182] 纤维素酶是水解纤维素(β -1,4-葡聚糖或 β -D-糖苷键)的酶,其导致葡萄糖、纤维二糖、纤维寡糖等的形成。纤维素酶传统上分为三大类:内切葡聚糖酶(EC 3.2.1.4) (“EG”)、外切葡聚糖酶或纤维二糖水解酶(EC3.2.1.91) (“CBH”)和 β -葡糖苷酶($[\beta]$ -D-葡萄糖苷葡萄糖水解酶;EC3.2.1.21) (“BG”)。根据一个实施例,失活的纤维素酶基因是**bgLC**内切葡聚糖酶。

[0183] 甘露聚糖酶

[0184] 甘露聚糖酶是一种酶,其分解称为甘露聚糖的化合物。这些多糖由单糖甘露糖构成,并且在自然界中广泛发现。在许多植物(并且,例如,在其种子中),甘露聚糖用作碳水化合物储备。在一方面,酶选自由以下各项组成的组:甘露聚糖酶,特别是内切 β -甘露聚糖酶、

酯酶、外切甘露聚糖酶、半乳聚糖酶。根据一个实施例，失活的甘露聚糖酶基因是gmuG甘露聚糖酶。

[0185] 果胶酶

[0186] 果胶酶是分解果胶(植物细胞壁中发现的多糖)的酶。通常被称为果胶酶,它们包括果胶降解酶、果胶糖酶(pectozyme)和多聚半乳糖醛酸酶。研究最多并最广泛使用的商业果胶酶之一是多聚半乳糖醛酸酶。根据一个实施例,失活的果胶酶基因是pe1果胶酸裂解酶。

[0187] 淀粉酶

[0188] 淀粉酶是一种催化淀粉水解成糖的酶。淀粉酶存在于人类和一些其他哺乳动物的唾液中,其中它开始消化的化学过程。胰腺和唾液腺产生淀粉酶(α 淀粉酶)来将膳食淀粉水解成二糖和三糖,这些二糖和三糖通过其他酶转化成葡萄糖,为身体提供能量。植物和一些细菌也产生淀粉酶。特定的淀粉酶蛋白由不同的希腊字母指定。所有淀粉酶都是糖苷水解酶,并且作用于 α -1,4-糖苷键。本发明优选另外涉及(EC 3.2.1.1)(CAS#9014-71-5)(替代名称:1,4- α -D-葡聚糖葡聚糖水解酶;糖原酶)的失活。根据一个实施例,失活的淀粉酶基因是amyE α -淀粉酶。

[0189] 在一方面,在此所用的术语“淀粉酶”是指淀粉酶,如 $[\alpha]$ -淀粉酶(EC 3.2.1.1)、 $[\beta]$ -淀粉酶(EC 3.2.1.2)和 $[\gamma]$ -淀粉酶(EC 3.2.1.3)。

[0190] 同一性程度

[0191] 两个氨基酸序列之间或两个核苷酸序列之间的关联性由参数“同一性”描述。

[0192] 在一个实施例中,通过以下方式来确定查询序列与参考序列之间的序列同一性程度:1)通过使用默认评分矩阵和默认空位罚分的任何合适的比对程序来比对两种序列,2)鉴别精确匹配的数量,其中精确匹配是比对程序已经在比对的给定位置处鉴别出两个比对序列中的同一的氨基酸或核苷酸,以及3)将精确匹配的数目除以参考序列的长度。

[0193] 在一个实施例中,通过以下方式来确定查询序列与参考序列之间的序列同一性程度:1)通过使用默认评分矩阵和默认空位罚分的任何合适的比对程序来比对两种序列,2)鉴别精确匹配的数量,其中精确匹配是其中比对程序已经在比对的给定位置处鉴别出两个比对序列中的同一的氨基酸或核苷酸,以及3)将精确匹配的数目除以两个序列中最长的长度。

[0194] 在另一个实施例中,通过以下方式来确定查询序列与参考序列之间的序列同一性程度:1)通过使用默认评分矩阵和默认空位罚分的任何合适的比对程序来比对两种序列,2)鉴别精确匹配的数量,其中精确匹配是其中比对程序已经在比对的给定位置处鉴别出两个比对序列中的同一的氨基酸或核苷酸,以及3)将精确匹配数目除以“比对长度”,其中该比对长度是包括序列的空位和突出部分的整个比对的长度。

[0195] 序列同一性比较可以通过眼睛,或更通常地借助于容易获得的序列比较程序进行。这些商业可得的计算机程序使用复杂的比较算法来比对两个或更多个序列,其最好地反映出可能导致两个或更多个序列之间的一个或多个差异的进化事件。因此,这些算法使用评分系统进行操作,评分系统奖励同一或相似氨基酸的比对,并惩罚空位的插入,空位延伸和不相似氨基酸的比对。比较算法的评分系统包括:

[0196] i) 每当插入空位时指定罚分(空位罚分),

[0197] ii) 每当现有的空位用额外的位置扩大时指定罚分(延伸罚分),

[0198] iii) 当比对同一的氨基酸时指定高分,以及

[0199] iv) 当比对非同一的氨基酸时指定可变分数。

[0200] 大多数比对程序允许修改空位罚分。然而,当使用这种序列比较软件时优选使用默认值。

[0201] 对于非同一的氨基酸的比对给出的得分是根据评分矩阵(也称作取代矩阵)指定的。在所述取代矩阵中提供的得分所反映的事实是:进化过程中一个氨基酸被另一个所取代的可能性是变化的并且其取决于将被取代的氨基酸的物理/化学性质。例如,与被疏水性氨基酸取代相比,极性氨基酸被另一极性氨基酸取代的可能性更高。因此,评分矩阵将为同一的氨基酸指定最高分数,为非同一但相似的氨基酸指定较低分数,并且为非同一的不相似氨基酸指定甚至更低分数。最常用的评分矩阵是PAM矩阵(戴霍夫(Dayhoff)等人,(1978),琼斯(Jones)等人,(1992))、BLOSUM矩阵(亨尼科夫(Henikoff)和亨尼科夫(1992))和贡内特(Gonnet)矩阵(贡内特等人,(1992))。

[0202] 用于进行这种比对的合适的计算机程序包括但不限于Vector NTI(英杰公司(Invitrogen Corp.))和ClustalV,ClustalW和ClustalW2程序(希金斯(Higgins)DG和夏普(Sharp)PM(1988),希金斯等人(1992),汤普森(Thompson)等人(1994),拉金(Larkin)等人(2007))。可从位于www.expasy.org的ExPASy蛋白质组服务器(Proteomics server)获得不同比对工具的选择。可以执行序列比对的软件的另一个实例是BLAST(基本局部比对搜索工具(Basic Local Alignment Search Tool)),该BLAST可从美国国家生物技术信息中心的网页获得,该网页目前可以在<http://www.ncbi.nlm.nih.gov/>找到,并且该BLAST首先在阿尔丘尔等人(1990)分子生物学杂志(J.Mol.Biol.)215:403-410中进行了描述。

[0203] 在本发明的一个优选实施例中,该比对程序正在执行全局比对程序,其优化了全长序列的比对。在另一个优选实施例中,全局比对程序是基于尼德尔曼-翁施(Needleman-Wunsch)算法(尼德尔曼,索尔(Saul)B.;和翁施,克里斯蒂安(Christian)D.(1970),“适用于搜索两种蛋白质氨基酸序列相似性的一般方法(A general method applicable to the search for similarities in the amino acid sequence of two proteins)”,分子生物学杂志(Journal of Molecular Biology)48(3):443-53)。使用尼德尔曼-翁施算法执行全局比对的当前程序的实例是EMBOSS Needle和EMBOSS Stretcher程序,两者均可以在<http://www.ebi.ac.uk/Tools/psa/>上获得。

[0204] EMBOSS Needle使用德尔曼-翁施比对算法执行最佳全局序列比对,以沿其整个长度找到两个序列的最佳比对(包括空位)。

[0205] EMBOSS Stretcher使用德尔曼-翁施算法的修改版,该修改版允许更大序列进行全局比对。

[0206] 在一个实施例中,通过全局比对程序对序列进行比对,并且通过鉴别由程序鉴别的精确匹配数目除以“比对长度”来计算序列同一性,其中比对长度是包括序列的空位和突出部分的整个比对的长度。

[0207] 在另一个实施例中,全局比对程序使用德尔曼-翁施算法,并且通过鉴别由程序鉴别的精确匹配数目除以“比对长度”来计算序列同一性,其中比对长度是包括序列的空位和突出部分的整个比对的长度。

[0208] 在又一个实施例中,全局比对程序选自由以下各项组成的组:EMBOSS Needle和EMBOSS stretcher,并且通过鉴别由程序鉴别的精确匹配数目除以“比对长度”来计算序列同一性,其中比对长度是包括序列的空位和突出部分的整个比对的长度。

[0209] 一旦软件生成了比对,就可以计算出相似度%和序列同一性%。该软件通常进行这种计算作为序列比较的一部分,并产生数值结果。

[0210] 在一个实施例中,优选使用ClustalW软件来执行序列比对。优选地,使用以下成对比对的参数进行采用ClustalW的比对:

[0211]	取代矩阵:	Gonnet 250
	空位开放罚分:	20
	空位延伸罚分:	0.2
	空位终止罚分:	无

[0212] 例如,ClustalW2可以在互联网上由欧洲生物信息学研究所EMBL-EBI网页www.ebi.ac.uk的工具-序列分析-ClustalW2下获得。目前,ClustalW2工具的确切地址是www.ebi.ac.uk/Tools/clustalw2。

[0213] 在另一个实施例中,优选使用Vector NTI (英杰公司)中的程序Align X进行序列比对。在一个实施例中,Exp10可以与默认设置一起使用:

[0214] 空位开放罚分:10

[0215] 空位延伸罚分:0.05

[0216] 空位分离罚分范围:8

[0217] 在另一个实施例中,通过使用分数矩阵:blosum62mt2和VectorNTI成对比对设置来确定一个氨基酸序列与,或相对于,另一个氨基酸序列的比对

[0218]	设置	K-元组	1
		最佳对角线数	5
		窗口大小	5
		空位罚分	3
		空位开放罚分	10
		空位延伸罚分	0,1

[0219] 在一个实施例中,通过使用字大小为3并且BLOSUM 62作为具有转半乳糖基化活性的取代矩阵的Blast来确定一个氨基酸序列与,或相对于,另一个氨基酸序列的同一性的百分比。

[0220] 多肽

[0221] 在一方面,在此披露的本发明使用等于或高于3%w/w初始乳糖浓度浓度的具有如下转半乳糖基化活性: β -半乳糖苷酶活性的比率的多肽:至少0.5、至少1、至少2、至少2.5、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、或至少12。

[0222] 在一方面,在此披露的本发明使用多肽,其中糖苷水解酶催化核心具有SEQ ID NO:7的氨基酸序列。

[0223] 在一方面,在此披露的本发明使用包含Glyco_hydro2N (PF02837)、Glyco_hydro (PF00703) 和/或Glyco_hydro 2C (PF02836) 结构域的多肽。

[0224] 在一方面,在此披露了包含细菌Ig样结构域(组4)(PF07532)的多肽。

[0225] 在一方面,在此披露了具有选自下组的转半乳糖基化活性的多肽,该组由以下各项组成:

[0226] g.包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列的多肽,其中所述多肽由至多980个氨基酸残基组成,

[0227] h.包含与SEQ ID NO:2具有至少97%序列同一性的氨基酸序列的多肽,其中所述多肽由至多975个氨基酸残基组成,

[0228] i.包含与SEQ ID NO:3具有至少96.5%序列同一性的氨基酸序列的多肽,其中所述多肽由至多1300个氨基酸残基组成,

[0229] j.由多核苷酸编码的多肽,该多核苷酸在至少低严格条件下与以下各项杂交:i)包含在SEQ ID NO:9、10、11、12或13中的编码SEQ ID NO:1、2、3、4或5的多肽的核酸序列;或ii) i)的互补链,

[0230] k.由包含如下各项的多核苷酸编码的多肽:与编码SEQ ID NO:1、2、3、4或5的多肽的核苷酸序列具有至少70%同一性的核苷酸序列,或包含在SEQ ID NO:9、10、11、12或13中的编码成熟多肽的核苷酸序列,和

[0231] l.包括SEQ ID NO:1、2、3、4或5的一个或多个氨基酸残基的缺失、插入和/或保守取代的多肽。

[0232] 在另一方面,在此披露的本发明使用具有选自下组的转半乳糖基化活性的多肽,该组由以下各项组成:

[0233] a.包含与SEQ ID NO:3具有至少96.5%序列同一性的氨基酸序列的多肽,其中所述多肽由至多1300个氨基酸残基组成,

[0234] b.包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列的多肽,其中所述多肽由至多980个氨基酸残基组成,

[0235] c.由多核苷酸编码的多肽,该多核苷酸在至少低严格条件下与以下各项杂交:i)包含在SEQ ID NO:9、10、11、12或13中的编码SEQ ID NO:1、2、3、4或5的多肽的核酸序列;或ii) i)的互补链,

[0236] d.由包含如下各项的多核苷酸编码的多肽:与编码SEQ ID NO:1、2、3、4或5的多肽的核苷酸序列具有至少70%同一性的核苷酸序列,或包含在SEQ ID NO:9、10、11、12或13中的编码成熟多肽的核苷酸序列,和

[0237] e.包括SEQ ID NO:1、2、3、4或5的一个或多个氨基酸残基的缺失、插入和/或保守取代的多肽。

[0238] 在一方面,在此披露的本发明使用多肽,其中氨基酸序列与SEQ ID NO:1、2、3、4或5的成熟氨基酸序列具有至少68%、70%、72%、74%、76%、78%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%序列同一性。

[0239] 在一方面,在此披露的本发明使用与SEQ ID NO:1的成熟氨基酸序列具有90%序列同一性的多肽。

[0240] 在一方面,在此披露的本发明使用与SEQ ID NO:2的成熟氨基酸序列具有90%序列同一性的多肽。

[0241] 在一方面,在此披露的本发明使用与SEQ ID NO:3的成熟氨基酸序列具有96.5%

序列同一性的多肽。

[0242] 在一方面,在此披露的本发明使用与SEQ ID NO:4的成熟氨基酸序列具有96.5%序列同一性的多肽。

[0243] 在一方面,在此披露的本发明使用与SEQ ID NO:5的成熟氨基酸序列具有96.5%序列同一性的多肽。

[0244] 在一方面,在此披露的本发明使用包括SEQ ID NO:1、2、3、4或5的氨基酸序列的或由其组成的多肽。

[0245] 在一方面,在此披露的本发明使用衍生自两歧双歧杆菌的多肽。

[0246] 在一方面,在此披露的本发明使用pH最佳值为6.5-7.5的多肽。

[0247] 在一方面,在此披露的本发明使用温度最佳值为30-60如42-60摄氏度的多肽。

[0248] 可以将对碳水化合物具有活性的多肽使用基于其底物特异性的IUBMB分类系统或CaZy指定分类为当前125个糖苷水解酶家族之一。在CaZy数据库中,指定是基于序列和结构信息,结合底物和产品的立体化学知识而进行的。

[0249] 在此披露了多肽的用途,当在包括编码所述多肽的核酸序列的合适的宿主菌株(例如,枯草芽孢杆菌)中是表达产物时,该多肽是表现出转半乳糖基化活性的所述核酸序列的唯一多肽表达产物。这可以通过使用本领域技术人员已知的以下技术来评估。使待评估的样品经受SDS-PAGE,并使用适合于蛋白质定量的染料例如像伯乐标准(Bio-Rad Criterion)系统进行可视化。然后使用适当的光密度扫描仪(例如像伯乐标准系统)扫描凝胶,并确保所得图像处于动态范围内。对与衍生自SEQ ID NO:8的任何变体/片段相对应的条带进行定量,并且如下计算多肽的百分比:所讨论的多肽的百分比=所讨论的多肽/(表现出转半乳糖基化活性的所有多肽的总和)*100。组合中衍生自SEQ ID NO:8的多肽变体/片段的总数可以通过本领域技术人员已知的方法使用多克隆抗体通过蛋白质印迹检测衍生自SEQ ID NO:8的片段来确定。

[0250] 在此披露的多肽包括酶内包含的至少两个分离的功能结构域。首先,如下所述,多肽应包含糖苷水解酶催化核。催化核心应属于相关糖苷水解酶家族的GH-A族。GH-A族的特征在于通过保留机制切割糖苷键,并具有基于TIM桶折叠的催化结构域(维伦加(Wierenga),2001,欧洲生化学会联合会快报(FEBS Letters),492(3),第193-8页)。催化结构域包含作为质子供体和亲核体的两个谷氨酸残基,这两个谷氨酸残基源于桶结构域的链4和7(詹金斯(Jenkins),1995,欧洲生化学会联合会快报,362(3),第281-5页)。TIM桶的总体结构是由8个 β 链和8个 α -螺旋组成的(β/α)8折叠。在一方面,在此披露的糖苷水解酶催化核心属于糖苷水解酶家族GH-2和-35之一,这些糖苷水解酶家族GH-2和-35都是属于GH-A族的TIM-桶酶。在另一方面,糖苷水解酶催化核心属于家族GH-2或GH-35。在另一方面,糖苷水解酶催化核心属于家族GH-2。一个共同特点是这些酶是所谓的保留酶,使得底物的立体化学在产品中是保守的(亨利萨特(Henrissat),1997,结构生物学当前观点(Curr Opin Struct Biol),7(5),637-44)。

[0251] 在一方面,在此披露的多肽对具有 $\beta(1\rightarrow4)$ 构象的碳水化合物键具有活性。这有效地将该酶归入 β -半乳糖苷酶的IUBMB EC 3.2.1.23类。该活性可以但不限于通过利用合成底物如对硝基苯酚- β -D-吡喃半乳糖苷(PNPG)、邻硝基苯酚- β -D-吡喃半乳糖苷(ONPG)或具有显色糖苷配基的 β -D-吡喃半乳糖苷(XGal)来确定。作为确定酶是否属于 β -半乳糖苷酶的

EC 3.2.1.23类的替代方法是与底物如乳糖一起孵育,并通过方法如酶测定、HPLC、TLC或本领域技术人员已知的其他方法测量葡萄糖的释放。

[0252] 为了预测多肽的功能实体,可以应用若干个可用的公共储存库,例如像Pfam(核酸研究(Nucl.Acids Res.) (2010) 38 (增刊1):D211-D222.doi:10.1093/nar/gkp985)和Interpro(核酸研究(2009) 37 (增刊1):D211-D215.doi:10.1093/nar/gkn785)。应当指出,当进行这种分析时,应该在从公共储存库数据库中获得的多肽的全长序列上进行分析。

[0253] 在另一方面,提供了包含选自以下的一种或多种Pfam结构域的多肽:Glyco_hydro2N (PF02837)、Glyco_hydro (PF00703)、Glyco_hydro 2C (PF02836) 和细菌Ig样结构域(组4) (PF07532)。在又另一方面,提供包含Pfam结构域Glyco_hydro2N (PF02837)、Glyco_hydro (PF00703)、Glyco_hydro 2C (PF02836) 和细菌Ig样结构域(组4) (PF07532)的多肽。在又另一方面,使用包含构成多肽催化结构域的Glyco_hydro2N (PF02837)、Glyco_hydro (PF00703) 和Glyco_hydro 2C (PF02836) 结构域的多肽。

[0254] 在另一方面,使用如在此披露的并具有如下转半乳糖基化活性: β -半乳糖苷酶活性比率的多肽,如在100ppm浓度下,在乳基测定中,在37°C和5w/w%乳糖下,在15、30或180如180分钟反应后所测量的,该比率为至少1、至少2.5、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、或至少12。在另一方面,该多肽衍生自两歧双歧杆菌。

[0255] 在一方面,在此所披露的一种或多种多肽具有转半乳糖基化活性,这样使得如在100ppm的浓度下,在乳基测定中,在37°C和5w/w%乳糖下,在15、30或180如180分钟反应后所测量的,超过20%、超过30%、超过40%、高达50%的初始乳糖发生了转半乳糖基化。

[0256] 在另一方面,在此披露的一种或多种多肽具有 β -半乳糖苷酶活性,这样使得如在100ppm的浓度下,在乳基测定中,在37°C和5w/w%乳糖下,在15、30或180如180分钟反应后所测量的,小于80%、小于70%、小于60%、小于50%、小于40%、小于30%、小于20%的乳糖已被水解。

[0257] 在一方面,按照方法4中规定的,在对应于2.13LAU的100ppm的浓度下测量 β -半乳糖苷酶活性和/或转半乳糖基化活性。通常,酶的活性单位可以根据W0 2003/186286中披露为方法4的测定来测量,并在以下实例部分4中重现。

[0258] 在另一方面,在此披露的一种或多种多肽具有以下特征中的一个或多个:

[0259] a) 如在100ppm浓度下,在乳基测定中,在37°C和5w/w%乳糖下,在15、30或180如180分钟反应后所测量的,为至少1、至少2.5、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、或至少12的转半乳糖基化活性: β -半乳糖苷酶活性比率,和/或

[0260] b) 具有转半乳糖基化活性,这样使得如在100ppm的浓度下,在乳基测定中,在37°C和5w/w%乳糖下,在15、30或180如180分钟反应后所测量的,超过20%、超过30%、超过40%、以及高达50%的初始乳糖已经发生了转半乳糖基化。

[0261] 在一方面,提供了包含与SEQ ID NO:3具有至少96.5%序列同一性的氨基酸序列的多肽,其中所述多肽由至多1300个氨基酸残基组成。在另一方面,提供了包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列的多肽,如其中所述序列同一性是至少95%,例如像至少96%、至少97%、至少98%、至少99%或至少100%序列同一性,并且其中所述多肽由至多980个氨基酸残基组成。在另一方面,提供了包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列的多肽,其中所述多肽由至多980个氨基酸残基组成。在又另一方

面,提供了一种多肽,其中所述多肽与SEQ ID NO:1具有至少90%序列同一性,如其中所述多肽与SEQ ID NO:1具有至少90%,例如像至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、或至少99%序列同一性。在另一方面,提供了与SEQ ID NO:2具有至少96.5%序列同一性的多肽,如其中所述多肽与SEQ ID NO:2具有至少97%,例如像至少98%或至少99%序列同一性。在一方面,提供了在此披露的多肽,这些多肽由至多975个氨基酸残基,例如像,至多970个氨基酸残基、如至多950个氨基酸残基、如至多940个氨基酸残基、至多930个氨基酸残基、至多920个氨基酸残基、至多910个氨基酸残基、至多900个氨基酸残基、至多895个氨基酸残基或至多890个氨基酸残基组成。在一方面,提供了由887或965个氨基酸残基组成的特定的多肽。在一方面,提供了一种多肽,该多肽包含与SEQ ID NO:2具有至少97%序列同一性的氨基酸序列,如其中所述序列同一性为至少98%,例如像至少99%或至少100%序列同一性,其中所述多肽由至多975个氨基酸残基,例如像,至多970个或至少965个氨基酸残基组成。在一方面,使用包含与SEQ ID NO:2具有至少97%序列同一性的氨基酸序列的多肽,其中所述多肽由至多975个氨基酸残基组成。

[0262] 在另一优选的方面,提供了包括SEQ ID NO:1、2、3、4或5的多肽。在又一优选方面,使用由SEQ ID NO:1、2、3、4或5的氨基酸序列组成的多肽,特别是由SEQ ID NO:1或2的氨基酸序列组成的多肽。

[0263] 在另一方面,使用包含与SEQ ID NO:3具有至少96.5%序列同一性的氨基酸序列的多肽,如其中所述序列同一性是至少97%,例如像至少98%、至少99或至少100%序列同一性,其中所述多肽由至多1300个氨基酸残基组成。

[0264] 在另一方面,提供一种多肽,其中所述多肽与SEQ ID NO:5具有至少98.5%,如至少99%或至少99.5%序列同一性。在一方面,这样的多肽由至多1290个氨基酸残基,例如像,至多1280、至多1270、至多1260、至多1250、至多1240、至多1230、至多1220或至多1215个氨基酸残基组成。在一优选方面,使用由1211个氨基酸残基组成的多肽。

[0265] 在另一方面,提供一种多肽,其中所述多肽与SEQ ID NO:4具有至少96%,如至少97%,例如像至少98%或至少99%序列同一性。在一方面,使用一种多肽,该多肽由至多1210个氨基酸残基,例如像至多1200、至多1190、至多1180、至多1170、至多1160、至多1150或至多1145个氨基酸残基,如1142个氨基酸残基组成。

[0266] 在另一方面,提供一种多肽,其中所述多肽与SEQ ID NO:3具有至少96.5%,如至少97%,例如像至少98%或至少99%序列同一性。在一方面,提供了一种多肽,该多肽由至多1130个氨基酸残基,例如像至多1120、至多1110、至多1100、至多1090、至多1080、至多1070、至多1060、至多1050、至多1055或至多1040个氨基酸残基组成。在一优选方面,使用由1038个氨基酸残基组成的多肽。

[0267] 在另一方面,在此披露的多肽具有高于100%,如高于150%,175%或200%的转半乳糖基化活性比率。

[0268] 蛋白质通常由一个或多个功能区域(通常称为结构域)组成。不同结构域在不同蛋白质的不同组合中的存在产生了在自然界中发现的蛋白质的多样性谱。描述结构域的一种方式借助于Pfam数据库,该Pfam数据库是如以下描述的蛋白质结构域家族的大集合:“Pfam蛋白家族数据库”:R.D. 芬恩, J. 米斯特里, J. 塔特, P. 柯吉尔, A. 赫格尔, J.E. 普林顿, O.L. 加文, P. 古恩赛卡伦, G. 赛利克, K. 福斯隆德, L. 霍尔姆, E.L. 松哈默, S.R. 艾迪, A. 贝特

曼,核酸研究(2010),数据库专辑38:D211-222。每个家族由多个序列比对和隐马尔可夫模型(HMM)表示。在此提供的一种或多种多肽优选包含Pfam结构域Glyco_hydro2N(PF02837)、Glyco_hydro(PF00703)、Glyco_hydro 2C(PF02836)和细菌Ig样结构域(组4)(PF07532)中的一种或多种。在一方面,在此提供的一种或多种多肽包含Glyco_hydro2N(PF02837)、Glyco_hydro(PF00703)、Glyco_hydro 2C(PF02836)和细菌Ig样结构域(组4)(PF07532)。

[0269] 在一方面,在此使用的多肽在4-9,如5-8、如5.5-7.5、如6.5-7.5的pH范围内具有有益的转半乳糖基化活性。

[0270] 本发明涵盖与在此定义的一种或多种氨基酸序列或与具有如在此定义的具体特性的多肽具有一定程度的序列同一性或序列同源性的多肽的用途。具体地而言,本发明涵盖与以下定义的SEQ ID NO:1、2、3、4或5中任一个或其同系物具有一定程度的序列同一性的肽的用途。

[0271] 同源氨基酸序列和/或核苷酸序列应提供和/或编码一种多肽,与SEQ ID NO:1、2、3、4或5的多肽相比,该多肽保留了功能性转半乳糖基化活性和/或增强了转半乳糖基化活性。

[0272] 在上下文中,同源序列被认为包括一种氨基酸序列,该氨基酸序列可以与主题序列具有至少66%、70%、75%、78%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%同一性。通常,同系物将包括与主题氨基酸序列相同的活性位点等。尽管同源性也可以根据相似性(即具有相似化学特性/功能的氨基酸残基)来考虑,但在本发明的上下文中,优选的是根据序列同一性来表示同源性。

[0273] 因此,本发明还涵盖如在此定义的蛋白质或多肽的任何氨基酸序列的变体、同系物和衍生物的用途,特别是以下定义的SEQ ID NO:1、2、3、4或5的那些。

[0274] 这些序列,特别是以下定义的SEQ ID NO:1、2、3、4或5的变体、同系物和衍生物,还可以具有氨基酸残基的缺失、插入或取代,这产生沉默变化并且形成功能上等效的物质。可以基于残基的极性、电荷、溶解度、疏水性、亲水性和/或两亲性性质的相似性进行有意的氨基酸取代,只要保留底物的次要结合活性即可。例如,带负电荷的氨基酸包括天冬氨酸和谷氨酸;带正电荷的氨基酸包括赖氨酸和精氨酸;并且具有不带电极性头部基团(该基团具有类似的亲水性值)的氨基酸包括亮氨酸、异亮氨酸、缬氨酸、甘氨酸、丙氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、苯丙氨酸和酪氨酸。

[0275] 本发明还涵盖了可能出现的保守取代(取代和置换在此均用于意指现有氨基酸残基与替代残基的交换),即对等取代,如碱性对碱性、酸性对酸性、极性对极性等。也可以发生非保守取代,即从一类残基到另一类残基或者可替代地涉及包括非天然氨基酸如鸟氨酸(在下文中称为Z)、二氨基丁酸鸟氨酸(在下文中称为B)、正亮氨酸鸟氨酸(在下文中称为O)、吡啶基丙氨酸、噻吩丙氨酸、萘基丙氨酸以及苯基甘氨酸。

[0276] 可以进行的保守取代是例如在如下组内:碱性氨基酸(精氨酸、赖氨酸和组氨酸)、酸性氨基酸(谷氨酸和天冬氨酸)、脂肪族氨基酸(丙氨酸、缬氨酸、亮氨酸、异亮氨酸)、极性氨基酸(谷氨酰胺、天冬酰胺、丝氨酸、苏氨酸)、芳香族氨基酸(苯丙氨酸、色氨酸和酪氨酸)、羟基氨基酸(丝氨酸、苏氨酸)、大氨基酸(苯丙氨酸和色氨酸)和小氨基酸(甘氨酸、丙氨酸)。

[0277] 在一方面,本发明中使用的多肽序列处于纯化形式。

[0278] 在一方面,用于本发明的多肽或蛋白质处于分离形式。

[0279] 在一方面,本发明的多肽是重组生产的。

[0280] 变体多肽包含与SEQ ID NO:1或2具有一定百分比,例如至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的多肽。

[0281] 变体多肽包含与SEQ ID NO:3、4或5具有一定百分比,例如至少96%、97%、98%或99%序列同一性的多肽。

[0282] 在一方面,在此使用的多肽包括与由编码转半乳糖基酶的核苷酸序列编码的、包含在两歧双歧杆菌DSM20215中、在此示为SEQ ID NO:22的成熟多肽的氨基酸序列具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%序列同一性的氨基酸序列。根据SEQ ID NO:1、2、3、4或5讨论的与序列同一性和功能有关的所有考虑和限制比照适用于这些多肽和核苷酸的序列同一性和官能度。

[0283] 在一方面,主题氨基酸序列是SEQ ID NO:1、2、3、4或5,并且主题核苷酸序列优选地是SEQ ID NO:9、10、11、12或13。

[0284] 在一方面,多肽是从SEQ ID NO:1、2、3、4或5的多肽的氨基和/或羧基末端缺失的一个或多个(若干个)氨基酸的片段;其中该片段具有转半乳糖基化活性。

[0285] 在一方面,片段包含至少500、550、600、650、700、750、800、850、900、950或1000个氨基酸残基。

[0286] 在另一方面,多肽变体的长度为500至1300个氨基酸残基。在另一方面,多肽变体的长度为600至1300个氨基酸。在另一方面,多肽变体的长度为700至1300个氨基酸。在另一方面,多肽变体的长度为800至1300个氨基酸。在另一方面,多肽变体的长度为800至1300个氨基酸。

[0287] 为了评估变体在宿主细胞中的表达,测定可以测量所表达的蛋白质、相应的mRNA或β-半乳糖苷酶活性。例如,合适的测定包括使用适当标记的杂交探针的RNA印迹法和DNA印迹法、RT-PCR(逆转录酶聚合酶链式反应)和原位杂交。合适的测定还包括测量样品中的活性。该变体的活性的合适的测定包括但不限于基于ONPG的测定或确定例如像在此的方法和实例中所描述的反应混合物中的葡萄糖。

[0288] SEQ ID NO:1、2、3、4或5的多肽变体

[0289] 在一方面,使用相对于SEQ ID NO:1、2、3、4或5在一个或多个位置处具有取代的SEQ ID NO:1、2、3、4或5的变体,该取代影响改变的特性,如改进的转半乳糖基化。在本文件中为了方便起见,这样的变体多肽又称为“变体多肽”、“多肽变体”或“变体”。在一方面,与SEQ ID NO:1、2、3、4或5的多肽相比,在此定义的多肽具有改进的转半乳糖基化活性。在另一方面,与SEQ ID NO:1、2、3、4或5的多肽相比,如在此所定义的多肽具有改进的反应速度。

[0290] 在此使用的多肽和变体多肽包括转半乳糖基化活性。

[0291] 在一方面,在如高于3%w/w初始乳糖浓度的浓度下30min反应后,转半乳糖基化活性:β-半乳糖苷酶活性的比率为至少0.5,如至少1,如至少1.5,或如至少2。

[0292] 在一方面,在如高于3%w/w初始乳糖浓度的浓度下30min反应后,转半乳糖基化活性:β-半乳糖苷酶活性的比率为至少2.5、如至少3、如至少4、如至少5、如至少6、如至少7、如至少8、如至少9、如至少10、如至少11、或如至少12。

[0293] 在一方面,如在此定义的多肽和变体可衍生自微生物来源,特别是可衍生自丝状

真菌或酵母,或可衍生自细菌。该酶可以例如衍生自如下各项的菌株:伞菌属,例如双孢蘑菇;Ascovaginospora;曲霉属,例如黑曲霉、泡盛曲霉、臭曲霉、日本曲霉、米曲霉;念珠菌属;毛壳菌属;Chaetotomastia;网柄菌属,例如盘基网柄菌;克鲁维酵母菌属,例如脆壁克鲁维酵母、乳酸克鲁维酵母;毛霉属,例如爪哇毛霉、高大毛霉、细小毛霉;脉孢菌属,例如粗糙脉孢菌;根毛霉属,例如微小根毛霉;根霉属,例如少根根霉、日本根霉、匍枝根霉;核盘霉属,例如白腐核盘霉;圆酵母属;球拟酵母属;毛癣菌属,例如红色毛癣菌;维氏核盘菌属,例如大豆维氏核盘菌;芽孢杆菌属,例如凝结芽孢杆菌、环状芽孢杆菌、巨大芽孢杆菌、诺瓦利斯芽孢杆菌(B.novalis)、枯草芽孢杆菌、短小芽孢杆菌、嗜热脂肪芽孢杆菌、苏云金芽孢杆菌;双歧杆菌属,例如长双歧杆菌、两歧双歧杆菌、动物双歧杆菌;金黄杆菌属;柠檬酸杆菌属,例如弗氏柠檬酸杆菌;梭菌属,例如产气荚膜梭菌;色二孢属,例如棉色二孢;肠杆菌属,例如产气肠杆菌,阴沟肠杆菌;爱德华菌属,迟钝爱德华菌;欧文氏菌属,例如草生欧文氏菌;埃希氏菌属,例如大肠杆菌;克雷伯氏菌属,例如肺炎克雷伯氏菌;多球菌属(Mirriococcus);漆斑菌属;毛霉属;脉孢菌属,例如粗糙脉孢菌;变形杆菌属,例如普通变形杆菌;普罗维登斯菌属,例如斯氏普罗威登斯菌;密孔菌属,例如朱红密孔菌、血红密孔菌;瘤胃球菌属,例如扭链瘤胃球菌;沙门氏菌属,例如鼠伤寒沙门菌;沙雷氏菌属,例如液化沙雷氏菌、粘质沙雷氏菌;志贺菌属,例如弗氏志贺菌;链霉菌属,例如抗生链霉菌、栗色球孢链霉菌、紫红链霉菌;栓菌属;木霉属,例如里氏木霉、绿色木霉;耶尔森氏菌属,例如小肠结肠炎耶尔森氏菌。

[0294] 提供包括如在此定义的多肽或变体多肽的分离的和/或纯化的多肽。在一个实施例中,变体多肽是多肽(SEQ ID NO:1、2、3、4或5)的成熟形式。在一方面,这些变体包括C端结构域。

[0295] 在一方面,如在此定义的变体多肽包括变体,其中相对于SEQ ID NO:1、2、3、4或5已经添加或缺失了在一个和约25个之间的氨基酸残基。在一方面,如在此定义的变体多肽包括变体,其中相对于SEQ ID NO:1、2、3、4或5已经取代、添加或缺失了在一个和25个之间的氨基酸残基。在一方面,该变体具有SEQ ID NO:1、2、3、4或5的氨基酸序列,其中已经取代了在一个和约25个之间任何数目的氨基酸。在另一方面,该变体具有SEQ ID NO:1、2、3、4或5的氨基酸序列,其中已经取代了在三个和十二个之间任何数目的氨基酸。在另一方面,该变体具有SEQ ID NO:1、2、3、4或5的氨基酸序列,其中已经取代了在五个和九个之间任何数目的氨基酸。

[0296] 在一方面,SEQ ID NO:1、2、3、4或5中的至少两个,在另一方面,至少三个,并且在又另一方面,至少五个氨基酸已经被取代。

[0297] 在一方面,在此披露的一个或多个多肽具有1、2、3、4或5的序列。

[0298] 在一方面,在此披露的一个或多个多肽具有SEQ ID NO:1、2、3、4或5的序列,其中N末端中的10,如9、如8、如7、如6、如5、如4、如3、如2、如1个氨基酸被取代和/或缺失。

[0299] 其酶和酶变体可以通过以下表征:其核酸和初级多肽序列、三维结构建模和/或其比活性。如在此所定义的多肽或多肽变体的另外的特征包括例如稳定性、pH范围、氧化稳定性和热稳定性。可以使用本领域技术人员已知的标准测定来评估表达水平和酶活性。在另一方面,变体表现出相对于具有SEQ ID NO:1、2、3、4或5的多肽的改进的性能特征,如在高温,例如,65°C-85°C下改进的稳定性。

[0300] 提供如在此定义的具有如下氨基酸序列的多肽变体,该氨基酸序列与SEQ ID NO: 1、2、3、4或5的多肽具有至少约66%、68%、70%、72%、74%、78%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%同一性。

[0301] 核苷酸

[0302] 在一方面,本发明使用如上所述的具有转半乳糖基化活性的分离的多肽,该多肽由如下多核苷酸编码,这些多核苷酸在非常低的严格条件,优选地低严格条件,更优选地中严格条件,更优选地中-高严格条件,甚至更优选地高严格条件,以及最优选地非常高严格条件下与i)包含在SEQ ID NO:9、10、11、12或13中的编码SEQ ID NO:1、2、3、4或5的成熟多肽的核酸序列;ii) i)的cDNA序列;或iii) i)或ii)的互补链杂交(J.萨姆布鲁克,E.F.弗里奇(Fritsch),和T.马尼亚蒂斯(Maniatis),1989,分子克隆,实验室手册(Molecular Cloning, A Laboratory Manual),第2版,冷泉港,纽约州)。SEQ ID NO:9、10、11、12或13的子序列包含至少100个连续核苷酸或优选至少200个连续核苷酸。此外,子序列可以编码具有乳糖酶活性的多肽片段。

[0303] 可以使用SEQ ID NO:9、10、11、12或13的核苷酸序列或其子序列、连同SEQ ID NO: 1、2、3、4或5的氨基酸序列或其片段来设计核酸探针,以根据本领域众所周知的方法来鉴定并克隆对来自不同属或物种的菌株的具有转半乳糖基酶活性的多肽进行编码的DNA。具体地说,这类探针可以用于按照标准DNA印迹程序与感兴趣的属或物种的基因组或cDNA杂交,以便鉴定并分离其中的相应基因。这些探针可以比完整序列短得多,但是长度应当是至少14个、优选至少25个、更优选至少35个、且最优选至少70个核苷酸。然而,优选的是核酸探针的长度是至少100个核苷酸。例如,核酸探针的长度可以是至少200个核苷酸,优选至少300个核苷酸,更优选至少400个核苷酸,或最优选至少500个核苷酸。可使用甚至更长的探针,例如长度为至少600个核苷酸,至少优选至少700个核苷酸,更优选至少800个核苷酸,或最优选至少900个核苷酸的核酸探针。DNA和RNA探针都可以使用。通常将探针标记以探测相应的基因(例如,用³²P、³H、³⁵S、生物素或抗生物素蛋白标记)。这类探针涵盖于本发明中。

[0304] 因此,可以针对与以上所描述的探针杂交并且编码具有乳糖酶活性的多肽的DNA对由这类其他生物体制备的基因组DNA文库进行筛选。可通过琼脂糖或聚丙烯酰胺凝胶电泳、或其他分离技术分离来自这类其他生物的基因组或其他DNA。可以将来自文库的DNA或分离的DNA转移并且固定在硝酸纤维素(nitrocellulose)或其他适合的载体材料上。为鉴定与SEQ ID NO:9、10、11、12或13或其子序列同源的克隆或DNA,在DNA印迹法中使用载体材料。

[0305] 出于本发明的目的,杂交表明该核苷酸序列在非常低到非常高严格条件下与一种被标记的核酸探针杂交,该探针对应于SEQ ID NO:9、10、11、12或13中所示的核苷酸序列、其互补链、或其子序列。可以使用X射线胶片检测在这些条件下与核酸探针杂交的分子。

[0306] 该核酸探针可以是SEQ ID NO:9、10、11、12或13的成熟多肽编码区域。

[0307] 对于至少100个核苷酸长度的长探针,非常低至非常高严格性条件被定义为最佳地按照标准DNA印迹程序,在42℃下在5X SSPE、0.3% SDS、200g/ml剪切和变性的鲑鱼精子DNA中,以及对于非常低和低严格性在25%甲酰胺中,对于中和中-高严格性在35%甲酰胺中,或对于高和非常高严格性在50%甲酰胺中预杂交和杂交12至24小时。

[0308] 对于长度为至少100个核苷酸的长探针而言,载体材料最终使用2X SSC、0.2%

SDS, 优选至少在45℃下(非常低严格性), 更优选至少在50℃下(低严格性), 更优选至少在55℃下(中严格性), 更优选至少在60℃下(中-高严格性), 甚至更优选至少在65℃下(高严格性), 并且最优选至少在70℃下(非常高严格性)洗涤三次, 每次15分钟。

[0309] 在一个具体实施例中, 洗涤使用0.2X SSC、0.2% SDS, 优选至少在45℃下(非常低严格性), 更优选至少在50℃下(低严格性), 更优选至少在55℃下(中严格性), 更优选至少在60℃下(中高严格性), 甚至更优选至少在65℃下(高严格性), 并且最优选至少在70℃下(非常高严格性)进行。在另一个具体实施例中, 洗涤使用0.1X SSC、0.2% SDS, 优选至少在45℃下(非常低严格性), 更优选至少在50℃下(低严格性), 更优选至少在55℃下(中严格性), 更优选至少在60℃下(中高严格性), 甚至更优选至少在65℃下(高严格性), 并且最优选至少在70℃下(非常高严格性)进行。

[0310] 对于长度为约15个核苷酸至约70个核苷酸的短探针而言, 严格条件被定义为根据标准DNA印迹程序, 在比使用根据博尔顿(Bolton)和麦卡锡(McCarthy) (1962, 美国国家科学院院刊(Proceedings of the National Academy of Sciences USA) 48:1390)的计算计算的 T_m 低约5℃至约10℃下, 在0.9M NaCl、0.09M Tris-HCl (pH 7.6)、6mM EDTA、0.5% NP-40、1X登哈特氏溶液(Denhardt's solution)、1mM焦磷酸钠、1mM磷酸二氢钠、0.1mM ATP、以及每ml 0.2mg的酵母RNA中预杂交、杂交及杂交后洗涤。

[0311] 对于长度为约15个核苷酸至约70个核苷酸的短探针而言, 将载体材料在比计算的 T_m 低5℃至10℃下, 在6X SSC加0.1% SDS中洗涤一次, 持续15分钟, 并且使用6X SSC洗涤两次, 每次15分钟。

[0312] 在包含盐的杂交条件下, 有效 T_m 控制了用于成功杂交的探针和滤器结合的DNA之间所需的同一性程度。可以使用以下公式确定有效 T_m , 以确定两种DNA在各种严格条件下杂交所需的同一性程度。

[0313] 有效 $T_m = 81.5 \pm 16.6 (\log M [Na^+]) + 0.41 (G+C\%) - 0.72 (\text{甲酰胺}\%)$

[0314] (参见www.ndsu.nodak.edu/instruct/mcclean/plsc731/dna/dna6.htm)

[0315] SEQ ID NO:10的G+C含量为42%, 并且SEQ ID NO:11的G+C含量为44%。对于中严格性, 甲酰胺为35%, 并且5X SSPE的 Na^+ 浓度为0.75M。

[0316] 另一个相关关系是两个DNA的1%错配将 T_m 降低1.4℃。为了确定两个DNA在42℃的中严格条件下杂交所需的同一性程度, 使用下列公式:

[0317] 同源性% = $100 - [(有效T_m - 杂交温度) / 1.4]$

[0318] (参见www.ndsu.nodak.edu/instruct/mcclean/plsc731/dna/dna6.htm)

[0319] 变体核酸包括与编码SEQ ID NO:1、2、3、4或5的核酸具有一定百分比, 例如60%、65%、70%、75%、80%、85%、90%、95%、或99%序列同一性的多核苷酸。在一方面, 提供了能够编码如在此披露的多肽的核酸。在另一方面, 在此披露的核酸具有与SEQ ID NO:9、10、11、12或13具有至少60%, 如至少65%, 如至少70%, 如至少75%, 如至少80%, 如至少85%, 如至少90%, 如至少95%, 如至少99%同一性的核酸序列。

[0320] 可以使用包括如在此所述的核酸的质粒。

[0321] 可以使用包括在此所述的核酸或能够表达如在此所述的多肽的表达载体。

[0322] 提供了与编码本文中所定义的任意多肽变体的核酸互补的核酸。另外, 提供了能够与补体杂交的核酸。在另一个实施例中, 用于本文所述方法和组合物的序列是合成序列。

它包括但不限于：用最佳密码子使用制备的用于在宿主生物体如酵母中表达的序列。

[0323] 根据本领域众所周知的程序，如在此提供的多肽变体可以经合成或通过在宿主细胞中重组表达来产生。在一方面，在此披露的一种或多种多肽是一种或多种重组多肽。任选地将如在此定义的所表达的多肽变体在使用前分离。

[0324] 在另一个实施例中，如在此定义的多肽变体在表达后进行纯化。例如，在美国专利号7,371,552、7,166,453、6,890,572、和6,667,065；以及美国公开申请号2007/0141693、2007/0072270、2007/0020731、2007/0020727、2006/0073583、2006/0019347、2006/0018997、2006/0008890、2006/0008888、和2005/0137111中描述了多肽变体的遗传修饰和重组生产方法。这些披露的相关教导通过引用并入本文，这些相关教导包括编码多肽的多核苷酸序列、引物、载体、选择方法、宿主细胞、所表达的多肽变体的纯化和重构以及如在此定义的多肽变体的表征，包括有用的缓冲液、pH范围、Ca²⁺浓度、底物浓度和用于酶法测定的酶浓度。

[0325] 提供编码SEQ ID NO:1、2、3、4或5的蛋白质的核酸序列或与编码SEQ ID NO:1、2、3、4或5的蛋白质的核酸具有至少约66%、68%、70%、72%、74%、78%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在一个实施例中，核酸序列与SEQ ID NO:9、10、11、12或13的核酸具有至少约60%、66%、68%、70%、72%、74%、78%、80%、85%、90%、95%、96%、97%、98%或99%序列同一性。

[0326] 载体

[0327] 在一方面，本发明使用包括多核苷酸的载体。在一方面，细菌细胞包括载体。在一些实施例中，将包括编码变体的核酸的DNA构建体在包括与编码序列可操作地连接的调节序列的表达载体中转移到宿主细胞。载体可以是整合到真菌宿主细胞基因组中并且当被引入宿主细胞时进行复制的任何载体。密苏里大学FGSC菌株目录列出合适的载体。萨姆布鲁克等人，分子克隆：实验室手册，第3版，冷泉港实验室出版社，冷泉港，纽约州（2001）；班尼特（Bennett）等人，真菌中更多的基因操作（More Gene Manipulations in Fungi），学术出版社，圣迭哥（1991），第396-428页；和美国专利号5,874,276中提供了合适的表达和/或整合载体的另外的实例。示例性载体包括pFB6、pBR322、PUC18、pUC100和pENTR/D、pDONTM201、pDONRTM221、pENTRTM、pGEM[®]3Z以及pGEM[®]4Z。用于细菌细胞的示例包括允许在大肠杆菌中复制的pBR322和pUC19，以及例如允许在芽孢杆菌属中复制的pE194。

[0328] 在一些实施例中，编码变体的核酸可操作地连接到合适的启动子，其允许在宿主细胞中转录。启动子可以衍生自编码与宿主细胞同源或异源的蛋白质的基因。启动子的合适的非限制性实例包括cbh1、cbh2、egl1、和egl2启动子。在一个实施例中，该启动子是对宿主细胞天然的启动子。例如，当宿主为嗜糖假单胞菌时，启动子为天然的嗜糖假单胞菌启动子。“诱导型启动子”是指在环境调节或发育调节下具有活性的启动子。在另一个实施例中，该启动子是对宿主细胞异源的启动子。

[0329] 在一些实施例中，将编码序列可操作地连接到编码信号序列的DNA序列上。在另一方面，代表性的信号肽是SEQ ID NO:27。代表性的信号肽是SEQ ID NO:9，该SEQ ID NO:9是枯草芽孢杆菌aprE前体的天然信号序列。在其他实施例中，将编码信号序列的DNA用编码来自其他细胞外枯草芽孢杆菌前体的信号序列的核苷酸序列替代。在一个实施例中，编码信号序列的多核苷酸紧邻编码多肽的多核苷酸的上游并在框内。信号序列可以选自与宿主细

胞相同的物种。

[0330] 在另外的实施例中，包扩待引入真菌宿主细胞的DNA构建体或载体的信号序列和启动子序列衍生自相同的来源。在一些实施例中，表达载体还包括终止序列。在一个实施例中，终止序列和启动子序列衍生自相同的来源。在另一个实施例中，终止序列与宿主细胞是同源的。

[0331] 在一些实施例中，表达载体包括选可选择的标志物。合适的可选择的标志物的实例包括赋予抗微生物剂抗性的那些，例如潮霉素或腐草霉素。营养选择性标志物也是合适的，并且包括amdS、argB和pyr4。在一个实施例中，选择性标志物是编码乙酰胺酶的amdS基因；它允许转化的细胞在作为氮源的乙酰胺上生长。在凯利 (Kelley) 等人，欧洲分子生物学学会杂志 (EMBO J.) 4:475-479 (1985) 和彭蒂拉 (Penttila) 等人，基因 (Gene) 61:155-164 (1987) 中描述了构巢曲霉amdS基因作为选择性标志物的用途。

[0332] 包含具有编码变体的多核苷酸的DNA构建体的合适的表达载体可以是能够在给定宿主生物体中自主复制的或整合到宿主DNA中的任何载体。在一些实施例中，表达载体是质粒。在一些实施例中，考虑了两种类型的表达载体，用于获得基因表达。第一表达载体包括DNA序列，其中启动子、编码区和终止子全部来源于待表达基因。在一些实施例中，在其自身的转录和翻译调节序列的控制下，通过删除不想要的DNA序列以留下待表达的结构域而获得基因截短。预先组装第二种类型表达载体，并且包含高水平转录所需的序列和可选择的标志物。在一些实施例中，将基因或其部分的编码区插入到该通用表达载体中，这样使得它处于表达构建体启动子和终止子序列的转录控制下。在一些实施例中，将基因或其部分插入强cbh1启动子的下游。

[0333] 表达宿主/宿主细胞

[0334] 在另一方面，使用包括如在此所述的质粒或如在此所述的表达载体的，优选经其转化的，宿主细胞。

[0335] 在另一方面，使用能够表达如在此所述的多肽的细胞。

[0336] 在一方面，如在此所述的宿主细胞或如在此所述的细胞是细菌、真菌或酵母细胞。

[0337] 在另一方面，宿主细胞选自由以下各项组成的组：瘤胃球菌属、双歧杆菌属、乳球菌属、乳杆菌属、链球菌属、明串珠菌属、埃希氏杆菌属、芽孢杆菌属、链霉菌属、酵母菌属、克鲁维酵母菌属、念珠菌属、圆酵母属、球拟酵母属和曲霉属。

[0338] 在另一方面，该宿主细胞选自由以下各项组成的组：汉逊瘤胃球菌 (*Ruminococcus hansenii*)、短双歧杆菌、长双歧杆菌、婴儿双歧杆菌、两歧双歧杆菌和乳酸乳球菌。

[0339] 在另一个实施例中，合适的宿主细胞包括选自下组的革兰氏阳性细菌，该组由以下各项组成：枯草芽孢杆菌、地衣芽孢杆菌、迟缓芽孢杆菌、短芽孢杆菌、嗜热芽孢杆菌、嗜碱芽孢杆菌、解淀粉芽孢杆菌、凝结芽孢杆菌、环状芽孢杆菌、灿烂芽孢杆菌、苏云金芽孢杆菌、变铅青链霉菌、或鼠灰链霉菌；或革兰氏阴性细菌，其中所述革兰氏阴性细菌为大肠杆菌或假单胞菌属物种。在一方面，宿主细胞是枯草芽孢杆菌或地衣芽孢杆菌。在一个实施例中，宿主细胞是枯草芽孢杆菌，并且所表达的蛋白质被工程化以包括枯草芽孢杆菌信号序列，如以下进一步详细阐述的。

[0340] 在一些实施例中，宿主细胞被遗传工程化以表达如在此定义的具有如下氨基酸序列的多肽变体，该氨基酸序列与SEQ ID NO:1、2、3、4或5的多肽具有至少约66%、68%、

70%、72%、74%、78%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%同一性。在一些实施例中,编码如在此定义的多肽变体的多核苷酸将具有编码SEQ ID NO:1、2、3、4或5的蛋白质的核酸序列或与编码SEQ ID NO:1、2、3、4或5的蛋白质的核酸具有至少约66%、68%、70%、72%、74%、78%、80%、85%、90%、95%、96%、97%、98%、99%或100%序列同一性的核酸序列。在一个实施例中,核酸序列与SEQ ID NO:9、10、11、12或13的核酸具有至少约60%、66%、68%、70%、72%、74%、78%、80%、85%、90%、95%、96%、97%、98%或99%序列同一性。

[0341] 诱变

[0342] 可以通过使用重组遗传操作技术的基因工程、使宿主经受诱变、或两者来获得纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型宿主细胞或其中这些酶基本上是无活性的宿主细胞。编码本发明的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶的基因的修饰或失活可能是由于使亲本细胞经受诱变和选择突变细胞造成的,在突变细胞中与亲本细胞相比,表达这些酶的能力已经降低。诱变可能是特异性的或随机的,可以通过例如使用适合的物理或化学诱变剂进行,通过使用适合的寡核苷酸进行,或通过所述DNA序列经受PCR产生的诱变来进行。此外,诱变可通过使用这些诱变剂的任意组合来进行。

[0343] 适用于本发明目的的物理或化学诱变剂的实例包括 γ 或紫外线(UV)照射、羟胺、N-甲基-N'-硝基-N-亚硝基胍(MNNG)、邻甲基羟胺、亚硝酸、乙基甲磺酸(EMS)、亚硫酸氢钠、甲酸和核苷酸类似物。当使用这些试剂时,通常通过如下方法来进行诱变:在合适的条件下在选择的诱变剂的存在下孵育待诱变的亲本细胞,并选择表现出基因表达降低的突变细胞。可替代地,可以使用遗传技术如杂交或交配、和原生质体融合或任何其他经典遗传技术来诱导遗传多样性来分离这些菌株。随后可以通过监测酶的表达水平来选择所获得的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型菌株。任选地,随后通过测量宿主细胞中待表达的给定的感兴趣基因的表达水平来选择纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型菌株。具有降低的酶活性的菌株的选择可以通过直接测量培养肉汤、培养上清液、透化细胞或细胞裂解物中的酶活性来进行。

[0344] 重组DNA技术

[0345] 可替代地,可以使用重组DNA技术构建具有减少量的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的宿主细胞或其中这些酶基本上无活性的宿主细胞。本领域中描述了基因失活或基因破坏的若干种技术,例如一步基因破坏法、标志物插入、定点诱变、缺失、RNA干扰、反义RNA等,并且都可以用于降低、抑制或干扰纤维素酶、甘露聚糖酶和果胶酶的合成以及任选地干扰淀粉酶活性,以便获得具有降低的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的工业生产菌株。此外,通过改变指导纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶基因的表达的一个或多个控制序列使纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶失活是本发明的一部分。其实例是通过基因破坏降低启动子活性。

[0346] 使用现代遗传修饰技术,人们可以获得重组纤维素酶、甘露聚糖酶和果胶酶,以及任选地淀粉酶缺陷型菌株,优选通过干扰编码纤维素酶、甘露聚糖酶和果胶酶,以及任选地淀粉酶活性的基因,更优选地通过将标记基因插入编码酶活性的基因,最优选地通过从基因组去除部分或全部酶编码区进行。已经针对许多不同的微生物描述了进行这种基因失活的方法,并且是本领域技术人员已知的(即,参见EP 357127)。因此可以减少或消除纤维素

酶、甘露聚糖酶和果胶酶以及任选地淀粉酶在突变细胞中表达。取决于使用这些技术修饰的宿主菌株,该过程可重复若干次以除去所有或大部分纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶编码序列。

[0347] 宿主基因如纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶的修饰或失活可以通过已建立的反义技术、使用与该基因的核苷酸序列互补的核苷酸序列进行。更具体地,可以通过引入与核苷酸序列互补的核苷酸序列来减少或消除基因的表达,该核苷酸序列可以在细胞中转录并且能够与细胞中产生的mRNA杂交。在允许互补的反义核苷酸序列与mRNA杂交的条件下,由此降低或消除了所翻译的蛋白质的量。尼安(Ngiam)等人(应用与环境微生物学(Appl. Environ. Microbiol.) 66:775-782, 2000)和茨伦纳(Zrenner)等人(普兰塔(Planta) 190:247-252, 1993)提供了表达反义RNA的实例。

[0348] 可以通过RNA干扰(RNAi)技术(欧洲微生物学会联合会微生物学快报(FEMS Microb. Lett.) 237:317-324, 2004)获得宿主基因的修饰、下调或失活。更具体地说,丝状真菌细胞基因表达可以通过如下来减少或消除:克隆其表达待受影响的核苷酸序列的同一的正义和反义部分(其后各自具有位于其间的核苷酸间隔物),插入表达载体,并且将表达载体引入细胞中,在该细胞中双链RNA(dsRNA)可以被转录,并且然后加工成能够杂交至靶mRNA的较短的siRNA。dsRNA转录后,小(21-23)核苷酸siRNA片段的形成将导致待受影响的mRNA的靶向降解。特异性mRNA的消除可以是各种程度的。WO 2005/05672和WO 2005/026356中描述的RNA干扰技术可用于宿主基因的修饰、下调或失活。

[0349] 纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型宿主细胞可能具有另一个核苷酸序列,该宿主细胞已经通过上述任何方法修饰或失活,并且在相同条件下培养时产生比亲本细胞更少的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性,如使用与之前定义的相同测定所测量的。

[0350] 具有降低的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的,通过经典遗传技术或重组DNA技术分离的或构建的这些工业生产菌株可以用于需要最终产品包含膳食纤维的相关工业方法。优选地,这些菌株用于生产具有转半乳糖基化活性的工业上相关的酶。更优选地,这些菌株用于生产食品工业中使用的酶,甚至更优选地,这些酶用于奶产品的加工。最优选地,将具有降低的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的这种工业生产菌株用于从乳糖生产GOS。

[0351] 优选地,本发明的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型宿主细胞是如在模式反应中所检测的具有少于50%的可检测的细胞内或细胞外纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的菌株(参见实例2、3或4中的实验信息)。更优选地,本发明的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型菌株是具有少于50%的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的菌株。更优选地,本发明的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型菌株是具有如下纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性的菌株,如在模式反应中所检测的,该活性少于其来源的宿主细胞的纤维素酶、甘露聚糖酶和果胶酶及任选地淀粉酶活性的25%,优选地少于10%,更优选地少于5%,更优选地少于1%,并且最优选地该纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶活性在本发明的缺陷型宿主细胞中是不可检测的。

[0352] 多种用于检测多肽的系统是本领域技术人员已知的。检测系统包括用于检测多肽

或酶活性的任何可能的测定。举例来说,这些测定系统包括但不限于基于以下测定:比色、光度、荧光、比浊、粘度、免疫学、生物学、色谱和其他可用测定。

[0353] 优选地,如果产生的多肽是酶,则通过在模式反应中测量其活性来确定产生的活性酶的量(参见实例2、3或4)。

[0354] 根据另一个优选的实施例,本发明的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型宿主细胞的特征在于如下事实:当该菌株已经用包括编码具有转半乳糖基化活性的多肽的基因的表达式构建体转化时,所述菌株至少产生其来源的野生型菌株在相同培养条件下产生的多肽的量(当野生型菌株也已经用与纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型宿主细胞相同的表达式构建体转化时)。

[0355] 优选地,本发明的纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型菌株是在相同培养条件下产生比其来源的野生型菌株相同或更多量的具有转半乳糖基化活性的多肽的菌株。更优选地,纤维素酶、甘露聚糖酶和果胶酶以及任选地淀粉酶缺陷型菌株比在相同培养条件下其来源的野生型菌株产生更多的给定多肽。

[0356] 具有转半乳糖基化活性的多肽的生产

[0357] 根据又一个实施例,本发明涉及在缺乏纤维素、甘露聚糖酶、果胶酶和/或淀粉酶活性的宿主细胞中转录核苷酸序列的方法,其中该转录序列编码具有转半乳糖基化活性的多肽,该方法包括:

[0358] (a) 在培养基中培养本发明的宿主细胞,该宿主细胞包括(i) 启动子,

[0359] (iv) 编码多肽的下游核苷酸序列,(iii) 翻译停止信号和(iv) 转录终止信号,

[0360] (b) 在宿主细胞中表达多肽,并且

[0361] (c) 任选地,从培养基或从宿主细胞中回收多肽。

[0362] 优选地,根据本发明的方法产生缺陷型菌株。可以使用本领域已知的方法将缺陷型菌株在适于产生所需多肽的营养介质中生长或维持。例如,可以将细胞接种在固体基质上,在烧瓶中摇动,在小规模或大规模发酵(包括连续、分批、补料分批或固态发酵)中在实验室或工业发酵罐中在适合的介质中并且在允许表达和/或分离多肽的条件下培养。使用本领域已知的程序,在包括碳源和氮源和无机盐的合适营养介质中进行培养(参见,例如,班尼特和拉热(LaSure)编辑,真菌中更多的基因操作,学术出版社,加利福尼亚州,1991)。适合的介质可从商业供应商获得或可以使用公开的组合物(例如,在美国典型培养物保藏中心的目录中)来制备。如果多肽分泌到该营养介质中,那么可直接从介质中回收多肽。如果多肽不被分泌,那么其可从细胞裂解液中回收。

[0363] 可以通过本领域已知的方法来分离该所得的多肽。例如,可以通过常规程序从营养介质中分离多肽,这些常规程序包括但不限于离心、过滤、提取、喷雾干燥、蒸发或沉淀。然后可以通过本领域已知的多种程序进一步纯化分离的多肽,这些程序包括但不限于色谱法(例如,离子交换、亲和力、疏水性、色谱聚焦或尺寸排阻),电泳(例如,制备型等电聚焦),差异溶解度(例如,丙酮或硫酸铵沉淀)或萃取(例如离液剂、盐或pH)。参见,例如詹森(Janson)和赖登(Ryden)编辑,蛋白质纯化(Protein Purification),VCH出版商,纽约,1989。

[0364] 可以使用特异性针对具有转半乳糖基化活性的多肽的本领域已知的方法来检测该多肽。这些检测方法可以包括使用特异性抗体、酶产物的形成、酶底物的消失或SDS-

PAGE。例如,可以使用酶测定来确定多肽的活性。对于许多酶,用于确定酶活性的程序在本领域中是已知的。

[0365] 用于生产多肽的方法

[0366] 在另一方面,如在此所述的表达多肽的方法包括获得如在此所述的宿主细胞或细胞,并从该细胞或宿主细胞表达具有转半乳糖基化活性的多肽,以及任选地纯化该多肽。可以在本发明中使用这样的多肽。

[0367] 宿主细胞的转化、表达和培养

[0368] 将DNA构建体或载体引入宿主细胞中包括如下技术如:转化;电穿孔;核显微注射;转导;转染,例如脂质介导的转染和DEAE-糊精介导的转染;用磷酸钙DNA沉淀孵育;用DNA包被的微粒高速轰击;和原生质体融合。通用转化技术是本领域已知的。参见,例如,奥苏贝尔等人(1987),同上,第9章;萨姆布鲁克等人(2001),同上;和坎贝尔(Campbell)等人,当代遗传学(Curr.Genet.) 16:53-56(1989)。

[0369] 本领域已知的方法可用于选择转化体。

[0370] 用于固定和配制多肽和多肽组合物的方法

[0371] 多肽组合物可以根据本领域已知的方法制备,并可以是液体或干组合物的形式。例如,该多肽组合物可以呈颗粒或微粒的形式。包括在该组合物中的多肽可以根据本领域中已知的方法稳定化。

[0372] 应用

[0373] 以下给出了本发明的多肽或多肽组合物的优选用途的实例。

[0374] 在一方面,在此披露了通过用在如此所述的多肽或多肽组合物处理包括乳糖的底物来生产食品的方法。

[0375] 在一方面,在此披露了通过用如在此所述的多肽或多肽组合物处理包括乳糖的乳基底物来生产奶产品的方法。

[0376] 在一方面,将包括乳糖的底物进一步用水解的 β -半乳糖苷酶处理。

[0377] 酶制剂(如处于根据本发明制备的食物成分的形式)可以是溶液形式或作为固体-这取决于用途和/或应用模式和/或给予模式。固体形式可以是呈干燥酶粉末或呈造粒酶。

[0378] 干酶配制品的实例包括喷雾干燥的产物,混合器造粒产物,层状产物如流化床颗粒,挤出或颗粒状颗粒造粒产物,冻干产物。

[0379] 酶制剂(如处于根据本发明制备的食物成分的形式)可以是溶液形式或作为固体-这取决于用途和/或应用模式和/或给予模式。固体形式可以是呈干燥酶粉末或呈造粒酶。

[0380] 在一方面,提供组合物,优选食物组合物,更优选包括如在此所述的细胞或多肽或多肽组合物的奶产品。

[0381] 此外,基于在组合物中与SEQ ID NO:22具有至少70%,例如像72%、74%、74%、78%、80%、82%、84%、86%、88%、90%序列同一性的多肽的总量,在此披露了包括至少5%,例如像,10%、15%、20%、25%、30%、35%、40%、45%、50%w/w的一种或多种多肽的组合物。这可以通过使用本领域技术人员已知的以下技术来评估。使待评估的样品进行SDS-PAGE,并使用适合于蛋白质定量的染料例如像伯乐标准(Bio-Rad Criterion)系统进行可视化。然后使用适当的光密度扫描仪(例如像伯乐标准系统)扫描凝胶,并确保所得图像处于动态范围内。对与衍生自SEQ ID NO:8的任何变体/片段相对应的条带进行定量,并

且如下计算多肽的百分比:所讨论的多肽的百分比=所讨论的多肽/(表现出转半乳糖基化活性的所有多肽的总和)*100。组合物中衍生自SEQ ID NO:8的多肽变体/片段的总数可以通过本领域技术人员已知的方法使用多克隆抗体通过蛋白质印迹检测衍生自SEQ ID NO:8的片段来确定。

[0382] 在一方面,根据本发明的组合物包括选自下组的一种或多种多肽,该组由SEQ ID NO:1、2、3、4和5组成的多肽组成。在另一方面,组合物包括一种或多种选自下组的多肽,该组由SEQ ID NO:1、2和3组成的多肽组成。在又另一方面,组合物包括一种或多种选自下组的多肽,该组由SEQ ID NO:1和2组成的多肽组成。

[0383] 在一方面,本发明提供了包括根据本发明所述的酶复合物、酶载体和任选地稳定剂和/或防腐剂的酶复合物制剂。

[0384] 在本发明的又另一方面,酶载体选自下组,该组由甘油或水组成。在一个实施例中,酶载体不包括多元醇(例如,甘油、丙二醇或山梨糖醇)。

[0385] 在另一方面,制剂/组合物包括稳定剂。在一方面,该稳定剂选自由以下各项组成的组:无机盐、多元醇、糖及其组合。在一方面,该稳定剂是无机盐如氯化钾。在另一方面,该多元醇是甘油、丙二醇或山梨糖醇。在另一方面,该稳定剂不是多元醇,如甘油、丙二醇或山梨糖醇。在又另一方面,糖是小分子碳水化合物,特别是若干种甜味碳水化合物如葡萄糖、半乳糖、果糖和蔗糖中的任何一种。

[0386] 在又另一方面,制剂包括防腐剂。在一方面,该防腐剂是对羟基苯甲酸甲酯、对羟基苯甲酸丙酯、苯甲酸酯、山梨酸酯或其他食品认可的防腐剂或其混合物。

[0387] 本发明的方法可以用固定化酶实施,例如,固定化的乳糖酶或其他产生半乳寡糖的酶。可以将酶固定在任何有机或无机支持物上。示例性无机支持物包括氧化铝、硅藻土、Dowex-1氯化物、玻璃珠和硅胶。示例性有机支持物包括DEAE-纤维素、藻酸盐水凝胶或藻酸盐珠或等效物。在本发明的各个方面,可以通过物理吸附到无机支持物上来优化乳糖酶的固定化。可以将用于实施本发明的酶固定在不同的介质中,包括水、Tris-HCl缓冲液和磷酸盐缓冲溶液。可以将酶固定在任何类型的底物上,例如滤器、纤维、柱、珠、胶体、凝胶、水凝胶、网眼等。

[0388] 在一方面,提供了通过用如在此所述的多肽或多肽组合物处理包括乳糖的乳基底物来生产奶产品的方法。在另一方面,提供了通过将包括乳糖的乳基底物用一种多肽处理用于生产奶产品的方法,该多肽在15min反应后具有高于60%、如高于70%、如高于75%的相对转半乳糖基化活性。在一方面,在30min反应后相对转半乳糖基化活性高于3。在另一方面,在30min反应后相对转半乳糖基化活性高于6。在又另一方面,在30min反应后相对转半乳糖基化活性高于12。在一方面,提供了一种方法,其中在酶的活性的最佳温度下进行如在此所述的多肽或多肽组合物的处理。在另一方面,将多肽或多肽组合物以0.01-1000ppm的浓度添加到乳基底物中。在又另一方面,将多肽或多肽组合物以0.1-100ppm的浓度添加到乳基底物中。在另一方面,将多肽或多肽组合物以1-10ppm的浓度添加到乳基底物中。在一方面,提供了一种进一步包括用微生物发酵底物如奶产品的方法。在另一方面,该奶产品是酸奶。在另一方面,基本上同时进行多肽或多肽组合物和微生物的处理。在一方面,基本上同时将多肽或多肽组合物和微生物添加到乳基底物中。

[0389] 在一方面,提供包括如在此所述的细胞或多肽或多肽组合物的奶产品。在一方面,

以0.01-1000ppm的浓度添加如在此定义的多肽或多肽组合物。

[0390] 在一方面,提供包括由如在此定义的多肽或多肽组合物原位形成的GOS的奶产品。在一方面,提供了包括如在此定义的细胞的奶产品。

[0391] 如在此所述的奶产品可以是例如脱脂乳、低脂乳、全脂乳、奶油、UHT乳、具有延长的保质期的乳、发酵的奶产品、奶酪、酸奶、黄油、乳质涂抹料(dairy spread)、黄油乳、酸化乳饮料、酸奶油、乳清基饮料、冰淇淋、炼乳、牛奶焦糖酱(dulce de leche)或风味乳饮料。奶产品可以通过本领域已知的任何方法制造。

[0392] 奶产品还可以包括非乳组分,例如,植物组分如植物油,植物蛋白和/或植物碳水化合物。奶产品还可以包括另外的添加剂,如酶、调味剂、微生物培养物如益生菌培养物、盐、甜味剂、糖、酸、水果、果汁或本领域已知的可作为奶产品的组分或添加到奶产品中的任何其他组分。

[0393] 在本发明的一个实施例中,一种或多种乳组分和/或乳馏分占奶产品的至少50%(重量/重量),如至少70%,例如至少80%,优选地至少90%。

[0394] 在本发明的一个实施例中,已经用如在此定义的具有转半乳糖基化活性的酶处理的一种或多种乳基底物占奶产品的至少50%(重量/重量),如至少70%,例如至少80%,优选地至少90%。

[0395] 在本发明的一个实施例中,该奶产品是未通过添加预先生成的半乳寡糖而被浓缩的奶产品。

[0396] 在本发明的一个实施例中,多肽处理的乳基底物在用作奶产品中的成分之前未进行干燥。

[0397] 在本发明的一个实施例中,该奶产品是冰淇淋。在上下文中,冰淇淋可以是任何种类的冰淇淋,如全脂冰淇淋、低脂冰淇淋或基于酸奶或其他发酵的乳产品的冰淇淋。冰淇淋可以通过本领域已知的任何方法制造。

[0398] 在本发明的一个实施例中,该奶产品是乳或炼乳。

[0399] 在本发明的一个实施例中,该奶产品是UHT乳。在本发明的上下文中的UHT乳是已经进行灭菌程序的乳,其旨在杀死包括细菌孢子在内的所有微生物。UHT(超高温)处理可以是例如在130℃下热处理30秒,或者在145℃下热处理1秒。

[0400] 在本发明的一个优选实施例中,该奶产品是ESL乳。在上下文中,ESL乳是由于微量过滤和/或热处理而具有延长的保质期,并且能够在2℃-5℃的商店架上保持新鲜至少15天,优选至少20天的乳。

[0401] 在本发明的另一个优选实施例中,该奶产品是发酵的奶产品,例如酸奶。

[0402] 用于大多数发酵的乳产品的微生物选自下组,该组由通常称为乳酸菌的细菌组成。如在此所用,术语“乳酸菌”是指革兰氏阳性、微量需氧或厌氧细菌,该细菌发酵糖,并且产生酸,这些酸包括作为主要产生的酸的乳酸、乙酸和丙酸。工业上最有用的乳酸菌可在“乳杆菌(Lactobacillales)”目中找到,其中包括乳球菌属物种、链球菌属物种、乳杆菌属物种、明串珠菌属物种、假明串珠菌属物种、片球菌属物种、短杆菌属物种、肠球菌属物种和丙酸菌属物种。此外,乳酸菌组中通常包括属于厌氧细菌的组的产乳酸细菌,双歧杆菌,即双歧杆菌属物种,这些产乳酸细菌常单独或与乳酸菌组合用作食品培养物。

[0403] 乳酸菌通常作为冷冻或冻干培养物供应给奶业,用于进行起子增殖(bulk

starter propagation) 或者所谓的“直投式”(DVS) 培养物,旨在直接接种入发酵容器或桶中用于产生发酵奶产品。这种培养物通常称作“起子培养物”或“起子”。

[0404] 常用的乳酸菌的起子培养物菌株通常分成嗜常温生物体(其最适生长温度约30℃),和嗜热生物体(其最适生长温度在约40℃至约45℃的范围内)。属于嗜常温组的典型生物体包括乳酸乳球菌、乳酸乳球菌乳脂亚种(*Lactococcus lactis* subsp.cremoris)、肠膜状明串珠菌乳脂亚种(*Leuconostoc mesenteroides* subsp.cremoris)、肠膜状假明串珠菌乳脂亚种(*Pseudoleuconostoc mesenteroides* subsp.cremoris)、戊糖片球菌(*Pediococcus pentosaceus*)、乳酸乳球菌乳酸亚种双乙酰生物突变株(*Lactococcus lactis* subsp.lactis biovar.diacetylactis)、干酪乳杆菌干酪亚种(*Lactococcus casei* subsp.casei)和副干酪乳杆菌副干酪亚种(*Lactococcus paracasei* subsp.paracasei)。嗜热乳酸细菌物种包括例如嗜热链球菌、尿肠球菌、德氏乳杆菌乳酸亚种、瑞士乳杆菌、德氏乳杆菌保加利亚亚种和嗜酸乳杆菌。

[0405] 此外,属于双歧杆菌属的厌氧细菌,包括两歧双歧杆菌、动物双歧杆菌和长双歧杆菌常用作乳品起子培养物,并且通常包括在乳酸菌组中。此外,丙酸杆菌属物种用作乳品起子培养物,特别是在奶酪的制造中。此外,属于双歧杆菌属的生物体通常用作食物起子培养物。

[0406] 另一组微生物起子培养物是真菌培养物,包括丝状真菌的酵母培养物和培养物,其具体用于制造某些类型的奶酪和饮料。真菌的实例包括娄地青霉、白青霉(*Penicillium candidum*)、白地霉、开菲尔圆酵母、开菲尔酵母和酿酒酵母。

[0407] 在本发明的一个实施例中,用于乳基底物发酵的微生物是干酪乳杆菌、或嗜热链球菌和德氏乳酸杆菌保加利亚亚种的混合物。

[0408] 本发明的方法中使用的发酵方法是众所周知的,并且本领域的技术人员知道如何选择合适的工艺条件,如温度、氧、微生物的量和特征、添加剂如碳水化合物、调味剂、矿物质、酶、和处理时间。显然,选择发酵条件以便支持本发明的实现。

[0409] 作为发酵的结果,将降低乳基底物的pH。本发明的发酵奶产品的pH可以是,例如,在3.5-6的范围内,如在3.5-5的范围内,优选在3.8-4.8的范围内。

[0410] 在一方面,提供了使用多肽或多肽组合物或使用上述细胞类型中的任何一种或多种用于产生寡糖的方法。寡糖包括但不限于果寡糖、半乳寡糖、异麦芽寡糖、麦芽寡糖、乳果糖和木寡糖。

[0411] 在本发明的一个实施例中,寡糖通过在包括二糖底物的介质中孵育表达多肽的细胞而产生,该二糖底物例如像乳果糖、海藻糖、鼠李糖、麦芽糖、蔗糖、乳糖或纤维二糖。孵育在生产寡糖的条件下进行。细胞可以是选自下组的产物的一部分,该组由以下各项组成:酸奶、奶酪、发酵的乳产品、膳食补充剂和益生菌食物产品。可替代地,可以回收寡糖并随后在其制备之前或之后将其添加到感兴趣的产品中。

[0412] 在一方面,提供了在此披露的细胞用于生产选自下组的产品的用途,该组由以下各项组成:酸奶、奶酪、发酵的乳产品、膳食补充剂和益生菌食物产品。

[0413] 在一方面,可以将在此所述的多肽或多肽组合物用于制备奶酪产品并且可以在制备奶酪产品的方法中使用。例如,奶酪产品可以选自由以下各项组成的组:奶油奶酪、松软奶酪、和精制奶酪。通过添加多肽或多肽组合物,奶酪可以包含显著增加的半乳寡糖水平和

降低的乳糖水平。在一方面,最终奶酪产品中的乳糖水平可以降低至少约25%,优选至少约50%,并且更优选至少约75%。多肽或多肽组合物可以用于将乳糖产品中的乳糖降低至少于约1克/份,这是大多数乳糖不耐个体可以耐受的一个量。

[0414] 在此提供的奶酪产品是具有增加的可溶性纤维含量、降低的热量含量、优异的感官特性、改进的质地和风味的营养增强的奶酪产品。此外,在此所述的多肽可以降低奶酪产品的血糖指数,因为GOS比乳糖或其水解产物更慢地吸收。最后,多肽或多肽组合物可以降低奶酪产品,特别是奶油奶酪产品的生产成本,因为GOS令人惊奇地为奶油奶酪产品提供了改进的质地,从而允许减少稳定剂的使用,或通过允许增加水分含量而无需脱水。

[0415] 在另一方面,提供了包括如在此所述的多肽或多肽组合物和碳水化合物底物的组合物。在另一方面,该碳水化合物底物是二糖。另一方面,该二糖是例如乳果糖、海藻糖、鼠李糖、麦芽糖、蔗糖、乳糖或纤维二糖。在又另一方面,该碳水化合物底物是乳糖。制备组合物,这样使得产生寡糖。如在此所述的多肽可以是选自下组的产物的一部分,该组由以下各项组成:酸奶、奶酪、发酵的乳产品、膳食补充剂和益生菌食物产品。在一方面,提供了包括如在此所述的多肽和稳定剂的组合物。稳定剂的实例是例如多元醇,例如像甘油或丙二醇、糖或糖醇、乳酸、硼酸或硼酸衍生物(例如芳族硼酸酯)。

[0416] 在一方面,提供了如在此披露的转半乳糖基化多肽或多肽组合物或如在此披露的细胞用于产生半乳寡糖的用途。在一方面,提供了如在此披露的转半乳糖基化多肽或多肽组合物或如在此披露的细胞用于生产半乳寡糖的用途,该半乳寡糖将成为选自下组的产品的一部分,该组由以下各项组成:酸奶、奶酪、发酵的奶产品、膳食补充剂和益生菌食物产品。在一方面,该产品是酸奶、奶酪或发酵的奶产品。在一方面,提供了如在此披露的转半乳糖基化多肽或多肽组合物或如在此披露的细胞用于产生半乳寡糖以增强双歧杆菌生长的用途。在一方面,提供了如在此披露的转半乳糖基化多肽或多肽组合物或如在此披露的细胞用于产生半乳寡糖以增强双歧杆菌在混合培养物发酵中的生长的用途。

[0417] 在一方面,提供了用于生产如在此披露的转半乳糖基化多肽或多肽组合物的方法,该方法包括在允许表达所述多肽的条件下在适合的培养基中培养如在此披露的细胞,并且从培养物回收得到的多肽。提供了一种用于生产半乳寡糖的方法,该方法包括使如在此披露的多肽或多肽组合物或如在此披露的细胞与包括乳糖的乳基溶液接触。

[0418] 添加寡糖可以增强单独的双歧杆菌或在混合培养物中的双歧杆菌的生长。

[0419] 用将乳糖转化为单糖或GOS的酶处理乳产品具有若干个优点。首先,产品可以被患有乳糖不耐症的人消耗,否则会出现症状如胃肠气胀和腹泻。其次,与乳糖相比,用乳糖酶处理的奶产品将比类似的未处理产品具有更高的甜度,这是因为葡萄糖和半乳糖的更高的感知甜度。这种效果对于应用如酸奶和冰淇淋尤为有益,其中最终产品的高甜度是所希望的,并且这允许所消耗的产品中的碳水化合物的净减少。第三,在冰淇淋生产中,经常看到称为沙质的现象,其中乳糖分子由于乳糖的相对低的溶解度而结晶。当乳糖转化为单糖或GOS时,冰淇淋的口感比未处理产品大大改进。可以消除由于乳糖结晶引起的沙感的存在,并且可以通过用乳清粉替代脱脂乳粉来降低原料成本。酶处理的主要作用是增加甜味。

[0420] 在一方面,如在此披露的转半乳糖基化多肽或多肽组合物可以与其他酶如蛋白酶如凝乳酶或肾素,脂酶如磷脂酶,淀粉酶,转移酶和乳糖酶一起使用。在一方面,如在此披露的一种或多种转半乳糖基化多肽可与乳糖酶一起使用。当用如在此披露的一种或多种转半

乳糖基化多肽处理后,特别是在低乳糖水平下希望减少残留的乳糖时,这可以是特别有用的。

[0421] 在一个实施例中,酶是来自细菌的乳糖酶,例如,来自双歧杆菌科,如来自双歧杆菌属如特别是在WO 2009/071539和WO 2013/182686中描述的乳糖酶。

[0422] 实例1-根据本发明的宿主细胞的构建

[0423] 本实例的宿主细胞衍生自枯草芽孢杆菌。

[0424] 去除amyE α -淀粉酶基因:通过重组DNA技术将野生型基因(amyE)的体外产生的缺失引入枯草芽孢杆菌菌株中。通过活性以及DNA印迹法监测amyE缺失。宿主菌株中不存在异源DNA。

[0425] 通过壮观霉素标志物替代bg1C内切葡聚糖酶基因:为了使bg1C基因缺失,使用了采用cre-lox方法的策略。克隆上游和下游bg1C序列,并且将中间(待删除)部分替代为允许在转化中进行选择的loxP-壮观霉素盒。用构建体转化枯草芽孢杆菌菌株,选择壮观霉素,将基因组bg1C基因用克隆的变体替代。将来自包含壮观霉素标志物的菌株的基因组DNA整合到内切葡聚糖酶bg1C基因座的缺失中,用于通过自然感受态转化菌株。阳性克隆由壮观霉素抗性确定。壮观霉素抗性基因的侧翼是能够除去基因的loxP序列。

[0426] 用四环素标志物替代gmuG甘露聚糖酶基因:使用与针对bg1C基因相同的技术使甘露聚糖酶gmuG缺失,尽管这次使用四环素标志物而不是壮观霉素标志物。在包含四环素的介质上选择转化体,并且用PCR验证阳性克隆。

[0427] 用壮观霉素标志物替代pel果胶酸裂解酶基因:使用与用于bg1C基因相同的技术,通过用壮观霉素标志物替代pel ORF,使果胶酸裂解酶基因pel缺失。在壮观霉素平板上选择转化体,并通过cPCR筛选。

[0428] BIF917基因的引入:使用本出版物中阐述的方法将在WO 2013/182626中披露为“BIF917”的 β -半乳糖苷酶的编码序列转化到枯草芽孢杆菌中。一般来说,这描述了使用合成基因生产多肽,该合成基因被设计编码两歧双歧杆菌全长(1752个残基)基因,针对在枯草芽孢杆菌中的表达所优化的密码子购自GeneART(雷根斯堡,德国)SEQ ID NO.8。

[0429] 采用反向引物,使用聚合酶链式反应构建两歧双歧杆菌截短突变体,其允许合成基因的所选择的区域的特异性扩增。

[0430] 正向引物:GGGGTAACTAGTGGAAGATGCAACAAGAAG(加下划线的SpeI)(SEQ ID NO: 15)。

[0431] 截短突变体和相应的反向引物的SEQ ID如下表2所示。

[0432]	截短突变体	引物序列
	BIF917 (SEQ ID NO: 9)	GCGCTTAATTAATTATGTTTTTCTGTGCTTGTTT SEQ ID NO: 16
	BIF995 (SEQ ID NO: 10)	GCGCTTAATTAATTACAGTGCGCCAATTCATCAATCA SEQ ID NO: 17
	BIF1068 (SEQ ID NO: 11)	GCGCTTAATTAATTATTGAACTCTAATTGTCGCTG SEQ ID NO: 18
	BIF1172 (SEQ ID NO: 12)	
	BIF1241 (SEQ ID NO: 13)	GCGCTTAATTAATTATGTCGCTGTTTTCAGTTCAAT SEQ ID NO: 19
	BIF1326 (SEQ ID NO: 14)	GCGCTTAATTAATTAAAATTCTTGTTCTGTGCCCA SEQ ID NO: 20
	BIF 1478	GCGCTTAATTAATTATCTCAGTCTAATTCGCTTGCGC SEQ ID NO: 21

[0433] 表2

[0434] 使用独特的限制性位点SpeI和PacI将合成基因克隆到pBNspe枯草芽孢杆菌表达载体中,并且将分离的质粒转化到枯草芽孢杆菌菌株中。将转化体作为选择重新划线到包含10μg/mL新霉素的LB平板上。

[0435] 测量β-半乳糖苷酶活性

[0436] 使用市售的底物2-硝基苯基-β-D-吡喃半乳糖苷 (ONPG) (西格玛 (Sigma) N1127) 测量酶活性。

[0437] ONPG w/o受体

[0438] 100mM KP04pH 6.0

[0439] 12,3mM ONPG

[0440] 补充有受体的ONPG

[0441] 100mM KP04pH 6.0

[0442] 20mM纤维二糖

[0443] 12,3mM ONPG

[0444] 终止溶液

[0445] 10%Na₂CO₃

[0446] 将10μl稀释系列的纯化的酶添加到包含90μl ONPG缓冲液的具有或没有受体的微量滴定板的孔中。将样品混合并在37℃下孵育10min,然后向每个孔中添加100μl终止溶液以终止反应。在由Softmax软件包控制的分子装置SpectraMax平板读数器上在420nm下记录吸光度测量值。

[0447] 如下计算转半乳糖基化活性的比率:

[0448] 对于其中吸光度在0.5和1.0之间的稀释液,转半乳糖基化活性比率 = (Abs₄₂₀^{纤维二糖}/Abs₄₂₀^{纤维二糖})*100。

[0449] LAU活性的确定

[0450] 原理:

[0451] 该测定方法的原理是乳糖酶在37℃下将2-邻硝基苯基-β-D-吡喃半乳糖苷(ONPG)水解成2-邻硝基苯酚(ONP)和半乳糖。用碳酸钠停止反应,并在分光光度计或色度计在420nm下测量释放的ONP。

[0452] 试剂:

[0453] MES缓冲液pH 6.4(100mM MES pH 6.4,10mM CaCl₂):将19.52g MES水合物(Mw:195.2g/mol,西格玛-奥德里奇(Sigma-aldrich) #M8250-250G)和1.470g CaCl₂二水合物(Mw:147.01g/mol,西格玛-奥德里奇)溶解在1000ml ddH₂O中,通过10M NaOH将pH调节至6.4。通过0.2μm滤器过滤溶液,并在4℃下储存1个月。

[0454] ONPG底物pH 6.4(12.28mM ONPG,100mM MES pH 6.4,10mM CaCl₂):将0.370g 2-邻硝基苯基-β-D-吡喃半乳糖苷(ONPG,Mw:301.55g/mol,西格玛-奥德里奇#N1127)溶解在100ml MES缓冲液pH 6.4中,并在4℃下黑暗保存长达7天。

[0455] 终止试剂(10%Na₂CO₃):将20.0g Na₂CO₃溶解在200ml ddH₂O中,通过0.2μm滤器过滤溶液,并在室温下保存1个月。

[0456] 程序:

[0457] 在MES缓冲液(pH 6.4)中制备稀释系列的酶样品,并且将10μL每个样品稀释液转移到包含90μl ONPG底物(pH 6.4)的微量滴定板(96孔格式)的孔中。将样品混合并使用恒温混匀仪(舒适型恒温混匀仪(Comfort Thermomixer),艾本德(Eppendorf))在37℃下孵育5min,并且随后向每个孔中添加100μl终止试剂以终止反应。使用MES缓冲液(pH 6.4)代替酶样品构建空白。在ELISA读数器(SpectraMax平板读数器,分子装置公司(Molecular Device))上针对空白测量在420nm处的吸光度的增加。

[0458] 酶活性计算:

[0459] 确定MES缓冲液(pH 6.4)中2-邻硝基苯酚(西格玛-奥德里奇#33444-25G)的摩尔消光系数(0.5998x 10⁻⁶M⁻¹x cm⁻¹)。一个单位(U)的乳糖酶活性(LAU)被定义为对应于每分钟1nmol的ONPG水解的量。使用总反应体积为200μL(光路径为0.52cm)的微量滴定板,可以使用以下公式计算每mL酶样品的乳糖酶活性:

$$[0460] \quad LAU/ml \left(\frac{nmol}{min \cdot mL} \right) = \frac{Abs_{420} \times 200\mu L \times \text{稀释因子}}{0.5998 \cdot 10^3 \cdot nM^{-1} \cdot cm^{-1} \times 0.52cm \times 5min \times 0.01mL}$$

[0461] 计算在此示为SEQ ID NO:1的BIF917的比活度:

[0462] BIF917浓度的确定:

[0463] 使用Criterion Stain free SDS-page系统(伯乐公司(BioRad))确定靶酶(BIF917)和截短产物的定量。将任何kD Stain free预制凝胶4%-20%Tris-HCl,18孔(康贝公司(Comb) #345-0418)与塞尔瓦(Serva)Tris-甘氨酸/SDS缓冲液(伯乐公司目录号42529)一起使用。使用以下参数运行凝胶:200V、120mA、25W、50min。使用BSA(1,43mg/ml)(西格玛-奥德里奇,目录号500-0007)作为蛋白质标准品,并且将Criterion Stain Free成像仪(伯乐公司)与Image Lab软件(伯乐公司)一起使用,用于使用与色氨酸含量相关的带强度进行定量。

[0464] 从两种独立发酵(如方法1中所述)的粗制酵素(超滤浓缩物)并且使用5种不同稀释度(参见表1)确定BIF917的特异性LAU活性。

[0465] 发现BIF917比活度为21.3LAU/mg或0.0213LAU/ppm。

[0466] 表1:BIF917比活度的确定

样品 ID	酶	发酵	稀 释 因子	活性	蛋 白 (BIF917) 浓度	蛋 白 (BIF917) 浓度	比活度	比活度
				LAU/ml	mg/ml	ppm	LAU/mg	LAU/ppm
1	BIF 917	a	5	26.9	1.23	1232	21.9	0.0219
2	BIF 917	a	10	53.9	2.44	2437	22.1	0.0221
3	BIF 917	a	10	75.4	3.56	3556	21.2	0.0212
4	BIF 917	a	20	163.9	7.78	7778	21.1	0.0211
5	BIF 917	a	30	233.6	11.06	11065	21.1	0.0211
6	BIF 917	b	5	30.26825	1.34	1342	22.6	0.0226
7	BIF 917	b	10	55.91536	2.61	2607	21.4	0.0214
8	BIF 917	b	10	76.96056	3.70	3697	20.8	0.0208
9	BIF 917	b	20	156.986	7.75	7755	20.2	0.0202
10	BIF 917	b	30	236.9734	11.45	11452	20.7	0.0207
						Arg	21.3	0.0213
						Std	0.700976	0.000701

[0468] 实例2

[0469] 测试:1

[0470] 测定:P-甘露聚糖酶活性,还原糖(MU)

[0471] 该测试方法用于以MU(甘露聚糖酶单位)单位确定甘露聚糖酶活性。

[0472] 目的

[0473] 该测定适用于甘露聚糖酶活性的QA/QC监测。

[0474] 原理

[0475] 该测定通过内切-1,4-P-D-甘露聚糖酶对槐豆胶(LBG)底物的作用测量还原糖的释放。将二硝基水杨酸反应用于确定还原糖含量的增加,这与所添加的酶活性成正比。

[0476] 程序

[0477] 1.材料和设备

[0478] 1.1能够达到3500rpm的离心机

[0479] 1.2设置为40℃的水浴

[0480] 1.3设置为60℃的水浴

[0481] 1.4沸水浴

[0482] 1.5外置活塞式移液器(Positive displacement pipet)和吸头(瑞宁公司(Ranin Inc))

[0483] 1.6涡旋

[0484] 1.7 16x 100mm玻璃试管(带帽)

[0485] 1.8计时器

[0486] 1.9容量瓶、刻度量筒、烧杯

[0487] 1.10磁力搅拌棒和搅拌板/加热板

[0488] 1.11垂熔玻璃滤器漏斗

- [0489] 1.12pH计
- [0490] 1.13分析天平
- [0491] 1.14制备平衡
- [0492] 1.15温度计
- [0493] 1.16冰水浴
- [0494] 1.17分光光度计,能够读取540nm
- [0495] 1.18具有一次性吸头的可变移液装置(1ml)
- [0496] 1.20玻璃器皿的烤箱或高压釜
- [0497] 2.试剂:使用以下试剂或等效物
- [0498] 2.1氢氧化铵,浓缩的(28%-30%)
- [0499] 2.2Tris-羟甲基氨基甲烷(Tris)盐酸化物
- [0500] 2.3氢氧化钠颗粒
- [0501] 2.4 3,5-二硝基水杨酸,98%,(DNS),西格玛-奥德里奇目录号128848
- [0502] 2.5槐豆胶,西格玛化学#G0753
- [0503] 2.6酒石酸钾钠,四水合物,西格玛-奥德里奇目录号217255
- [0504] 2.7去离子水
- [0505] 2.8 1-丙醇
- [0506] 3.试剂制备
- [0507] 3.1氢氧化铵1.5%
- [0508] 3.1.1使用去离子水,在容量瓶中将5ml 30%氢氧化铵稀释至100ml。
- [0509] 3.2Tris-盐酸盐缓冲液
- [0510] 3.2.1在约1900ml去离子水中溶解15.67gm Tris-盐酸盐。用氢氧化铵(1.5%)调节至pH 7.5±0.05,并用去离子水稀释至2000ml。该溶液可以在室温下保持至少两周。
- [0511] 3.3氢氧化钠10.67%
- [0512] 3.3.1添加32.0gm氢氧化钠颗粒至200ml去离子水中。搅拌至溶解并冷却。用去离子水使达到300ml
- [0513] 3.4DNS试剂
- [0514] (a) 在2L烧杯中将20gm的DNS悬浮于1000g去离子水中。添加300ml的10.67%氢氧化钠溶液。
- [0515] (b) 将悬浮液在加热搅拌板上加热,直至溶液澄清。温度不应超过50℃。
- [0516] (c) 逐渐添加600gm酒石酸钾钠四水合物至溶液中,并持续混合。使溶液达到室温。
- [0517] (d) 用去离子水将溶液稀释至2000ml,并且如果需要,通过烧结玻璃滤器过程过滤。在室温下存放在深琥珀色瓶中。该溶液持续至少两个月保持良好。
- [0518] 3.5槐豆胶(LBG)底物溶液
- [0519] 3.5.1确保与底物接触的所有玻璃器皿都非常干净,并且没有可能的甘露聚糖酶污染。推荐在120℃下进行1-丙醇洗涤或烘烤。
- [0520] 5.1.1将500ml Tris-盐酸盐缓冲液放入1000ml烧杯中。将烧杯放在沸水浴或加热搅拌板中,并且使烧杯中的温度达到80℃。快速搅拌溶液,同时非常缓慢地添加1.4gm的槐豆胶。减少混合并在水浴中将溶液保持在60℃持续60分钟。冷却至室温。用去离子水调节至

500ml。以3500rpm离心10分钟。使用澄清的上清液作为底物。

[0521] 3.6酶标准品制备

[0522] 3.6.1选择大量的甘露聚糖酶最终产品作为标准品。将其设置为等于批次分析证书上所报告的MU值。使用这种材料,使用以MU/升计的已知浓度的样品制作3点标准曲线。相应地使用Tris-盐酸盐缓冲液稀释该标准品,这样使得其净吸光度在减去试剂空白后落入测定的线性范围内。该测定的线性范围为0.17和0.52AA。0.050-0.140MU/升之间的标准曲线和样品浓度通常落入该测定的线性范围内。

[0523] 3.7液体样品制备

[0524] 3.7.1在Tris-盐酸盐缓冲液中稀释每个样品(w/v),这样使得部分4中概述的测定反应在540nm下落入17和0.52AA之间。

[0525] 3.7.2将稀释的酶溶液储存在冰上。为最佳结果,应在1小时内测定稀释的酶溶液。

[0526] 4.测定程序

[0527] 4.1酶样品-应将酶样品一式两份进行检测

[0528] 4.1.1将2ml LBG底物在40℃下在16×100玻璃试管中平衡20分钟。

[0529] 4.1.2添加0.5ml酶样品稀释液,混合并且精确孵育10分钟。

[0530] 4.1.3通过添加3.0ml的DNS溶液并混合来终止反应。

[0531] 4.1.4在沸水浴中将样品煮沸15分钟,用帽覆盖每个试管的顶部以防止蒸发。在冰水浴中冷却50分钟。使样品平衡至室温持续10分钟。

[0532] 5.1.1针对去离子水空白,读取540nm处的吸光度。在减去空白试剂后,吸光度应在0.17和0.52AA之间。

[0533] 4.2空白试剂

[0534] 4.2.1可以单独运行空白。该反应对照用于LBG底物中存在的或酶样品中存在的还原糖。

[0535] 5.1.1将2ml LBG底物在40℃下在16×100玻璃试管中平衡20分钟。

[0536] 5.1.2孵育另外10分钟。

[0537] 5.1.3添加3.0ml DNS溶液并混合。添加0.5ml酶稀释液并再次混合。

[0538] 5.1.4在沸水浴中煮沸15分钟,用帽覆盖试管的顶部以防止蒸发。在冰水浴中冷却5分钟。使样品平衡至室温持续10分钟。

[0539] 5.1.5针对去离子水空白,读取540nm处的吸光度。

[0540] 5.计算

[0541] 5.1为了确定净吸光度,从所有标准品和样品的吸光度读数中减去平均试剂/酶空白。

[0542] 5.2使用线性回归制备标准曲线,其中净吸光度在y轴上,并且浓度(MU/升)在x轴上。

[0543] 5.3相关系数必须是>0.998。

[0544] 5.4从线性回归确定每个样品的浓度。

[0545] 5.5对于液体:MU甘露聚糖酶/kg样品=来自曲线的值(MU/升)*样品稀释因子(总体积(升)/样品重量(kg))

[0546] 实例3

- [0547] 测试2:内切葡聚糖酶活性,羧甲基纤维素(CMC)活性
- [0548] 目的该测定方法用于以CMC单位计确定纤维素酶的内切葡聚糖酶活性。
- [0549] 注意:以IU/g或IU/ml(国际单位)所报告的纤维素酶分别相当于CMC/g或CMC/ml。
- [0550] 原理该测定通过纤维素酶对CMC底物的作用来测量还原糖的释放。如通过与3,5-二硝基水杨酸(DNS)的反应所测量的,还原糖释放速率与酶活性成正比。一个CMC单位定义为在测定条件下每分钟产生1⁻摩尔葡萄糖还原糖当量所需的酶量。相对于酶标准品,以指定的CMC单位测量此程序中的活性。
- [0551] 程序
- [0552] 1.0材料和设备
- [0553] 1.18 18x 150mm试管
- [0554] 1.19石珠
- [0555] 1.20比色皿
- [0556] 1.21外置活塞式移液器和吸头,瑞宁公司
- [0557] 1.22pH计
- [0558] 1.23设置为50℃的水浴
- [0559] 1.24冰浴
- [0560] 1.25沸水浴
- [0561] 1.26分光光度计
- [0562] 1.27磁力搅拌器
- [0563] 1.28秒表
- [0564] 1.29Coarse glass filter©,VWR KT93700-47
- [0565] 1.30深琥珀色瓶
- [0566] 1.31外置活塞式移液器
- [0567] 1.32涡旋振荡器
- [0568] 1.33试管架
- [0569] 2.0试剂:使用以下试剂或等效物。
- [0570] 2.9羧甲基纤维素钠盐Fluka 21900取代度必须为0.70-0.85
- [0571] 2.10 3,5-二硝基水杨酸,钠盐(DNS),默克(Merck) 10846
- [0572] 2.11酒石酸钾钠,四水合物,默克8087
- [0573] 2.12氢氧化钠,试剂级别
- [0574] 2.13冰乙酸,试剂级别
- [0575] 3.0试剂制备
- [0576] 3.1 0.05M乙酸钠缓冲液,pH 4.8
- [0577] 3.1.1向900mL蒸馏水中添加2.85mL冰乙酸。在用磁力搅拌棒搅拌的同时,用50%氢氧化钠调节pH至4.8。用蒸馏水使总体积为一升。
- [0578] 3.4 10.67% (w/v) 氢氧化钠溶液
- [0579] 3.2.1将32.0g氢氧化钠颗粒添加到200mL蒸馏水中。搅拌至溶解并冷却。用蒸馏水使达到300mL。
- [0580] 3.5 1% CMC底物溶液

[0581] 3.3.1向99mL乙酸钠缓冲液中添加1.00g CMC。搅拌来彻底混合,并且保持在4℃下持续至少1小时,然后再使用。将该溶液在4℃下稳定3天

[0582] 3.6 1%3,-5二硝基水杨酸(DNS)

[0583] (e) 将20.0g DNS悬浮于1000mL蒸馏水中,并在混合的同时逐渐添加300mL的10.67%氢氧化钠溶液。

[0584] (f) 在设置为50℃ (122°F) 的水浴中加热该悬浮液直到该溶液澄清。水浴温度应当不超过50℃。

[0585] 3.4.3逐渐向溶液中添加600g酒石酸钾钠四水合物,并连续混合。

[0586] 3.4.4用蒸馏水将溶液稀释至2000mL,并通过粗玻璃滤器过滤。在室温下存放在深琥珀色瓶中。该溶液稳定2个月。溶液必须澄清才能使用。

[0587] 4.1颗粒样品制备

[0588] 3.7.2将颗粒样品溶解在乙酸盐缓冲液中,这样使得从部分4.5中概述的测定反应产生的净吸光度落入0.4和0.5之间。约0.2CMC/mL的制剂通常会产生所需的吸光度。

[0589] 3.7.3在制备样品时,称重至少100mg颗粒。相应地按4.1.1用乙酸盐缓冲液通过磁力搅拌溶解。针对5.1中的计算以mg/mL记录颗粒浓度。以中等速度混合持续至少20分钟。将稀释的样品储存在冰上。稳定至少2小时。

[0590] 3.7.4注意一些样品在混合时间后会形成细浆。应取代表性的浆液样品进行4.5中的测定。

[0591] 4.1.4一式三份执行每个颗粒样品制备。

[0592] 4.2液体样品制备

[0593] 4.1.3将每个样品稀释在乙酸盐缓冲液中,这样使得从部分4.5中概述的测定反应产生的净吸光度落入0.4和0.5之间。约0.2CMC/mL的制剂通常会产生所需的吸光度。

[0594] 4.1.4当制备样品时,使用外置活塞式移液器等分出至少0.10mL的液体样品。将稀释的样品储存在冰上。稳定至少2小时。

[0595] 4.1.5一式三份执行每个液体样品制备。

[0596] 4.3工作标准品准备

[0597] 4.4

[0598] 4.3.1选择大量的纤维素酶最终产品(液体或颗粒)作为标准品。将其设置为等于批次分析证书上所报告的CMC值。使用此材料,按4.1或4.2制备已知CMC/mL的工作标准品。使用乙酸盐缓冲液稀释标准品,这样使得来自部分4.5中概述的测定反应的净吸光度产率落入0.4和0.5之间。制备约0.2CMC/mL的标准品溶液通常会产生所需的吸光度。

[0599] 4.3.2以CMC/mL记录工作标准品的活性。将其在计算部分5.1和5.2中使用。

[0600] 4.3.3一式三份准备标准品。

[0601] 4.4空白溶液

[0602] 4.4.1在部分4.5中使用乙酸盐缓冲液作为空白酶溶液。一式两份运行每个空白酶溶液。

[0603] 4.5酶测定反应

[0604] 4.5.1在50℃下,在18x 150mm玻璃试管中孵育1.00mL的CMC底物10-15分钟,这些玻璃试管竖立在试管架中。

[0605] 4.5.2以20秒的间隔,使用外置活塞式移液器将1.00mL的酶稀释液和空白添加到CMC底物中。在50℃ (122°F) 下,将每个反应混合并孵育10分钟。

[0606] 4.5.3以与4.5.2中相同的时间间隔,添加3.0mL DNS溶液并且进行混合。

[0607] 4.5.4通过将试管和支架置于沸水浴中,将所有反应混合物+DNS精确煮沸5分钟。用石珠盖住试管的顶部,以防止沸腾期间蒸发。所有样品,标准品和空白应一起煮沸。煮沸后,在冰浴中将管冷却。

[0608] 4.5.5针对作为零吸光度的蒸馏水,测量酶样品和空白样品在540nm处的吸光度。

[0609] 4.5.6从平均样品和标准品吸光度中减去平均空白吸光度。此净吸光度应在0.4-0.5之间。如果不是,则应重复该测定。

[0610] 4.5.7在测定中一式三份运行每个酶样品。

[0611] 5.0计算

[0612] 5.1如下计算颗粒样品中的活性:

[0613] $\text{CMC/g (或IU/g)} = (\text{样品的净吸光度}) (\text{CMC/mL的工作标准品}) (1000\text{mg/g}) (\text{工作标准品的净吸光度}) (\text{mg颗粒/mL乙酸盐缓冲液})$

[0614] 5.2如下计算液体样品中的活性:

[0615] $\text{CMC/mL (或IU/mL)} = (\text{样品的净吸光度}) (\text{CMC/mL的工作标准品}) (\text{样品稀释液})$

[0616] (工作标准品的净吸光度)

[0617] 注意:以IU/g或IU/mL (国际单位) 所报告的纤维素酶分别相当于CMC/g或CMC/mL。

[0618] 实例4

[0619] 测试3

[0620] 使用基于Viscoman的粘度法。该方法目前可用于检测非常低水平的甘露聚糖酶、淀粉酶、CMC酶和果胶酶活性。(有用的粘度方法也在US2013/0045498中公开,其通过引用结合在此)。水状胶体被酶切割,从而导致水状胶体溶液的粘度降低。粘度的降低表示为从添加酶的水状胶体样品的粘度相对于添加H₂O的水状胶体样品的粘度所计算出的相对粘度值。因此,该值始终将在1和0之间,其中1意指样品的粘度没有降低。

[0621] 底物为GUAR、淀粉、CMC或果胶。

[0622] 根据底物使用两种缓冲液;

[0623] McIlvaine的柠檬酸磷酸盐:

[0624] A) 0.1M柠檬酸;21.0g C6H8O7, H2O/1L去离子水

[0625] B) 0.2M磷酸氢二钠;35.6g Na2HPO4, 2H2O/1L去离子水。保质期:3个月

[0626] 缓冲液 (pH 4.0) = 71.50mL A+28.50mL B (在带刻度的量筒中)

[0627] 缓冲液 (pH 6.7) = 25.00mL A+75.00mL B

[0628] 底物浓度

	果胶	2.0%	pH 4.0
[0629]	CMC	0.5%	pH 6.7
	GUAR	0.5%	pH 6.7
	淀粉	3.0%	pH 6.7

[0630] 底物溶液的制备:

[0631] 将用于特定底物的缓冲液用水稀释五次,并在带有磁铁的电热板上在Blue Cap瓶

中加热。称重水状胶体,并且当缓冲液冒泡时,粉末在剧烈搅拌下撒入直至稀释(约30分钟)。对于淀粉,建议在添加底物时关掉加热,并且取下瓶盖,并且避免底物沸溢-以防止形成块状物等。将水状胶体溶液冷却至40℃,淀粉溶液除外。必须将淀粉储存在50℃下直到使用,以便防止老化。

[0632] 样品处理

[0633] 将包括空白的每个样品的5ml底物转移到Wheaton玻璃杯。将100μl酶样品或样品的稀释液(用水制成的)添加到5mL底物溶液中并混合。最多可添加500μL的样品,只要将相同的体积应用于空白样品。将样品在40℃下孵育20h,并且然后通过将样品放在冰上终止反应。然后使用Brookfield将样品回火至10℃进行粘度测量。

[0634] 实例5

[0635] 酶副活性筛选

[0636] 对样品1和商业GODO YNL2乳糖酶进行筛选,用于进行比较。样品1是包括由枯草芽孢杆菌宿主细胞产生的具有转半乳糖基化活性的酶的多肽组合物,其中纤维素、甘露聚糖酶和果胶酶未失活。测定和结果列于表2中。观察到,与GODO YNL2乳糖酶相比,样品1具有显著更高水平的甘露聚糖酶和CMC酶活性。

[0637] 表2

[0638]	样品 ID	样品 1	GoDo-YNL2
	测试 3 果胶酶, 相对活性	Uf: 0,552 F10: 0,895 F100: 0,964	Uf: 0,806 F10: 1,016 F100: 0,989
	测试 3 甘露聚糖酶, 相对活性	Uf: 0,006, F10: 0,009 F100: 0,011	Uf: 1,010 F10: 1,011 F100: 1,001
	测试 3 CMC 酶, 相对活性	Uf: 0,010 F10: 0,008 F100: 0,010	Uf: 0,989 F10: 0,995 F100: 1,021

[0639] 甘露聚糖酶和纤维素酶水平的半定量

[0640] 通过稀释样品直到几乎不能再观察到粘度变化来评估样品1中存在的纤维素酶和甘露聚糖酶的水平。测试浸解酶以确定当在发酵期间在大豆介质中使用时,它是否对存在于最终产品中的甘露聚糖酶和纤维素酶的水平有影响。浸解酶是一种已知的纤维素酶,并且因此它只测试甘露聚糖酶活性,因为纤维素酶活性将被预期。稀释度(参见表3)表明在浸解酶中高水平的甘露聚糖酶活性和样品中高水平的甘露聚糖酶和纤维素酶活性。存在于浸解酶中的甘露聚糖酶活性水平不能仅仅解释存在的甘露聚糖酶水平,因为在浸解酶的F10.000稀释液中相对粘度接近1,而样品1必须稀释超过F100.000以便相对粘度接近1。

[0641] 添加100μL所指示浸解酶原料或样品1的稀释液的GUAR或CMC溶液的相对粘度。浸解酶原料是浸渍样品的54倍稀释液,该浸渍样品对应于在发酵期间向大豆介质中添加酶时所制备的稀释液。

[0642] 表3

样品稀 释液	浸解酶原料 54 倍稀释液 GUAR	BIF 917 GUAR	BIF 917 CMC
1x	0,034		0,009
10x	0,235	0,044	
100x	0,708	0,054	
1.000x		0,123	
10.000x	0,957	0,429	0,302
100.000x		0,900	0,874

[0644] 在敲除果胶酶、甘露聚糖酶和纤维素酶的新宿主菌株中的副活性筛选

[0645] 决定从枯草芽孢杆菌宿主菌株(样品2)中敲除果胶酶、纤维素酶和甘露聚糖酶,并且然后重新测试副活性的存在。收到使用新菌株的第一次3K发酵样品,用于分析。将来自3K样品的结果与使用样品1的结果进行比较,显然使用样品2,果胶酶、纤维素酶和甘露聚糖酶的副活性水平降低(参见表4)。

[0646] 表4

	果胶酶相对活性	甘露聚糖酶相对活性	CMC 酶相对活性
样品 1 乳糖酶	Uf: 0,552 F10: 0,895 F100: 0,964	Uf: 0,006, F10: 0,009 F100: 0,011	Uf: 0,010 F10: 0,008 F100: 0,010
样品 2 乳糖酶	F1: 0,607 F10: 1,020 F100: 1,027	Uf: 0,009 F10: 0,223 F100: 0,875	Uf: 0,025 F10: 0,030 F100: 0,081

[0648] 使用Gilson Viscoman的粘度测定规范

[0649] 为了确定酸奶应用中的甘露聚糖酶和纤维素酶活性的阈值水平,使用了基于粘度降低来定量甘露聚糖酶和纤维素酶活性的测定。对于该测定,使用来自Biolab A/S的Viscoman自动移液器来确定每个样品中的粘度。

[0650] 为了确定使用Viscoman的样品中的变化,在根据测试3制备的0.5%GUAR底物上测量粘度。在6种不同的底物制剂上进行三次样品测量。6次运行中最大变化为约4%。为了安全起见,将样品中的变化限定为5%。为了使空白样品的变化不与添加酶的样品的变化重叠,我们需要观察粘度的变化处于大于10%。这意味着为了说样品包含甘露聚糖酶活性,相对粘度应该小于0.9。因此,将标准曲线的相对粘度的上限设定为0.85。

[0651] 表5

[0652]

运行编 码	平均粘度 (Cp)	STDEV	CV%
1	51,67	0,23	0,45
2	64,57	0,81	1,25
3	68,00	0,82	1,20
4	68,30	2,62	3,84
5	77,10	1,51	1,96
6	78,80	1,40	1,78

[0653] 使用标准品 (699MU/Kg) 测试适用于甘露聚糖酶粘度测定的酶稀释范围。还测试了包含甘露聚糖酶活性的样品, 样品1。通过最小二乘法发现当相对粘度相对于LN (MU/Kg) 进行绘制时, 数据呈现直线 (参见图2)。从在GUAR底物中具有标准酶稀释液的6次运行得出结论, $R^2 \geq 0.98$ 的直线呈现在0.25-0.85的范围内 (数据未显示)。为了限制测试量, 对于使用的所有底物, 相对粘度值的规定设定为 ≥ 0.25 且 ≤ 0.85 。

[0654] 图2中的曲线图显示样品1 (标准品) 的稀释液相对于LN (MU/Kg) 绘制的相对粘度, 其中基于Viscoman使用速度2确定粘度。它还显示了相对于LN (MU/Kg) 所绘制的, 以速度1所确定的样品1的1-4MU/Kg样品的相对粘度, 和以速度2所确定的样品1的5-80MU/Kg样品的相对粘度。

[0655] 另外, 还显而易见的是, 仅当基于Viscoman使用相同速度获得数据时才能获得线性函数 (参见图1)。基于Viscoman使用相同速度获得数据。针对该测定, 低粘度样品使最相关的, 并且因此所有样品应使用速度2测量, 这意味着粘度低于400mPa.s。因此, 定义了当以速度2用Viscoman测量时, 底物空白样品应具有在60-130mPa.s内的粘度。

[0656] 甘露聚糖酶定量测定:

[0657] 为了建立基于相对粘度的甘露聚糖酶定量测定, 我们使用来自甘露聚糖酶比色测定的建立的对照和标准品450MU/Kg。确定样品1、样品2和浸解酶的甘露聚糖酶活性。外推该线, 发现粉末的检测限为约0.003MU/Kg (在10mL中最大1g), 并且液体的检测限为0.0003MU/Kg。

[0658] 每个样品如测试3中所述进行处理, 但使用Viscoman确定粘度。使用四点标准曲线计算每个样品的甘露聚糖酶活性。下表涉及图3。

[0659]

样品 ID	活性 (MU/Kg)
对照	671,51
标准品	742,62
样品 1	9,14
样品 2	0,062
浸解酶	3,089

[0660] 计算表明, 样品2中的甘露聚糖酶活性主要来源于浸解酶 (3MU/Kg除以54, 相当于在大豆介质中添加浸解酶时的稀释度)。因此可以确认甘露聚糖酶基因已被有效地从宿主菌株中敲除。

[0661] 纤维素酶定量测定

[0662] 基于相对粘度,使用对照2222CMCU/g和标准品2306CMCU/g来建立的纤维素酶定量测定。将标准品的相对粘度值相对于LN (CMCU/g) 进行绘制,并在相对粘度为0,242-0,885 (参见图4) 内,被证实是线性的。外推该线,发现该测定的液体检测限为约0.002CMCU/g,并且粉末检测限为0.02CMCU/g。

[0663] 将来自发酵的具有或没有浸解酶的两个额外样品与现有样品一起进行纤维素酶活性测试。结果显示,从宿主菌株敲除纤维素酶是有效的 (34CMCU/g对比约4CMCU/g),尽管由于在发酵期间在大豆介质中浸解酶的添加,仍存在一定活性 (参见表6和图4)。

[0664] 表6

[0665]	样品ID	CMCU/g
	对照	2414
	样品1	34,38
	样品2	4,1
	标准品:用浸解酶发酵,806LAU/g	3,4
	标准品:用浸解酶发酵,971LAU/g	<0,002

[0666] 在酸奶应用中的纤维素酶阈值水平

[0667] 为了建立酸奶应用的纤维素酶阈值水平,采用用于粘度测定的纤维素酶对照处理饮用酸奶。将一份饮用酸奶 (MILRAM,开菲尔饮品 (Kefir Drink),草莓 (Erdbeere);BB 14-02-14) 等分5mL在Wheaton玻璃杯中,每份样品一式两份。将原料稀释液 (0,081mL,在1000mL中) 进一步以10倍因子稀释为四个新样品。将每个样品的100μL以及空白样品的100μL ddH₂O添加到每个5mL酸奶中。然后将样品储存在4℃下或40℃下。在各个时间点,使用Viscoman以速度2测量粘度。

[0668] 将相对粘度相对于在a) 40℃或b) 4℃下稀释样品的孵育时间 (a) 小时b) 天) 进行绘制。对于每次稀释,注意CMCU/mL酸奶。

[0669] 这些结果呈现为相对于在40℃ (图5a) 或4℃ (图5b) 下孵育的时间所绘制的相对粘度。曲线显示随时间的推移类似的发展,表明40℃孵育可用作加速测定,代替耗时的4℃孵育测定。

[0670] 为了计算阈值水平,定义了纤维素酶活性的可接受水平在储存期间内会降低粘度小于10%。基于图5b中所示的数据,计算每个数据点的比活度 (相对粘度降低/CMCU/每mL酸奶/h),假设观察到的最高比活度将代表初始速率。

[0671] 发现最高的比活度为9.6相对粘度降低/CMCU/每mL乳酸/h,意味着在一mL酸奶中每小时粘度将降低960%每CMCU。然后可以计算各储存时期允许10%粘度降低的CMCU/mL阈值水平 (参见表7)。

[0672] 阈值水平=0,1/(比活度*储存时间)

[0673] 对于1个月的储存,如下计算阈值水平: $0.1/(9.6*720) = 1,45 \times 10^{-5}$

[0674] 表7

[0675] 纤维素酶对照960%粘度降低/h/CMCU/mL

[0676]	在 4°C 下的存储 时间 (月)	小时	粘度降低 x 小时	阈值 (CMCU/ml 酸奶) (最大 10%粘度降低)
	9	6480	62208	1,61E-06
	8	5760	55296	1,81E-06
	7	5040	48384	2,07E-06
	6	4320	41472	2,41E-06
	5	3600	34560	2,89E-06
	4	2880	27648	3,62E-06
	3	2160	20736	4,82E-06
	2	1440	13824	7,23E-06
	1	720	6912	1,45E-05
	14 天	336	3226	3,10E-05

[0677] 将计算与图5b中的数据进行比较,发现它们与在4°C下孵育1个月的所有稀释液所获得数据相关。 3.6×10^{-6} CMCU/mL或更低的水平在一个月没有降低粘度超过10%,而超过 3.6×10^{-6} CMCU/mL则使粘度降低了10%以上。

[0678] 结论

[0679] 果胶酶、纤维素酶和甘露聚糖酶基因已被有效地从宿主菌株中敲除。

[0680] 上述说明书中提及的所有出版物通过引用结合在此。本发明的所描述的方法和系统的各种修改和变化对于本领域的技术人员将是显而易知的,而不背离本发明的范围和精神。虽然已结合特定优选实施例描述了本发明,但应当理解,要求保护的本发明不应不适当地限于这类特定实施例。事实上,对于化学、生物化学、生物学、或相关领域的技术人员而言显而易见的、用于执行本发明的所述模式的各种修改预期在以下权利要求的范围内。

[0681] 序列表

[0682] >SEQ ID NO:1 (BIF_917)

[0683] vedatrsdsttqmsstpevyssavdskqrntsdffdanwkmfmlsdsvqaqdpafddsawqqvdlphdy
sitqkysqsneaesaylpgggtgwyksftidrdlagkriainfdgvymnatvwfngvklgthpygyspfsfdltgn
akfggentivvkvenrlpssrwysgsgiyrdvtltvtdgvhvgngvaiktpslatqnggdvtmnlttkvandtea
aanitlkqtvfpkggktdaaigtvttasksiaagasadvsttitaasplksiknplytvrtevlnggkvldtyd
teygrwtgfdatsgfslngekvklkgvsmhhdqgslgavanrraierqveilqkmgvnsirthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavlbgdkdetwakfdltstinrdnapsvimwslgnem
megisgsvsgfpatsaklvawtkaadstrpmtgydnkikanwnesntmgdnltaggvvgtnysdganydkirtth
pswaiygsetasainsrgiynrttggagssdkqltsydnasvawgavassawydvvqrdfvagtyvwtgfdylgep
tpwngtgsgavgswwpspknsyfgivdtagfpkdttyfyqsqwnddvhtlhilpawnenvvakgsgnnvpvvytda
akvklyftpkgstekrligeksftkkttaaagytyqvyegsdkdstahknmyltnvpwaegtisaeaydenrrip
egstegnasvtttgaaklkadadrktitadgkdlsyievdvtdanghivpdaanrvtfdvkgagklvgvngssp
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvkiaattavpgtstekt

[0684] >SEQ ID NO:2 (BIF_995)

[0685] vedatrsdsttqmsstpevyssavdskqrntsdffdanwkmfmlsdsvqaqdpafddsawqqvdlphdy
sitqkysqsneaesaylpgggtgwyksftidrdlagkriainfdgvymnatvwfngvklgthpygyspfsfdltgn

akfggentivvkvenrlpssrwysgsgiyrdvtltvtdgvhvgngvaiktpslatqnggdvtmnlttkvandtea
aanitlkqtvfpkggktdaaigtvttasksiaagasadvttitaaspklwsiknpnlytvrtevlnggkvldtyd
teygrwtgfdatsgfslngekvklkgvsmhhdqgslgavanrraierqveilqkmgvnsirtthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavlbgdkdetwakfdltstinrdnapsvimwslgnem
megisgsvsgfpatsaklvawtkaadstrpmtgydnkikanwnesntmgdnl tanggvvgt nysdganydkirtth
pswaiygsetasainsrgiynrttggaqssdkqltsydnasvlgwagavassawydvvqrdfvagtyvwtgfdylgep
tpwngtgsgavgswwpspknsyfgivdtagfpkdtyyfyqsqwnddvhtlhilpawnenvvakgsgnnvpvvvytda
akvklyftpkgstekrligeksftkkttagytyqvyegsddstahknmyltnvpwaegtisaeaydenrnlip
egstegnasvtttgaaklkadadrktitadgkdlseyevdvdanghivpdaanrvtfdvkgagklvgvngssp
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvkiaattavpgtstektvrsfyysrnyyvktnk
pilpsdvevrysdgtsdrqnvtwdavsdqiaagsfsvagt vagqkisvrvtmideigal

[0686] >SEQ ID NO:3 (BIF_1068)

[0687] Vedatrdssttqmsstpevyssavdskqnrtsd f danw k f m l s d s v q a q d p a f d d s a w q q v d l p h d y
sitqkysqsneaesaylp g g t g w y r k s f t i d r d l a g k r i a i n f d g v y m n a t v w f n g v k l g t h p y g s p f s f d l t g n
akfggentivvkvenrlpssrwysgsgiyrdvtltvtdgvhvgngvaiktpslatqnggdvtmnlttkvandtea
aanitlkqtvfpkggktdaaigtvttasksiaagasadvttitaaspklwsiknpnlytvrtevlnggkvldtyd
teygrwtgfdatsgfslngekvklkgvsmhhdqgslgavanrraierqveilqkmgvnsirtthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavlbgdkdetwakfdltstinrdnapsvimwslgnem
megisgsvsgfpatsaklvawtkaadstrpmtgydnkikanwnesntmgdnl tanggvvgt nysdganydkirtth
pswaiygsetasainsrgiynrttggaqssdkqltsydnasvlgwagavassawydvvqrdfvagtyvwtgfdylgep
tpwngtgsgavgswwpspknsyfgivdtagfpkdtyyfyqsqwnddvhtlhilpawnenvvakgsgnnvpvvvytda
akvklyftpkgstekrligeksftkkttagytyqvyegsddstahknmyltnvpwaegtisaeaydenrnlip
egstegnasvtttgaaklkadadrktitadgkdlseyevdvdanghivpdaanrvtfdvkgagklvgvngssp
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvkiaattavpgtstektvrsfyysrnyyvktnk
pilpsdvevrysdgtsdrqnvtwdavsdqiaagsfsvagt vagqkisvrvtmideigallnysastpvgt pavl
pgsrpavl pdgtvtsanfavhwtkpadtyvntagtvkvpgtatvf g k e f k v t a t i r v q

[0688] >SEQ ID NO:4 (BIF_1172)

[0689] vedatrdssttqmsstpevyssavdskqnrtsd f danw k f m l s d s v q a q d p a f d d s a w q q v d l p h d y
sitqkysqsneaesaylp g g t g w y r k s f t i d r d l a g k r i a i n f d g v y m n a t v w f n g v k l g t h p y g s p f s f d l t g n
akfggentivvkvenrlpssrwysgsgiyrdvtltvtdgvhvgngvaiktpslatqnggdvtmnlttkvandtea
aanitlkqtvfpkggktdaaigtvttasksiaagasadvttitaaspklwsiknpnlytvrtevlnggkvldtyd
teygrwtgfdatsgfslngekvklkgvsmhhdqgslgavanrraierqveilqkmgvnsirtthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavlbgdkdetwakfdltstinrdnapsvimwslgnem
megisgsvsgfpatsaklvawtkaadstrpmtgydnkikanwnesntmgdnl tanggvvgt nysdganydkirtth
pswaiygsetasainsrgiynrttggaqssdkqltsydnasvlgwagavassawydvvqrdfvagtyvwtgfdylgep
tpwngtgsgavgswwpspknsyfgivdtagfpkdtyyfyqsqwnddvhtlhilpawnenvvakgsgnnvpvvvytda
akvklyftpkgstekrligeksftkkttagytyqvyegsddstahknmyltnvpwaegtisaeaydenrnlip
egstegnasvtttgaaklkadadrktitadgkdlseyevdvdanghivpdaanrvtfdvkgagklvgvngssp
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvkiaattavpgtstektvrsfyysrnyyvktnk

pilpsdvevrysdgtsdrqnvtdavsdqdiakagsfsvagt vagqkisvrvtmideigallnysastpvgtpavl
pgsrpavlpdgtvtsanfahwtkpadtyntagtvkvpgtatvfgkefkvtatirvqrsqvtigssvsgnalrlt
qnipadkqsdtdaikdgsttdantggganpsawtnwayskaghntaeitfeyateqqlgqivmyffrdsnavrf
pdagktkiqi

[0690] >SEQ ID NO:5 (BIF_1241)

[0691] vedatrdssttqmsstpevyssavdskqnrtsd f danw k f m l s d s v q a q d p a f d d s a w q q v d l p h d y
sitqkysqsneaesaylp g g t g w y r k s f t i d r d l a g k r i a i n f d g v y m n a t v w f n g v k l g t h p y g s p f s f d l t g n
akfggentivvkvenrlpssrwysgsgiyrdvtltvtdgvhvgngvaiktpslatqnggdvtmnl t t k v a n d t e a
aanitlkqtvfpkggktdaaigtvttasksiaagasadv t s t i t a a s p k l w s i k n p n l y t v r t e v l n g g k v l d t y d
teygrwtgfdatsgfslngekvklkgvsmhhdqsglgavanrraierqveilqkmgvnsirtthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavl g g d k d e t w a k f d l t s t i n r d n a p s v i m w s l g n e m
megisgsvsgfpatsaklvawtkaadstrpmtygdnkikanwnesntmgdnl tanggvvgt n y s d g a n y d k i r t t h
pswaiygsetasainsrgiynrttggagssdkqltsydn savgw g a v a s s a w y d v v q r d f v a g t y v w t g f d y l g e p
tpwngtgsgavgs w p s p k n s y f g i v d t a g f p k d t y y f y q s q w n d d v h t l h i l p a w n e n v v a k g s g n n p v v v y t d a
akvklyftpkgstekrlige k s f t k k t t a a g y t y q v y e g s d k d s t a h k n m y l t w n v p w a e g t i s a e y d e n n r l i p
egstegnasvtttgaaklkadadrktitadgkdlsyievdvtdanghi v p d a a n r v t f d v k g a g k l v g v d n g s s p
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvk i a t t a v p g t s t e k t v r s f y y s r n y y v k t g n k
pilpsdvevrysdgtsdrqnvtdavsdqdiakagsfsvagt vagqkisvrvtmideigallnysastpvgtpavl
pgsrpavlpdgtvtsanfahwtkpadtyntagtvkvpgtatvfgkefkvtatirvqrsqvtigssvsgnalrlt
qnipadkqsdtdaikdgsttdantggganpsawtnwayskaghntaeitfeyateqqlgqivmyffrdsnavrf
pdagktkiqi s a d g k n w t d l a a t e t i a a q e s s d r v k p y t y d f a p v g a t f v k v t v t n a d t t t p s g v v c a g l t e i e l k
t a t

[0692] >SEQ ID NO:6 (BIF_1326)

[0693] vedatrdssttqmsstpevyssavdskqnrtsd f danw k f m l s d s v q a q d p a f d d s a w q q v d l p h d y
sitqkysqsneaesaylp g g t g w y r k s f t i d r d l a g k r i a i n f d g v y m n a t v w f n g v k l g t h p y g s p f s f d l t g n
akfggentivvkvenrlpssrwysgsgiyrdvtltvtdgvhvgngvaiktpslatqnggdvtmnl t t k v a n d t e a
aanitlkqtvfpkggktdaaigtvttasksiaagasadv t s t i t a a s p k l w s i k n p n l y t v r t e v l n g g k v l d t y d
teygrwtgfdatsgfslngekvklkgvsmhhdqsglgavanrraierqveilqkmgvnsirtthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavl g g d k d e t w a k f d l t s t i n r d n a p s v i m w s l g n e m
megisgsvsgfpatsaklvawtkaadstrpmtygdnkikanwnesntmgdnl tanggvvgt n y s d g a n y d k i r t t h
pswaiygsetasainsrgiynrttggagssdkqltsydn savgw g a v a s s a w y d v v q r d f v a g t y v w t g f d y l g e p
tpwngtgsgavgs w p s p k n s y f g i v d t a g f p k d t y y f y q s q w n d d v h t l h i l p a w n e n v v a k g s g n n p v v v y t d a
akvklyftpkgstekrlige k s f t k k t t a a g y t y q v y e g s d k d s t a h k n m y l t w n v p w a e g t i s a e y d e n n r l i p
egstegnasvtttgaaklkadadrktitadgkdlsyievdvtdanghi v p d a a n r v t f d v k g a g k l v g v d n g s s p
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvk i a t t a v p g t s t e k t v r s f y y s r n y y v k t g n k
pilpsdvevrysdgtsdrqnvtdavsdqdiakagsfsvagt vagqkisvrvtmideigallnysastpvgtpavl
pgsrpavlpdgtvtsanfahwtkpadtyntagtvkvpgtatvfgkefkvtatirvqrsqvtigssvsgnalrlt
qnipadkqsdtdaikdgsttdantggganpsawtnwayskaghntaeitfeyateqqlgqivmyffrdsnavrf
pdagktkiqi s a d g k n w t d l a a t e t i a a q e s s d r v k p y t y d f a p v g a t f v k v t v t n a d t t t p s g v v c a g l t e i e l k

tatskfvtntsaaalssltnvgtkvsdsvlaagsyntpaiiadvkaegegnasvtvlpahdnvirvitesedhvtrk
tftinlgteqef

[0694] >SEQ ID NO:7两歧双歧杆菌糖苷水解酶催化核心

[0695] qnrtsdfdanwkwfmlsdsvqaqdpafddsawqqvdlphdysitqkysqsneaesaylpgggtgwyksf
tidrdlagkriainfdgvymnatvwfngvklgthpygyspfsfdltgnakfggentivkvlenrlpssrwysgsgi
yrdvtltvtldgvhvgngvaiktpslatqnggdvtmnlttkvandteaaanitlkqtfvpkggktdaaigtvtas
ksiaagasadvttitaaspklwsiknpnlytvrtevlnggkvldtydteygfrwtgfdatsgfslngekvklkgv
smhhdqgslgavanrraierqveilqkmgvnsirtthnpaakalidvcnekglvveevfdmwnrskngntedygk
wfgqaiagdnavlggdkdetwakfdltstinrdnapsvimwslgnemmegisgsvsgfpatsaklvawtkaadst
rpmtty

[0696] >SEQ ID NO:8编码全长的核苷酸序列

[0697] gcagttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagttgtttattc
atcagcggctgatagcaacaaaatcgcacaagcgattttgatgcgaactggaaatttatgtctgcagatagcgtt
caagcacaagatccggcatttgatgattcagcatggcaacaagttgatctgccgatgattatagcatcacacaga
aatatagccaaagcaatgaagcagaatcagcatatcttccgggaggcacaggctggtatagaaaaagctttacaat
tgatagagatctggcaggcaaacgcattgcgattaattttgatggcgtctatatgaatgcaacagtctggttta
ggcgtttaaactgggcacacatccgtatggctattcacggttttcatttgatctgacaggcaatgcaaaatttggcg
gagaaaacacaattgtcgtcaaagttgaaaatagactgccgtcatcaagatggatttcaggcagcggcatttatag
agatgttacactgacagttacagatggcgttcattgttggaataatggcgtcgcaattaaaacaccgtcactggca
acacaaaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacatta
cactgaaacagacagttttccgaaaggcggaaaaacggatgcagcaattggcacagttacaacagcatcaaaatc
aattgcagcaggcgcacatcagcagatgttacaagcacaattacagcagcaagcccgaactgttggtcaattaaaaac
ccgaacctgtatacagttagaacagaagttctgaacggaggcaaagttctggatacatatgatacagaatatggct
ttcgtctggacaggctttgatgcaacatcaggcttttctactgaatggcgaaaaagtcaaactgaaaggcgttagcat
gcatcatgatcaaggctcacttggcgcagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaa
atgggcgtcaatagcattcgcacaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaaggcg
ttctggttgctgaagaagtctttgatatgtggaaccgcagcaaaaatggcaacacggaagattatggcaaatggtt
tgccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttaca
tcaacaattaaccgcgatagaaatgcaccgtcagttattatgttggtcactgggcaatgaaatgatggaaggcattt
caggctcagtttcaggctttccggcaacatcagcaaaactggttgcatggacaaaagcagcagattcaacaagacc
gatgacatatggcgataacaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaat
ggcggagttgttggcacaaattattcagatggcgcaaaactatgataaaattcgtacaacacatccgtcatgggcaa
tttatggctcagaaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatc
agataaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatggtatgatgtt
gttcagagagattttgtcgcaggcacatatgtttggacaggatttgattatctgggcgaaccgacaccgtggaatg
gcacaggctcaggcgcagttggctcatggccgtcaccgaaaaatagctattttggcatcgttgatacagcaggctt
tccgaaagatacatattatttttatcagagccagtggaaatgatgatgttcatacactgcatattcttccggcatgg
aatgaaaatgttgttgcaaaaaggctcaggcaataatgttccggttgctgtttatacagatgcagcgaaagtgaaac
tgtattttacaccgaaaggctcaacagaaaaaagactgatcggcgaaaaatcatttacaaaaaaaacaacagcggc

aggctatacatatcaagtcctatgaaggcagcgataaagattcaacagcgcataaaaacatgtatctgacatggaat
gttccgtgggcagaaaggcacaatttcagcgggaagcgtatgatgaaaataatcgccgtgattccggaaggcagcacag
aaggcaacgcacatcagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagc
ggatggcaaagatctgtcatatattgaagtcgatgtcacagatgcaaatggccatattgttccggatgcagcaaat
agagtcacatttgatgttaaaggcgcaggcaaactggttggcggttgataatggctcatcacggatcatgattcat
atcaagcggataaccgcaaagcattttcaggcaaagtcttggaattgttcagtcaacaaaagaagcaggcgaaat
tacagttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagc
acagaaaaaacagtcgagcgtttttattacagccgcaactattatgtcaaaacaggcaacaaaccgattctgcccgt
cagatgttgaagttcgctattcagatggaacaagcagatagacaaaacggttacatgggatgcagtttcagatgatca
aattgcaaaagcaggctcattttcagttgcaggcacagttgcaggccaaaaaattagcgttcgcgtcacaaatgatt
gatgaaattggcgactgtgaattattcagcaagcacaccggttgccacaccggcagttcttccgggatcaagac
cggcagtcctgcccgtatggcacagtcacatcagcaaatgttcagtcattggacaaaaccggcagatacagtcata
taatacagcaggcacagtcacaaagtaccgggaacagcaacagtttttgcaaaagaatttaaagtcacagcgacaatt
agagttcaaagaagccaagttacaattggctcatcagtttcaggaaatgcactgagactgacacaaaatattccgg
cagataaacaatcagatacactggatgcgattaaagatggctcaacaacagttgatgcaaatcaggcggaggcgc
aaatccgtcagcatggacaaaattgggcatattcaaaagcaggccataacacagcggaaattacatttgaatatgcg
acagaacaacaactggggcagatcgctcatgtatttttttcgcgatagcaatgcagttagatttccggatgctggca
aaacaaaaattcagatcagcgcagatggcaaaaattggacagatctggcagcaacagaaacaattgcagcgcaga
atcaagcgatagagtcacaaaccgtatacatatgattttgcaccggttgccgcaacatttgttaaagtgcagtcaca
aacgcagatacaacaacaccgtcaggcgttggttgcgcaggcctgacagaaattgaactgaaaacagcgacaagca
aatttgtcacaatacatcagcagcactgtcatcacttacagtcaatggcacaaaagtttcagattcagttctggc
agcaggctcatataacacaccggcaattatcgcatgtttaaagcgggaaggcgaaggcaatgcaagcgttacagtc
cttccggcatgataatgttattcgcgtcattacagaaagcgaagatcatgtcacacgcaaaacatttacaatca
acctgggcacagaacaagaatttccggctgattcagatgaaagagattatccggcagcagatatgacagtcacagt
tggtcagaacaacatcaggcacagcaacagaaggaccgaaaaaatttgcagtcgatggcaacacatcaacatat
tggtcatagcaattggacaccgacaacagttaatgatctgttgatcgcggttgactgcaaaaaccgacaaaactgg
atgcactgagatatcttccgcgtccggcaggctcaaaaaatggcagcgtcacagaatataaagttcaggtgtcaga
tgatggaacaaactggacagatgcaggctcaggcacatggacaacaggattatggctggaaactggcggaaattta
caaccggtcacaacaaaacatgttagactgaaagcgggtcagatacaacggatatttcaggcgcacagttacagttccggcaaa
actgacagttgatagagttgatgcagatcatccggcaacatttgcaacaaaagatgtcacagttacactgggagat
gcaacactgagatatggcggtgattatctgctggattatgcaggcaatacagcagttggcaagcaacagtgacag
ttagaggcattgataaatattcaggcacagtcgcgaaaacatttacaattgaactgaaaaatgcaccggcaccgga
accgacactgacatcagttagcgtcaaaacaaaaccgagcaaaactgacatatgttgcggagatgcattttgatccg
gcaggcctgggttctgcaacatgatagacaagcagatagacctccgcaaccgctgggtggcgaacaagcggatgaac
gcggactgacatgcggcacaagatgcgatagagttgaacaactgcgcaaacatgaaaaatagagaagcgcatagaac
aggcctggatcatctggaatttgttggcgcagcagatggcgcagttggagaacaagcaacatttaaagtccatgtc
catgcagatcaggagatggcagacatgatgatgcagatgaacgcgatattgatccgatgttccggctgatcatg
cagttggcgaactggcaagagcagcatgccatcatgttattggcctgagagtcgatacacatagacttaaagcaag

cggtttc aaattccggctgatgatattggcagaaatcgatcg cattacaggctttcatcg ttttgaacgccatgtc

[0698] >SEQ ID NO:9 编码BIF_917的核苷酸序列

[0699] gttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagtgtttattcatc
agcggtcgatagcaaacaaaaatcgcaacagcgattttgatgcgaactggaaatttatgctgtcagatagcgttcaa
gcacaagatccggcatttgatgattcagcatggcaacaagttgatctgccgcatgattatagcatcacacagaaat
atagccaaagcaatgaagcagaatcagcatatcttccgggaggcacaggctggtatagaaaaagctttacaattga
tagagatctggcaggcaaacgcattgcgattaattttgatggcgtctatatgaatgcaacagtcctggtttaatggc
gttaaactgggcacacatccgtatggctattcacgcttttcatcttgatctgacaggcaatgcaaaatttggcggag
aaaacacaattgtcgtcaaagttgaaaatagactgccgtcatcaagatggatttcaggcagcggcatttatagaga
tgttacactgacagttacagatggcgttcattgttggcaataatggcgtcgcaattaaaacaccgtcactggcaaca
caaaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacattacac
tgaaacagacagtttttccgaaaggcggaaaaacggatgcagcaattggcacagttacaacagcatcaaaatcaat
tgcagcaggcgcacatcagcagatgttacaagcacaattacagcagcaagcccgaaactgtggtcaattaaaaacccg
aacctgtatacagttagaacagaagttctgaacggaggcaagttctggatacatatgatacagaatatggctttc
gctggacaggctttgatgcaacatcaggcttttactgaatggcgaaaaagtcaaactgaaaggcgttagcatgca
tcatgatcaaggctcacttggcgcagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaaatg
ggcgtcaatagcattcgcaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaggcgttc
tggttgctgaagaagtccttgatatgtggaaccgcagcaaaaatggcaacacggaagattatggcaaatggtttgg
ccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttacatca
acaattaaccgcgatagaaatgcaccgtcagttattatgttggtcactgggcaatgaaatgatggaaggcatttcag
gctcagtttcaggctttccggcaacatcagcaaaactggttgcatggacaaaagcagcagattcaacaagaccgat
gacatatggcgcataacaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaatggc
ggagttgttggcacaattattcagatggcgcaaaactatgataaaattcgtacaacacatccgtcatgggcaattt
atggctcagaaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatcaga
taaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatggtatgatgttggt
cagagagattttgtcgcaggcacatatgtttggacaggatttgattatctgggcgaaccgacaccgtggaatggca
caggctcaggcgcagttggctcatggcgtcacggaaaaatagctattttggcatcggttgatacagcaggctttcc
gaaagatacatattatttttatcagagccagtggaatgatgatgttcatacactgcatattcttccggcatggaat
gaaaatgttggttgcaaaaggctcaggcaataatgttccggttgctgctttatacagatgcagcgaagtgaaactgt
attttacaccgaaaggctcaacagaaaaaagactgatcggcgaaaaatcatttacaaaaaaaacaacagcggcagg
ctatacatatcaagcttatgaaggcagcgataaagattcaacagcgcataaaaaacatgtatctgacatggaatggt
ccgtgggcgagaaggcacaaatttcagcgggaagcgtatgatgaaaataatcgctgattccggaaggcagcacagaag
gcaacgcatcagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagcggga
tggaagatctgtcatatatgaagtcgatgtcacagatgcaaatggccatatgttccggatgcagcaaataga
gtcacatttgatgttaaaggcgcaggcaaaactggttggcgttgataatggctcatcaccggatcatgattcataatc
aagcggataaccgcaagcattttcaggcaagtcctggcaattgttcagtcacaaaaagaagcaggcgaatttac
agttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagcaca
gaaaaaaca

[0700] >SEQ ID NO:10 编码BIF_995的核苷酸序列

[0701] gttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagttgtttattcatc
agcggtcgatagcaaaacaaaatcgcaacgcatTTTTgatgcgaactggaaatttatgtctgtcagatagcgttcaa
gcacaagatccggcatttgatgattcagcatggcaacaagttgatctgccgcatgattatagcatcacacagaaat
atagccaaagcaatgaagcagaatcagcatatcttccgggaggcacaggctggatatagaaaaagctttacaattga
tagagatctggcaggcaaacgcattgcgattaattttgatggcgctctatatgaatgcaacagtcctggtttaattggc
gttaaactggggcacacatccgtatggctattcaccgttttcatTTTgatctgacaggcaatgcaaaatttggcggag
aaaacacaattgtcgtcaaagttgaaaatagactgccgtcatcaagatggatttcaggcagcggcatttatagaga
tgttacactgacagttacagatggcgttcatgttggcaataatggcgctcgcaattaaaacaccgtcactggcaaca
caaaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacattacac
tgaaacagacagtttttccgaaaggcggaaaaacggatgcagcaattggcacagttacaacagcatcaaaatcaat
tgcagcaggcgcatcagcagatgttacaagcacaattacagcagcaagcccgaaactgttggtcaattaaaaacccg
aacctgtatacagttagaacagaagttctgaacggaggcaaagttctggatacatatgatacagaatatggctttc
gctggacaggctttgatgcaacatcaggcttttactgaatggcgaaaaagtcaaactgaaaggcgtagcatgca
tcatgatcaaggctcacttggcgcagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaaatg
ggcgtaatagcattcgcaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaggcgttc
tggttgctgaagaagtccttgatatgtggaaccgcagcaaaaaatggcaacacggaagattatggcaaatggtttgg
ccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttaccatca
acaattaaccgcgatagaaatgcaccgtcagttattatgttggtcactgggcaatgaaatgatggaaggcatttcag
gctcagtttcaggctttccggcaacatcagcaaaactggttgcatggacaaaagcagcagattcaacaagaccgat
gacatatggcgataacaaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaatggc
ggagttgttggcacaattattcagatggcgcaaaactatgataaaattcgtacaacacatccgtcatgggcaattt
atggctcagaaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatcaga
taaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatggatatgatgttgtt
cagagagattttgtcgcaggcacatatgtttggacaggatttgattatctgggcgaaccgacaccgtggaatggca
caggctcaggcgagttggctcatggcgtcaccgaaaaatagctatTTTtgcatcgttgatacagcaggctttcc
gaaagatacatattatttttatcagagccagtggaaatgatgatgttcatacactgcatattcttccggcatggaat
gaaaatgttggttgcaaaaggctcaggcaataatgttccggttgctgTTTatacagatgcagcgaaagtgaaactgt
atTTTtacaccgaaaggctcaacagaaaaaagactgatcggcgaaaaatcatttacaaaaaaacaacagcggcagg
ctatacatatcaagcttatgaaggcagcgataaagattcaacagcgcataaaaaacatgtatctgacatggaatgtt
ccgtgggcagaaggcacaatttcagcggaaagcgatatgatgaaaataatcgctgattccggaaggcagcacagaag
gcaacgcatcagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagcggga
tggaagatctgtcatatatgaagtcgatgtcacagatgcaaatggccatatgttccggatgcagcaaataga
gtcacatttgatgttaaaggcgcaggcaaactggttggcgTTTgataatggctcatcaccggatcatgattcatatc
aagcggataaccgcaaagcattttcaggcaaagtcctggcaattgttcagtcaacaaaagaagcaggcgaaattac
agttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagcaca
gaaaaaacagtcgcagcttttattacagccgcaactattatgtcaaaacaggcaacaaaccgattctgccgtcag
atgttgaagttcgctattcagatggaacaagcgatagacaaaacgttacatgggatgcagtttcagatgatcaaat
tgcaaaagcaggctcattttcagttgcaggcacagttgcaggccaaaaaattagcgttcgcgtcacaatgattgat
gaaattggcgactg

[0702] >SEQ ID NO:11编码BIF_1068的核苷酸序列

[0703] gttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagttgtttattcatc
agcggtcgatagcaaacaaaatcgcaacagcattttgatgcgaactggaaatttatgctgtcagatagcgttcaa
gcacaagatccggcatttgatgattcagcatggcaacaagttgatctgccgcatgattatagcatcacacagaaat
atagccaaagcaatgaagcagaatcagcatatcttccgggaggcacaggctggatatagaaaaagctttacaattga
tagagatctggcaggcaaacgcattgcgattaattttgatggcgctctatatgaatgcaacagtcctggtttaatggc
gttaaactgggcacacatccgtatggctattcaccgttttcatttgatctgacaggcaatgcaaaatttggcggag
aaaacacaattgtcgtcaaagttgaaaatagactgccgtcatcaagatggatttcaggcagcggcatttatagaga
tgttacactgacagttacagatggcggttcattgttggaataatggcgctcgcaattaaaacaccgtcactggcaaca
caaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacattacac
tgaaacagacagtttttccgaaaggcggaaaaacggatgcagcaattggcacagttacaacagcatcaaaatcaat
tgcagcaggcgcacatcagcagatgttacaagcacaattacagcagcaagcccgaaactgtggtcaattaaaaacccg
aacctgtatacagttagaacagaagttctgaacggaggcaagttctggatacatatgatacagaatatggccttcc
gctggacaggctttgatgcaacatcaggcttttactgaatggcgaaaaagtcaaaactgaaaggcgttagcatgca
tcatgatcaaggctcacttggcgcagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaaatg
ggcggtcaatagcattcgcaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaaggcgttc
tggttgctgaagaagtctttgatattgtggaaccgcagcaaaaatggcaacacggaagattatggcaaatggtttgg
ccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttacatca
acaattaaccgcgatagaaatgcaccgtcagttattatgtggtcactgggcaatgaaatgatggaaggcatttcag
gctcagtttcaggctttccggcaacatcagcaaaactggttgcatggacaaaagcagcagattcaacaagaccgat
gacatatggcgataacaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaatggc
ggagttgttggcacaattattcagatggcgcaaaactatgataaaattcgtacaacacatccgtcatgggcaattt
atggctcagaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatcaga
taaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatggatatgatgttgtt
cagagagattttgtcgcaggcacatatgtttggacaggatttgattatctgggcgaaccgacaccgtggaatggca
caggctcaggcgcagttggctcatggcgtcaccgaaaaatagctattttggcatcgttgatacagcaggctttcc
gaaagatacatattatttttatcagagccagtggaatgatgatgttcatacactgcatattcttccggcatggaat
gaaaatgttgttgcaaaaggctcaggcaataatgttccggttgctgtttatacagatgcagcgaagtgaaactgt
attttacaccgaaaggctcaacagaaaaagactgatcggcgaaaaatcatttcaaaaaaaacaacagcggcagg
ctatacatatcaagtcattgaaggcagcgataaagattcaacagcgcataaaaaacatgtatctgacatggaatgtt
ccgtgggcagaaggcacaatttcagcgggaagcgtatgatgaaaataatcgctgattccggaaggcagcacagaag
gcaacgcatcagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagcggg
tggaagatctgtcatatattgaagtcgatgtcacagatgcaaatggccatatgttccggatgcagcaaataga
gtcacatttgatgttaaaggcgcaggcaaaactggttgcggttgataatggctcatcaccggatcatgattcatatc
aagcggataaccgcaaagcattttcaggcaaagtcctggcaattgttcagtcaacaaaagaagcaggcgaattac
agttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagcaca
gaaaaaacagtcgcagcttttattacagccgcaactattatgtcaaaacaggcaacaaaccgattctgccgtcag
atgttgaagttcgctattcagatggaacaagcgatagacaaaacgttacatgggatgcagtttcagatgatcaaat
tgcaaaagcaggctcattttcagttgcaggcacagttgcaggccaaaaaattagcgttcgcgtcacaatgattgat

gaaattggcgactgctgaattattcagcaagcacaccggttggcacaccggcagttcttccgggatcaagaccgg
cagtcctgccggatggcacagtcacatcagcaattttgcagtcattggacaaaaccggcagatacagtcataa
tacagcaggcacagtcaaagtaccgggaacagcaacagtttttggcaaagaatttaaagtcacagcgacaattaga
gttcaa

[0704] >SEQ ID NO:12编码BIF_1172的核苷酸序列

[0705] gttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagttgtttattcatc
agcggctgatagcaaacaaaatcgcacaaagcattttgatgcgaactggaaatttatgctgtcagatagcgttcaa
gcacaagatccggcattttgatgattcagcatggcaacaagttgatctgccgatgattatagcatcacacagaaat
atagccaaagcaatgaagcagaatcagcatatcttccgggaggcacaggctggtatagaaaaagctttacaattga
tagagatctggcaggcaaacgcattgcatgtaattttgatggcgctctatatgaatgcaacagtcctggtttaatggc
gttaaactgggcacacatccgtatggctattcacggttttcatttgatctgacaggcaatgcaaaatttggcggag
aaaacacaattgtcgtcaaagttgaaatagactgccgtcatcaagatggtattcaggcagcggcatttatagaga
tgtttacactgacagttacagatggcgcttcatgttggcaataatggcgctgcaattaaaacaccgtcactggcaaca
caaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacattacac
tgaaacagacagtttttccgaaaggcggaaaaacggatgcagcaattggcacagttacaacagcatcaaaatcaat
tgcagcaggcgcacatcagcagatgttacaagcacaattacagcagcaagcccgaaactgtggtcaattaaaaaccg
aacctgtatacagttagaacagaagttctgaacggaggcaaaagttctggatacatatgatacagaatatggccttc
gctggacaggctttgatgcaacatcaggcttttactgaatggcgaaaaagtcaaactgaaaggcgttagcatgca
tcatgatcaaggctcacttggcgagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaaatg
ggcgctcaatagcattcgcacaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaaggcgcttc
tggttgctgaagaagtctttgatatgtggaaccgcagcaaaaatggcaacacggaagattatggcaaatggtttgg
ccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttacatca
acaattaaccgcgatagaaatgcaccgtcagttattatgtggtcactgggcaatgaaatgatggaaggcatttcag
gctcagtttcaggctttccggcaacatcagcaaaactgggttgcatggacaaaagcagcagattcaacaagaccgat
gacatatggcgataacaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaatggc
ggagttgttggcacaattattcagatggcgcaaaactatgataaaattcgtacaacacatccgtcatgggcaattt
atggctcagaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatcaga
taaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatggtatgatgttggt
cagagagattttgtcgcaggcacatatgtttggacaggatttgattatctgggcgaaccgacaccgtggaatggca
caggctcaggcgcagttggctcatggcgtcacggaaaaatagctattttggcatcgttgatacagcaggctttcc
gaaagatacatattatttttatcagagccagtggaaatgatgatgttcatacactgcataattcttccggcatggaat
gaaaatgttggttgcaaaaggctcaggcaataatgttccggttgctgtttatacagatgcagcgaaagtgaactgt
attttacaccgaaaggctcaacagaaaaagactgatcggcgaaaaatcatttacaaaaaaacaacagcggcagg
ctatacatatcaagtcctatgaaggcagcgataaagattcaacagcgcataaaaaacatgtatctgacatggaatgtt
ccgtgggcagaaggcacaatttcagcggaaagcgtatgatgaaaataatcgctgattccggaaggcagcacagaag
gcaacgcacagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagcgga
tggcaaagatctgtcatatattgaagtcgatgtcacagatgcaaatggccatattgttccggatgcagcaaataga
gtcacatttgatgttaaaggcgcaggcaaaactggttggcgcttgataatggctcatcaccggatcatgattcatatc
aagcggataaccgcaaagcattttcaggcaaagtcctggcaattgttcagtcacaacaaaagaagcaggcgaaattac

agttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagcaca
gaaaaaacagtcgcgagcttttattacagccgcaactattatgtcaaaacaggcaacaaaccgattctgccgtcag
atgttgaagttcgctattcagatggaacaagcgatagacaaaacgttacatgggatgcagtttcagatgatcaaat
tgcaaaagcaggctcattttcagttgcaggcacagttgcaggccaaaaaattagcgttcgcgtcacaatgattgat
gaaattggcgactgctgaattattcagcaagcacaccggttggcacaccggcagttcttcgggatcaagaccgg
cagtcctgccggatggcacagtcacatcagcaaattttgagtcattggacaaaaccggcagatacagtcataa
tacagcaggcacagtc aaagtaccgggaacagcaacagtttttgcaaaagaatttaaagtcacagcgacaattaga
gttcaaagaagccaagttacaattgggtcatcagtttcaggaaatgcactgagactgacacaaaatattccggcag
ataacaatcagatacactggatgcgattaaagatggctcaacaacagttgatgcaaatacaggcggaggcgcaaa
tccgtcagcatggacaaattgggcatattcaaaagcaggccataacacagcggaaattacatttgaatatgcgaca
gaacaacaactgggccagatcgctcatgtatttttttcgcgatagcaatgcagttagatttccggatgctggcaaaa
caaaaattcagatc

[0706] >SEQ ID NO:13编码BIF_1241的核苷酸序列

[0707] gttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagttgtttattcattc
agcggtcgatagcaaaacaaaatcgcaacgcattttgatgcgaactggaaatttatgtgtcagatagcgttcaa
gcacaagatccggcatttgatgattcagcatggcaacaagttgatctgccgcatgattatagcatcacacagaat
atagccaaagcaatgaagcagaatcagcatatcttcggggaggcacaggctggtatagaaaaagctttacaattga
tagagatctggcaggcaaacgcattgcgattaattttgatggcgtctatatgaatgcaacagtcctggtttaatggc
gttaaactgggcacacatccgtatggctattcaccgttttcatattgatctgacaggcaatgcaaaatttggcggag
aaaacacaattgtcgtcaaagttgaaaatagactgccgtcatcaagatggtattcaggcagcggcatttatagaga
tgtttacactgacagttacagatggcgttcatgttggcaataatggcgtcgcaattaaaacaccgtcactggcaaca
caaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacattcac
tgaaacagacagtttttccgaaaggcggaacacggatgcagcaattggcacagttacaacagcatcaaaatcaat
tgcagcaggcgcattcagcagatgttacaagcacaattacagcagcaagcccgaaactgtggtcaattaaaaaccg
aacctgtatacagttagaacagaagttctgaacggaggcaagttctggatacatatgatacagaatatggccttc
gctggacaggctttgatgcaacatcaggcttttactgaatggcgaaaaagtcaactgaaaggcgttagcatgca
tcatgatcaaggctcacttggcgcagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaaatg
ggcgtcaatagcattcgcaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaggcgttc
tggttgctgaagaagtccttgatatgtggaaccgcagcaaaaatggcaacacggaagattatggcaaatggtttgg
ccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttacatca
acaattaaccgcgatagaaatgcaccgtcagttattatgttggtcactgggcaatgaaatgatggaaggcatttcag
gctcagtttcaggctttccggcaacatcagcaaaactggttgcatggacaaaagcagcagattcaacaagaccgat
gacatatggcgataacaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaatggc
ggagttgttggcacaattattcagatggcgcaaaactatgataaaattcgtaacacatccgtcatgggcaattt
atggctcagaaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatcaga
taaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatggtatgatgttgtt
cagagagattttgtcgcaggcacatatgtttggacaggatttgattatctgggcgaaccgacaccgtggaatggca
caggctcaggcgcagttggctcatggccgtcaccgaaaaatagctattttggcatcgttgatacagcaggctttcc
gaaagatacatattatttttatcagagccagtggaatgatgatgttcatacactgcatattcttcggcatggaat

gaaaatgttgttgcaaaaggctcaggcaataatgttccggttgctgtttatacagatgcagcgaaagtgaactgtt
attttacaccgaaaggctcaacagaaaaagactgatcggcgaaaaatcatttacaacacagcggcagg
ctatacatatcaagtctatgaaggcagcgataaagattcaacagcgcataaaaacatgtatctgacatggaatgtt
ccgtgggcagaaggcacaatttcagcgggaagcgtatgatgaaaataatcgctgattccggaaggcagcacagaag
gcaacgcatcagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagcgga
tggaagatctgtcatatattgaagtcgatgtcacagatgcaaattggccatattgttccggatgcagcaaataga
gtcacatttgatgttaaaggcgcaggcaaactggttggcgttgataatggctcatcaccggatcatgattcatatc
aagcggataaccgcaaagcattttcaggcaaagtcctggcaattgttcagtcaacaaaagaagcaggcgaaattac
agttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagcaca
gaaaaaacagtcgcagcttttattacagccgcaactattatgtcaaacaggcaacaaaccgattctgccgtcag
atgttgaaagttcgctattcagatggaacaagcgatagacaaaacgttacatgggatgcagtttcagatgatcaaat
tgcaaaagcaggtcattttcagttgcaggcacagttgcaggccaaaaattagcgttcgcgtcacaatgattgat
gaaattggcgcactgctgaattattcagcaagcacaccggttgccacaccggcagttcttcgggatcaagaccgg
cagtcctgccggatggcacagtcacatcagcaattttgcagttccattggacaaaaccggcagatacagtcctataa
tacagcaggcacagtcaaagtaccgggaacagcaacagtttttgcaaaagaatttaaagtcacagcgacaattaga
gttcaaagaagccaagttacaattggctcatcagtttcaggaaatgcactgagactgacacaaaatattccggcag
ataacaatcagatacactggatgcatgaaagatggctcaacaacagttgatgcaaatacaggcggaggcgcaaa
tccgtcagcatggacaaattgggcatattcaaaagcaggccataacacagcggaaattacatttgaaatgcgaca
gaacaacaactgggccagatcgctcatgtatttttttcgcatagcaatgcagttagatttcgggatgctggcaaaa
caaaaattcagatcagcgcagatggcaaaaattggacagatctggcagcaacagaaacaattgcagcgcagaatc
aagcgatagagtcaaaccgtatacatatgattttgcaccggttggcgcaacatttgtaaagtgcagtcacaaac
gcagatacaacaacaccgtcaggcgttggttgcgcaggcctgcagaaaattgaactgaaaacagcgaca

[0708] >SEQ ID NO:14编码BIF_1326的核苷酸序列

[0709] gttgaagatgcaacaagaagcgatagcacaacacaaatgtcatcaacaccggaagtgtttattcatc
agcggtcgatagcaaacaaaatcgcacaaagcattttgatgcgaactggaaatttatgctgtcagatagcgttcaa
gcacaagatccggcatttgatgattcagcatggcaacaagttgatctgccgcatgattatagcatcacacagaat
atagccaaagcaatgaagcagaatcagcatatcttcgggaggcacaggctggtatagaaaaagctttacaattga
tagagatctggcaggcaaacgcatttgcgattaattttgatggcgtctatatgaatgcaacagtcctggtttaattggc
gttaaactgggcacacatccgtatggctattaccgttttcatttgatctgacaggcaatgcaaaatttgccggag
aaaacacaattgtcgtaaaagttgaaatagactgccgtcatcaagatggattcaggcagcggcatttatagaga
tgttacactgacagttacagatggcgttcatgttggcaataatggcgtcgcaattaaaacaccgtcactggcaaca
caaaatggcggagatgtcacaatgaacctgacaacaaaagtcgcgaatgatacagaagcagcagcgaacattacac
tgaaacagacagtttttccgaaaggcggaaaaacggatgcagcaattggcacagttacaacagcatcaaaatcaat
tgcagcaggcgcacatcagcagatgttacaagcacaattacagcagcaagccgaaactgtggtcaattaaaaaccg
aacctgtatacagttagaacagaagttctgaacggaggcaaagttctggatacatatgatacagaatatggctttc
gctggacaggtttgatgcaacatcaggcttttactgaatggcgaaaaagtcaaactgaaaggcgttagcatgca
tcatgatcaaggctcacttggcgcagttgcaaatagacgcgcaattgaaagacaagtcgaaatcctgcaaaaaatg
ggcgtcaatagcattcgcaaacacataatccggcagcaaaagcactgattgatgtctgcaatgaaaaaggcgttc
tggttgctgaagaagtccttgatattgtgaaccgcagcaaaaatggcaacacggaagattatggcaaatggttttg

ccaagcaattgcaggcgataatgcagttctgggaggcgataaagatgaaacatgggcgaaatttgatcttacatca
acaattaaccgcgatagaaatgcaccgtcagttattatgtgtggtcactgggcaatgaaatgatggaaggcatttcag
gctcagtttcaggctttccggcaacatcagcaaaactggttgcatggacaaaagcagcagattcaacaagaccgat
gacatatggcgataacaaaattaaagcgaactggaacgaatcaaatacaatgggcgataatctgacagcaaatggc
ggagttgttggcacaattattcagatggcgcaaactatgataaaattcgtacaacacatccgtcatgggcaattt
atggctcagaaacagcatcagcgattaatagccgtggcattttataatagaacaacaggcggagcacaatcatcaga
taaacagctgacaagctatgataattcagcagttggctggggagcagttgcatcatcagcatgggtatgatgttgtt
cagagagattttgtcgcaggcacatatgtttggacaggatttgattatctggggaacccgacaccgtggaatggca
caggctcaggcgcagttggctcatggcgtcacggaaaaatagctattttggcatcggtgatacagcaggctttcc
gaaagatacatattatttttatcagagccagtggaaatgatgatgttcatacactgcatattcttccggcatggaat
gaaaatgttgttgcaaaaggctcaggcaataatgttccggttgctgtttatacagatgcagcgaaagtgaactgt
attttacaccgaaaggctcaacagaaaaagactgatcggcgaaaaatcatttacaaaaaaacaacagcggcagg
ctatacatatcaagtctatgaaggcagcgataaagattcaacagcgcataaaaacatgtatctgacatggaatgtt
ccgtgggcagaaggcacaatttcagcggaaagcgtatgatgaaaataatcgctgattccggaaggcagcacagaag
gcaacgcacatcagttacaacaacaggcaaagcagcaaaactgaaagcagatgcggatcgcaaaacaattacagcgga
tggcaaagatctgtcatatattgaagtcgatgtcacagatgcaaattggccatattgttccggatgcagcaaataga
gtcacatttgatgttaaaggcgcaggcaaactggttggcggttgataatggctcatcaccggatcatgattcatatc
aagcggataaccgcaaagcattttcaggcaaagtcctggcaattgttcagtcacaaaaagaagcaggcgaaattac
agttacagcaaaagcagatggcctgcaatcaagcacagttaaaattgcaacaacagcagttccgggaacaagcaca
gaaaaaacagtccgcagcttttattacagccgcaactattatgtcaaaacaggcaacaaaccgattctgccgtcag
atgttgaagttcgctattcagatggaacaagcgatagacaaaacggttacatgggatgcagtttcagatgatcaaat
tgcaaaagcaggctcattttcagttgcaggcacagttgcaggccaaaaaattagcgttcgcgtcacaatgattgat
gaaattggcgactgctgaattattcagcaagcacaccggttggcacaccggcagttcttccgggatcaagaccgg
cagtcctgccggatggcacagtcacatcagcaaattttgagttccattggacaaaaccggcagatacagtcataa
tacagcaggcacagtcaaagtaccgggaacagcaacagtttttggcaaagaatttaaagtcacagcgacaattaga
gttcaaagaagccaagttacaattggctcatcagtttcaggaaatgcactgagactgacacaaaatattccggcag
ataacaacatcagatacactggatgagattaaagatggctcaacaacagttgatgcaaatacaggcggaggcgcaaa
tccgtcagcatggacaaattgggcatattcaaaagcaggccataacacagcggaaattacatttgaatatgcgaca
gaacaacaactgggccagatcgctcatgtatttttttcgcatagcaatgcagtttagatttccggatgctggcaaaa
caaaaattcagatcagcgcagatggcaaaaattggacagatctggcagcaacagaaacaattgcagcgcaagaatc
aagcgatagagtcaaaccgtatacatatgattttgcaccggttggcgcaacatttgttaaagtgacagtcacaaac
gcagatacaacaacaccgtcaggcgttgtttgcgcaggcctgacagaaattgaactgaaaacagcgacaagcaaat
ttgtcacaatacatcagcagcactgtcatcacttacagtcattggcacaacaaagtttcagattcagttctggcagc
aggctcatataacacaccggcaattatcgcatgtttaaagcggaaggcgaaaggcaatgcaagcgttacagtcctt
ccggcacatgataatgttattcgctcattacagaaagcgaagatcatgtcacacgcaaaacatttacaatcaacc
tgggcacagaacaagaattt

[0710] >SEQ ID NO:15用于生成BIF变体的正向引物

[0711] GGGGTAAGTGGGAAGATGCAACAAGAAG

[0712] >SEQ ID NO:16BIF917的反向引物

- [0713] GCGCTTAATTAATTATGTTTTTCTGTGCTTGTTTC
- [0714] >SEQ ID NO:17BIF995的反向引物
- [0715] GCGCTTAATTAATTACAGTGCGCCAATTCATCAATCA
- [0716] >SEQ ID NO:18BIF1068的反向引物
- [0717] GCGCTTAATTAATTATTGAACCTAATTGTCGCTG
- [0718] >SEQ ID NO:19BIF1241的反向引物
- [0719] GCGCTTAATTAATTATGTCGCTGTTTCAGTTCAAT
- [0720] >SEQ ID NO:20BIF1326的反向引物
- [0721] GCGCTTAATTAATTAATAATTCTTGTCTGTGCCCA
- [0722] >SEQ ID NO:21BIF1478的反向引物
- [0723] GCGCTTAATTAATTATCTCAGTCTAATTCGCTTGCGC
- [0724] >SEQ ID NO:22两歧双歧杆菌BIF1750
- [0725] vedatrsdsttqmsstpevyssavdskqnrtssdfdanwkmfmlsdsvqaqdpafddsawqqvdlphdy
sitqkysqsneaesaylpggtgwyksftidrdlagkriainfdgvymnatvwfngvklgthpygyspfsfdltgn
akfggentivvkvenrlpssrwysgsgiyrdvtltvtgdgvhvgngngaiktptlatqnggdvtmnlrttkvandtea
aanitlkqtvfpkggktdaaigtvttasksiaagasadvsttitaasplwsiknpnlytvrtevlnggkvldtyd
teygfrrwtgfdatsgfslngekvklkgvsmhhdqgslgavanrraierqveilqkmgvnsirtthnpaakalidvc
nekgvlvveevfdmwnrskngntedygkwfgqaiagdnavlbgdkdetwakfdltstinrdnapsvimwslgnem
megisgsvsgfpatsaklvawtkaadstrpmtgydnkikanwnesntmgdnl tanggvgt nysdganydkirtth
pswaiygsetasainsrgiynrttggagssdkqltsydnasvlgwagavassawydvvqrdfvagtyvwtgfdylgep
tpwngtgsgavgswwpspknsyfgivdtagfpkdtyyfyqsqwnddvhtlhlipawnenvvaksggnnpvvvytda
akvklyftpkgstekrligeksftkkttagtytyqvyegsdkdstahknmyltnvpwaegtisaeaydenrnlip
egstegnasvtttgaaklkadadrktitadgkdlsyievdvtdanghivpdaanrvtfdvkgagklvgvngssp
dhdsyqadnrkafsgkvlaivqstkeageitvtakadglqsstvkiaattavpgtstektvrsfyysrnyyvtgnk
pilpsdvevrysdgtsdrqnvtdavsdqdiakagsfsvagt vagqkisvrvtmideigallnysastpvgtpavl
pgsrpavlpdgtvtsanfahwtkpadtvynntagtvkvpgtatvfgkefkvtatirvqrsqvtigssvsgnalrlt
qnipadkqsdltldaikdgsttvdantggganpsawtnwayskaghntaeitfeiateqqlgqivmyffrdsnavrf
pdagktkiqisadgknwtdlaaetiaaqessdrvkpytydfapvgatfvkvtvtnadtttsgvvcaglteielk
tatskfvtntsaalssltnvgtkvsvlaagsyntpaiiadvkaegegnasvtvlpahdnvirvitesedhvtrk
tftinlgteqefpadsderdypaadmvtvtvgseqtsgtategpkkfavdgntstywhsnwtpttndlwiafelqk
ptkladalrylprpagskngsvteykvqvsddgtnwtdagsgtwttdygwklaefnqpvttkhvrkavhtyadsgn
dkfmsaseirlrkavdttdisgatvtpakltvdrvdadhpafatkdvtvtlgdatlrygvdyllldyagntavgk
atvtvrgidkysgtvaktftielknapapeptltsvsvktkpskltyvvgdafdpaglvlqhdrqadrppqplvge
qadergltcgtrechrveqlrkhenreahrtgldhlefvgaadgavgeqatfkvhvhadqgdgrhddaderdidphv
pvdhavgelaraachhviglrvdthrlkasgfqipaddmaeidritgfhrferhvg
- [0726] >SEQ ID NO:23来自两歧双歧杆菌DSM20215的胞外乳糖酶的信号序列
- [0727] Vrskklwisllfalaliftmafgtssaq

序列表

<110> 杜邦营养生物科学有限公司 (DuPont Nutrition Biosciences ApS)
 <120> 宿主细胞
 <130> P105952PCT
 <150> GB 1419897.2
 <151> 2014-11-07
 <150> GB 1419894.9
 <151> 2014-11-07
 <150> GB 1515645.8
 <151> 2015-09-03
 <150> GB1419900.4
 <151> 2014-11-07
 <160> 23
 <170> PatentIn 3.5版
 <210> 1
 <211> 887
 <212> PRT
 <213> 人工序列
 <220>
 <223> 截短的乳糖酶片段
 <400> 1

Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
 1 5 10 15

[0001]

Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
 20 25 30

Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
 35 40 45

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
 50 55 60

Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
 65 70 75 80

Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
 85 90 95

Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
 100 105 110

Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
 115 120 125

Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp Leu Thr Gly Asn
 130 135 140

Ala Lys Phe Gly Gly Glu Asn Thr Ile Val Val Lys Val Glu Asn Arg
 145 150 155 160

Leu Pro Ser Ser Arg Trp Tyr Ser Gly Ser Gly Ile Tyr Arg Asp Val
 165 170 175

	Thr	Leu	Thr	Val	Thr	Asp	Gly	Val	His	Val	Gly	Asn	Asn	Gly	Val	Ala	
				180					185					190			
	Ile	Lys	Thr	Pro	Ser	Leu	Ala	Thr	Gln	Asn	Gly	Gly	Asp	Val	Thr	Met	
			195					200					205				
	Asn	Leu	Thr	Thr	Lys	Val	Ala	Asn	Asp	Thr	Glu	Ala	Ala	Ala	Asn	Ile	
		210					215						220				
	Thr	Leu	Lys	Gln	Thr	Val	Phe	Pro	Lys	Gly	Gly	Lys	Thr	Asp	Ala	Ala	
	225					230					235					240	
	Ile	Gly	Thr	Val	Thr	Thr	Ala	Ser	Lys	Ser	Ile	Ala	Ala	Gly	Ala	Ser	
				245						250					255		
	Ala	Asp	Val	Thr	Ser	Thr	Ile	Thr	Ala	Ala	Ser	Pro	Lys	Leu	Trp	Ser	
			260						265					270			
	Ile	Lys	Asn	Pro	Asn	Leu	Tyr	Thr	Val	Arg	Thr	Glu	Val	Leu	Asn	Gly	
			275					280					285				
	Gly	Lys	Val	Leu	Asp	Thr	Tyr	Asp	Thr	Glu	Tyr	Gly	Phe	Arg	Trp	Thr	
		290					295					300					
	Gly	Phe	Asp	Ala	Thr	Ser	Gly	Phe	Ser	Leu	Asn	Gly	Glu	Lys	Val	Lys	
	305					310					315					320	
[0002]	Leu	Lys	Gly	Val	Ser	Met	His	His	Asp	Gln	Gly	Ser	Leu	Gly	Ala	Val	
					325					330					335		
	Ala	Asn	Arg	Arg	Ala	Ile	Glu	Arg	Gln	Val	Glu	Ile	Leu	Gln	Lys	Met	
				340					345					350			
	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ala	Lys	Ala	Leu	
		355						360					365				
	Ile	Asp	Val	Cys	Asn	Glu	Lys	Gly	Val	Leu	Val	Val	Glu	Glu	Val	Phe	
		370					375					380					
	Asp	Met	Trp	Asn	Arg	Ser	Lys	Asn	Gly	Asn	Thr	Glu	Asp	Tyr	Gly	Lys	
	385					390					395					400	
	Trp	Phe	Gly	Gln	Ala	Ile	Ala	Gly	Asp	Asn	Ala	Val	Leu	Gly	Gly	Asp	
				405						410					415		
	Lys	Asp	Glu	Thr	Trp	Ala	Lys	Phe	Asp	Leu	Thr	Ser	Thr	Ile	Asn	Arg	
			420						425					430			
	Asp	Arg	Asn	Ala	Pro	Ser	Val	Ile	Met	Trp	Ser	Leu	Gly	Asn	Glu	Met	
			435					440					445				
	Met	Glu	Gly	Ile	Ser	Gly	Ser	Val	Ser	Gly	Phe	Pro	Ala	Thr	Ser	Ala	
		450					455					460					
	Lys	Leu	Val	Ala	Trp	Thr	Lys	Ala	Ala	Asp	Ser	Thr	Arg	Pro	Met	Thr	
	465					470					475					480	

	Tyr	Gly	Asp	Asn	Lys	Ile	Lys	Ala	Asn	Trp	Asn	Glu	Ser	Asn	Thr	Met	
					485					490						495	
	Gly	Asp	Asn	Leu	Thr	Ala	Asn	Gly	Gly	Val	Val	Gly	Thr	Asn	Tyr	Ser	
				500					505						510		
	Asp	Gly	Ala	Asn	Tyr	Asp	Lys	Ile	Arg	Thr	Thr	His	Pro	Ser	Trp	Ala	
			515					520					525				
	Ile	Tyr	Gly	Ser	Glu	Thr	Ala	Ser	Ala	Ile	Asn	Ser	Arg	Gly	Ile	Tyr	
		530					535					540					
	Asn	Arg	Thr	Thr	Gly	Gly	Ala	Gln	Ser	Ser	Asp	Lys	Gln	Leu	Thr	Ser	
	545					550					555					560	
	Tyr	Asp	Asn	Ser	Ala	Val	Gly	Trp	Gly	Ala	Val	Ala	Ser	Ser	Ala	Trp	
					565					570						575	
	Tyr	Asp	Val	Val	Gln	Arg	Asp	Phe	Val	Ala	Gly	Thr	Tyr	Val	Trp	Thr	
				580					585						590		
	Gly	Phe	Asp	Tyr	Leu	Gly	Glu	Pro	Thr	Pro	Trp	Asn	Gly	Thr	Gly	Ser	
			595					600					605				
	Gly	Ala	Val	Gly	Ser	Trp	Pro	Ser	Pro	Lys	Asn	Ser	Tyr	Phe	Gly	Ile	
		610					615					620					
[0003]	Val	Asp	Thr	Ala	Gly	Phe	Pro	Lys	Asp	Thr	Tyr	Tyr	Phe	Tyr	Gln	Ser	
	625					630					635					640	
	Gln	Trp	Asn	Asp	Asp	Val	His	Thr	Leu	His	Ile	Leu	Pro	Ala	Trp	Asn	
					645					650					655		
	Glu	Asn	Val	Val	Ala	Lys	Gly	Ser	Gly	Asn	Asn	Val	Pro	Val	Val	Val	
				660					665					670			
	Tyr	Thr	Asp	Ala	Ala	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Pro	Lys	Gly	Ser	
			675					680					685				
	Thr	Glu	Lys	Arg	Leu	Ile	Gly	Glu	Lys	Ser	Phe	Thr	Lys	Lys	Thr	Thr	
		690					695					700					
	Ala	Ala	Gly	Tyr	Thr	Tyr	Gln	Val	Tyr	Glu	Gly	Ser	Asp	Lys	Asp	Ser	
	705					710					715					720	
	Thr	Ala	His	Lys	Asn	Met	Tyr	Leu	Thr	Trp	Asn	Val	Pro	Trp	Ala	Glu	
					725					730					735		
	Gly	Thr	Ile	Ser	Ala	Glu	Ala	Tyr	Asp	Glu	Asn	Asn	Arg	Leu	Ile	Pro	
				740					745					750			
	Glu	Gly	Ser	Thr	Glu	Gly	Asn	Ala	Ser	Val	Thr	Thr	Thr	Gly	Lys	Ala	
			755					760					765				
	Ala	Lys	Leu	Lys	Ala	Asp	Ala	Asp	Arg	Lys	Thr	Ile	Thr	Ala	Asp	Gly	
		770					775					780					

Lys Asp Leu Ser Tyr Ile Glu Val Asp Val Thr Asp Ala Asn Gly His
785 790 795 800

Ile Val Pro Asp Ala Ala Asn Arg Val Thr Phe Asp Val Lys Gly Ala
805 810 815

Gly Lys Leu Val Gly Val Asp Asn Gly Ser Ser Pro Asp His Asp Ser
820 825 830

Tyr Gln Ala Asp Asn Arg Lys Ala Phe Ser Gly Lys Val Leu Ala Ile
835 840 845

Val Gln Ser Thr Lys Glu Ala Gly Glu Ile Thr Val Thr Ala Lys Ala
850 855 860

Asp Gly Leu Gln Ser Ser Thr Val Lys Ile Ala Thr Thr Ala Val Pro
865 870 875 880

Gly Thr Ser Thr Glu Lys Thr
885

<210> 2
<211> 965
<212> PRT
<213> 人工序列

<220>
<223> 截短的乳糖酶片段

[0004]

<400> 2

Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
1 5 10 15

Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
20 25 30

Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
35 40 45

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
50 55 60

Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
65 70 75 80

Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
85 90 95

Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
100 105 110

Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
115 120 125

Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp Leu Thr Gly Asn
130 135 140

	Ala	Lys	Phe	Gly	Gly	Glu	Asn	Thr	Ile	Val	Val	Lys	Val	Glu	Asn	Arg	
	145					150					155					160	
	Leu	Pro	Ser	Ser	Arg	Trp	Tyr	Ser	Gly	Ser	Gly	Ile	Tyr	Arg	Asp	Val	
					165					170					175		
	Thr	Leu	Thr	Val	Thr	Asp	Gly	Val	His	Val	Gly	Asn	Asn	Gly	Val	Ala	
				180					185					190			
	Ile	Lys	Thr	Pro	Ser	Leu	Ala	Thr	Gln	Asn	Gly	Gly	Asp	Val	Thr	Met	
			195					200					205				
	Asn	Leu	Thr	Thr	Lys	Val	Ala	Asn	Asp	Thr	Glu	Ala	Ala	Ala	Asn	Ile	
		210					215					220					
	Thr	Leu	Lys	Gln	Thr	Val	Phe	Pro	Lys	Gly	Gly	Lys	Thr	Asp	Ala	Ala	
	225					230					235					240	
	Ile	Gly	Thr	Val	Thr	Thr	Ala	Ser	Lys	Ser	Ile	Ala	Ala	Gly	Ala	Ser	
				245						250					255		
	Ala	Asp	Val	Thr	Ser	Thr	Ile	Thr	Ala	Ala	Ser	Pro	Lys	Leu	Trp	Ser	
				260					265					270			
	Ile	Lys	Asn	Pro	Asn	Leu	Tyr	Thr	Val	Arg	Thr	Glu	Val	Leu	Asn	Gly	
			275					280					285				
[0005]	Gly	Lys	Val	Leu	Asp	Thr	Tyr	Asp	Thr	Glu	Tyr	Gly	Phe	Arg	Trp	Thr	
		290					295					300					
	Gly	Phe	Asp	Ala	Thr	Ser	Gly	Phe	Ser	Leu	Asn	Gly	Glu	Lys	Val	Lys	
	305					310					315					320	
	Leu	Lys	Gly	Val	Ser	Met	His	His	Asp	Gln	Gly	Ser	Leu	Gly	Ala	Val	
					325					330					335		
	Ala	Asn	Arg	Arg	Ala	Ile	Glu	Arg	Gln	Val	Glu	Ile	Leu	Gln	Lys	Met	
				340					345					350			
	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ala	Lys	Ala	Leu	
		355					360						365				
	Ile	Asp	Val	Cys	Asn	Glu	Lys	Gly	Val	Leu	Val	Val	Glu	Glu	Val	Phe	
		370					375					380					
	Asp	Met	Trp	Asn	Arg	Ser	Lys	Asn	Gly	Asn	Thr	Glu	Asp	Tyr	Gly	Lys	
	385					390					395					400	
	Trp	Phe	Gly	Gln	Ala	Ile	Ala	Gly	Asp	Asn	Ala	Val	Leu	Gly	Gly	Asp	
				405						410					415		
	Lys	Asp	Glu	Thr	Trp	Ala	Lys	Phe	Asp	Leu	Thr	Ser	Thr	Ile	Asn	Arg	
				420					425					430			
	Asp	Arg	Asn	Ala	Pro	Ser	Val	Ile	Met	Trp	Ser	Leu	Gly	Asn	Glu	Met	
		435						440					445				

	Met	Glu	Gly	Ile	Ser	Gly	Ser	Val	Ser	Gly	Phe	Pro	Ala	Thr	Ser	Ala	
	450						455					460					
	Lys	Leu	Val	Ala	Trp	Thr	Lys	Ala	Ala	Asp	Ser	Thr	Arg	Pro	Met	Thr	
	465					470					475					480	
	Tyr	Gly	Asp	Asn	Lys	Ile	Lys	Ala	Asn	Trp	Asn	Glu	Ser	Asn	Thr	Met	
					485					490					495		
	Gly	Asp	Asn	Leu	Thr	Ala	Asn	Gly	Gly	Val	Val	Gly	Thr	Asn	Tyr	Ser	
				500					505					510			
	Asp	Gly	Ala	Asn	Tyr	Asp	Lys	Ile	Arg	Thr	Thr	His	Pro	Ser	Trp	Ala	
			515					520					525				
	Ile	Tyr	Gly	Ser	Glu	Thr	Ala	Ser	Ala	Ile	Asn	Ser	Arg	Gly	Ile	Tyr	
	530						535					540					
	Asn	Arg	Thr	Thr	Gly	Gly	Ala	Gln	Ser	Ser	Asp	Lys	Gln	Leu	Thr	Ser	
	545					550					555					560	
	Tyr	Asp	Asn	Ser	Ala	Val	Gly	Trp	Gly	Ala	Val	Ala	Ser	Ser	Ala	Trp	
					565					570					575		
	Tyr	Asp	Val	Val	Gln	Arg	Asp	Phe	Val	Ala	Gly	Thr	Tyr	Val	Trp	Thr	
				580					585					590			
[0006]	Gly	Phe	Asp	Tyr	Leu	Gly	Glu	Pro	Thr	Pro	Trp	Asn	Gly	Thr	Gly	Ser	
			595					600					605				
	Gly	Ala	Val	Gly	Ser	Trp	Pro	Ser	Pro	Lys	Asn	Ser	Tyr	Phe	Gly	Ile	
		610					615					620					
	Val	Asp	Thr	Ala	Gly	Phe	Pro	Lys	Asp	Thr	Tyr	Tyr	Phe	Tyr	Gln	Ser	
	625					630					635					640	
	Gln	Trp	Asn	Asp	Asp	Val	His	Thr	Leu	His	Ile	Leu	Pro	Ala	Trp	Asn	
					645					650					655		
	Glu	Asn	Val	Val	Ala	Lys	Gly	Ser	Gly	Asn	Asn	Val	Pro	Val	Val	Val	
				660					665				670				
	Tyr	Thr	Asp	Ala	Ala	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Pro	Lys	Gly	Ser	
			675					680					685				
	Thr	Glu	Lys	Arg	Leu	Ile	Gly	Glu	Lys	Ser	Phe	Thr	Lys	Lys	Thr	Thr	
		690					695					700					
	Ala	Ala	Gly	Tyr	Thr	Tyr	Gln	Val	Tyr	Glu	Gly	Ser	Asp	Lys	Asp	Ser	
	705					710					715					720	
	Thr	Ala	His	Lys	Asn	Met	Tyr	Leu	Thr	Trp	Asn	Val	Pro	Trp	Ala	Glu	
					725					730					735		
	Gly	Thr	Ile	Ser	Ala	Glu	Ala	Tyr	Asp	Glu	Asn	Asn	Arg	Leu	Ile	Pro	
				740					745					750			

[0007]

Glu Gly Ser Thr Glu Gly Asn Ala Ser Val Thr Thr Thr Gly Lys Ala
 755 760 765
 Ala Lys Leu Lys Ala Asp Ala Asp Arg Lys Thr Ile Thr Ala Asp Gly
 770 775 780
 Lys Asp Leu Ser Tyr Ile Glu Val Asp Val Thr Asp Ala Asn Gly His
 785 790 795 800
 Ile Val Pro Asp Ala Ala Asn Arg Val Thr Phe Asp Val Lys Gly Ala
 805 810 815
 Gly Lys Leu Val Gly Val Asp Asn Gly Ser Ser Pro Asp His Asp Ser
 820 825 830
 Tyr Gln Ala Asp Asn Arg Lys Ala Phe Ser Gly Lys Val Leu Ala Ile
 835 840 845
 Val Gln Ser Thr Lys Glu Ala Gly Glu Ile Thr Val Thr Ala Lys Ala
 850 855 860
 Asp Gly Leu Gln Ser Ser Thr Val Lys Ile Ala Thr Thr Ala Val Pro
 865 870 875 880
 Gly Thr Ser Thr Glu Lys Thr Val Arg Ser Phe Tyr Tyr Ser Arg Asn
 885 890 895
 Tyr Tyr Val Lys Thr Gly Asn Lys Pro Ile Leu Pro Ser Asp Val Glu
 900 905 910
 Val Arg Tyr Ser Asp Gly Thr Ser Asp Arg Gln Asn Val Thr Trp Asp
 915 920 925
 Ala Val Ser Asp Asp Gln Ile Ala Lys Ala Gly Ser Phe Ser Val Ala
 930 935 940
 Gly Thr Val Ala Gly Gln Lys Ile Ser Val Arg Val Thr Met Ile Asp
 945 950 955 960
 Glu Ile Gly Ala Leu
 965
 <210> 3
 <211> 1038
 <212> PRT
 <213> 人工序列
 <220>
 <223> 截短的乳糖酶片段
 <400> 3
 Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
 1 5 10 15
 Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
 20 25 30
 Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
 35 40 45

[0008]

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
 50 55 60
 Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
 65 70 75 80
 Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
 85 90 95
 Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
 100 105 110
 Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
 115 120 125
 Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp Leu Thr Gly Asn
 130 135 140
 Ala Lys Phe Gly Gly Glu Asn Thr Ile Val Val Lys Val Glu Asn Arg
 145 150 155 160
 Leu Pro Ser Ser Arg Trp Tyr Ser Gly Ser Gly Ile Tyr Arg Asp Val
 165 170 175
 Thr Leu Thr Val Thr Asp Gly Val His Val Gly Asn Asn Gly Val Ala
 180 185 190
 Ile Lys Thr Pro Ser Leu Ala Thr Gln Asn Gly Gly Asp Val Thr Met
 195 200 205
 Asn Leu Thr Thr Lys Val Ala Asn Asp Thr Glu Ala Ala Ala Asn Ile
 210 215 220
 Thr Leu Lys Gln Thr Val Phe Pro Lys Gly Gly Lys Thr Asp Ala Ala
 225 230 235 240
 Ile Gly Thr Val Thr Thr Ala Ser Lys Ser Ile Ala Ala Gly Ala Ser
 245 250 255
 Ala Asp Val Thr Ser Thr Ile Thr Ala Ala Ser Pro Lys Leu Trp Ser
 260 265 270
 Ile Lys Asn Pro Asn Leu Tyr Thr Val Arg Thr Glu Val Leu Asn Gly
 275 280 285
 Gly Lys Val Leu Asp Thr Tyr Asp Thr Glu Tyr Gly Phe Arg Trp Thr
 290 295 300
 Gly Phe Asp Ala Thr Ser Gly Phe Ser Leu Asn Gly Glu Lys Val Lys
 305 310 315 320
 Leu Lys Gly Val Ser Met His His Asp Gln Gly Ser Leu Gly Ala Val
 325 330 335
 Ala Asn Arg Arg Ala Ile Glu Arg Gln Val Glu Ile Leu Gln Lys Met
 340 345 350

	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ala	Lys	Ala	Leu	
			355					360							365		
	Ile	Asp	Val	Cys	Asn	Glu	Lys	Gly	Val	Leu	Val	Val	Glu	Glu	Val	Phe	
		370					375						380				
	Asp	Met	Trp	Asn	Arg	Ser	Lys	Asn	Gly	Asn	Thr	Glu	Asp	Tyr	Gly	Lys	
	385					390					395					400	
	Trp	Phe	Gly	Gln	Ala	Ile	Ala	Gly	Asp	Asn	Ala	Val	Leu	Gly	Gly	Asp	
				405						410					415		
	Lys	Asp	Glu	Thr	Trp	Ala	Lys	Phe	Asp	Leu	Thr	Ser	Thr	Ile	Asn	Arg	
				420					425					430			
	Asp	Arg	Asn	Ala	Pro	Ser	Val	Ile	Met	Trp	Ser	Leu	Gly	Asn	Glu	Met	
			435					440					445				
	Met	Glu	Gly	Ile	Ser	Gly	Ser	Val	Ser	Gly	Phe	Pro	Ala	Thr	Ser	Ala	
		450					455					460					
	Lys	Leu	Val	Ala	Trp	Thr	Lys	Ala	Ala	Asp	Ser	Thr	Arg	Pro	Met	Thr	
	465					470					475					480	
	Tyr	Gly	Asp	Asn	Lys	Ile	Lys	Ala	Asn	Trp	Asn	Glu	Ser	Asn	Thr	Met	
				485						490					495		
[0009]	Gly	Asp	Asn	Leu	Thr	Ala	Asn	Gly	Gly	Val	Val	Gly	Thr	Asn	Tyr	Ser	
			500						505					510			
	Asp	Gly	Ala	Asn	Tyr	Asp	Lys	Ile	Arg	Thr	Thr	His	Pro	Ser	Trp	Ala	
			515					520					525				
	Ile	Tyr	Gly	Ser	Glu	Thr	Ala	Ser	Ala	Ile	Asn	Ser	Arg	Gly	Ile	Tyr	
		530					535					540					
	Asn	Arg	Thr	Thr	Gly	Gly	Ala	Gln	Ser	Ser	Asp	Lys	Gln	Leu	Thr	Ser	
	545					550					555					560	
	Tyr	Asp	Asn	Ser	Ala	Val	Gly	Trp	Gly	Ala	Val	Ala	Ser	Ser	Ala	Trp	
				565						570					575		
	Tyr	Asp	Val	Val	Gln	Arg	Asp	Phe	Val	Ala	Gly	Thr	Tyr	Val	Trp	Thr	
				580					585					590			
	Gly	Phe	Asp	Tyr	Leu	Gly	Glu	Pro	Thr	Pro	Trp	Asn	Gly	Thr	Gly	Ser	
		595						600					605				
	Gly	Ala	Val	Gly	Ser	Trp	Pro	Ser	Pro	Lys	Asn	Ser	Tyr	Phe	Gly	Ile	
		610					615					620					
	Val	Asp	Thr	Ala	Gly	Phe	Pro	Lys	Asp	Thr	Tyr	Tyr	Phe	Tyr	Gln	Ser	
	625					630					635					640	
	Gln	Trp	Asn	Asp	Asp	Val	His	Thr	Leu	His	Ile	Leu	Pro	Ala	Trp	Asn	
				645						650					655		

[0010]

Glu Asn Val Val Ala Lys Gly Ser Gly Asn Asn Val Pro Val Val Val
 660 665 670
 Tyr Thr Asp Ala Ala Lys Val Lys Leu Tyr Phe Thr Pro Lys Gly Ser
 675 680 685
 Thr Glu Lys Arg Leu Ile Gly Glu Lys Ser Phe Thr Lys Lys Thr Thr
 690 695 700
 Ala Ala Gly Tyr Thr Tyr Gln Val Tyr Glu Gly Ser Asp Lys Asp Ser
 705 710 715 720
 Thr Ala His Lys Asn Met Tyr Leu Thr Trp Asn Val Pro Trp Ala Glu
 725 730 735
 Gly Thr Ile Ser Ala Glu Ala Tyr Asp Glu Asn Asn Arg Leu Ile Pro
 740 745 750
 Glu Gly Ser Thr Glu Gly Asn Ala Ser Val Thr Thr Thr Gly Lys Ala
 755 760 765
 Ala Lys Leu Lys Ala Asp Ala Asp Arg Lys Thr Ile Thr Ala Asp Gly
 770 775 780
 Lys Asp Leu Ser Tyr Ile Glu Val Asp Val Thr Asp Ala Asn Gly His
 785 790 795 800
 Ile Val Pro Asp Ala Ala Asn Arg Val Thr Phe Asp Val Lys Gly Ala
 805 810 815
 Gly Lys Leu Val Gly Val Asp Asn Gly Ser Ser Pro Asp His Asp Ser
 820 825 830
 Tyr Gln Ala Asp Asn Arg Lys Ala Phe Ser Gly Lys Val Leu Ala Ile
 835 840 845
 Val Gln Ser Thr Lys Glu Ala Gly Glu Ile Thr Val Thr Ala Lys Ala
 850 855 860
 Asp Gly Leu Gln Ser Ser Thr Val Lys Ile Ala Thr Thr Ala Val Pro
 865 870 875 880
 Gly Thr Ser Thr Glu Lys Thr Val Arg Ser Phe Tyr Tyr Ser Arg Asn
 885 890 895
 Tyr Tyr Val Lys Thr Gly Asn Lys Pro Ile Leu Pro Ser Asp Val Glu
 900 905 910
 Val Arg Tyr Ser Asp Gly Thr Ser Asp Arg Gln Asn Val Thr Trp Asp
 915 920 925
 Ala Val Ser Asp Asp Gln Ile Ala Lys Ala Gly Ser Phe Ser Val Ala
 930 935 940
 Gly Thr Val Ala Gly Gln Lys Ile Ser Val Arg Val Thr Met Ile Asp
 945 950 955 960

Glu Ile Gly Ala Leu Leu Asn Tyr Ser Ala Ser Thr Pro Val Gly Thr
965 970 975

Pro Ala Val Leu Pro Gly Ser Arg Pro Ala Val Leu Pro Asp Gly Thr
980 985 990

Val Thr Ser Ala Asn Phe Ala Val His Trp Thr Lys Pro Ala Asp Thr
995 1000 1005

Val Tyr Asn Thr Ala Gly Thr Val Lys Val Pro Gly Thr Ala Thr
1010 1015 1020

Val Phe Gly Lys Glu Phe Lys Val Thr Ala Thr Ile Arg Val Gln
1025 1030 1035

<210> 4

<211> 1142

<212> PRT

<213> 人工序列

<220>

<223> 截短的乳糖酶片段

<400> 4

Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
1 5 10 15

Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
20 25 30

Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
35 40 45

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
50 55 60

Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
65 70 75 80

Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
85 90 95

Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
100 105 110

Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
115 120 125

Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp Leu Thr Gly Asn
130 135 140

Ala Lys Phe Gly Gly Glu Asn Thr Ile Val Val Lys Val Glu Asn Arg
145 150 155 160

Leu Pro Ser Ser Arg Trp Tyr Ser Gly Ser Gly Ile Tyr Arg Asp Val
165 170 175

[0011]

	Thr	Leu	Thr	Val	Thr	Asp	Gly	Val	His	Val	Gly	Asn	Asn	Gly	Val	Ala	
				180					185					190			
	Ile	Lys	Thr	Pro	Ser	Leu	Ala	Thr	Gln	Asn	Gly	Gly	Asp	Val	Thr	Met	
			195					200					205				
	Asn	Leu	Thr	Thr	Lys	Val	Ala	Asn	Asp	Thr	Glu	Ala	Ala	Ala	Asn	Ile	
		210					215					220					
	Thr	Leu	Lys	Gln	Thr	Val	Phe	Pro	Lys	Gly	Gly	Lys	Thr	Asp	Ala	Ala	
	225					230					235					240	
	Ile	Gly	Thr	Val	Thr	Thr	Ala	Ser	Lys	Ser	Ile	Ala	Ala	Gly	Ala	Ser	
				245						250					255		
	Ala	Asp	Val	Thr	Ser	Thr	Ile	Thr	Ala	Ala	Ser	Pro	Lys	Leu	Trp	Ser	
			260						265					270			
	Ile	Lys	Asn	Pro	Asn	Leu	Tyr	Thr	Val	Arg	Thr	Glu	Val	Leu	Asn	Gly	
			275					280					285				
	Gly	Lys	Val	Leu	Asp	Thr	Tyr	Asp	Thr	Glu	Tyr	Gly	Phe	Arg	Trp	Thr	
		290					295					300					
	Gly	Phe	Asp	Ala	Thr	Ser	Gly	Phe	Ser	Leu	Asn	Gly	Glu	Lys	Val	Lys	
	305					310					315					320	
[0012]	Leu	Lys	Gly	Val	Ser	Met	His	His	Asp	Gln	Gly	Ser	Leu	Gly	Ala	Val	
					325					330					335		
	Ala	Asn	Arg	Arg	Ala	Ile	Glu	Arg	Gln	Val	Glu	Ile	Leu	Gln	Lys	Met	
				340					345					350			
	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ala	Lys	Ala	Leu	
		355						360					365				
	Ile	Asp	Val	Cys	Asn	Glu	Lys	Gly	Val	Leu	Val	Val	Glu	Glu	Val	Phe	
		370					375					380					
	Asp	Met	Trp	Asn	Arg	Ser	Lys	Asn	Gly	Asn	Thr	Glu	Asp	Tyr	Gly	Lys	
	385					390					395					400	
	Trp	Phe	Gly	Gln	Ala	Ile	Ala	Gly	Asp	Asn	Ala	Val	Leu	Gly	Gly	Asp	
				405						410					415		
	Lys	Asp	Glu	Thr	Trp	Ala	Lys	Phe	Asp	Leu	Thr	Ser	Thr	Ile	Asn	Arg	
			420						425					430			
	Asp	Arg	Asn	Ala	Pro	Ser	Val	Ile	Met	Trp	Ser	Leu	Gly	Asn	Glu	Met	
			435					440					445				
	Met	Glu	Gly	Ile	Ser	Gly	Ser	Val	Ser	Gly	Phe	Pro	Ala	Thr	Ser	Ala	
		450					455					460					
	Lys	Leu	Val	Ala	Trp	Thr	Lys	Ala	Ala	Asp	Ser	Thr	Arg	Pro	Met	Thr	
	465					470					475					480	

	Tyr	Gly	Asp	Asn	Lys	Ile	Lys	Ala	Asn	Trp	Asn	Glu	Ser	Asn	Thr	Met	
					485					490						495	
	Gly	Asp	Asn	Leu	Thr	Ala	Asn	Gly	Gly	Val	Val	Gly	Thr	Asn	Tyr	Ser	
				500					505						510		
	Asp	Gly	Ala	Asn	Tyr	Asp	Lys	Ile	Arg	Thr	Thr	His	Pro	Ser	Trp	Ala	
			515					520					525				
	Ile	Tyr	Gly	Ser	Glu	Thr	Ala	Ser	Ala	Ile	Asn	Ser	Arg	Gly	Ile	Tyr	
		530					535					540					
	Asn	Arg	Thr	Thr	Gly	Gly	Ala	Gln	Ser	Ser	Asp	Lys	Gln	Leu	Thr	Ser	
	545					550					555					560	
	Tyr	Asp	Asn	Ser	Ala	Val	Gly	Trp	Gly	Ala	Val	Ala	Ser	Ser	Ala	Trp	
					565					570					575		
	Tyr	Asp	Val	Val	Gln	Arg	Asp	Phe	Val	Ala	Gly	Thr	Tyr	Val	Trp	Thr	
				580					585					590			
	Gly	Phe	Asp	Tyr	Leu	Gly	Glu	Pro	Thr	Pro	Trp	Asn	Gly	Thr	Gly	Ser	
			595					600					605				
	Gly	Ala	Val	Gly	Ser	Trp	Pro	Ser	Pro	Lys	Asn	Ser	Tyr	Phe	Gly	Ile	
		610					615					620					
[0013]	Val	Asp	Thr	Ala	Gly	Phe	Pro	Lys	Asp	Thr	Tyr	Tyr	Phe	Tyr	Gln	Ser	
	625					630					635					640	
	Gln	Trp	Asn	Asp	Asp	Val	His	Thr	Leu	His	Ile	Leu	Pro	Ala	Trp	Asn	
				645						650					655		
	Glu	Asn	Val	Val	Ala	Lys	Gly	Ser	Gly	Asn	Asn	Val	Pro	Val	Val	Val	
				660					665					670			
	Tyr	Thr	Asp	Ala	Ala	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Pro	Lys	Gly	Ser	
			675					680					685				
	Thr	Glu	Lys	Arg	Leu	Ile	Gly	Glu	Lys	Ser	Phe	Thr	Lys	Lys	Thr	Thr	
		690					695					700					
	Ala	Ala	Gly	Tyr	Thr	Tyr	Gln	Val	Tyr	Glu	Gly	Ser	Asp	Lys	Asp	Ser	
	705					710					715					720	
	Thr	Ala	His	Lys	Asn	Met	Tyr	Leu	Thr	Trp	Asn	Val	Pro	Trp	Ala	Glu	
					725					730					735		
	Gly	Thr	Ile	Ser	Ala	Glu	Ala	Tyr	Asp	Glu	Asn	Asn	Arg	Leu	Ile	Pro	
				740					745					750			
	Glu	Gly	Ser	Thr	Glu	Gly	Asn	Ala	Ser	Val	Thr	Thr	Thr	Gly	Lys	Ala	
			755					760					765				
	Ala	Lys	Leu	Lys	Ala	Asp	Ala	Asp	Arg	Lys	Thr	Ile	Thr	Ala	Asp	Gly	
		770					775					780					

	Lys	Asp	Leu	Ser	Tyr	Ile	Glu	Val	Asp	Val	Thr	Asp	Ala	Asn	Gly	His
	785					790					795					800
	Ile	Val	Pro	Asp	Ala	Ala	Asn	Arg	Val	Thr	Phe	Asp	Val	Lys	Gly	Ala
					805					810					815	
	Gly	Lys	Leu	Val	Gly	Val	Asp	Asn	Gly	Ser	Ser	Pro	Asp	His	Asp	Ser
				820					825					830		
	Tyr	Gln	Ala	Asp	Asn	Arg	Lys	Ala	Phe	Ser	Gly	Lys	Val	Leu	Ala	Ile
			835					840					845			
	Val	Gln	Ser	Thr	Lys	Glu	Ala	Gly	Glu	Ile	Thr	Val	Thr	Ala	Lys	Ala
		850					855					860				
	Asp	Gly	Leu	Gln	Ser	Ser	Thr	Val	Lys	Ile	Ala	Thr	Thr	Ala	Val	Pro
	865					870					875					880
	Gly	Thr	Ser	Thr	Glu	Lys	Thr	Val	Arg	Ser	Phe	Tyr	Tyr	Ser	Arg	Asn
					885					890					895	
	Tyr	Tyr	Val	Lys	Thr	Gly	Asn	Lys	Pro	Ile	Leu	Pro	Ser	Asp	Val	Glu
				900					905					910		
	Val	Arg	Tyr	Ser	Asp	Gly	Thr	Ser	Asp	Arg	Gln	Asn	Val	Thr	Trp	Asp
			915					920					925			
[0014]	Ala	Val	Ser	Asp	Asp	Gln	Ile	Ala	Lys	Ala	Gly	Ser	Phe	Ser	Val	Ala
		930					935					940				
	Gly	Thr	Val	Ala	Gly	Gln	Lys	Ile	Ser	Val	Arg	Val	Thr	Met	Ile	Asp
	945					950					955					960
	Glu	Ile	Gly	Ala	Leu	Leu	Asn	Tyr	Ser	Ala	Ser	Thr	Pro	Val	Gly	Thr
					965					970					975	
	Pro	Ala	Val	Leu	Pro	Gly	Ser	Arg	Pro	Ala	Val	Leu	Pro	Asp	Gly	Thr
				980					985					990		
	Val	Thr	Ser	Ala	Asn	Phe	Ala	Val	His	Trp	Thr	Lys	Pro	Ala	Asp	Thr
			995				1000						1005			
	Val	Tyr	Asn	Thr	Ala	Gly	Thr	Val	Lys	Val	Pro	Gly	Thr	Ala	Thr	
		1010					1015					1020				
	Val	Phe	Gly	Lys	Glu	Phe	Lys	Val	Thr	Ala	Thr	Ile	Arg	Val	Gln	
		1025					1030					1035				
	Arg	Ser	Gln	Val	Thr	Ile	Gly	Ser	Ser	Val	Ser	Gly	Asn	Ala	Leu	
		1040					1045					1050				
	Arg	Leu	Thr	Gln	Asn	Ile	Pro	Ala	Asp	Lys	Gln	Ser	Asp	Thr	Leu	
		1055					1060					1065				
	Asp	Ala	Ile	Lys	Asp	Gly	Ser	Thr	Thr	Val	Asp	Ala	Asn	Thr	Gly	
		1070					1075					1080				

Gly Gly Ala Asn Pro Ser Ala Trp Thr Asn Trp Ala Tyr Ser Lys
1085 1090 1095

Ala Gly His Asn Thr Ala Glu Ile Thr Phe Glu Tyr Ala Thr Glu
1100 1105 1110

Gln Gln Leu Gly Gln Ile Val Met Tyr Phe Phe Arg Asp Ser Asn
1115 1120 1125

Ala Val Arg Phe Pro Asp Ala Gly Lys Thr Lys Ile Gln Ile
1130 1135 1140

<210> 5
<211> 1211
<212> PRT
<213> 人工序列

<220>
<223> 截短的乳糖酶片段

<400> 5

Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
1 5 10 15

Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
20 25 30

Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
35 40 45

[0015]

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
50 55 60

Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
65 70 75 80

Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
85 90 95

Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
100 105 110

Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
115 120 125

Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp Leu Thr Gly Asn
130 135 140

Ala Lys Phe Gly Gly Glu Asn Thr Ile Val Val Lys Val Glu Asn Arg
145 150 155 160

Leu Pro Ser Ser Arg Trp Tyr Ser Gly Ser Gly Ile Tyr Arg Asp Val
165 170 175

Thr Leu Thr Val Thr Asp Gly Val His Val Gly Asn Asn Gly Val Ala
180 185 190

Ile Lys Thr Pro Ser Leu Ala Thr Gln Asn Gly Gly Asp Val Thr Met
195 200 205

[0016]

Asn Leu Thr Thr Lys Val Ala Asn Asp Thr Glu Ala Ala Ala Asn Ile
 210 215 220
 Thr Leu Lys Gln Thr Val Phe Pro Lys Gly Gly Lys Thr Asp Ala Ala
 225 230 235 240
 Ile Gly Thr Val Thr Thr Ala Ser Lys Ser Ile Ala Ala Gly Ala Ser
 245 250 255
 Ala Asp Val Thr Ser Thr Ile Thr Ala Ala Ser Pro Lys Leu Trp Ser
 260 265 270
 Ile Lys Asn Pro Asn Leu Tyr Thr Val Arg Thr Glu Val Leu Asn Gly
 275 280 285
 Gly Lys Val Leu Asp Thr Tyr Asp Thr Glu Tyr Gly Phe Arg Trp Thr
 290 295 300
 Gly Phe Asp Ala Thr Ser Gly Phe Ser Leu Asn Gly Glu Lys Val Lys
 305 310 315 320
 Leu Lys Gly Val Ser Met His His Asp Gln Gly Ser Leu Gly Ala Val
 325 330 335
 Ala Asn Arg Arg Ala Ile Glu Arg Gln Val Glu Ile Leu Gln Lys Met
 340 345 350
 Gly Val Asn Ser Ile Arg Thr Thr His Asn Pro Ala Ala Lys Ala Leu
 355 360 365
 Ile Asp Val Cys Asn Glu Lys Gly Val Leu Val Val Glu Glu Val Phe
 370 375 380
 Asp Met Trp Asn Arg Ser Lys Asn Gly Asn Thr Glu Asp Tyr Gly Lys
 385 390 395 400
 Trp Phe Gly Gln Ala Ile Ala Gly Asp Asn Ala Val Leu Gly Gly Asp
 405 410 415
 Lys Asp Glu Thr Trp Ala Lys Phe Asp Leu Thr Ser Thr Ile Asn Arg
 420 425 430
 Asp Arg Asn Ala Pro Ser Val Ile Met Trp Ser Leu Gly Asn Glu Met
 435 440 445
 Met Glu Gly Ile Ser Gly Ser Val Ser Gly Phe Pro Ala Thr Ser Ala
 450 455 460
 Lys Leu Val Ala Trp Thr Lys Ala Ala Asp Ser Thr Arg Pro Met Thr
 465 470 475 480
 Tyr Gly Asp Asn Lys Ile Lys Ala Asn Trp Asn Glu Ser Asn Thr Met
 485 490 495
 Gly Asp Asn Leu Thr Ala Asn Gly Gly Val Val Gly Thr Asn Tyr Ser
 500 505 510

[0017]

Asp Gly Ala Asn Tyr Asp Lys Ile Arg Thr Thr His Pro Ser Trp Ala
 515 520 525

Ile Tyr Gly Ser Glu Thr Ala Ser Ala Ile Asn Ser Arg Gly Ile Tyr
 530 535 540

Asn Arg Thr Thr Gly Gly Ala Gln Ser Ser Asp Lys Gln Leu Thr Ser
 545 550 555 560

Tyr Asp Asn Ser Ala Val Gly Trp Gly Ala Val Ala Ser Ser Ala Trp
 565 570 575

Tyr Asp Val Val Gln Arg Asp Phe Val Ala Gly Thr Tyr Val Trp Thr
 580 585 590

Gly Phe Asp Tyr Leu Gly Glu Pro Thr Pro Trp Asn Gly Thr Gly Ser
 595 600 605

Gly Ala Val Gly Ser Trp Pro Ser Pro Lys Asn Ser Tyr Phe Gly Ile
 610 615 620

Val Asp Thr Ala Gly Phe Pro Lys Asp Thr Tyr Tyr Phe Tyr Gln Ser
 625 630 635 640

Gln Trp Asn Asp Asp Val His Thr Leu His Ile Leu Pro Ala Trp Asn
 645 650 655

Glu Asn Val Val Ala Lys Gly Ser Gly Asn Asn Val Pro Val Val Val
 660 665 670

Tyr Thr Asp Ala Ala Lys Val Lys Leu Tyr Phe Thr Pro Lys Gly Ser
 675 680 685

Thr Glu Lys Arg Leu Ile Gly Glu Lys Ser Phe Thr Lys Lys Thr Thr
 690 695 700

Ala Ala Gly Tyr Thr Tyr Gln Val Tyr Glu Gly Ser Asp Lys Asp Ser
 705 710 715 720

Thr Ala His Lys Asn Met Tyr Leu Thr Trp Asn Val Pro Trp Ala Glu
 725 730 735

Gly Thr Ile Ser Ala Glu Ala Tyr Asp Glu Asn Asn Arg Leu Ile Pro
 740 745 750

Glu Gly Ser Thr Glu Gly Asn Ala Ser Val Thr Thr Thr Gly Lys Ala
 755 760 765

Ala Lys Leu Lys Ala Asp Ala Asp Arg Lys Thr Ile Thr Ala Asp Gly
 770 775 780

Lys Asp Leu Ser Tyr Ile Glu Val Asp Val Thr Asp Ala Asn Gly His
 785 790 795 800

Ile Val Pro Asp Ala Ala Asn Arg Val Thr Phe Asp Val Lys Gly Ala
 805 810 815

	Gly	Lys	Leu	Val	Gly	Val	Asp	Asn	Gly	Ser	Ser	Pro	Asp	His	Asp	Ser	
				820					825						830		
	Tyr	Gln	Ala	Asp	Asn	Arg	Lys	Ala	Phe	Ser	Gly	Lys	Val	Leu	Ala	Ile	
			835					840					845				
	Val	Gln	Ser	Thr	Lys	Glu	Ala	Gly	Glu	Ile	Thr	Val	Thr	Ala	Lys	Ala	
		850					855					860					
	Asp	Gly	Leu	Gln	Ser	Ser	Thr	Val	Lys	Ile	Ala	Thr	Thr	Ala	Val	Pro	
	865					870					875					880	
	Gly	Thr	Ser	Thr	Glu	Lys	Thr	Val	Arg	Ser	Phe	Tyr	Tyr	Ser	Arg	Asn	
				885						890					895		
	Tyr	Tyr	Val	Lys	Thr	Gly	Asn	Lys	Pro	Ile	Leu	Pro	Ser	Asp	Val	Glu	
			900						905					910			
	Val	Arg	Tyr	Ser	Asp	Gly	Thr	Ser	Asp	Arg	Gln	Asn	Val	Thr	Trp	Asp	
			915					920					925				
	Ala	Val	Ser	Asp	Asp	Gln	Ile	Ala	Lys	Ala	Gly	Ser	Phe	Ser	Val	Ala	
		930				935						940					
	Gly	Thr	Val	Ala	Gly	Gln	Lys	Ile	Ser	Val	Arg	Val	Thr	Met	Ile	Asp	
	945					950					955				960		
[0018]	Glu	Ile	Gly	Ala	Leu	Leu	Asn	Tyr	Ser	Ala	Ser	Thr	Pro	Val	Gly	Thr	
				965						970					975		
	Pro	Ala	Val	Leu	Pro	Gly	Ser	Arg	Pro	Ala	Val	Leu	Pro	Asp	Gly	Thr	
				980					985					990			
	Val	Thr	Ser	Ala	Asn	Phe	Ala	Val	His	Trp	Thr	Lys	Pro	Ala	Asp	Thr	
			995				1000						1005				
	Val	Tyr	Asn	Thr	Ala	Gly	Thr	Val	Lys	Val	Pro	Gly	Thr	Ala	Thr		
		1010					1015					1020					
	Val	Phe	Gly	Lys	Glu	Phe	Lys	Val	Thr	Ala	Thr	Ile	Arg	Val	Gln		
		1025					1030					1035					
	Arg	Ser	Gln	Val	Thr	Ile	Gly	Ser	Ser	Val	Ser	Gly	Asn	Ala	Leu		
		1040					1045					1050					
	Arg	Leu	Thr	Gln	Asn	Ile	Pro	Ala	Asp	Lys	Gln	Ser	Asp	Thr	Leu		
		1055					1060					1065					
	Asp	Ala	Ile	Lys	Asp	Gly	Ser	Thr	Thr	Val	Asp	Ala	Asn	Thr	Gly		
		1070					1075					1080					
	Gly	Gly	Ala	Asn	Pro	Ser	Ala	Trp	Thr	Asn	Trp	Ala	Tyr	Ser	Lys		
		1085					1090					1095					
	Ala	Gly	His	Asn	Thr	Ala	Glu	Ile	Thr	Phe	Glu	Tyr	Ala	Thr	Glu		
		1100					1105					1110					

Gln Gln Leu Gly Gln Ile Val Met Tyr Phe Phe Arg Asp Ser Asn
1115 1120 1125

Ala Val Arg Phe Pro Asp Ala Gly Lys Thr Lys Ile Gln Ile Ser
1130 1135 1140

Ala Asp Gly Lys Asn Trp Thr Asp Leu Ala Ala Thr Glu Thr Ile
1145 1150 1155

Ala Ala Gln Glu Ser Ser Asp Arg Val Lys Pro Tyr Thr Tyr Asp
1160 1165 1170

Phe Ala Pro Val Gly Ala Thr Phe Val Lys Val Thr Val Thr Asn
1175 1180 1185

Ala Asp Thr Thr Thr Pro Ser Gly Val Val Cys Ala Gly Leu Thr
1190 1195 1200

Glu Ile Glu Leu Lys Thr Ala Thr
1205 1210

<210> 6
<211> 1296
<212> PRT
<213> 人工序列

<220>
<223> 截短的乳糖酶片段

[0019]

<400> 6

Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
1 5 10 15

Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
20 25 30

Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
35 40 45

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
50 55 60

Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
65 70 75 80

Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
85 90 95

Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
100 105 110

Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
115 120 125

Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp Leu Thr Gly Asn
130 135 140

	Ala	Lys	Phe	Gly	Gly	Glu	Asn	Thr	Ile	Val	Val	Lys	Val	Glu	Asn	Arg	
	145					150					155					160	
	Leu	Pro	Ser	Ser	Arg	Trp	Tyr	Ser	Gly	Ser	Gly	Ile	Tyr	Arg	Asp	Val	
					165					170					175		
	Thr	Leu	Thr	Val	Thr	Asp	Gly	Val	His	Val	Gly	Asn	Asn	Gly	Val	Ala	
				180					185					190			
	Ile	Lys	Thr	Pro	Ser	Leu	Ala	Thr	Gln	Asn	Gly	Gly	Asp	Val	Thr	Met	
			195					200					205				
	Asn	Leu	Thr	Thr	Lys	Val	Ala	Asn	Asp	Thr	Glu	Ala	Ala	Ala	Asn	Ile	
		210					215					220					
	Thr	Leu	Lys	Gln	Thr	Val	Phe	Pro	Lys	Gly	Gly	Lys	Thr	Asp	Ala	Ala	
	225					230					235					240	
	Ile	Gly	Thr	Val	Thr	Thr	Ala	Ser	Lys	Ser	Ile	Ala	Ala	Gly	Ala	Ser	
				245						250					255		
	Ala	Asp	Val	Thr	Ser	Thr	Ile	Thr	Ala	Ala	Ser	Pro	Lys	Leu	Trp	Ser	
				260					265					270			
	Ile	Lys	Asn	Pro	Asn	Leu	Tyr	Thr	Val	Arg	Thr	Glu	Val	Leu	Asn	Gly	
			275					280					285				
[0020]	Gly	Lys	Val	Leu	Asp	Thr	Tyr	Asp	Thr	Glu	Tyr	Gly	Phe	Arg	Trp	Thr	
		290					295					300					
	Gly	Phe	Asp	Ala	Thr	Ser	Gly	Phe	Ser	Leu	Asn	Gly	Glu	Lys	Val	Lys	
	305					310					315					320	
	Leu	Lys	Gly	Val	Ser	Met	His	His	Asp	Gln	Gly	Ser	Leu	Gly	Ala	Val	
					325					330					335		
	Ala	Asn	Arg	Arg	Ala	Ile	Glu	Arg	Gln	Val	Glu	Ile	Leu	Gln	Lys	Met	
				340					345					350			
	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ala	Lys	Ala	Leu	
		355					360						365				
	Ile	Asp	Val	Cys	Asn	Glu	Lys	Gly	Val	Leu	Val	Val	Glu	Glu	Val	Phe	
		370					375					380					
	Asp	Met	Trp	Asn	Arg	Ser	Lys	Asn	Gly	Asn	Thr	Glu	Asp	Tyr	Gly	Lys	
	385					390					395					400	
	Trp	Phe	Gly	Gln	Ala	Ile	Ala	Gly	Asp	Asn	Ala	Val	Leu	Gly	Gly	Asp	
				405						410					415		
	Lys	Asp	Glu	Thr	Trp	Ala	Lys	Phe	Asp	Leu	Thr	Ser	Thr	Ile	Asn	Arg	
				420					425					430			
	Asp	Arg	Asn	Ala	Pro	Ser	Val	Ile	Met	Trp	Ser	Leu	Gly	Asn	Glu	Met	
		435						440					445				

	Met	Glu	Gly	Ile	Ser	Gly	Ser	Val	Ser	Gly	Phe	Pro	Ala	Thr	Ser	Ala	
	450						455					460					
	Lys	Leu	Val	Ala	Trp	Thr	Lys	Ala	Ala	Asp	Ser	Thr	Arg	Pro	Met	Thr	
	465					470					475					480	
	Tyr	Gly	Asp	Asn	Lys	Ile	Lys	Ala	Asn	Trp	Asn	Glu	Ser	Asn	Thr	Met	
					485					490					495		
	Gly	Asp	Asn	Leu	Thr	Ala	Asn	Gly	Gly	Val	Val	Gly	Thr	Asn	Tyr	Ser	
				500					505					510			
	Asp	Gly	Ala	Asn	Tyr	Asp	Lys	Ile	Arg	Thr	Thr	His	Pro	Ser	Trp	Ala	
			515					520					525				
	Ile	Tyr	Gly	Ser	Glu	Thr	Ala	Ser	Ala	Ile	Asn	Ser	Arg	Gly	Ile	Tyr	
	530						535					540					
	Asn	Arg	Thr	Thr	Gly	Gly	Ala	Gln	Ser	Ser	Asp	Lys	Gln	Leu	Thr	Ser	
	545					550					555					560	
	Tyr	Asp	Asn	Ser	Ala	Val	Gly	Trp	Gly	Ala	Val	Ala	Ser	Ser	Ala	Trp	
					565					570					575		
	Tyr	Asp	Val	Val	Gln	Arg	Asp	Phe	Val	Ala	Gly	Thr	Tyr	Val	Trp	Thr	
				580					585					590			
[0021]	Gly	Phe	Asp	Tyr	Leu	Gly	Glu	Pro	Thr	Pro	Trp	Asn	Gly	Thr	Gly	Ser	
			595					600					605				
	Gly	Ala	Val	Gly	Ser	Trp	Pro	Ser	Pro	Lys	Asn	Ser	Tyr	Phe	Gly	Ile	
		610					615					620					
	Val	Asp	Thr	Ala	Gly	Phe	Pro	Lys	Asp	Thr	Tyr	Tyr	Phe	Tyr	Gln	Ser	
	625					630					635					640	
	Gln	Trp	Asn	Asp	Asp	Val	His	Thr	Leu	His	Ile	Leu	Pro	Ala	Trp	Asn	
					645					650					655		
	Glu	Asn	Val	Val	Ala	Lys	Gly	Ser	Gly	Asn	Asn	Val	Pro	Val	Val	Val	
				660					665				670				
	Tyr	Thr	Asp	Ala	Ala	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Pro	Lys	Gly	Ser	
			675					680					685				
	Thr	Glu	Lys	Arg	Leu	Ile	Gly	Glu	Lys	Ser	Phe	Thr	Lys	Lys	Thr	Thr	
		690					695					700					
	Ala	Ala	Gly	Tyr	Thr	Tyr	Gln	Val	Tyr	Glu	Gly	Ser	Asp	Lys	Asp	Ser	
	705					710					715					720	
	Thr	Ala	His	Lys	Asn	Met	Tyr	Leu	Thr	Trp	Asn	Val	Pro	Trp	Ala	Glu	
					725					730					735		
	Gly	Thr	Ile	Ser	Ala	Glu	Ala	Tyr	Asp	Glu	Asn	Asn	Arg	Leu	Ile	Pro	
				740					745					750			

	Glu	Gly	Ser	Thr	Glu	Gly	Asn	Ala	Ser	Val	Thr	Thr	Thr	Gly	Lys	Ala	
			755					760						765			
	Ala	Lys	Leu	Lys	Ala	Asp	Ala	Asp	Arg	Lys	Thr	Ile	Thr	Ala	Asp	Gly	
		770					775					780					
	Lys	Asp	Leu	Ser	Tyr	Ile	Glu	Val	Asp	Val	Thr	Asp	Ala	Asn	Gly	His	
	785					790					795					800	
	Ile	Val	Pro	Asp	Ala	Ala	Asn	Arg	Val	Thr	Phe	Asp	Val	Lys	Gly	Ala	
					805					810					815		
	Gly	Lys	Leu	Val	Gly	Val	Asp	Asn	Gly	Ser	Ser	Pro	Asp	His	Asp	Ser	
				820					825					830			
	Tyr	Gln	Ala	Asp	Asn	Arg	Lys	Ala	Phe	Ser	Gly	Lys	Val	Leu	Ala	Ile	
		835						840					845				
	Val	Gln	Ser	Thr	Lys	Glu	Ala	Gly	Glu	Ile	Thr	Val	Thr	Ala	Lys	Ala	
		850					855					860					
	Asp	Gly	Leu	Gln	Ser	Ser	Thr	Val	Lys	Ile	Ala	Thr	Thr	Ala	Val	Pro	
	865					870					875					880	
	Gly	Thr	Ser	Thr	Glu	Lys	Thr	Val	Arg	Ser	Phe	Tyr	Tyr	Ser	Arg	Asn	
					885					890					895		
[0022]	Tyr	Tyr	Val	Lys	Thr	Gly	Asn	Lys	Pro	Ile	Leu	Pro	Ser	Asp	Val	Glu	
				900					905					910			
	Val	Arg	Tyr	Ser	Asp	Gly	Thr	Ser	Asp	Arg	Gln	Asn	Val	Thr	Trp	Asp	
			915					920					925				
	Ala	Val	Ser	Asp	Asp	Gln	Ile	Ala	Lys	Ala	Gly	Ser	Phe	Ser	Val	Ala	
		930					935					940					
	Gly	Thr	Val	Ala	Gly	Gln	Lys	Ile	Ser	Val	Arg	Val	Thr	Met	Ile	Asp	
	945					950					955				960		
	Glu	Ile	Gly	Ala	Leu	Leu	Asn	Tyr	Ser	Ala	Ser	Thr	Pro	Val	Gly	Thr	
					965					970					975		
	Pro	Ala	Val	Leu	Pro	Gly	Ser	Arg	Pro	Ala	Val	Leu	Pro	Asp	Gly	Thr	
				980					985					990			
	Val	Thr	Ser	Ala	Asn	Phe	Ala	Val	His	Trp	Thr	Lys	Pro	Ala	Asp	Thr	
			995				1000						1005				
	Val	Tyr	Asn	Thr	Ala	Gly	Thr	Val	Lys	Val	Pro	Gly	Thr	Ala	Thr		
		1010					1015					1020					
	Val	Phe	Gly	Lys	Glu	Phe	Lys	Val	Thr	Ala	Thr	Ile	Arg	Val	Gln		
		1025					1030					1035					
	Arg	Ser	Gln	Val	Thr	Ile	Gly	Ser	Ser	Val	Ser	Gly	Asn	Ala	Leu		
		1040					1045					1050					

[0023]

Arg Leu Thr Gln Asn Ile Pro Ala Asp Lys Gln Ser Asp Thr Leu
 1055 1060 1065
 Asp Ala Ile Lys Asp Gly Ser Thr Thr Val Asp Ala Asn Thr Gly
 1070 1075 1080
 Gly Gly Ala Asn Pro Ser Ala Trp Thr Asn Trp Ala Tyr Ser Lys
 1085 1090 1095
 Ala Gly His Asn Thr Ala Glu Ile Thr Phe Glu Tyr Ala Thr Glu
 1100 1105 1110
 Gln Gln Leu Gly Gln Ile Val Met Tyr Phe Phe Arg Asp Ser Asn
 1115 1120 1125
 Ala Val Arg Phe Pro Asp Ala Gly Lys Thr Lys Ile Gln Ile Ser
 1130 1135 1140
 Ala Asp Gly Lys Asn Trp Thr Asp Leu Ala Ala Thr Glu Thr Ile
 1145 1150 1155
 Ala Ala Gln Glu Ser Ser Asp Arg Val Lys Pro Tyr Thr Tyr Asp
 1160 1165 1170
 Phe Ala Pro Val Gly Ala Thr Phe Val Lys Val Thr Val Thr Asn
 1175 1180 1185
 Ala Asp Thr Thr Thr Pro Ser Gly Val Val Cys Ala Gly Leu Thr
 1190 1195 1200
 Glu Ile Glu Leu Lys Thr Ala Thr Ser Lys Phe Val Thr Asn Thr
 1205 1210 1215
 Ser Ala Ala Leu Ser Ser Leu Thr Val Asn Gly Thr Lys Val Ser
 1220 1225 1230
 Asp Ser Val Leu Ala Ala Gly Ser Tyr Asn Thr Pro Ala Ile Ile
 1235 1240 1245
 Ala Asp Val Lys Ala Glu Gly Glu Gly Asn Ala Ser Val Thr Val
 1250 1255 1260
 Leu Pro Ala His Asp Asn Val Ile Arg Val Ile Thr Glu Ser Glu
 1265 1270 1275
 Asp His Val Thr Arg Lys Thr Phe Thr Ile Asn Leu Gly Thr Glu
 1280 1285 1290
 Gln Glu Phe
 1295

<210> 7
 <211> 453
 <212> PRT
 <213> 两岐双歧杆菌

<400> 7

Gln Asn Arg Thr Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser

[0024]

1	5	10	15
Asp Ser Val Gln Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln	20	25	30
Gln Val Asp Leu Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln	35	40	45
Ser Asn Glu Ala Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr	50	55	60
Arg Lys Ser Phe Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala	65	70	75
Ile Asn Phe Asp Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly	85	90	95
Val Lys Leu Gly Thr His Pro Tyr Gly Tyr Ser Pro Phe Ser Phe Asp	100	105	110
Leu Thr Gly Asn Ala Lys Phe Gly Gly Glu Asn Thr Ile Val Val Lys	115	120	125
Val Glu Asn Arg Leu Pro Ser Ser Arg Trp Tyr Ser Gly Ser Gly Ile	130	135	140
Tyr Arg Asp Val Thr Leu Thr Val Thr Asp Gly Val His Val Gly Asn	145	150	155
Asn Gly Val Ala Ile Lys Thr Pro Ser Leu Ala Thr Gln Asn Gly Gly	165	170	175
Asp Val Thr Met Asn Leu Thr Thr Lys Val Ala Asn Asp Thr Glu Ala	180	185	190
Ala Ala Asn Ile Thr Leu Lys Gln Thr Val Phe Pro Lys Gly Gly Lys	195	200	205
Thr Asp Ala Ala Ile Gly Thr Val Thr Thr Ala Ser Lys Ser Ile Ala	210	215	220
Ala Gly Ala Ser Ala Asp Val Thr Ser Thr Ile Thr Ala Ala Ser Pro	225	230	235
Lys Leu Trp Ser Ile Lys Asn Pro Asn Leu Tyr Thr Val Arg Thr Glu	245	250	255
Val Leu Asn Gly Gly Lys Val Leu Asp Thr Tyr Asp Thr Glu Tyr Gly	260	265	270
Phe Arg Trp Thr Gly Phe Asp Ala Thr Ser Gly Phe Ser Leu Asn Gly	275	280	285
Glu Lys Val Lys Leu Lys Gly Val Ser Met His His Asp Gln Gly Ser	290	295	300
Leu Gly Ala Val Ala Asn Arg Arg Ala Ile Glu Arg Gln Val Glu Ile			

305	310	315	320
Leu Gln Lys Met Gly Val Asn Ser Ile Arg Thr Thr His Asn Pro Ala			
	325	330	335
Ala Lys Ala Leu Ile Asp Val Cys Asn Glu Lys Gly Val Leu Val Val			
	340	345	350
Glu Glu Val Phe Asp Met Trp Asn Arg Ser Lys Asn Gly Asn Thr Glu			
	355	360	365
Asp Tyr Gly Lys Trp Phe Gly Gln Ala Ile Ala Gly Asp Asn Ala Val			
	370	375	380
Leu Gly Gly Asp Lys Asp Glu Thr Trp Ala Lys Phe Asp Leu Thr Ser			
	385	390	395
Thr Ile Asn Arg Asp Arg Asn Ala Pro Ser Val Ile Met Trp Ser Leu			
	405	410	415
Gly Asn Glu Met Met Glu Gly Ile Ser Gly Ser Val Ser Gly Phe Pro			
	420	425	430
Ala Thr Ser Ala Lys Leu Val Ala Trp Thr Lys Ala Ala Asp Ser Thr			
	435	440	445
Arg Pro Met Thr Tyr			
	450		

[0025]

<210> 8
 <211> 5160
 <212> DNA
 <213> 人工序列

 <220>
 <223> 编码细胞外乳糖酶的核苷酸序列

 <400> 8
 gcagttgaag atgcaacaag aagcgatagc acaacacaaa tgtcatcaac accggaagtt 60
 gtttattcat cagcggctga tagcaaaaa aatcgacaaa gcgattttga tgcaactgg 120
 aaatttatgc tgcagatag cgttcaagca caagatccgg catttgatga ttcagcatgg 180
 caacaagttg atctgccgca tgattatagc atcacacaga aatatagcca aagcaatgaa 240
 gcagaatcag catatcttcc gggaggcaca ggctggtata gaaaaagctt tacaattgat 300
 agagatctgg caggcaaacg cattgcgatt aattttgatg gcgtctatat gaatgaaca 360
 gtctggttta atggcgtaa actgggcaca catccgatg gctattcacc gttttcattt 420
 gatctgacag gcaatgcaaa atttggcgga gaaaacacaa ttgtcgtcaa agttgaaaat 480
 agactgccgt catcaagatg gtattcaggc agcggcattt atagagatgt tacactgaca 540
 gttacagatg gcgttcattg ttgcaataat ggcgtcgcaa ttaaaacacc gtcactggca 600
 acacaaaatg gcggagatgt cacaatgaac ctgacaacaa aagtcgcaaa tgatacagaa 660
 gcagcagcga acattacact gaaacagaca gtttttccga aaggcgcaaa aacggatgca 720
 gcaattggca cagttacaac agcatcaaaa tcaattgcag caggcgcatc agcagatgtt 780
 acaagcacia ttacagcagc aagccccaaa ctgtgggtcaa ttaaaaaccc gaacctgtat 840

[0026]

```

acagttagaa cagaagtctt gaacggaggg aaagttctgg atacatatga tacagaatat 900
ggcttttcgt ggacaggctt tgatgcaaca tcaggctttt cactgaatgg cgaaaaagtc 960
aaactgaaag gcgttagcat gcatcatgat caaggctcac ttggcgagct tgcaaataga 1020
cgcgcaattg aaagacaagt cgaatcctg caaaaaatgg gcgtcaatag cattcgaca 1080
acacataatc cggcagcaaa agcactgatt gatgtctgca atgaaaaagg cgttctggtt 1140
gtcgaagaag tctttgatat gtggaaccgc agcaaaaatg gcaacacgga agattatggc 1200
aaatggtttg gccaaagcaat tgcaggcgat aatgcagttc tgggaggcga taaagatgaa 1260
acatggcgga aatttgatct tacatcaaca attaacccgc atagaaatgc accgtcagtt 1320
attatgtggt cactgggcaa tgaatgatg gaaggcattt caggctcagt ttcaggcttt 1380
ccggcaacat cagcaaaact ggttgcattg acaaaagcag cagattcaac aagaccgatg 1440
acatatggcg ataacaaaat taaagcaac tggaaacgaat caaatacaat gggcgataat 1500
ctgacagcaa atggcggagt tgttggcaca aattattcag atggcgcaaa ctatgataaa 1560
attcgtacaa cacatccgct atgggcaatt tatggctcag aaacagcatc agcgattaat 1620
agccgtggca ttataatag aacaacaggc ggagcacaat catcagataa acagctgaca 1680
agctatgata attcagcagt tggtgggga gcagttgcat catcagcatg gtatgatgtt 1740
gttcagagag attttgcgc aggcacatat gtttgacag gatttgatta tctggcgcaa 1800
ccgacaccgt ggaatggcac aggcacaggc gcagttggct catggccgct accgaaaaat 1860
agctattttg gcatcgttga tacagcaggc ttccgaaag atacatatta tttttatcag 1920
agccagtgga atgatgatgt tcataactg catattcttc cgcatggaa tgaaaatgtt 1980
gttgcaaaag gtcaggcaa taatgttccg gttgtcgtt atacagatgc agcgaaagtg 2040
aaactgtatt ttacaccgaa aggcctcaaca gaaaaagac tgatcgcgca aaaatcattt 2100
acaaaaaaaa caacagcggc aggcctataca tatcaagtct atgaaggcag cgataaagat 2160
tcaacagcgc ataaaaacat gtatctgaca tggaatgttc cgtgggcaga aggcacaatt 2220
tcagcggaag cgtatgatga aaataatcgc ctgattccgg aaggcagcac agaaggcaac 2280
gcatcagtta caacaacagg caaagcagca aaactgaaag cagatgcgga tcgcaaaaca 2340
attacagcgg atggcaaga tctgtcatat attgaagtcg atgtcacaga tgcaaatggc 2400
catattgttc cggatgcagc aaatagagtc acatttgatg ttaaaggcgc aggcacaaactg 2460
gttggcgttg ataattggctc atcaccggat catgattcat atcaagcgga taaccgcaaa 2520
gcattttcag gcaaagtcct ggcaattgtt cagtcaaca aagaagcagg cgaaattaca 2580
gttacagcaa aagcagatgg cctgcaatca agcacagtta aaattgcaac aacagcagtt 2640
ccgggaacaa gcacagaaaa aacagtccgc agcttttatt acagccgcaa ctattatgct 2700
aaaaacagga acaaacccgat tctgccgtca gatgttgaag ttcgctatc agatggaaca 2760
agcgatagac aaaacgttac atgggatgca gtttcagatg atcaaattgc aaaagcaggc 2820
tcattttcag ttgcaggcac agttgcaggc caaaaaatta gcgttcgcgt cacaatgatt 2880
gatgaaattg gcgcactgct gaattattca gcaagcacac cggttggcac accggcagtt 2940
cttcgggat caagaccggc agtcctgccg gatggcacag tcacatcagc aaattttgca 3000
gtccattgga caaacccggc agatacagtc tataatacag caggcacagt caaagtaccg 3060
ggaacagcaa cagtttttgg caaagaattt aaagtcacag cgacaattag agttcaaaga 3120

```

	agccaagtta caattggctc atcagtttca ggaaatgcac tgagactgac acaaaatatt	3180
	ccggcagata aacaatcaga tacactggat gcgattaaag atggctcaac aacagttgat	3240
	gcaaatacag gcggaggcgc aaatccgtca gcatggacaa attgggcata ttcaaaagca	3300
	ggccataaca cagcggaaat tacatttgaa tatgcgacag aacaacaact gggccagatc	3360
	gtcatgtatt tttttcgcga tagcaatgca gttagatttc cggatgctgg caaaacaaaa	3420
	attcagatca gcgcagatgg caaaaattgg acagatctgg cagcaacaga aacaattgca	3480
	gcgcaagaat caagcgatag agtcaaacgg tatacatatg attttgcacc ggttggcgca	3540
	acattttgta aagtgcagct cacaaacgca gatacaaca caccgtcagg cgttgtttgc	3600
	gcaggcctga cagaaattga actgaaaaca gcgacaagca aatttgtcac aaatacatca	3660
	gcagcactgt catcacttac agtcaatggc aaaaaagttt cagattcagt tctggcagca	3720
	ggctcatata acacaccggc aattatcgca gatgttaaag cggaaggcga aggcaatgca	3780
	agcgttacag tccttcggc acatgataat gttattcgcg tcattacaga aagcgaagat	3840
	catgtcacac gaaaacatt tacaatcaac ctgggcacag aacaagaatt tccggctgat	3900
	tcagatgaaa gagattatcc ggcagcagat atgacagtca cagttggctc agaacaaca	3960
	tcaggcacag caacagaagg accgaaaaaa ttgcagtcg atggcaacac atcaacatat	4020
	tggcatagca attggacacc gacaacagtt aatgatctgt ggatcgcgtt tgaactgcaa	4080
	aaaccgacaa aactggatgc actgagatat ctccgcgtc cggcaggctc aaaaaatggc	4140
	agcgtcacag aatataaagt tcagggtgca gatgatggaa caaactggac agatgcaggc	4200
	tcaggccat ggacaacgga ttatggctgg aaactggcgg aatttaatca accggtcaca	4260
[0027]	acaaaacatg ttagactgaa agcggttcat acatatgcag atagcggcaa cgataaattt	4320
	atgagcgcaa gcgaaattag actgagaaaa gcggtcgata caacggatat ttcaggcgca	4380
	acagttacag ttccggcaaa actgacagtt gatagagttg atgcagatca tccggcaaca	4440
	tttgcaaca aagatgtcac agttacactg ggagatgcaa cactgagata tggcgttgat	4500
	tatctgtggt attatgcagg caatacagca gttggcaaag caacagtgc agttagaggc	4560
	attgataaat attcaggcac agtcgcgaaa acatttaca ttgaactgaa aaatgcaccg	4620
	gcaccggaac cgacactgac atcagttagc gtcaaaaca aaccgagcaa actgacatat	4680
	gttgtcggag atgcatttga tccggcaggc ctggttctgc aacatgatag acaagcagat	4740
	agacctccgc aaccgctggt tggcgaacaa gcggtgaac gcggactgac atgcggcaca	4800
	agatgcgata gagttgaaca actgcgcaaa catgaaaata gagaagcgca tagaacaggc	4860
	ctggatcatc tggaatttgt tggcgcagca gatggcgagc ttggagaaca agcaacattt	4920
	aaagtccatg tccatgcaga tcaggagat ggcagacatg atgatgcaga tgaacgcgat	4980
	attgatccgc atgttccggt cgatcatgca gttggcgaac tggcaagagc agcatgcat	5040
	catgttattg gcttagagat cgatacacat agacttaaag caagcggtt tcaaattccg	5100
	gctgatgata tggcagaaat cgatcgcat acaggctttc atcgttttga acgcatgtc	5160
<210>	9	
<211>	2661	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	编码BIF_917的核苷酸序列	

[0028]	<400> 9	
	gttgaagatg caacaagaag cgatagcaca acacaaatgt catcaacacc ggaagtgtgt	60
	tattcatcag cggtcgatag caaacaataat cgcacaagcg attttgatgc gaactggaaa	120
	tttatgctgt cagatagcgt tcaagcaca gatccggcat ttgatgattc agcatggcaa	180
	caagttgatc tgccgcatga ttatagcatc acacagaaat atagccaaag caatgaagca	240
	gaatcagcat atcttccggg aggcacagcg tggatataga aaagctttac aattgataga	300
	gatctggcag gcaaacgcat tgcgattaat ttgatggcg tctatatgaa tgcaacagtc	360
	tggtttaagt gcgttaaact gggcacacat ccgtatggct attcaccgtt ttcatttgat	420
	ctgacagcca atgcaaaatt tggcggagaa aacacaattg tcgtcaaagt tgaataataga	480
	ctgccgcat caagatggta ttcaggcagc ggcatattata gagatgttac actgacagtt	540
	acagatggcg ttcatgttgg caataatggc gtcgcaatta aaacaccgtc actggcaaca	600
	caaaatggcg gagatgtcac aatgaacctg acaacaaaag tcgcgaatga tacagaagca	660
	gcagcgaaca ttacactgaa acagacagtt ttccgaaaag gcggaanaac ggatgcagca	720
	attggcacag ttacaacagc atcaaaatca attgcagcag gcgcatcagc agatgttaca	780
	agcacaatta cagcagcaag cccgaaactg tggtaatta aaaaccgaa cctgtataca	840
	gttagaacag aagtctgaa cggaggcaaa gttctggata catatgatac agaataatggc	900
	tttcgctgga caggctttga tgcaacatca ggcttttcac tgaatggcga aaaagtcaaa	960
	ctgaaaggcg ttagcatgca tcatgatcaa ggctcacttg gcgcagttgc aaatagacgc	1020
	gcaattgaaa gacaagtcca aatcctgcaa aaaatgggcg tcaatagcat tcgcacaaca	1080
	cataatccgg cagcaaaagc actgattgat gtctgcaatg aaaaaggcgt tctggttgtc	1140
	gaagaagtct ttgatattgt gaaccgcagc aaaaatggca acacggaaga ttatggcaaa	1200
	tggtttgcc aagcaattgc aggcgataat gcagttctgg gaggcgataa agatgaacaa	1260
	tggcgcaaat ttgatcttac atcaacaatt aaccgcgata gaaatgcacc gtcagttatt	1320
	atgtggtcac tgggcaatga aatgatggaa ggcatctcag gctcagtttc aggccttccg	1380
	gcaacatcag caaaactggt tgcatggaca aaagcagcag attcaacaag accgatgaca	1440
	tatggcgata acaaaattaa agcgaactgg aacgaatcaa atacaatggg cgataatctg	1500
	acagcaaatg gcggagtgtg tggcacaat tattcagatg gcgcaacta tgataaaatt	1560
	cgtacaacac atccgtcatg ggcaatttat ggctcagaaa cagcatcagc gattaatagc	1620
	cgtggcattt ataatagaac aacaggcgga gcacaatcat cagataaaca gctgacaagc	1680
	tatgataatt cagcagttgg ctggggagca gttgcatcat cagcatggta tgatgttgtt	1740
	cagagagatt ttgtcgcagg cacatatgtt tggacaggat ttgattatct gggcgaaccg	1800
	acaccgtgga atggcacagg ctcaggcgca gttggctcat ggccgtcacc gaaaaatagc	1860
	tattttggca tcgttgatac agcaggcttt ccgaaagata catattattt ttatcagagc	1920
	cagtggaaat atgatgttca tacactgcat attcttccgg catggaatga aatgttgtt	1980
	gcaaaaggct caggcaataa tgttccggtt gtcgtttata cagatgcagc gaaagtga	2040
	ctgtatttta caccgaaagg ctcaacagaa aaaagactga tcggcgaaaa atcatttaca	2100
	aaaaaaacaa cagcggcagg ctatacatat caagtctatg aaggcagcga taaagattca	2160
	acagcgcata aaaacatgta tctgacatgg aatgttccgt gggcagaagg cacaatttca	2220

gcggaagcgt atgatgaaa taatcgctg attccggaag gcagcacaga aggcaacgca 2280
 tcagttacaa caacaggcaa agcagcaaaa ctgaaagcag atgcggatcg caaacaatt 2340
 acagcggatg gcaaagatct gtcatatatt gaagtcgatg tcacagatgc aaatggccat 2400
 attgttccgg atgcagcaaa tagagtcaca ttgatgtta aaggcgcagg caaactgggt 2460
 ggctgtgata atggctcatc accgatcat gattcatatc aagcggataa ccgcaaagca 2520
 ttttcaggca aagtcttggc aattgttcag tcaacaaaag aagcaggcga aattacagtt 2580
 acagcaaaaag catagtcctt gcaatcaagc acagttaaaa ttgcaacaac agcagttccg 2640
 ggaacaagca cagaaaaaac a 2661

<210> 10
 <211> 2895
 <212> DNA
 <213> 人工序列

<220>
 <223> 编码BIF_995的核苷酸序列

<400> 10
 gttgaagatg caacaagaag cgatagcaca acacaaatgt catcaacacc ggaagttgtt 60
 tattcatcag cggtcgatag caaacaaaat cgcacaagcg attttgatgc gaactggaaa 120
 tttatgtctg catagatcgt tcaagcaca gatccggcat ttgatgattc agcatggcaa 180
 caagttgatc tgccgatga ttatagcatc acacagaaat atagccaaag caatgaagca 240
 gaatcagcat atcttccggg aggcacagcg tggatataga aaagctttac aattgataga 300
 gatctggcag gcaaacgcat tgcgattaat ttgatggcg tctatatgaa tgcaacagtc 360
 tggtttaaat gcgttaaact gggcacacat ccgtatggct attcaccgtt ttcatttgat 420
 ctgacaggca atgcaaaatt tggcggagaa aacacaattg tcgtcaaaat tgaaaataga 480
 ctgccgtcat caagatggta ttcaggcagc ggcatattata gagatgttac actgacagtt 540
 acagatggcg ttcatgttgg caataatggc gtcgcaatta aaacaccgtc actggcaaca 600
 caaaatggcg gagatgtcac aatgaacctg acaacaaaag tcgcgaatga tacagaagca 660
 gcagcgaaca ttacttgaa acagacagtt tttccgaaag gcggaanaac ggatgcagca 720
 attggcacag ttacaacagc atcaaaatca attgcagcag gcgcatcagc agatgttaca 780
 agcacaatta cagcagcaag cccgaaactg tggtaatta aaaacccgaa cctgtatata 840
 gttagaacag aagtcttgaa cggaggcaaa gttctggata catatgatac agaatatggc 900
 tticgctgga caggcttga tgcaacatca ggcttttcac tgaatggcga aaaagtcaaa 960
 ctgaaaggcg ttagcatgca tcatgatcaa ggctcacttg gcgcagttgc aaatagacgc 1020
 gcaattgaaa gacaagtga aatcctgcaa aaaatggcg tcaatagcat tcgcacaaca 1080
 cataatccgg cagcaaaagc actgattgat gtctgcaatg aaaaaggcgt tctggttgtc 1140
 gaagaagtct ttgatattg gaaccgcagc aaaaatggca acacggaaga ttatggcaaa 1200
 tggtttggcc aagcaattgc aggcgataat gcagttctgg gaggcgataa agatgaaaca 1260
 tgggcgaaat ttgatcttac atcaacaatt aaccgcgata gaaatgcacc gtcagttatt 1320
 atgttggtcac tgggcaatga aatgatggaa ggcatttcag gtcagtttc aggccttccg 1380
 gcaacatcag caaaactggg tgcatggaca aaagcagcag attcaacaag accgatgaca 1440
 tatggcgata acaaaattaa agcgaactgg aacgaatcaa atacaatggg cgataatctg 1500

[0029]

[0030]

```

acagcaaatg gcggagtgtg tggcacaatg tttcagatg gcgcaacta tgataaaatt 1560
cgtacaacac atccgtcatg ggcaatttat ggctcagaaa cagcatcagc gattaatagc 1620
cgtggcattt ataatagaac aacaggcgga gcacaatcat cagataaaca gctgacaagc 1680
tatgataatt cagcagttgg ctggggagca gttgcatcat cagcatggta tgatgttgtt 1740
cagagagatt ttgtcgagg cacatatgtt tggacaggat ttgattatct gggcgaaccg 1800
acaccgtgga atggcacagg ctgaggcgca gttggctcat ggccgtcacc gaaaaatagc 1860
tatttttgca tcgttgatac agcaggcttt ccgaaagata catattatit ttatcagagc 1920
cagtggaaatg atgatgttca taaactgcat attcttccgg catggaatga aaatgttgtt 1980
gcaaaaggct caggcaataa tgttccggtt gtcgtttata cagatgcagc gaaagtgaaa 2040
ctgtatttta caccgaaagg ctcaacagaa aaaagactga tcggcgaaaa atcattttaca 2100
aaaaaaacaa cagcggcagg ctatacatat caagtctatg aaggcagcga taaagattca 2160
acagcgcata aaaacatgta tctgacatgg aatgttccgt gggcagaagg cacaatttca 2220
gcggaagcgt atgatgaaa taatcgctgt attccggaag gcagcacaga aggcaacgca 2280
tcagttacaa caacaggcaa agcagcaaaa ctgaaagcag atgcggatcg caaaacaatt 2340
acagcggatg gcaaagatct gtcatatatt gaagtcgatg tcacagatgc aaatggccat 2400
attgttccgg atgcagcaaa tagagtcaca tttgatgtta aaggcgcagg caaactgggt 2460
ggcgttgata atggctcatc accggatcat gattcatatc aagcggataa ccgcaaagca 2520
ttttcaggca aagtcctggc aattgttcag tcaacaaaag aagcaggcga aattacagtt 2580
acagcaaaag cagatggcct gcaatcaagc acagttaaaa ttgcaacaac agcagttccg 2640
ggaacaagca cagaaaaaac agtccgcagc ttttattaca gccgcaacta ttatgtcaaa 2700
acaggcaaca aaccgattct gccgtcagat gttgaagttc gctattcaga tggaacaagc 2760
gatagacaaa acgttacatg ggatgcagtt tcagatgacg aaattgcaaa agcaggetca 2820
ttttcagttg caggcacagt tgcaggccaa aaaattagcg ttcgcgtcac aatgattgat 2880
gaaattggcg cactg 2895

```

<210> 11
 <211> 3114
 <212> DNA
 <213> 人工序列

<220>
 <223> 编码BIF_1068的核苷酸序列

```

<400> 11
gttgaagatg caacaagaag cgatagcaca acacaaatgt catcaacacc ggaagtgtgt 60
tattcatcag cggtcgatag caaacaataa cgcacaagcg attttgatgc gaactggaaa 120
tttatgtctg cagatagcgt tcaagcaca gatccggcat ttgatgatgc agcatggcaa 180
caagttgate tgccgcatga ttatagcatc acacagaaat atagccaaag caatgaagca 240
gaatcagcat atcttccggg aggcacagcg tggatatgaa aaagctttac aattgataga 300
gatctggcag gcaaacgcat tgcgattaat tttgatggcg tctatatgaa tgcaacagtc 360
tggtttaatg gcgttaaact gggcacacat ccgtatggct attcaccgtt ttcatattgat 420
ctgacaggca atgcaaaatt tggcggagaa aacacaattg tcgtcaaagt tgaaaaataga 480
ctgccgtcat caagatggta ttcaggcagc ggcatttata gagatgttac actgacagtt 540

```

[0031]

```

acagatggcg ttcatgttgg caataatggc gtcgcaatta aaacaccgtc actggcaaca 600
caaaatggcg gagatgtcac aatgaacctg acaacaaaag tcgcaatga tacagaagca 660
gcagcgaaca ttacactgaa acagacagtt ttccgaaag gcggaaaaac ggatgcagca 720
attggcacag ttacaacagc atcaaaatca attgcagcag gcgcatcagc agatgttaca 780
agcacaatta cagcagcaag cccgaaactg tggtaaatla aaaacccgaa cctgtatata 840
gttagaacag aagtcttgaa cggaggcaaa gttctggata catatgatac agaatatggc 900
tttcgctgga caggctttga tgcaacatca ggcttttcac tgaatggcga aaaagtcaaa 960
ctgaaaggcg ttagcatgca tcatgatcaa ggctcacttg gcgcagttgc aaatagacgc 1020
gcaattgaaa gacaagtcca aatcctgcaa aaaatggcg tcaatagcat tcgcacaaca 1080
cataatccgg cagcaaaagc actgattgat gtctgcaatg aaaaaggcgt tctggttgtc 1140
gaagaagtct ttgatatttg gaaccgcagc aaaaatggca acacggaaga ttatggcaaa 1200
tggtttggcc aagcaattgc aggcgataat gcagttcttg gaggcgataa agatgaaaca 1260
tgggcgaaat ttgatcttac atcaacaatt aaccgcgata gaaatgcacc gtcagttatt 1320
atgtggtcac tgggcaatga aatgatggaa ggcatctcag gctcagtttc aggtttccg 1380
gcaacatcag caaaactggt tgcatggaca aaagcagcag attcaacaag accgatgaca 1440
tatggcgata acaaaattaa agcgaactgg aacgaatcaa atacaatggg cgataatctg 1500
acagcaaatg gcggagtgtg tggcacaat tattcagatg gcgcaacta tgataaaatt 1560
cgtacaacac atccgtcatg ggcaatttat ggctcagaaa cagcatcagc gattaatagc 1620
cgtggcattt ataatagaac aacaggcgga gcacaatcat cagataaaca gctgacaagc 1680
tatgataatt cagcagttgg ctggggagca gttgcatcat cagcatggta tgatgttgtt 1740
cagagagatt ttgtcgcagg cacatatgtt tggacaggat ttgattatct gggcgaaccg 1800
acaccgtgga atggcacagg ctcaggcgca gttggtcat ggccgtcacc gaaaaatagc 1860
tatttttgca tcgttgatac agcaggcttt ccgaaagata catattatit ttatcagagc 1920
cagtggaaat atgatgttca taaactgcat attcttccgg catggaatga aaatgttgtt 1980
gcaaaaggct caggcaataa tgttccggtt gtcgtttata cagatgcagc gaaagtgaag 2040
ctgtatttta caccgaaagg ctcaacagaa aaaagactga tcggcgaaaa atcatttaca 2100
aaaaaaacaa cagcggcagg ctatacatat caagtctatg aaggcagcga taaagattca 2160
acagcgcata aaaaatgta tctgacatgg aatgttccgt gggcagaagg cacaatttca 2220
gcggaagcgt atgatgaaa taatcgcttg attccggaag gcagcacaga aggcaacgca 2280
tcagttacaa caacaggcaa agcagcaaaa ctgaaagcag atcgcgatcg caaaacaatt 2340
acagcggatg gcaaagatct gtcatatatt gaagtcgatg tcacagatgc aaatggccat 2400
attgttccgg atgcagcaaa tagagtcaca ttgatgtta aaggcgcagg caaactgggt 2460
ggcgttgata atggctcatc accggatcat gattcatatc aagcggataa ccgcaaagca 2520
ttttcaggca aagtcctgga aattgttcag tcaacaaaag aagcagcgca aattacagtt 2580
acagcaaaag cagatggcct gcaatcaagc acagttaaaa ttgcaacaac agcagttccg 2640
ggaacaagca cagaaaaaac agtccgcagc ttttattaca gccgcaacta ttatgtcaaa 2700
acaggcaaca aaccgattct gccgtcagat gttgaagttc gctattcaga tggaacaagc 2760
gatagacaaa acgttacatg ggatgcagtt tcagatgac aaattgcaaa agcaggctca 2820

```

	ttttcagttg caggcacagt tgcaggccaa aaaattagcg ttcgcgtcac aatgattgat	2880
	gaaattggcg cactgctgaa ttattcagca agcacaccgg ttggcacacc ggcagttctt	2940
	ccgggatcaa gaccggcagt cctgccggat ggcacagtca catcagcaaa ttttcagtc	3000
	cattggacaa aaccggcaga tacagtctat aatacagcag gcacagtcaa agtaccggga	3060
	acagcaacag tttttggcaa agaattttaa gtcacagcga caattagagt tcaa	3114
	<210> 12	
	<211> 3426	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 编码BIF_1172的核苷酸序列	
	<400> 12	
	gttgaagatg caacaagaag cgatagcaca acacaaatgt catcaacacc ggaagttgtt	60
	tattcatcag cggctgatag caaacaaaaat cgcacaagcg attttgcagc gaactggaaa	120
	tttatgctgt cagatagcgt tcaagcaca gatccggcat ttgatgattc agcatggcaa	180
	caagttgacg tgcgcgatga ttatagcatc acacagaaat atagccaaag caatgaagca	240
	gaatcagcat atcttccggg aggcacaggc tggatagaaa aaagctttac aattgataga	300
	gatctggcag gcaaacgcat tgcgattaat tttgatggcg tctatatgaa tgcaacagtc	360
	tggtttaatg gcgttaaaact gggcacacat ccgtatggct attcaccgtt ttcatttgat	420
	ctgacaggca atgcaaaatt tggcggagaa aacacaattg tcgtcaaagt tgaaaataga	480
	ctgccgtcat caagatggta ttcaggcagc ggcatattata gagatgttac actgacagtt	540
[0032]	acagatggcg ttcatgttgg caataatggc gtcgcaatta aaacaccgtc actggcaaca	600
	caaaatggcg gagatgtcac aatgaacctg acaacaaaag tcgcgaatga tacagaagca	660
	gcagcgaaca ttacactgaa acagacagtt tttccgaaag gcggaaaaac ggatgcagca	720
	attggcacag ttacaacagc atcaaaatca attgcagcag gcgcatcagc agatgttaca	780
	agcacaatta cagcagcaag cccgaaactg tggtaatta aaaacccgaa cctgtatata	840
	gttagaacag aagtcttgaa cggaggcaaa gttctggata catatgatac agaatatggc	900
	tttcgctgga caggctttga tgcaacatca ggcttttcac tgaatggcga aaaagtcaaa	960
	ctgaaaggcg ttagcatgca tcatgatcaa ggctcacttg gcgcagttgc aaatagacgc	1020
	gcaattgaaa gacaagtcca aatcctgcaa aaaatggcg tcaatagcat tcgcacaaca	1080
	cataatccgg cagcaaaagc actgattgat gtctgcaatg aaaaaggcgt tctggttgc	1140
	gaagaagtct ttgatatgtg gaaccgcagc aaaaatggca acacggaaga ttatggcaaa	1200
	tggtttggcc aagcaattgc aggcgataat gcagttctgg gaggcgataa agatgaaaca	1260
	tgggcgaaat ttgatcttac atcaacaatt aaccgcgata gaaatgcacc gtcagttatt	1320
	atgtggtcac tgggcaatga aatgatggaa ggcatttcag gctcagtttc aggtttccg	1380
	gcaacatcag caaaactggt tgcatggaca aaagcagcag attcaacaag accgatgaca	1440
	tatggcgata acaaaattaa agcgaactgg aacgaatcaa atacaatggg cgataatctg	1500
	acagcaaatg gcggagtgtg tggcacaat tttcagatg gcgcaacta tgataaaatt	1560
	cgtacaacac atccgtcatg ggcaatttat ggctcagaaa cagcatcagc gattaatagc	1620
	cgtggcattt ataatagaac aacaggcgga gcacaatcat cagataaaca gctgacaagc	1680

[0033]

```

taigataatt cagcagttgg ctggggagca gttgcatcat cagcatggta tgatgttgtt 1740
cagagagatt ttgtcgcagg cacatatgtt tggacaggat ttgattatct gggcgaaccg 1800
acaccgtgga atggcacagg ctcaggcgca gttggctcat ggccgtcacc gaaaaatagc 1860
tatttttgga tcgttgatac agcaggtttt ccgaaagata catattatit ttatcagagc 1920
cagtggaaatg atgatgttca tacactgcat attcttccgg catggaatga aaatgttgtt 1980
gcaaaaggct caggcaataa tgttccggtt gtcgtttata cagatgcagc gaaagtgaaa 2040
ctgtatttta caccgaaagg ctcaacagaa aaaagactga tcggcgaaaa atcatttaca 2100
aaaaaaacaa cagcggcagg ctatacatat caagtctatg aaggcagcga taaagattca 2160
acagcgcata aaaacatgta tctgacatgg aatgttccgt gggcagaagg cacaatttca 2220
gcggaagcgt atgatgaaaa taatcgctg attccggaag gcagcacaga aggcaacgca 2280
tcagttacaa caacaggcaa agcagcaaaa ctgaaagcag atgcggatcg caaaacaatt 2340
acagcggatg gcaaagatct gtcatatatt gaagtcgatg tcacagatgc aaatggccat 2400
attgttccgg atgcagcaaa tagagtcaca ttgatgtta aaggcgcagg caaactggtt 2460
ggcgttgata atggctcatc accggatcat gattcatatc aagcggataa ccgcaaagca 2520
ttttcaggca aagtcttggc aattgttcag tcaacaaaag aagcaggcga aattacagtt 2580
acagcaaaag cagatggcct gcaatcaagc acagttaaaa ttgcaacaac agcagttccg 2640
ggaacaagca cagaaaaaac agtccgcagc ttttattaca gccgcaacta ttatgtcaaa 2700
acaggcaaca aaccgattct gccgtcagat gttgaagttc gctattcaga tggaacaagc 2760
gatagacaaa acgttacatg ggatgcagtt tcagatgac aaattgcaaa agcaggctca 2820
ttttcagttg caggcacagt tgcaggccaa aaaattagcg ttcgcgtcac aatgattgat 2880
gaaattggcg cactgctgaa ttattcagca agcacaccgg ttggcacacc ggcagttctt 2940
ccgggatcaa gaccggcagt cctgccggat ggacacagtc catcagcaaa ttttgcagtc 3000
cattggacaa aaccggcaga tacagtctat aatacagcag gcacagtc aaataccggga 3060
acagcaacag tttttgcaa agaattttaa gtcacagcga caattagagt tcaaagaagc 3120
caagttacaa ttggctcatc agtttcagga aatgcactga gactgacaca aaatattccg 3180
gcagataaac aatcagatac actggatgcg attaaagatg gctcaacaac agttgatgca 3240
aatacaggcg gaggcgcaaa tccgtcagca tggacaaatt gggcatattc aaaagcaggc 3300
cataacacag cggaaattac atttgaatat gcgacagaac aacaactggg ccagatcgtc 3360
atgtattttt ttcgcgatag caatgcagtt agatttccgg atgctggcaa aacaaaaatt 3420
cagatc 3426

```

<210> 13
 <211> 3633
 <212> DNA
 <213> 人工序列

<220>
 <223> 编码BIF_1241的核苷酸序列

```

<400> 13
gttgaagatg caacaagaag cgatagcaca acacaaatgt catcaacacc ggaagttgtt 60
tattcatcag cggctgatag caaacaaaaat cgcacaagcg attttgatgc gaactggaaa 120
tttatgctgt cagatagcgt tcaagcacia gatccggcat ttgatgattc agcatggcaa 180

```

[0034]	caagttgcatc tgccgcatga ttatagcatc acacagaaat atagccaaag caatgaagca	240
	gaatcagcat atcttccggg aggcacaggc tggatagaa aaagctttac aattgataga	300
	gaictggcag gcaaacgcat tgcgattaat ttgatggcg tctatatgaa tgcaacagtc	360
	tggtttaatg gcgttaaact gggcacacat ccgtatggct attcaccgtt ttcatttgat	420
	ctgacaggca atgcaaaatt tggcggagaa aacacaattg tcgtcaaagt tgaaaataga	480
	ctgccgtcat caagatggta ttcaggcagc ggcatittata gagatgttac actgacagtt	540
	acagatggcg ttcattgttg caataatggc gtcgcaatta aaacaccgtc actggcaaca	600
	caaaatggcg gagatgtcac aatgaacctg acaacaaaag tcgcgaatga tacagaagca	660
	gcagcgaaca ttacactgaa acagacagtt tttccgaaag gcggaaaaac ggatgcagca	720
	attggcacag ttacaacagc atcaaaatca attgcagcag gcgcatcagc agatgttaca	780
	agcacaatta cagcagcaag cccgaaactg tggtaatta aaaacccgaa cctgtatata	840
	gttagaacag aagtcttgaa cggaggcaaa gtcttgata catatgatac agaatatggc	900
	tttcgtgga caggctttga tgcaacatca ggcttttcac tgaatggcga aaaagtcaaa	960
	ctgaaaggcg ttagcatgca tcatgatcaa ggctcacttg gcgcagttgc aaatagacgc	1020
	gcaattgaaa gacaagtcca aatcctgcaa aaaatggcg tcaatagcat tcgcacaaca	1080
	cataatccgg cagcaaaagc actgattgat gtctgcaatg aaaaaggcgt tctggttgtc	1140
	gaagaagtct ttgatattg gaaccgcagc aaaaatggca acacggaaga ttatggcaaa	1200
	tggtttggcc aagcaattgc aggcgataat gcagtcttg gaggcgataa agatgaaaca	1260
	tgggcgaaat ttgatcttac atcaacaatt aaccgcgata gaaatgcacc gtcagttatt	1320
	atgtgttcac tgggcaatga aatgatggaa ggcatttcag gtcagtttc aggccttccg	1380
	gcaacatcag caaaactggt tgcatggaca aaagcagcag attcaacaag accgatgaca	1440
	tatggcgata acaaaattaa agcgaactgg aacgaatcaa atacaatggg cgataatctg	1500
	acagcaaatg gcggagttgt tggcacaat tttcagatg gcgcaaaacta tgataaaatt	1560
	cgtacaacac atccgtcatg ggcaatttat ggctcagaaa cagcatcagc gattaatagc	1620
	cgtggcattt ataatagaac aacaggcgga gcacaatcat cagataaaca gctgacaagc	1680
	tatgataatt cagcagttgg ctggggagca gttgcatcat cagcatggta tgatgttgtt	1740
	cagagagatt ttgtcgcagg cacatatgtt tggacaggat ttgattatct gggcgaaccg	1800
	acaccgtgga atggcacagg ctcaggcgca gttggctcat ggccgtcacc gaaaaatagc	1860
	tattttggca tcgttgatac agcaggcttt ccgaaagata catattatit ttatcagagc	1920
	cagtggaaatg atgatgttca tacactgcat attcttccgg catggaatga aaatgttgtt	1980
	gcaaaaggct caggcaataa gtttccggtt gtcgtttata cagatgcagc gaaagtgaaa	2040
	ctgtatttta caccgaaagg ctcaacagaa aaaagactga tcggcgaaaa atcattttaca	2100
	aaaaaaacaa cagcggcagg ctatacatat caagtctatg aaggcagcga taaagattca	2160
	acagcgcata aaaacatgta tctgacatgg aatgttccgt gggcagaagg cacaatttca	2220
	gcggaagcgt atgatgaaaa taatcgctg attccggaag gcagcacaga aggcaacgca	2280
	tcagttacaa caacaggcaa agcagcaaaa ctgaaagcag atgcggatcg caaaacaatt	2340
	acagcggatg gcaaagatct gtcatatatt gaagtcgatg tcacagatgc aaatggccat	2400
	attgttccgg atgcagcaaa tagagtcaca ttgatgtta aaggcgcagg caaactgggt	2460

[0035]

```

ggcgttgata atggctcatc accggatcat gattcatatc aagcggataa ccgcaaagca 2520
ttttcaggca aagtcctggc aattgttcag tcaacaaaag aagcaggcga aattacagtt 2580
acagcaaaaag cagatggcct gcaatcaagc acagttaaaa ttgcaacaac agcagttccg 2640
ggaacaagca cagaaaaaac agtccgcagc ttttattaca gccgcaacta ttatgtcaaa 2700
acaggcaaca aaccgattct gccgtcagat gttgaagttc gctattcaga tggaacaagc 2760
gatagacaaa acgttacatg ggatgcagtt tcagatgac aaattgcaaa agcaggctca 2820
ttttcagttg caggcacagt tgcaggccaa aaaattagcg ttcgcgtcac aatgattgat 2880
gaaattggcg cactgctgaa ttattcagca agcacaccgg ttggcacacc ggcagttctt 2940
ccgggatcaa gaccggcagt cctgccgat ggcacagtca catcagcaa ttttcagtc 3000
cattggacaa aaccggcaga tacagtctat aatacagcag gcacagtcaa agtaccggga 3060
acagcaacag tttttggcaa agaatttaaa gtcacagcga caattagagt tcaaagaagc 3120
caagttacaa ttggctcatc agtttcagga atgcactga gactgacaca aaatattccg 3180
gcagataaac aatcagatac actggatgcg attaaagatg gctcaacaac agttgatgca 3240
aatacaggcg gaggcgcaaa tccgtcagca tggacaaatt gggcatattc aaaagcaggc 3300
cataacacag cgaaattac attgaatat gcgacagaac aacaactggg ccagatcgtc 3360
atgtattttt ttcgcgatag caatgcagtt agatttccgg atcttgcaa acaaaaaatt 3420
cagatcagcg cagatggcaa aaattggaca gatctggcag caacagaaac aattgcagcg 3480
caagaatcaa gcgatagagt caaacgtat acatatgatt ttgcaccggt tggcgcaaca 3540
tttgttaaaag tgacagtcac aaacgcagat acaacaacac cgtcaggcgt tgtttgcgca 3600
ggcctgacag aaattgaact gaaaacagcg aca 3633

```

<210> 14

<211> 3888

<212> DNA

<213> 人工序列

<220>

<223> 编码BIF_1326的核苷酸序列

<400> 14

```

gttgaagatg caacaagaag cgatagcaca acacaaatgt catcaacacc ggaagttgtt 60
tattcatcag cggctcgatg caacaaaat cgcaacagcg attttgatgc gaactggaaa 120
tttatgctgt cagatagcgt tcaagcaca gatccggcat ttgatgattc agcatggcaa 180
caagttgatc tgccgcatga ttatagcatc acacagaaat atagccaaag caatgaagca 240
gaatcagcat atcttccggg aggcacagcg tggatagaa aaagctttac aattgataga 300
gatctggcag gcaaacgcatt tgcgattaat ttgatggcg tctatatgaa tgcaacagtc 360
tggtttaatg gcgttaaac gggcacacat ccgtatggct attcaccgtt ttcatttgat 420
ctgacaggca atgcaaaatt tggcggagaa aacacaattg tcgtcaaagt tgaataataga 480
ctgccgtcat caagatggta ttcaggcagc ggcatttata gagatgttac actgacagtt 540
acagatggcg ttcattgttg caataatggc gtcgcaatta aaacaccgtc actggcaaca 600
caaatggcg gagatgtcac aatgaacctg acaacaaaag tcgcgaatga tacagaagca 660
gcagcgaaca ttacttgaa acagacagtt ttccgaaaag gcggaaaaac ggatgcagca 720
attggcacag ttacaacagc atcaaatca attgcagcag gcgcatcagc agatgttaca 780

```

[0036]

```

agcacaatta cagcagcaag cccgaaactg tggtaatta aaaacccgaa cctgtatata 840
gttagaacag aagttctgaa cggaggcaaa gtcttgata catatgatac agaataatggc 900
tttcgttga caggctttga tgcaacatca ggcttttcac tgaatggcga aaaagtcaaa 960
ctgaaaggcg ttagcatgca tcatgatcaa ggctcacttg gcgcagttgc aaatagacgc 1020
gcaattgaaa gacaagtcca aatcctgcaa aaaaatggcg tcaatagcat tcgcacaaca 1080
cataatccgg cagcaaaagc actgattgat gtctgcaatg aaaaaggcgt tctggttgtc 1140
gaagaagtct ttgatattgt gaaccgcagc aaaaatggca acacggaaga ttatggcaaa 1200
tggtttggcc aagcaattgc aggcgataat gcagttctgg gaggcgataa agatgaaaca 1260
tgggcgaaat ttgatcttac atcaacaatt aaccgcgata gaaatgcacc gtcagttatt 1320
atgttggtcac tgggcaatga aatgatggaa ggcatctcag gctcagtttc aggccttccg 1380
gcaacatcag caaaactggt tgcatggaca aaagcagcag attcaacaag accgatgaca 1440
tatggcgata acaaaattaa agcgaactgg aacgaatcaa atacaatggg cgataatctg 1500
acagcaaatg gcggagtgtg tggcacaatg tttcagatg gcgcaacta tgataaaatt 1560
cgtacaacac atccgtcatg ggcaatttat ggctcagaaa cagcatcagc gattaatagc 1620
cgtggcattt ataatagaac aacaggcgga gcacaatcat cagataaaca gctgacaagc 1680
tatgataatt cagcagttgg ctggggagca gttgcatcat cagcatggta tgatgttgtt 1740
cagagagatt ttgtcgagg cacatatgtt tggacaggat ttgattatct gggcgaaccg 1800
acaccgtgga atggcacagg ctcaggcgca gttggctcat ggccgtcacc gaaaaatagc 1860
tattttggca tcgttgatac agcaggcttt ccgaaagata catattatct ttatcagagc 1920
cagtggaaat atgatgttca tacactgcat attcttccgg catggaatga aaatgttgtt 1980
gcaaaaggct caggcaataa tgttccggtt gtcgtttata cagatgcagc gaaagtgaag 2040
ctgtatttta caccgaaagg ctcaacagaa aaaagactga tcggcgaaaa atcatttaca 2100
aaaaaaacaa cagcggcagg ctatacatat caagtctatg aaggcagcga taaagattca 2160
acagcgcata aaaacatgta tctgacatgg aatgttccgt gggcagaagg cacaatttca 2220
gcggaagcgt atgatgaaa taatcgctg attccggaag gcagcacaga aggcaacgca 2280
tcagttacaa caacaggcaa agcagcaaaa ctgaaagcag atgcggatcg caaaacaatt 2340
acagcggatg gcaagatct gtcatatatt gaagtcgatg tcacagatgc aaatggccat 2400
attgttccgg atgcagcaaa tagagtcaca ttgatgtta aaggcgcagg caaactggtt 2460
ggcgttgata atggctcatc accggatcat gattcatatc aagcggataa ccgcaaagca 2520
ttttcaggca aagtcctggc aattgttcag tcaacaaaag aagcaggcga aattacagtt 2580
acagcaaaag cagatggcct gcaatcaagc acagttaaaa ttgcaacaac agcagttccg 2640
ggaacaagca cagaaaaaac agtccgcagc ttttattaca gccgcaacta ttatgtcaaa 2700
acaggcaaca aaccgattct gccgtcagat gttgaagttc gctattcaga tggaacaagc 2760
gatagacaaa acgttacatg ggatgcagtt tcagatgac aaattgcaaa agcaggctca 2820
ttttcagttg caggcacagt tgcaggccaa aaaattagcg ttcgcgtcac aatgattgat 2880
gaaattggcg cactgctgaa ttattcagca agcacaccgg ttggcacacc ggcagttctt 2940
ccgggatcaa gaccggcagt cctgccgat ggcacagtca catcagcaaa ttttcagtc 3000
cattggacaa aaccggcaga tacagtctat aatcacagcag gcacagtcaa agtaccggga 3060

```

	acagcaacag tttttggcaa agaatttaaa gtcacagcga caattagagt tcaaagaagc	3120
	caagttacaa ttggctcatc agtttcagga aatgcactga gactgacaca aaatattccg	3180
	gcagataaac aatcagatac actggatgcg attaaagatg gctcaacaac agttgatgca	3240
	atacaggcg gaggcgcaaa tccgtcagca tggacaaatt gggcatattc aaaagcaggc	3300
	cataacacag cggaattac atttgaatat gcgacagaac aacaactggg ccagatcgtc	3360
	atgtattttt ttgcgatag caatgcagtt agatttccgg atgctggcaa aacaaaaatt	3420
	cagatcagcg cagatggcaa aaattggaca gatctggcag caacagaaac aattgcagcg	3480
	caagaatcaa gcgatagagt caaacgtat acatatgatt ttgcaccggt tggcgcaaca	3540
	tttgttaag tgacagtcac aaacgcagat acaacaacac cgtcaggcgt tgtttgcgca	3600
	ggcctgacag aaattgaact gaaaacagcg acaagcaaatt ttgtcacaaa tacatcagca	3660
	gcactgtcat cacttacagt caatggcaca aaagtttcag attcagttct ggcagcaggc	3720
	tcatataaca caccggcaat tatcgcatg gttaaagcgg aaggcgaagg caatgcaagc	3780
	gttacagtcc ttccggcaca tgataatggt attcgcgtca ttacagaaag cgaagatcat	3840
	gtcacacgca aaacatttac aatcaacctg ggcacagaac aagaattt	3888
	<210> 15	
	<211> 30	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 用于生成BIF变体的正向引物	
[0037]	<400> 15	
	ggggtaacta gtggaagatg caacaagaag	30
	<210> 16	
	<211> 35	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> BIF917的反向引物	
	<400> 16	
	gcgcttaatt aattatgttt tttctgtgct tgttc	35
	<210> 17	
	<211> 38	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> BIF995的反向引物	
	<400> 17	
	gcgcttaatt aattacagtg cgccaatttc atcaatca	38
	<210> 18	
	<211> 35	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> BIF1068的反向引物	
	<400> 18	
	gcgcttaatt aattattgaa ctctaattgt cgctg	35

<210> 19
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> BIF1241的反向引物

<400> 19
 gcgcttaatt aattatgtcg ctgttttcag ttcaat 36

<210> 20
 <211> 35
 <212> DNA
 <213> 人工序列

<220>
 <223> BIF1326的反向引物

<400> 20
 gcgcttaatt aattaaaatt ctgtttctgt gccca 35

<210> 21
 <211> 38
 <212> DNA
 <213> 人工序列

<220>
 <223> BIF1478的反向引物

<400> 21
 gcgcttaatt aattatctca gtctaatttc gcttgccg 38

[0038]

<210> 22
 <211> 1720
 <212> PRT
 <213> 两歧双歧杆菌

<400> 22

Val Glu Asp Ala Thr Arg Ser Asp Ser Thr Thr Gln Met Ser Ser Thr
 1 5 10 15

Pro Glu Val Val Tyr Ser Ser Ala Val Asp Ser Lys Gln Asn Arg Thr
 20 25 30

Ser Asp Phe Asp Ala Asn Trp Lys Phe Met Leu Ser Asp Ser Val Gln
 35 40 45

Ala Gln Asp Pro Ala Phe Asp Asp Ser Ala Trp Gln Gln Val Asp Leu
 50 55 60

Pro His Asp Tyr Ser Ile Thr Gln Lys Tyr Ser Gln Ser Asn Glu Ala
 65 70 75 80

Glu Ser Ala Tyr Leu Pro Gly Gly Thr Gly Trp Tyr Arg Lys Ser Phe
 85 90 95

Thr Ile Asp Arg Asp Leu Ala Gly Lys Arg Ile Ala Ile Asn Phe Asp
 100 105 110

Gly Val Tyr Met Asn Ala Thr Val Trp Phe Asn Gly Val Lys Leu Gly
 115 120 125

	Thr	His	Pro	Tyr	Gly	Tyr	Ser	Pro	Phe	Ser	Phe	Asp	Leu	Thr	Gly	Asn	
	130						135					140					
	Ala	Lys	Phe	Gly	Gly	Glu	Asn	Thr	Ile	Val	Val	Lys	Val	Glu	Asn	Arg	
	145					150					155					160	
	Leu	Pro	Ser	Ser	Arg	Trp	Tyr	Ser	Gly	Ser	Gly	Ile	Tyr	Arg	Asp	Val	
					165					170					175		
	Thr	Leu	Thr	Val	Thr	Asp	Gly	Val	His	Val	Gly	Asn	Asn	Gly	Val	Ala	
				180					185					190			
	Ile	Lys	Thr	Pro	Ser	Leu	Ala	Thr	Gln	Asn	Gly	Gly	Asp	Val	Thr	Met	
			195					200					205				
	Asn	Leu	Thr	Thr	Lys	Val	Ala	Asn	Asp	Thr	Glu	Ala	Ala	Ala	Asn	Ile	
	210						215					220					
	Thr	Leu	Lys	Gln	Thr	Val	Phe	Pro	Lys	Gly	Gly	Lys	Thr	Asp	Ala	Ala	
	225					230					235					240	
	Ile	Gly	Thr	Val	Thr	Thr	Ala	Ser	Lys	Ser	Ile	Ala	Ala	Gly	Ala	Ser	
				245						250					255		
	Ala	Asp	Val	Thr	Ser	Thr	Ile	Thr	Ala	Ala	Ser	Pro	Lys	Leu	Trp	Ser	
				260					265					270			
[0039]	Ile	Lys	Asn	Pro	Asn	Leu	Tyr	Thr	Val	Arg	Thr	Glu	Val	Leu	Asn	Gly	
			275					280					285				
	Gly	Lys	Val	Leu	Asp	Thr	Tyr	Asp	Thr	Glu	Tyr	Gly	Phe	Arg	Trp	Thr	
		290					295					300					
	Gly	Phe	Asp	Ala	Thr	Ser	Gly	Phe	Ser	Leu	Asn	Gly	Glu	Lys	Val	Lys	
	305					310					315					320	
	Leu	Lys	Gly	Val	Ser	Met	His	His	Asp	Gln	Gly	Ser	Leu	Gly	Ala	Val	
					325					330					335		
	Ala	Asn	Arg	Arg	Ala	Ile	Glu	Arg	Gln	Val	Glu	Ile	Leu	Gln	Lys	Met	
				340					345						350		
	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ala	Lys	Ala	Leu	
			355					360					365				
	Ile	Asp	Val	Cys	Asn	Glu	Lys	Gly	Val	Leu	Val	Val	Glu	Glu	Val	Phe	
		370					375					380					
	Asp	Met	Trp	Asn	Arg	Ser	Lys	Asn	Gly	Asn	Thr	Glu	Asp	Tyr	Gly	Lys	
	385					390					395					400	
	Trp	Phe	Gly	Gln	Ala	Ile	Ala	Gly	Asp	Asn	Ala	Val	Leu	Gly	Gly	Asp	
				405						410					415		
	Lys	Asp	Glu	Thr	Trp	Ala	Lys	Phe	Asp	Leu	Thr	Ser	Thr	Ile	Asn	Arg	
				420					425					430			

	Asp	Arg	Asn	Ala	Pro	Ser	Val	Ile	Met	Trp	Ser	Leu	Gly	Asn	Glu	Met
			435					440					445			
	Met	Glu	Gly	Ile	Ser	Gly	Ser	Val	Ser	Gly	Phe	Pro	Ala	Thr	Ser	Ala
	450					455						460				
	Lys	Leu	Val	Ala	Trp	Thr	Lys	Ala	Ala	Asp	Ser	Thr	Arg	Pro	Met	Thr
	465					470					475					480
	Tyr	Gly	Asp	Asn	Lys	Ile	Lys	Ala	Asn	Trp	Asn	Glu	Ser	Asn	Thr	Met
				485						490					495	
	Gly	Asp	Asn	Leu	Thr	Ala	Asn	Gly	Gly	Val	Val	Gly	Thr	Asn	Tyr	Ser
				500					505					510		
	Asp	Gly	Ala	Asn	Tyr	Asp	Lys	Ile	Arg	Thr	Thr	His	Pro	Ser	Trp	Ala
			515					520					525			
	Ile	Tyr	Gly	Ser	Glu	Thr	Ala	Ser	Ala	Ile	Asn	Ser	Arg	Gly	Ile	Tyr
	530						535					540				
	Asn	Arg	Thr	Thr	Gly	Gly	Ala	Gln	Ser	Ser	Asp	Lys	Gln	Leu	Thr	Ser
	545					550					555					560
	Tyr	Asp	Asn	Ser	Ala	Val	Gly	Trp	Gly	Ala	Val	Ala	Ser	Ser	Ala	Trp
					565					570					575	
[0040]	Tyr	Asp	Val	Val	Gln	Arg	Asp	Phe	Val	Ala	Gly	Thr	Tyr	Val	Trp	Thr
				580					585					590		
	Gly	Phe	Asp	Tyr	Leu	Gly	Glu	Pro	Thr	Pro	Trp	Asn	Gly	Thr	Gly	Ser
			595					600					605			
	Gly	Ala	Val	Gly	Ser	Trp	Pro	Ser	Pro	Lys	Asn	Ser	Tyr	Phe	Gly	Ile
	610						615					620				
	Val	Asp	Thr	Ala	Gly	Phe	Pro	Lys	Asp	Thr	Tyr	Tyr	Phe	Tyr	Gln	Ser
	625					630					635					640
	Gln	Trp	Asn	Asp	Asp	Val	His	Thr	Leu	His	Ile	Leu	Pro	Ala	Trp	Asn
				645					650						655	
	Glu	Asn	Val	Val	Ala	Lys	Gly	Ser	Gly	Asn	Asn	Val	Pro	Val	Val	Val
				660					665					670		
	Tyr	Thr	Asp	Ala	Ala	Lys	Val	Lys	Leu	Tyr	Phe	Thr	Pro	Lys	Gly	Ser
			675					680					685			
	Thr	Glu	Lys	Arg	Leu	Ile	Gly	Glu	Lys	Ser	Phe	Thr	Lys	Lys	Thr	Thr
	690						695					700				
	Ala	Ala	Gly	Tyr	Thr	Tyr	Gln	Val	Tyr	Glu	Gly	Ser	Asp	Lys	Asp	Ser
	705					710					715					720
	Thr	Ala	His	Lys	Asn	Met	Tyr	Leu	Thr	Trp	Asn	Val	Pro	Trp	Ala	Glu
				725						730					735	

	Gly	Thr	Ile	Ser	Ala	Glu	Ala	Tyr	Asp	Glu	Asn	Asn	Arg	Leu	Ile	Pro	
				740					745					750			
	Glu	Gly	Ser	Thr	Glu	Gly	Asn	Ala	Ser	Val	Thr	Thr	Thr	Gly	Lys	Ala	
			755					760						765			
	Ala	Lys	Leu	Lys	Ala	Asp	Ala	Asp	Arg	Lys	Thr	Ile	Thr	Ala	Asp	Gly	
		770					775						780				
	Lys	Asp	Leu	Ser	Tyr	Ile	Glu	Val	Asp	Val	Thr	Asp	Ala	Asn	Gly	His	
	785					790					795					800	
	Ile	Val	Pro	Asp	Ala	Ala	Asn	Arg	Val	Thr	Phe	Asp	Val	Lys	Gly	Ala	
				805						810					815		
	Gly	Lys	Leu	Val	Gly	Val	Asp	Asn	Gly	Ser	Ser	Pro	Asp	His	Asp	Ser	
			820						825					830			
	Tyr	Gln	Ala	Asp	Asn	Arg	Lys	Ala	Phe	Ser	Gly	Lys	Val	Leu	Ala	Ile	
		835						840					845				
	Val	Gln	Ser	Thr	Lys	Glu	Ala	Gly	Glu	Ile	Thr	Val	Thr	Ala	Lys	Ala	
		850					855					860					
	Asp	Gly	Leu	Gln	Ser	Ser	Thr	Val	Lys	Ile	Ala	Thr	Thr	Ala	Val	Pro	
	865					870					875					880	
[0041]	Gly	Thr	Ser	Thr	Glu	Lys	Thr	Val	Arg	Ser	Phe	Tyr	Tyr	Ser	Arg	Asn	
				885						890					895		
	Tyr	Tyr	Val	Lys	Thr	Gly	Asn	Lys	Pro	Ile	Leu	Pro	Ser	Asp	Val	Glu	
			900						905					910			
	Val	Arg	Tyr	Ser	Asp	Gly	Thr	Ser	Asp	Arg	Gln	Asn	Val	Thr	Trp	Asp	
		915						920					925				
	Ala	Val	Ser	Asp	Asp	Gln	Ile	Ala	Lys	Ala	Gly	Ser	Phe	Ser	Val	Ala	
		930				935						940					
	Gly	Thr	Val	Ala	Gly	Gln	Lys	Ile	Ser	Val	Arg	Val	Thr	Met	Ile	Asp	
	945					950					955				960		
	Glu	Ile	Gly	Ala	Leu	Leu	Asn	Tyr	Ser	Ala	Ser	Thr	Pro	Val	Gly	Thr	
				965						970					975		
	Pro	Ala	Val	Leu	Pro	Gly	Ser	Arg	Pro	Ala	Val	Leu	Pro	Asp	Gly	Thr	
			980						985					990			
	Val	Thr	Ser	Ala	Asn	Phe	Ala	Val	His	Trp	Thr	Lys	Pro	Ala	Asp	Thr	
		995					1000						1005				
	Val	Tyr	Asn	Thr	Ala	Gly	Thr	Val	Lys	Val	Pro	Gly	Thr	Ala	Thr		
		1010					1015					1020					
	Val	Phe	Gly	Lys	Glu	Phe	Lys	Val	Thr	Ala	Thr	Ile	Arg	Val	Gln		
		1025					1030					1035					

	Arg	Ser	Gln	Val	Thr	Ile	Gly	Ser	Ser	Val	Ser	Gly	Asn	Ala	Leu
	1040						1045					1050			
	Arg	Leu	Thr	Gln	Asn	Ile	Pro	Ala	Asp	Lys	Gln	Ser	Asp	Thr	Leu
	1055						1060					1065			
	Asp	Ala	Ile	Lys	Asp	Gly	Ser	Thr	Thr	Val	Asp	Ala	Asn	Thr	Gly
	1070						1075					1080			
	Gly	Gly	Ala	Asn	Pro	Ser	Ala	Trp	Thr	Asn	Trp	Ala	Tyr	Ser	Lys
	1085						1090					1095			
	Ala	Gly	His	Asn	Thr	Ala	Glu	Ile	Thr	Phe	Glu	Tyr	Ala	Thr	Glu
	1100						1105					1110			
	Gln	Gln	Leu	Gly	Gln	Ile	Val	Met	Tyr	Phe	Phe	Arg	Asp	Ser	Asn
	1115						1120					1125			
	Ala	Val	Arg	Phe	Pro	Asp	Ala	Gly	Lys	Thr	Lys	Ile	Gln	Ile	Ser
	1130						1135					1140			
	Ala	Asp	Gly	Lys	Asn	Trp	Thr	Asp	Leu	Ala	Ala	Thr	Glu	Thr	Ile
	1145						1150					1155			
	Ala	Ala	Gln	Glu	Ser	Ser	Asp	Arg	Val	Lys	Pro	Tyr	Thr	Tyr	Asp
	1160						1165					1170			
[0042]	Phe	Ala	Pro	Val	Gly	Ala	Thr	Phe	Val	Lys	Val	Thr	Val	Thr	Asn
	1175						1180					1185			
	Ala	Asp	Thr	Thr	Thr	Pro	Ser	Gly	Val	Val	Cys	Ala	Gly	Leu	Thr
	1190						1195					1200			
	Glu	Ile	Glu	Leu	Lys	Thr	Ala	Thr	Ser	Lys	Phe	Val	Thr	Asn	Thr
	1205						1210					1215			
	Ser	Ala	Ala	Leu	Ser	Ser	Leu	Thr	Val	Asn	Gly	Thr	Lys	Val	Ser
	1220						1225					1230			
	Asp	Ser	Val	Leu	Ala	Ala	Gly	Ser	Tyr	Asn	Thr	Pro	Ala	Ile	Ile
	1235						1240					1245			
	Ala	Asp	Val	Lys	Ala	Glu	Gly	Glu	Gly	Asn	Ala	Ser	Val	Thr	Val
	1250						1255					1260			
	Leu	Pro	Ala	His	Asp	Asn	Val	Ile	Arg	Val	Ile	Thr	Glu	Ser	Glu
	1265						1270					1275			
	Asp	His	Val	Thr	Arg	Lys	Thr	Phe	Thr	Ile	Asn	Leu	Gly	Thr	Glu
	1280						1285					1290			
	Gln	Glu	Phe	Pro	Ala	Asp	Ser	Asp	Glu	Arg	Asp	Tyr	Pro	Ala	Ala
	1295						1300					1305			
	Asp	Met	Thr	Val	Thr	Val	Gly	Ser	Glu	Gln	Thr	Ser	Gly	Thr	Ala
	1310						1315					1320			

	Thr	Glu	Gly	Pro	Lys	Lys	Phe	Ala	Val	Asp	Gly	Asn	Thr	Ser	Thr
	1325						1330					1335			
	Tyr	Trp	His	Ser	Asn	Trp	Thr	Pro	Thr	Thr	Val	Asn	Asp	Leu	Trp
	1340						1345					1350			
	Ile	Ala	Phe	Glu	Leu	Gln	Lys	Pro	Thr	Lys	Leu	Asp	Ala	Leu	Arg
	1355						1360					1365			
	Tyr	Leu	Pro	Arg	Pro	Ala	Gly	Ser	Lys	Asn	Gly	Ser	Val	Thr	Glu
	1370						1375					1380			
	Tyr	Lys	Val	Gln	Val	Ser	Asp	Asp	Gly	Thr	Asn	Trp	Thr	Asp	Ala
	1385						1390					1395			
	Gly	Ser	Gly	Thr	Trp	Thr	Thr	Asp	Tyr	Gly	Trp	Lys	Leu	Ala	Glu
	1400						1405					1410			
	Phe	Asn	Gln	Pro	Val	Thr	Thr	Lys	His	Val	Arg	Leu	Lys	Ala	Val
	1415						1420					1425			
	His	Thr	Tyr	Ala	Asp	Ser	Gly	Asn	Asp	Lys	Phe	Met	Ser	Ala	Ser
	1430						1435					1440			
	Glu	Ile	Arg	Leu	Arg	Lys	Ala	Val	Asp	Thr	Thr	Asp	Ile	Ser	Gly
	1445						1450					1455			
[0043]	Ala	Thr	Val	Thr	Val	Pro	Ala	Lys	Leu	Thr	Val	Asp	Arg	Val	Asp
	1460						1465					1470			
	Ala	Asp	His	Pro	Ala	Thr	Phe	Ala	Thr	Lys	Asp	Val	Thr	Val	Thr
	1475						1480					1485			
	Leu	Gly	Asp	Ala	Thr	Leu	Arg	Tyr	Gly	Val	Asp	Tyr	Leu	Leu	Asp
	1490						1495					1500			
	Tyr	Ala	Gly	Asn	Thr	Ala	Val	Gly	Lys	Ala	Thr	Val	Thr	Val	Arg
	1505						1510					1515			
	Gly	Ile	Asp	Lys	Tyr	Ser	Gly	Thr	Val	Ala	Lys	Thr	Phe	Thr	Ile
	1520						1525					1530			
	Glu	Leu	Lys	Asn	Ala	Pro	Ala	Pro	Glu	Pro	Thr	Leu	Thr	Ser	Val
	1535						1540					1545			
	Ser	Val	Lys	Thr	Lys	Pro	Ser	Lys	Leu	Thr	Tyr	Val	Val	Gly	Asp
	1550						1555					1560			
	Ala	Phe	Asp	Pro	Ala	Gly	Leu	Val	Leu	Gln	His	Asp	Arg	Gln	Ala
	1565						1570					1575			
	Asp	Arg	Pro	Pro	Gln	Pro	Leu	Val	Gly	Glu	Gln	Ala	Asp	Glu	Arg
	1580						1585					1590			
	Gly	Leu	Thr	Cys	Gly	Thr	Arg	Cys	Asp	Arg	Val	Glu	Gln	Leu	Arg
	1595						1600					1605			

Lys His Glu Asn Arg Glu Ala His Arg Thr Gly Leu Asp His Leu
1610 1615 1620

Glu Phe Val Gly Ala Ala Asp Gly Ala Val Gly Glu Gln Ala Thr
1625 1630 1635

Phe Lys Val His Val His Ala Asp Gln Gly Asp Gly Arg His Asp
1640 1645 1650

Asp Ala Asp Glu Arg Asp Ile Asp Pro His Val Pro Val Asp His
1655 1660 1665

Ala Val Gly Glu Leu Ala Arg Ala Ala Cys His His Val Ile Gly
1670 1675 1680

Leu Arg Val Asp Thr His Arg Leu Lys Ala Ser Gly Phe Gln Ile
1685 1690 1695

[0044]

Pro Ala Asp Asp Met Ala Glu Ile Asp Arg Ile Thr Gly Phe His
1700 1705 1710

Arg Phe Glu Arg His Val Gly
1715 1720

<210> 23

<211> 30

<212> PRT

<213> 两歧双歧杆菌

<400> 23

Val Arg Ser Lys Lys Leu Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu
1 5 10 15

Ile Phe Thr Met Ala Phe Gly Ser Thr Ser Ser Ala Gln Ala
20 25 30

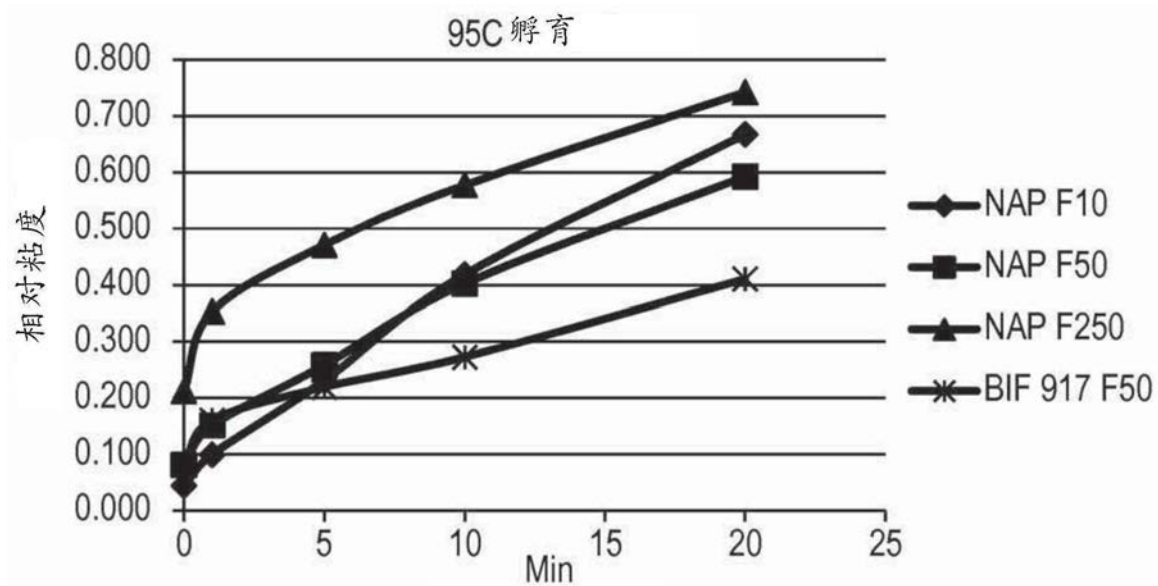


图1

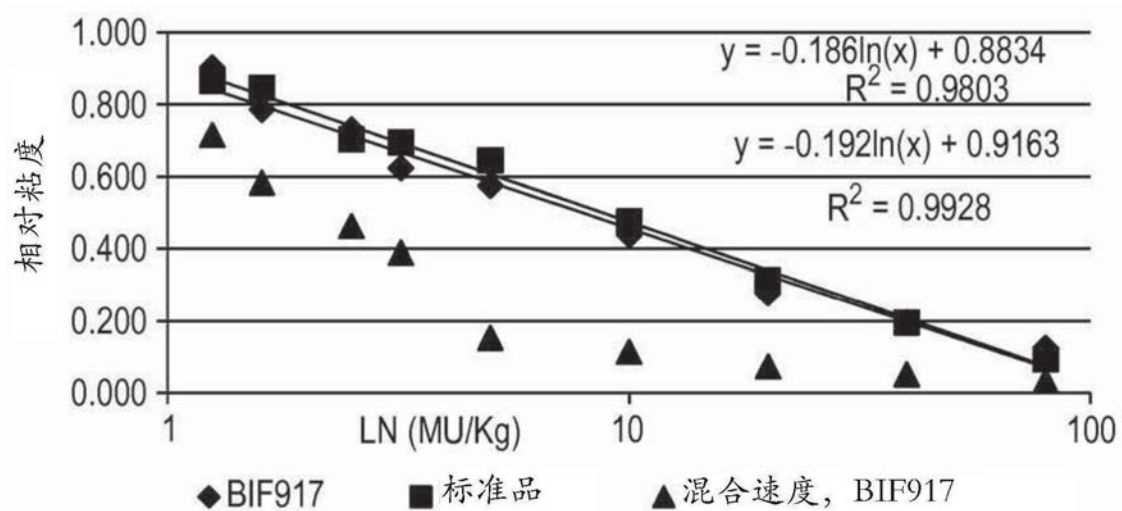


图2

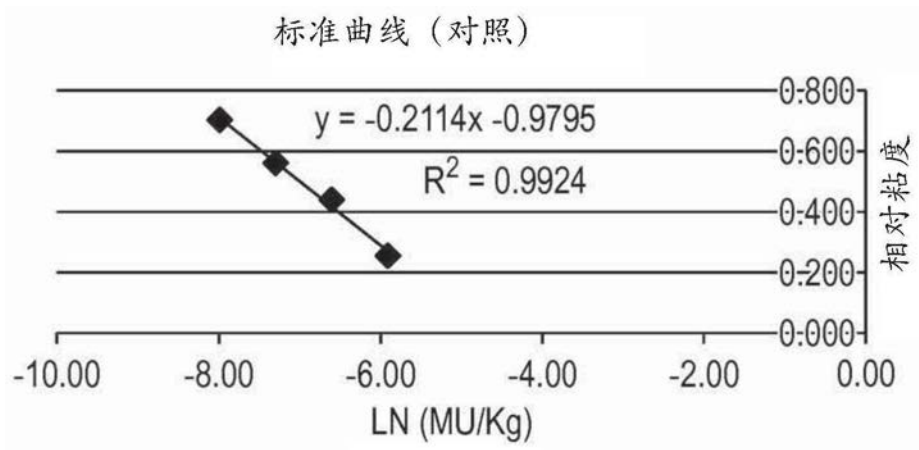


图3

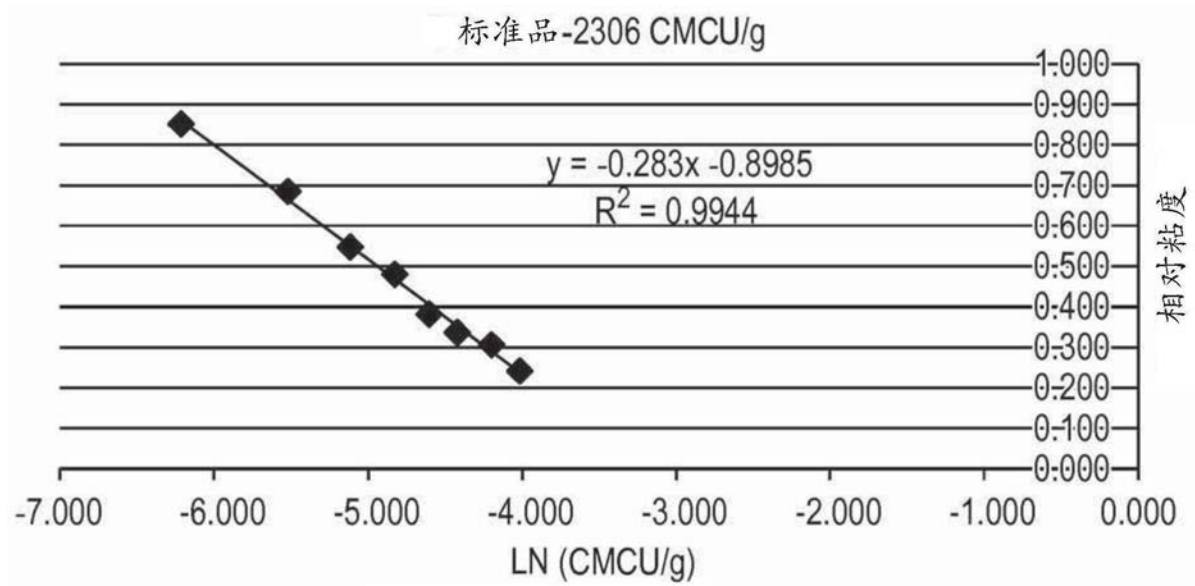


图4

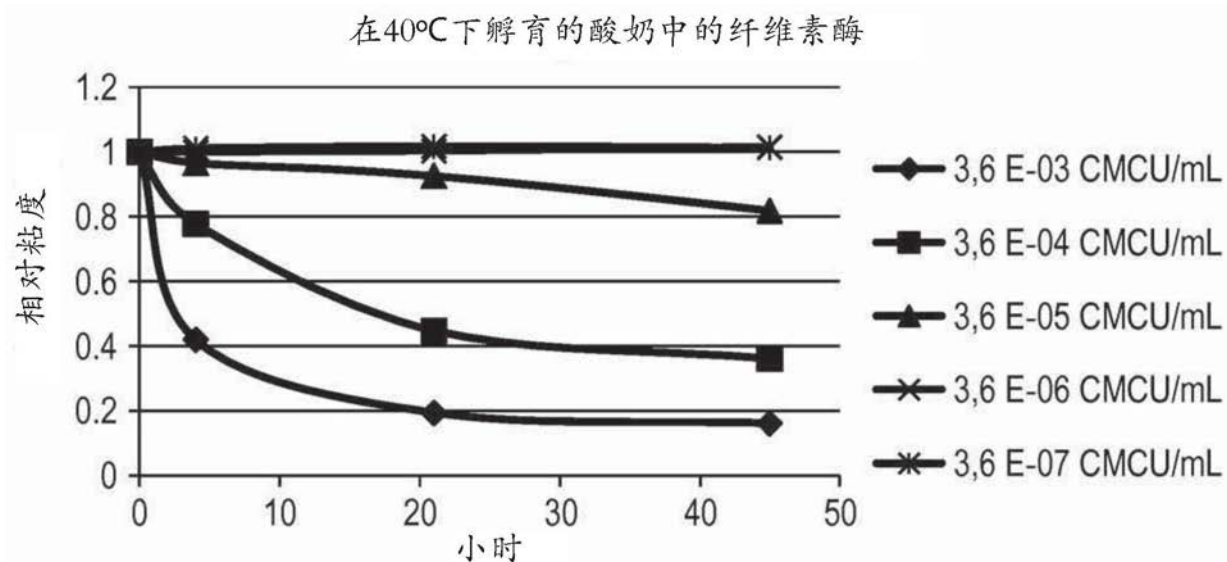


图5a

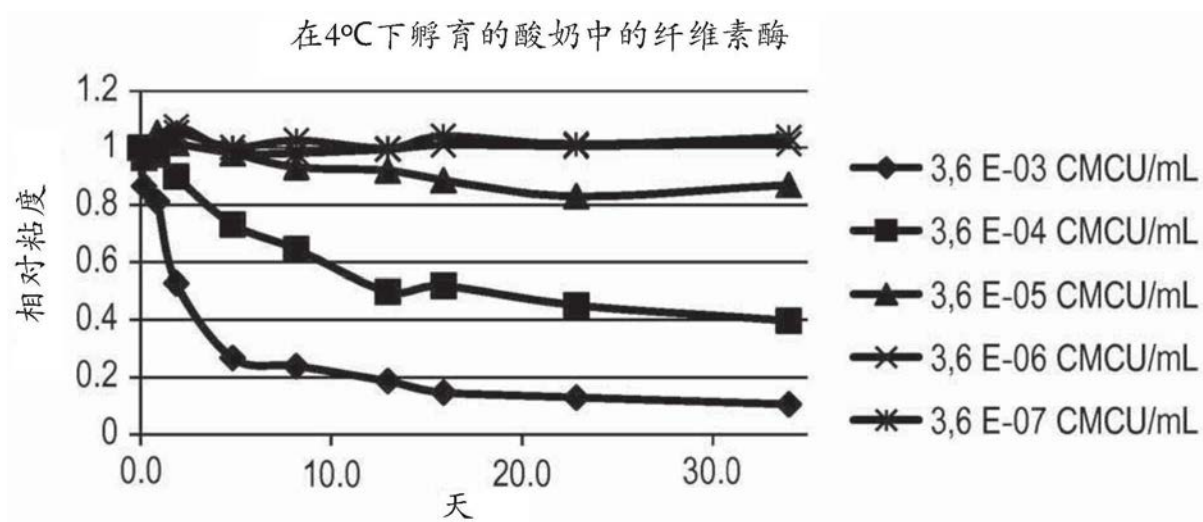


图5b