



(12) 发明专利申请

(10) 申请公布号 CN 115698280 A

(43) 申请公布日 2023. 02. 03

(21) 申请号 202280004542.9

(22) 申请日 2022.03.11

(30) 优先权数据

10-2021-0032885 2021.03.12 KR

(85) PCT国际申请进入国家阶段日

2022.11.24

(86) PCT国际申请的申请数据

PCT/KR2022/003415 2022.03.11

(87) PCT国际申请的公布数据

W02022/191653 K0 2022.09.15

(71) 申请人 CJ第一制糖株式会社

地址 韩国首尔市

(72) 发明人 李周姬 曹阿拉 池昶俊

(74) 专利代理机构 北京英赛嘉华知识产权代理
有限责任公司 11204

专利代理师 王达佐 洪欣

(51) Int.Cl.

C12N 9/52 (2006.01)

C12N 15/52 (2006.01)

C12N 15/75 (2006.01)

A23K 10/14 (2006.01)

A23K 10/16 (2006.01)

权利要求书1页 说明书22页

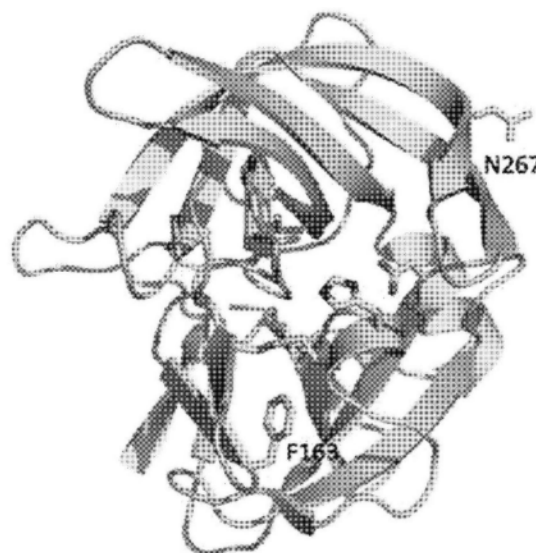
序列表61页 附图1页

(54) 发明名称

新的丝氨酸蛋白酶变体

(57) 摘要

本申请涉及新的丝氨酸蛋白酶变体。



1. 丝氨酸蛋白酶变体,所述丝氨酸蛋白酶变体与SEQ ID NO:54的氨基酸序列具有至少60%且小于100%的同源性或同一性,并包含对应于SEQ ID NO:54的氨基酸序列的第12位的氨基酸的取代。

2. 根据权利要求1所述的丝氨酸蛋白酶变体,其中所述丝氨酸蛋白酶变体进一步包含对应于SEQ ID NO:54的氨基酸序列的第116位的氨基酸的取代。

3. 根据权利要求1所述的丝氨酸蛋白酶变体,其中所述丝氨酸蛋白酶变体与SEQ ID NO:54的氨基酸序列具有至少75%且小于100%的同源性或同一性。

4. 根据权利要求1所述的丝氨酸蛋白酶变体,其中对应于第12位的氨基酸被亲水性氨基酸、非极性氨基酸或碱性氨基酸取代。

5. 根据权利要求1所述的丝氨酸蛋白酶变体,其中对应于第12位的氨基酸被酪氨酸(Y)、丙氨酸(A)、丝氨酸(S)或精氨酸(R)取代。

6. 根据权利要求1所述的丝氨酸蛋白酶变体,其中所述丝氨酸蛋白酶变体包含对应于SEQ ID NO:71或72的氨基酸序列的第201位的氨基酸的取代。

7. 根据权利要求1所述的丝氨酸蛋白酶变体,其中所述丝氨酸蛋白酶变体包含对应于SEQ ID NO:70的氨基酸序列的第203位的氨基酸的取代。

8. 根据权利要求1所述的丝氨酸蛋白酶变体,其中所述丝氨酸蛋白酶变体与SEQ ID NO:70至72中任一氨基酸序列具有至少60%且小于100%的序列同源性。

9. 根据权利要求2所述的丝氨酸蛋白酶变体,其中对应于第116位的氨基酸被亲水性氨基酸、非极性氨基酸或酸性氨基酸取代。

10. 根据权利要求2所述的丝氨酸蛋白酶变体,其中对应于第116位的氨基酸被天冬氨酸(D)、丝氨酸(S)、苏氨酸(T)或甘氨酸(G)取代。

11. 组合物,其包含权利要求1至10中任一项所述的丝氨酸蛋白酶变体。

12. 多核苷酸,其编码权利要求1至10中任一项所述的丝氨酸蛋白酶变体。

13. 载体,其包含权利要求12所述的多核苷酸。

14. 宿主细胞,其包含权利要求1至10中任一项所述的丝氨酸蛋白酶变体、编码所述丝氨酸蛋白酶变体的多核苷酸以及包含所述多核苷酸的载体中的至少一者。

15. 组合物,其包含权利要求1至10中任一项所述的丝氨酸蛋白酶变体和表达所述丝氨酸蛋白酶变体的微生物中的至少一者。

新的丝氨酸蛋白酶变体

技术领域

[0001] 本申请涉及新的丝氨酸蛋白酶变体。

背景技术

[0002] 蛋白酶参与各种功能,例如生物体内的消化、吸收和防御,并且根据活性位点的结构分为丝氨酸蛋白酶、半胱氨酸蛋白酶、天冬氨酸蛋白酶和金属蛋白酶。在这些酶中,丝氨酸蛋白酶(或丝氨酸内肽酶)的特征在于在它们的活性位点中共同地具有切割蛋白中肽键的活性丝氨酸残基,其中丝氨酸在蛋白酶的活性位点作为亲核氨基酸(Hedstrom, 2002.Chem Rev 102:4501-4524)。

[0003] 丝氨酸蛋白酶已经在多种应用中使用。除了治疗人类疾病(例如,溶解血凝块)的治疗应用之外,丝氨酸蛋白酶不仅用作洗衣除垢剂和隐形眼镜清洁剂的组分,而且还用于乳蛋白的改性、蚕丝脱胶、皮革浸泡、脱毛、寡肽的合成、从肺X射线胶片中回收银、饲料和食品的生产 and 改进等(韩国专利公开号10-2005-0068750)。

发明内容

[0004] 技术问题

[0005] 本领域需要开发具有改进的热稳定性、增加的活性等的丝氨酸蛋白酶,以获得更高的工业成本效益和效率。

[0006] 技术方案

[0007] 本申请的一个目的是提供丝氨酸蛋白酶变体。

[0008] 本申请的另一个目的是提供编码丝氨酸蛋白酶变体的多核苷酸和包含所述多核苷酸的载体。

[0009] 本申请的另一个目的是提供包含丝氨酸蛋白酶变体、编码丝氨酸蛋白酶变体的多核苷酸和包含所述多核苷酸的载体中的至少一者的微生物。

[0010] 本申请的另一个目的是提供包含丝氨酸蛋白酶变体和表达丝氨酸蛋白酶变体的微生物中的至少一者的饲料组合物。

[0011] 有益效果

[0012] 与现有的丝氨酸蛋白酶相比,本申请的丝氨酸蛋白酶变体具有优异的活性,因此可以在工业上使用。

[0013] 附图简要说明

[0014] 图1展示了来源于*Thermobifida fusca*的丝氨酸蛋白酶变体的三级结构中引入突变的残基位置。

[0015] 发明详细描述

[0016] 下文详细描述本申请的内容。同时,在本申请中公开的每处描述和实施方案可以在本文应用于不同的描述和实施方案。换言之,本申请中公开的各种组分的所有组合都包括在本申请的范围内。此外,本申请的范围不应受下文提供的描述的限制。

[0017] 此外,本领域普通技术人员可以认识到或能够使用不超过常规实验来确认本申请的具体实施方案的多种等同方案。这种等同方案旨在包含在本申请的范围内。

[0018] 本申请的一个方面提供了丝氨酸蛋白酶变体。

[0019] 本文所用术语“丝氨酸蛋白酶”是指属于蛋白酶亚组并具有蛋白水解活性的酶。具体而言,丝氨酸蛋白酶可以通过水解肽键而降解蛋白并且基本上在其活性位点具有活性丝氨酸残基的酶,并且更具体而言,是具有组氨酸、天冬氨酸和丝氨酸的氨基酸残基的空间排列(可以被称为催化三联体)的酶,但不限于此。

[0020] 根据本申请的丝氨酸蛋白酶可以来源于*Thermobifida*属,拟诺卡氏菌(*Nocardiopsis*)属,*Actinorugispora*属或*Spinactinospira*属的微生物,但不限于此。具体而言,在本申请中,野生型丝氨酸蛋白酶可以是来源于以下的丝氨酸蛋白酶:*Thermobifida fusca*,*Thermobifida celulosilytica*,*Thermobifida halotolerans*,*Actinorugispora endohytica*,*Spinactinospira alkalitolerans*,堆肥拟诺卡氏菌(*Nocardiopsis composta*)或*Nocardiopsis potens*,但不限于此。

[0021] 在一个实施方案中,本申请的丝氨酸蛋白酶可以是多肽,其包含SEQ ID NO:31所示的氨基酸序列,基本上由SEQ ID NO:31所示的氨基酸序列组成,或由SEQ ID NO:31所示的氨基酸序列组成,但不限于此。在一个实施方案中,SEQ ID NO:31的氨基酸序列可以是来源于SEQ ID NO:40或SEQ ID NO:2的氨基酸序列,但不限于此。

[0022] 在一个实施方案中,本申请的丝氨酸蛋白酶可以包括SEQ ID NO:49至54中任一氨基酸序列,基本上由SEQ ID NO:49至54中任一氨基酸序列组成,或由SEQ ID NO:49至54中任一氨基酸序列组成,但不限于此。在一个实施方案中,SEQ ID NO:49至54的氨基酸序列可以来源于SEQ ID NO:67至72中任一氨基酸序列,但不限于此。

[0023] 本申请的丝氨酸蛋白酶可以包括与上文描述的氨基酸序列具有相同活性的任何序列,但不限于此。此外,丝氨酸蛋白酶可以包含下述序列,或基本上由下述序列组成,或由下述序列组成:SEQ ID NO:31和49至54中任一氨基酸序列或与SEQ ID NO:31和49至54中任一氨基酸序列具有至少60%同源性或同一性的氨基酸序列,但不限于此。具体而言,氨基酸序列可以包含SEQ ID NO:31和49至54所示的氨基酸序列中的任一个或与SEQ ID NO:31和49至54所示的氨基酸序列中的任一个具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更多的同源性或同一性的氨基酸序列。另外,能够清楚理解的是,对于在部分序列内具有包括缺失、修饰、取代或添加的氨基酸序列的任何蛋白而言,只要该氨基酸序列具有上述同源性或同一性以及蛋白的效果等同的效果,其落入本申请的范围内。

[0024] 换言之,尽管在本申请中使用的表述是“具有预定的SEQ ID NO:所示的氨基酸序列的蛋白或多肽”和“包含预定的SEQ ID NO:所示的氨基酸序列的蛋白或多肽”,但是能清楚判断的是,在本申请中也可以使用具有在部分序列中包括缺失、修饰、取代或添加的氨基酸序列的任何蛋白,只要该蛋白具有与由相应氨基酸序列组成的多肽相同或等同的活性。例如,“包含SEQ ID NO:31的氨基酸序列的多肽”显然属于“由SEQ ID NO:31的氨基酸序列组成的多肽”,只要前者具有与后者相同或等同的活性。

[0025] 本文所用术语“同源性”或“同一性”是指两个给定氨基酸序列或核苷酸序列之间的相关性程度,并且可以显示为百分比。术语同源性和同一性可以互换使用。

[0026] 保守多核苷酸或多肽的序列同源性或同一性可以通过标准比对算法确定,并且可以一起使用由要使用的程序建立的默认缺口罚分。基本上,同源或相同的序列通常可以沿着整个序列或整个序列的至少约50%、60%、70%、80%或90%在中等或高度严格条件下彼此杂交。在杂交中,也考虑含有简并密码子代替密码子的多核苷酸。

[0027] 任何两个给定多核苷酸或多肽之间的序列同源性、相似性或同一性可以使用已知的计算机算法,如“FASTA”程序,通过使用Pearson et al. (1988) Proc. Natl. Acad. Sci. USA 85:2444的默认参数来确定。或者,可以使用Needleman-Wunsch算法 (Needleman and Wunsch, 1970, J. Mol. Biol. 48:443-453), 其在欧洲分子生物学开放软件套件 (European Molecular Biology Open Software Suite (EMBOSS)) 包的Needleman程序中进行 (Rice et al., 2000, Trends Genet. 16:276-277) (5.0.0版或其后的版本) (包括GCG程序包 (Devereux, J., et al., Nucleic Acids Research 12:387 (1984))、BLASTP、BLASTN、FASTA (Atschul, S.F., et al., J MOLEC BIOL 215:403 (1990); Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994和CARILLO et al. (1988) SIAM J Applied Math 48:1073)。例如,可以使用来自国家生物技术信息中心数据库 (National Center for Biotechnology Information database) 的BLAST或ClustalW来确定同源性、相似性或同一性。

[0028] 多核苷酸或多肽的同源性、相似性或同一性可以通过使用例如GAP计算机程序比较序列信息来确定,例如Needleman et al., (1970), J Mol Biol. 48:443, 如Smith and Waterman, Adv. Appl. Math (1981) 2:482所公开的程序。简言之, GAP程序将相似性定义为通过将相似比对的符号 (即核苷酸或氨基酸) 的数目除以两个序列中较短序列中符号的总数而获得的值。用于GAP程序的默认参数可以包括: (1) 二进制比较矩阵 (包含相同的值1和不相同的值0) 和Gribskov et al. (1986) Nucl. Acids Res. 14:6745的加权比较矩阵, 公开于Schwartz and Dayhoff, eds., Atlas Of Protein Sequence And Structure, National Biomedical Research Foundation, pp. 353-358 (1979) (或EDNAFULL (NCBI NUC4.4的EMBOSS版本) 取代矩阵); (2) 每个缺口的罚分为3.0, 并且每个缺口中的每个符号的额外的0.10罚分 (或缺口开放罚分为10, 缺口延伸罚分为0.5); (3) 端部缺口无罚分。

[0029] 此外, 任何两个给定的多核苷酸或多肽之间的序列同源性、相似性或同一性可以通过在限定的严格条件下、通过Southern杂交测试比较它们的序列来确认, 并且限定的适当杂交条件在本申请的范围内并且可以通过本领域普通技术人员熟知的方法来确定。

[0030] 在上文描述的具有丝氨酸蛋白酶的蛋白的一个实施方案中, 本申请中提供的丝氨酸蛋白酶变体可以指这样的变体: 在该变体中, 在具体位置上的氨基酸被取代, 从而与突变前的蛋白相比具有超过100%的酶活性。

[0031] 在一个具体的实施方案中, 本申请提供的变体可以具有与包括SEQ ID NO: 31和49至54中任一氨基酸序列的野生型酶相比, 超过100%的酶活性, 特别是约110%, 约120%, 约130%, 约140%, 约150%, 约160%, 约170%, 约180%, 约190%, 或约200%或更高的增加的酶活性, 但不限于此。

[0032] 术语“约”是包括所有 ± 0.5 、 ± 0.4 、 ± 0.3 、 ± 0.2 、 ± 0.1 等的范围, 并且包括在等

于或类似于术语“约”之后的数值的范围内的所有数值,但不限于此。

[0033] 本文所用术语“变体”是指通过保守取代和/或修饰所述序列的至少一个不同氨基酸、同时保留蛋白质的功能或特性而获得的多肽。由于数个氨基酸的取代、缺失或添加,变体具有与所鉴定的序列不同的氨基酸序列。这种变体通常可以通过修饰上述多肽序列之一并评价修饰的多肽的性质来鉴定。也就是说,相对于天然蛋白,变体的能力可以增强、不改变或减弱。

[0034] 此外,一些变体可以包括其中至少一部分(如N末端前导序列或跨膜结构域)被去除的变体。其它变体可以包括从成熟蛋白的N末端和/或C末端去除一部分或向成熟蛋白的N末端和/或C末端添加一部分的变体。

[0035] 术语“变体”也可以与其它术语互换使用,例如修饰/突变的蛋白质,修饰,修饰的多肽,突变体,突变蛋白和发散体(divergent),并且也可以使用用于指示变异的任何术语,对此没有限制。

[0036] 与天然野生型或未修饰的蛋白质相比,变体可具有增强的活性,但不限于此。

[0037] 本文所用的术语“保守型取代”是指一个氨基酸被具有相似结构和/或化学性质的另一个氨基酸取代。例如,变体可以具有一个或多个保守型取代,同时仍然保留一项或多项生物活性。这种氨基酸取代通常可以基于残基的极性、电荷、溶解度、疏水性、亲水性和/或两性性质的相似性而发生。例如,在具有侧链的带电氨基酸中,带正电的(碱性)氨基酸包括精氨酸,赖氨酸和组氨酸,带负电的(酸性)氨基酸包括谷氨酸和天冬氨酸;在具有侧链的不带电荷的氨基酸中,非极性氨基酸包括甘氨酸,丙氨酸,缬氨酸,亮氨酸,异亮氨酸,甲硫氨酸,苯丙氨酸,色氨酸和脯氨酸,极性或亲水性氨基酸包括丝氨酸,苏氨酸,半胱氨酸,酪氨酸,天冬酰胺和谷氨酰胺;在氨基酸中,芳香氨基酸包括苯丙氨酸,色氨酸和酪氨酸。变体还可以包括对多肽的性质和二级结构具有最小影响的氨基酸的缺失或添加。例如,以共翻译的方式或翻译后方式,多肽可以与参与蛋白质转移的蛋白质的N末端的信号(或前导)序列缀合。所述多肽也可以与不同的序列或接头缀合,以鉴定、纯化或合成多肽。

[0038] 本文所用术语“丝氨酸蛋白酶变体”是指包括在具有丝氨酸蛋白酶活性的多肽的氨基酸序列中的至少一个氨基酸的取代的多肽。

[0039] 根据本申请的丝氨酸蛋白酶变体可以包括在对应于SEQ ID NO:31的氨基酸序列自N末端起的第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代。具体而言,丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:31的第12位氨基酸和/或第116位氨基酸的氨基酸的取代,并且可以包括与SEQ ID NO:31和49至54中任一氨基酸序列具有至少60%且小于100%同源性或同一性的氨基酸序列。

[0040] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括在对应于SEQ ID NO:31的第12位氨基酸和/或第116位氨基酸的位置上的氨基酸的取代,并且可以与SEQ ID NO:31和49至54中任一氨基酸序列具有至少60%且具有小于100%(例如,61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高)的同源性或同一性,但不限于此。

[0041] 同时,由于SEQ ID NO:31的自N末端起的第12位氨基酸和第116位氨基酸对应于SEQ ID NO:49至54的自N末端起的第12位氨基酸和第116位氨基酸,因此对基于SEQ ID NO:

31的氨基酸的位置的描述同样可应用于SEQ ID NO:49至54中任一氨基酸序列的第12位氨基酸和第116位氨基酸。

[0042] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括在对应于SEQ ID NO:54的第12位氨基酸和/或第116位氨基酸的位置上的氨基酸的取代,并且可以包括与SEQ ID NO:52至54中任一氨基酸序列具有至少60%且小于100%(例如61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高)的同源性或同一性的氨基酸序列。具体而言,丝氨酸蛋白酶变体可以包括在对应于SEQ ID NO:54的第12位氨基酸的位置上的氨基酸的取代,并且可以与SEQ ID NO:54的氨基酸序列具有至少60%且小于100%(例如61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高)的同源性或同一性,但不限于此。

[0043] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以是这样的蛋白:SEQ ID NO:31和49至54中任一氨基酸序列中的自N末端起的第12位,第116位,或第12和116位相对应的氨基酸全部被其它氨基酸取代。“其它氨基酸”是指与取代前的氨基酸不同的氨基酸,并且不受限制,只要它们是取代前的氨基酸以外的氨基酸即可。

[0044] 在一个实施方案中,根据本申请的丝氨酸蛋白酶变体可以是这样的变体:SEQ ID NO:31和49至51中任一氨基酸序列中第12位的苯丙氨酸被甘氨酸,丙氨酸,精氨酸,天冬氨酸,半胱氨酸,谷氨酸,天冬酰胺,谷氨酰胺,组氨酸,脯氨酸,丝氨酸,酪氨酸,异亮氨酸,亮氨酸,赖氨酸,色氨酸、缬氨酸,甲硫氨酸或苏氨酸取代;和/或第116位的天冬酰胺被甘氨酸,丙氨酸,精氨酸,天冬氨酸,半胱氨酸,谷氨酸,谷氨酰胺,组氨酸,脯氨酸,丝氨酸,酪氨酸,异亮氨酸,亮氨酸,赖氨酸,苯丙氨酸,色氨酸,缬氨酸,甲硫氨酸或苏氨酸取代,但不限于此。

[0045] 在一个实施方案中,根据本申请的丝氨酸蛋白酶变体可以是这样的变体:SEQ ID NO:52至54中任一氨基酸序列中第12位的脯氨酸被苯丙氨酸,甘氨酸,丙氨酸,精氨酸,天冬氨酸,半胱氨酸,谷氨酸,天冬酰胺,谷氨酰胺,组氨酸,丝氨酸,酪氨酸,异亮氨酸,亮氨酸,赖氨酸,色氨酸、缬氨酸,甲硫氨酸或苏氨酸取代;和/或第116位的天冬酰胺被甘氨酸,丙氨酸,精氨酸,天冬氨酸,半胱氨酸,谷氨酸,谷氨酰胺,组氨酸,脯氨酸,丝氨酸,酪氨酸,异亮氨酸,亮氨酸,赖氨酸,苯丙氨酸,色氨酸,缬氨酸,甲硫氨酸或苏氨酸取代,但不限于此。

[0046] 具体而言,变体可以是这样的蛋白:对应于SEQ ID NO:31和49至54中任一氨基酸序列中的第12位的氨基酸被酪氨酸(Y),丝氨酸(S),丙氨酸(A)或精氨酸(R)取代;对应于第116位的氨基酸被天冬氨酸(D),丝氨酸(S),苏氨酸(T)或甘氨酸(G)取代;或者SEQ ID NO:31的氨基酸序列中第12和116位的氨基酸分别被酪氨酸(Y)和天冬氨酸(D),酪氨酸(Y)和丝氨酸(S),丝氨酸(S)和天冬氨酸(D),丝氨酸(S)和苏氨酸(T),或丙氨酸(A)和甘氨酸(G)取代,但不限于此。在一个实施方案中,丝氨酸蛋白酶变体可以是这样的变体:SEQ ID NO:52至54中任一氨基酸序列中的第12位的脯氨酸被酪氨酸、丙氨酸、丝氨酸或精氨酸取代,但不限于此。

[0047] 能够清楚理解的是,SEQ ID NO:31和49-54中任一氨基酸序列中的第12位和/或第

116位氨基酸被其它氨基酸取代的变体包括其中对应于所述位置的氨基酸被其它氨基酸取代的变体。

[0048] 另外,变体还包括这样的变体:对应于SEQ ID NO:31和49至54中任一氨基酸序列的自N末端起的第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被上述选自SEQ ID NO:31和49至54中任一氨基酸序列或与选自SEQ ID NO:31和49至54的任一氨基酸序列具有至少60% (60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高) 同源性或同一性的氨基酸序列中的其它氨基酸取代。

[0049] 在变体的一个实施方案中,SEQ ID NO:31的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:32至39中任一氨基酸序列,基本上由SEQ ID NO:32至39中任一氨基酸序列组成,或由SEQ ID NO:32至39中任一氨基酸序列组成,但不限于此。

[0050] 在变体的一个实施方案中,SEQ ID NO:49的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:55或56的氨基酸序列,基本上由SEQ ID NO:55或56的氨基酸序列组成,或由SEQ ID NO:55或56的氨基酸序列组成,但不限于此。

[0051] 在变体的一个实施方案中,SEQ ID NO:50的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:57或58的氨基酸序列,基本上由SEQ ID NO:57或58的氨基酸序列组成,或由SEQ ID NO:57或58的氨基酸序列组成,但不限于此。

[0052] 在变体的一个实施方案中,SEQ ID NO:51的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:59或60的氨基酸序列,基本上由SEQ ID NO:59或60的氨基酸序列组成,或由SEQ ID NO:59或60的氨基酸序列组成,但不限于此。

[0053] 在变体的一个实施方案中,SEQ ID NO:52的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:61或62的氨基酸序列,基本上由SEQ ID NO:61或62的氨基酸序列组成,或由SEQ ID NO:61或62的氨基酸序列组成,但不限于此。

[0054] 在变体的一个实施方案中,SEQ ID NO:53的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:63或64的氨基酸序列,基本上由SEQ ID NO:63或64的氨基酸序列组成,或由SEQ ID NO:63或64的氨基酸序列组成,但不限于此。

[0055] 在变体的一个实施方案中,SEQ ID NO:54的氨基酸序列中对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代的变体可以包含SEQ ID NO:65或66的氨基酸序列,基本上由SEQ ID NO:65或66的氨基酸序列组成,或由SEQ ID NO:65或66的氨基酸序列组成,但不限于此。

[0056] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:31和49至54中任一氨基酸序列的第12位和/或第116位的位置被其它氨基酸取代;可以与SEQ ID

N0:31和49至54中任一氨基酸序列具有60%,65%,66%,67%,68%,69%,70%,71%,72%,73%,74%,75%,76%,77%,78%,79%,80%,81%,82%,83%,84%,85%,86%,87%,88%,89%,90%,91%,92%,93%,94%,95%,96%,97%,98%,或99%或更多,且小于100%的序列同源性;并且可以具有丝氨酸蛋白酶活性。

[0057] 本申请的丝氨酸蛋白酶变体与修饰前的多肽、天然野生型多肽或未修饰的多肽相比,可具有增强的活性,但不限于此。此外能够清楚理解的是,在部分氨基酸序列具有缺失、修饰、取代和添加的氨基酸序列的任何蛋白质都在本申请的范围内,只要该蛋白质具有上述同源性以及与所述蛋白等同的效果。

[0058] 此外能够清楚理解的是,除了在第12位氨基酸和/或第116位氨基酸中的突变或在其相应位置的突变之外,在相应的SEQ ID NO的氨基酸序列的正向或反向中添加无义序列、或具有天然存在的突变、或具有沉默突变的任何变体不被排除在本申请的范围之外,只要该变体具有与根据本申请的变体相同或等同的活性,也包括在本申请的范围内。

[0059] 同时,NCBI参考序列WP_016188200.1 (SEQ ID NO:40)的成熟区对应于本申请的SEQ ID NO:31的氨基酸序列,并且SEQ ID NO:40除信号肽之外的序列对应于本申请的SEQ ID NO:2。

[0060] 如上所述能够清楚理解的是,本申请的丝氨酸蛋白酶变体可以包括对丝氨酸蛋白酶的性质和二级结构具有很小影响的氨基酸的缺失或添加,其中在对应于SEQ ID NO:31的第12位和/或第116位氨基酸的位置上的氨基酸被不同的氨基酸取代。此外,通过本领域公知的序列比对,本领域技术人员能够清楚理解的是,本申请的SEQ ID NO:31的自N末端的第12位和第116位对应于SEQ ID NO:40的第193位和第297位以及SEQ ID NO:2的第163位和第267位,并且SEQ ID NO:31包括在SEQ ID NO:40和SEQ ID NO:2中。

[0061] 因此,关于各自包括SEQ ID NO:31的氨基酸序列的SEQ ID NO:2和40的氨基酸序列,本申请的丝氨酸蛋白酶变体包括这样的变体:在所述变体中,对应于SEQ ID NO:31的第12位和第116位的氨基酸(SEQ ID NO:2中的第163位和/或第267位氨基酸以及SEQ ID NO:40中的第193位和/或第297位氨基酸)的位置处的氨基酸各自被取代。另外,上文关于SEQ ID NO:31及第12位和第116位氨基酸给出的描述也适用于SEQ ID NO:2及其第163位和第267位氨基酸和SEQ ID NO:40及其第193位和第297位氨基酸。

[0062] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括这样的氨基酸序列:在对应于SEQ ID NO:31的第12位和第116位的位置处的氨基酸被其它氨基酸取代,并且可以与SEQ ID NO:2具有至少60%且小于100%(例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高)的序列同源性。在另一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括SEQ ID NO:2的第163位氨基酸和/或第266位氨基酸被其它氨基酸取代,并且可以与SEQ ID NO:2具有至少60%且小于100%的序列同源性,并且可以与SEQ ID NO:3至10中任一氨基酸序列具有至少60%的序列同源性,但不限于此。

[0063] 同时能够清楚理解的是,在包括SEQ ID NO:49至54中任一氨基酸序列的多肽中包括对应于SEQ ID NO:49至54的自N末端起的第12位和/或第116位的氨基酸的取代的变体也包括在本申请的丝氨酸蛋白酶的范围内。

[0064] 包含SEQ ID NO:49至54中任一氨基酸序列的多肽的序列可以是,例如,在GenBank登录号KUP96625.1中描述的氨基酸序列(SEQ ID NO:67),NCBI参考序列WP_068687914.1(SEQ ID NO:68),NCBI参考序列WP_133739400.1(SEQ ID NO:69),NCBI参考序列WP_179641868.1(SEQ ID NO:70),NCBI参考序列WP_184391208.1(SEQ ID NO:71),NCBI参考序列WP_017594871.1(SEQ ID NO:72)等。

[0065] 本领域技术人员可以通过本领域已知的序列比对,鉴定对应于SEQ ID NO:67-72中SEQ ID NO:49-54的自N末端起的第12位和/或第116位的氨基酸,并应用SEQ ID NO:49-54的自N末端起的第12位和/或第116位的描述。

[0066] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括用其它氨基酸取代对应于SEQ ID NO:49至54中任一氨基酸序列的自N末端起的第12位和/或第116位的氨基酸,并且可以与SEQ ID NO:67至72中任一氨基酸序列具有至少60%(例如61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高)的同源性或同一性。

[0067] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括SEQ ID NO:54所示的氨基酸序列中对应于第12位的氨基酸的取代,并且可以与SEQ ID NO:70至72任一项所示的氨基酸序列具有至少60%(例如61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高)的同源性或同一性。作为一个实例,丝氨酸蛋白酶变体可以进一步包括SEQ ID NO:54所示的氨基酸序列中对应于第116位的氨基酸的取代。

[0068] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:67所示的氨基酸序列第198位的氨基酸进行取代。变体可以与SEQ ID NO:67具有至少60%的同源性或同一性,例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高。例如,变体可以包括氨基酸序列,其中在对应于SEQ ID NO:49的第12位氨基酸的位置上的氨基酸被另一个氨基酸取代,并且与SEQ ID NO:49具有至少70%的序列同一性。作为实例,丝氨酸蛋白酶变体可以进一步包括对应于SEQ ID NO:67所示的氨基酸序列的第302位的氨基酸进行取代。

[0069] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:68所示的氨基酸序列的第178位的氨基酸进行取代。变体可以与SEQ ID NO:68具有至少60%的同源性或同一性,例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高。例如,变体可以包括氨基酸序列,其中在对应于SEQ ID NO:50的第12位氨基酸的位置上的氨基酸被另一个氨基酸取代,并且与SEQ ID NO:50具有至少70%的序列同一性。作为实例,丝氨酸蛋白酶变体可以进一步包括对应于SEQ ID NO:68所示的氨基酸序列的第282位的氨基酸进行取代。

[0070] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:69所示的氨基酸序列的第207位的氨基酸进行取代。变体可以与SEQ ID NO:69具有至少60%的同源性或同一性,例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高。例如,变体可以包括氨基酸序列,其中在对应于SEQ ID NO:51的第12位氨基酸的位置上的氨基酸被另一个氨基酸取代,并且与SEQ ID NO:51具有至少70%的序列同一性。作为实例,丝氨酸蛋白酶变体可以进一步包括对应于SEQ ID NO:69所示的氨基酸序列的第311位的氨基酸进行取代。

[0071] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:70所示的氨基酸序列的第203位的氨基酸进行取代。变体可以与SEQ ID NO:70具有至少60%的同源性或同一性,例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高。例如,变体可以包括氨基酸序列,其中在对应于SEQ ID NO:52的第12位氨基酸的位置上的氨基酸被另一个氨基酸取代,并且与SEQ ID NO:52具有至少70%的序列同一性。作为实例,丝氨酸蛋白酶变体可以进一步包括对应于SEQ ID NO:70所示的氨基酸序列的第303位的氨基酸进行取代。

[0072] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:71所示的氨基酸序列的第201位的氨基酸进行取代。变体可以与SEQ ID NO:71具有至少60%的同源性或同一性,例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高。例如,变体可以包括氨基酸序列,其中在对应于SEQ ID NO:53的第12位氨基酸的位置上的氨基酸被另一个氨基酸取代,并且与SEQ ID NO:53具有至少70%的序列同一性。作为实例,丝氨酸蛋白酶变体可以进一步包括对应于SEQ ID NO:71所示的氨基酸序列的第304位的氨基酸进行取代。

[0073] 在一个实施方案中,本申请的丝氨酸蛋白酶变体可以包括对应于SEQ ID NO:72所示的氨基酸序列的第201位的氨基酸进行取代。变体可以与SEQ ID NO:72具有至少60%的同源性或同一性,例如60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%或更高。例如,变体可以包括氨基酸序列,其中在对应于SEQ ID NO:54的第12位氨基酸的位置上的氨基酸被另一个氨基酸取代,并且与SEQ ID NO:54具有至少70%的序列同一性。作为实例,丝氨酸蛋白酶变体可以进一步包括对应于SEQ ID NO:72所示的氨基酸序列的第304位的氨基酸进行取代。

[0074] 然而,本申请的丝氨酸蛋白酶变体不限于上述那些。

[0075] 本文所用术语“对应于”是指蛋白或多肽中所述位置处的氨基酸残基,或与蛋白或多肽中所述位置相似、相同或同源的氨基酸残基。鉴定对应位置的氨基酸可以是确定参考

特定序列的序列的特定氨基酸。本文所用术语“对应区域”通常是指相关蛋白或参照蛋白中类似或对应的位置。例如,任何氨基酸序列可以与SEQ ID NO:31比对,并且基于此,氨基酸序列的每个氨基酸残基可以参考对应于SEQ ID NO:31的氨基酸残基的氨基酸残基的数字位置来编号。例如,序列比对算法(如本申请中描述的算法)可以鉴定氨基酸的位置或与查询序列(也称为“参考序列”)相比发生修饰(如取代、插入或缺失)的位置。

[0076] 对于这种比对,可使用Needleman-Wunsch算法(Needleman and Wunsch,1970, J.Mol.Biol.48:443-453),EMBOSS程序包(The European Molecular Biology Open Software Suite,Rice et al.,2000,Trends Genet.16:276-277)的Needle程序等,但比对程序不限于此。可以通过多序列比对来鉴定对应的氨基酸残基。本领域已知的多序列比对程序的实例包括,MUSCLE(通过对数期望进行多序列比较;版本3.5或更高;Edgar,2004, Nucleic Acids Research 32:1792-1797)和MAFFT(6.857版或更高版本;Katoh and Kuma, 2002,Nucleic Acids Research 30:3059-3066;Katoh et al.,2005,Nucleic Acids Research 33:511-518;Katoh and Toh,2007,Bioinformatics 23:372-374;Katoh et al.,2009,Methods in Molecular Biology 537:39-64;Katoh and Toh,2010, Bioinformatics 26:1899-1900)和EMBOSS EMMA,使用ClustalW(1.83或更高);Thompson et al.,1994,Nucleic Acids Research 22:4673-4680),并且可以使用每个程序的基本参数,但是程序不限于此。

[0077] 本申请的另一方面提供了编码丝氨酸蛋白酶变体的多核苷酸。

[0078] 本文所用术语“多核苷酸”是指核苷酸的聚合物,其中核苷酸单体通过共价键以长链形状彼此连接,并且通常是指具有一定或更长长度的DNA或RNA链,更具体而言,是指编码变体的多核苷酸片段。

[0079] 编码本申请的丝氨酸蛋白酶变体的多核苷酸可以包括编码根据本申请的具有增强的活性的丝氨酸蛋白酶变体的任何多核苷酸序列,但不限于此。在一个实施方案中,编码本申请中的野生型丝氨酸蛋白酶的基因可以来源于Thermobifida属,拟诺卡氏菌(Nocardiopsis)属,Actinorugispora属或Spinactinospira属的微生物,具体而言,可以来源于Thermobifida fusca,Thermobifida celulosilytica,Thermobifida halotolerans, Actinorugispora endohytica,Spinactinospira alkalitolerans,堆肥拟诺卡氏菌(Nocardiopsis composta)或Nocardiopsis potens,但不限于此。

[0080] 在不改变氨基酸序列的范围内,本申请的多核苷酸可以包括在多肽的氨基酸序列中的编码区中进行的各种修饰,这是由于密码子简并性或考虑了多肽将表达于的活生物体的偏好密码子。具体而言,可以包括编码变体的任何多核苷酸序列,在变体中,在对应于SEQ ID NO:31和49至54的任一氨基酸序列的自N末端的第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其他氨基酸取代,但不限于此。

[0081] 例如,本申请的多核苷酸可以是编码本申请的变体的多核苷酸序列,具体而言,变体是由SEQ ID NO:32至39和55至66中任一氨基酸序列组成的多肽,或者与所述多肽具有一定同源性的多肽,但不限于此。

[0082] 在一个实施方案中,编码由SEQ ID NO:32至39中任一氨基酸序列组成的多肽的多核苷酸序列可以由SEQ ID NO:41至48中任一多核苷酸序列组成,但不限于此。

[0083] 如上文所述,本申请的丝氨酸蛋白酶变体包括这样的变体,其中在包含SEQ ID

N0:31和49至54中任一氨基酸序列的多肽中,对应于序列第12位和/或第116位的位置的氨基酸被取代,因此能够清楚理解的是,编码这种丝氨酸蛋白酶变体的多核苷酸序列也包括在本申请的范围内。

[0084] 作为实例,在SEQ ID N0:2和SEQ ID N0:40中,在对应于SEQ ID N0:31的第12位和第116位的位置上的氨基酸(SEQ ID N0:2的第163位氨基酸和/或第267位氨基酸以及SEQ ID N0:40的第193位氨基酸和/或第297位氨基酸)被取代的变体也包括在本申请的丝氨酸蛋白酶的范围。因此,编码这些变体的多核苷酸序列也包括在本申请的范围内。例如,编码丝氨酸蛋白酶变体的多核苷酸序列可以是编码SEQ ID N0:3-10中任一氨基酸序列的序列,具体而言,是由SEQ ID N0:23-30中任一多核苷酸序列组成的序列,但不限于此。

[0085] 另外,可以包括编码具有变体活性的蛋白质的任何序列,其中在严格条件下,通过与可以由已知基因序列(例如,与核苷酸序列的全部或部分互补的序列)制备的探针杂交,在所述变体中SEQ ID N0:31和49至54中任一个的自N末端起的对应于第12位氨基酸和/或第116位氨基酸的位置上的氨基酸被其它氨基酸取代,但不限于此。

[0086] 术语“严格条件”是指能够在多核苷酸之间进行特异性杂交的条件。这样的条件具体公开于文献中。例如,严格条件可以包括如下所述的条件,具有高同源性的基因(至少40%、特别是至少90%、更特别是至少95%、还更特别是至少97%、并且甚至还更特别是至少99%的同源性的基因)实现杂交,但具有比上述同源性更低的同源性的基因不实现杂交的条件;或者用于Southern杂交的典型洗涤条件,即,洗涤在对应于60℃、1×SSC和0.1% SDS,特别是60℃、0.1×SSC和0.1% SDS,更特别是68℃、0.1×SSC和0.1% SDS的盐浓度和温度下进行一次,特别是两次或三次。然而,严格条件不限于此,而是可以由本领域技术人员根据其目的适当地调节。

[0087] 杂交需要两个核苷酸具有互补序列,但是根据杂交严格性,碱基之间的错配是可能的。术语“互补”用于描述可彼此杂交的核苷酸碱基之间的关系。例如,在DNA中,腺嘌呤与胸腺嘧啶互补,胞嘧啶与鸟嘌呤互补。因此,本申请不仅可以包括基本上相似的核苷酸序列,而且可以包括与整个序列互补的分离的多核苷酸片段。

[0088] 具体而言,具有同源性的多核苷酸可以在55℃的T_m值下通过使用上述杂交条件来检测。此外,T_m值可以是60℃、63℃或65℃,但不限于此,并且可以由本领域技术人员根据目的适当地控制。

[0089] 多核苷酸杂交的适当严格程度可以取决于多核苷酸的长度和互补程度,参数是本领域熟知的。

[0090] 本申请的又一方面提供了包含编码本申请的丝氨酸蛋白酶变体的多核苷酸的载体。

[0091] 本文所用术语“载体”是指含有编码目标蛋白的多核苷酸的核苷酸序列的DNA构建体,所述编码目标蛋白的多核苷酸与适当的控制序列可操作地连接,以在适当的宿主细胞中表达目标多核苷酸。控制序列可以包括能够启动转录的启动子,用于控制转录的任何操纵子序列,编码适当的mRNA核糖体结合位点的序列,和用于控制转录和翻译终止的序列。在转化入适当的宿主细胞后,载体可以独立于宿主的基因组复制或起作用,或者可以整合到基因组本身中。

[0092] 本文所用术语“可操作地连接”是指编码本申请的目标蛋白的多核苷酸序列与启

动和介导多核苷酸转录的启动子序列功能性连接。可操作的连接可以通过本领域已知的遗传重组技术制备，位点特异性DNA切割和连接可以使用本领域已知的限制酶、连接酶等制备，但不限于此。

[0093] 本申请中使用的载体没有特别限制，可以使用本领域已知的任何载体。常用载体的实例可包括天然或重组形式的质粒、粘粒、病毒和噬菌体。例如，pWE15、M13、MBL3、MBL4、IXII、ASHII、APII、t10、t11、Charon4A、Charon21A等可以用作噬菌体载体或粘粒载体，并且基于pBR、基于pUC、基于pBluescriptII、基于pGEM、基于pTZ、基于pCL、基于pET、基于pUB110的载体可以用作质粒载体。具体而言，可以使用pDZ、pACYC177、pACYC184、pCL、pECCG117、pUC19、pBR322、pMW118、pCC1BAC和pSM704等载体。可用于本申请的载体没有特别限制，而是可以使用任何已知的表达载体。

[0094] 在一个实施方案中，使用用于细胞中染色体插入的载体，可以用突变的多核苷酸取代编码染色体中的变体目标的多核苷酸。将多核苷酸插入染色体可以通过本领域已知的任何方法进行，例如同源重组，但不限于此。载体可以进一步包括选择标记，以确认染色体插入。选择标记用于选择被载体转化的细胞，即，确认是否插入目标核酸分子。选择标记的实例可以包括提供选择性表型的标记，例如耐药性，营养缺陷型，对细胞毒性剂的抗性，或表面突变多肽的表达。在被选择性试剂处理的环境中，只有表达选择标记的细胞可以存活或显示不同的表型，因此可以选择被转化的细胞。

[0095] 本申请的又一方面提供包含本申请的丝氨酸蛋白酶变体、编码所述变体的多核苷酸和包含所述多核苷酸的载体中的至少一者的宿主细胞。

[0096] 特别地，宿主细胞可以是微生物。

[0097] 包含丝氨酸蛋白酶变体、编码该变体的多核苷酸和包含该多核苷酸的载体中的至少一者的微生物可以特别是通过用包含编码变体的多核苷酸的载体转化而制备的微生物，但不限于此。

[0098] 微生物可以是表达丝氨酸蛋白酶变体的微生物。

[0099] 本文所用术语“待表达/所表达/表达”蛋白质是指将目标蛋白引入微生物或在微生物中表达的状态。考虑到本申请的目的，“目标蛋白”可以是上述丝氨酸蛋白酶变体。

[0100] 具体而言，术语“引入蛋白质”可以指在最初不具有具蛋白质的微生物中表现出该具体蛋白质的活性，或者与该蛋白质的内源性活性或修饰前的活性相比，表现出该蛋白质的增强活性。例如，蛋白质的引入可以指将编码具体蛋白质的多核苷酸引入微生物的染色体，或将包含编码具体蛋白质的多核苷酸的载体引入微生物，从而显示蛋白质的活性。

[0101] 微生物可以是重组微生物。重组可以通过遗传修饰(如转化)来实现。

[0102] 本文所用术语“转化”是指将含有编码目标蛋白的多核苷酸的载体引入宿主细胞，以允许所述多核苷酸编码的蛋白在宿主细胞中表达的过程。无论被转化的多核苷酸是插入宿主细胞的染色体中的形式还是位于染色体外的形式，两种形式的被转化的多核苷酸都属于本申请的范围内，只要被转化的多核苷酸在宿主细胞中表达。此外，多核苷酸包括编码目标蛋白的DNA和RNA。多核苷酸可以以任何形式引入宿主细胞，只要多核苷酸被引入宿主细胞中并在其中表达。例如，多核苷酸可以以表达盒的形式引入宿主细胞，表达盒是含有自身复制所需的所有必需元件的基因构建体。表达盒通常可以包括与多核苷酸可操作地连接的启动子，转录终止信号，核糖体结合位点和翻译终止信号。表达盒可以是能够自我复制的表

达载体。另外,多核苷酸可以以其原始形式被引入到宿主细胞中,并且可操作地连接到在宿主细胞中表达所需的序列,而不限于此。用于转化的方法包括用于将多核苷酸引入细胞的任何方法,并且可以通过本领域已知的合适的标准技术进行。例如,转化方法包括电穿孔法、磷酸钙($\text{Ca}(\text{H}_2\text{PO}_4)_2$ 、 CaHPO_4 或 $\text{Ca}_3(\text{PO}_4)_2$)沉淀法,氯化钙(CaCl_2)沉淀法,显微注射法,聚乙二醇(PEG)法,DEAE-葡聚糖法,阳离子脂质体法,天然感受(例如,参见Perry and Kuramitsu,1981,Infect.Immun.32:1295-1297]),和乙酸锂-DMSO方法,但不限于此。

[0103] 重组微生物可以是本申请的丝氨酸蛋白酶的活性得到增强的微生物。

[0104] “活性增强”可以指微生物的具体蛋白质的活性与内源性活性或修饰前的活性相比得到增强。术语“内源活性”可以指当微生物由天然或人工因素引起的遗传突变转化时,转化前微生物的亲本菌株所具有的具体蛋白质的活性。

[0105] 具体而言,本申请中的蛋白质变体的活性的增强可以通过以下至少一种方法实现:增加编码蛋白质变体的基因的细胞内拷贝数的方法,将突变引入编码蛋白质变体的基因的表达控制序列的方法,用具有更强活性的序列替换编码蛋白质变体的基因的表达控制序列的方法,用编码蛋白变体的基因替换编码具有丝氨酸蛋白酶活性的野生型蛋白的染色体基因的方法,和进一步将突变引入编码蛋白变体的基因中,以增强蛋白变体的活性的方法,但不限于此。

[0106] 接下来,通过缺失、插入、非保守取代、保守取代或上述方式的组合,在核酸序列中引入突变,以进一步增强表达控制序列的活性,或者通过用具有更强活性的核酸序列替换核酸序列,可以进行增加多核苷酸表达的表达控制序列的修饰,但不限于此。表达控制序列可以包括启动子、操纵子序列、核糖体结合位点编码序列、用于调节转录和翻译的序列等,但不限于此。

[0107] 用于替换内源启动子的强启动子可以连接在多核苷酸表达单元的上游,但启动子不限于此。已知强启动子的实例可包括c j1至c j7启动子(US7662943B2),lac启动子, trp启动子, trc启动子, tac启动子, λ 噬菌体PR启动子, P_L 启动子, tet启动子, gapA启动子, SPL7启动子, SPL13(sm3)启动子(US 10584338 B2), O2启动子(US 10273491 B2), tkt启动子, yccA启动子等,但不限于此。

[0108] 此外,染色体上多核苷酸序列的修饰可以通过核酸序列的缺失、插入、非保守取代、保守取代或以上方式的组合在表达控制序列中引入突变,以进一步增强多核苷酸序列的活性,或者通过用修饰为具有更强活性的多核苷酸序列替换该序列而进行,但不限于此。

[0109] 通常,蛋白质活性的引入和增强可以将对应蛋白质的活性或浓度从野生型或未修饰的微生物株中的活性或浓度增加1%、10%、25%、50%、75%、100%、150%、200%、300%、400%或500%至最大1000%或2000%,但不限于此。

[0110] 根据本申请的宿主细胞或微生物可以是通过包含本申请的多核苷酸或本申请的载体而表达丝氨酸蛋白酶变体的任何微生物。具体而言,宿主细胞或微生物的实例可包括属于埃希氏菌属(*Escherichia*),沙雷氏菌属(*Serratia*),欧文氏菌属(*Erwinia*),肠细菌属(*Enterobacteria*),普罗维登斯菌属(*Providencia*),沙门氏菌属(*Salmonella*),链霉菌属(*Streptomyces*),假单胞菌属(*Pseudomonas*),短杆菌属(*Brevibacterium*),棒状杆菌属(*Corynebacterium*)或芽孢杆菌属(*Bacillus*等)的微生物菌株,具体而言,宿主细胞或微生物可以是枯草芽孢杆菌(*Bacillus subtilis*)、地衣芽孢杆菌(*Bacillus licheniformis*),解淀

粉芽孢杆菌 (*Bacillus amyloliquefaciens*), 韦利茨芽孢杆菌 (*Bacillus velezensis*), 大肠杆菌 (*Escherichia coli*), 谷氨酸棒杆菌 (*Corynebacterium glutamicum*) 或米曲霉 (*Aspergillus oryzae*) 的菌株, 更具体而言, 可以是枯草芽孢杆菌 (*Bacillus subtilis*), 但不限于此。

[0111] 本申请的又一方面提供了制备本申请的丝氨酸蛋白酶变体的方法。

[0112] 制备本申请的变体的方法可以包括培养包含本申请的丝氨酸蛋白酶变体、编码所述变体的多核苷酸和包含所述多核苷酸的载体中的至少一者的微生物的步骤。

[0113] 本文所用术语“培养”是指在适当调节的环境条件下使宿主细胞生长。本申请中的培养方法可以在本领域已知的适当培养条件下使用适当的培养基进行。本领域技术人员可以根据所选择的菌株容易地调节和使用这种培养方法。具体而言, 培养可以是分批培养、连续培养和补料分批培养, 但不限于此。

[0114] 本文所用术语“培养基”是指培养宿主细胞所需的营养物作为主要成分混合并供给存活和生长必需的营养素和生长因子(包括水)的物质。具体而言, 作为用于培养本申请的宿主细胞的培养基和其它培养条件, 可以使用任何培养基而没有特别限制, 只要其是用于培养宿主细胞的常规培养基, 但是本申请的宿主细胞可以在有氧条件下、在含有适当碳源、氮源、磷源、无机化合物、氨基酸和/或维生素等的常规培养基中培养, 同时控制温度、pH 等。

[0115] 在一个实施方案中, 制备本申请的变体的方法可以进一步包括回收在培养步骤中表达的本申请的变体的步骤。

[0116] 在另一个实施方案中, 可以使用本申请所属领域已知的方法回收在培养步骤中表达的变体。例如, 可以通过常规方法从营养培养基中回收变体, 所述常规方法包括但不限于收集, 离心, 过滤, 提取, 喷雾干燥, 蒸发或沉淀。

[0117] 回收方法可以是使用本领域已知的合适方法收集变体, 方法取决于本申请的宿主细胞的培养方法, 例如分批、连续或补料分批培养方法。例如, 可以使用离心、过滤、用结晶蛋白沉淀剂处理(盐析方法)、提取、超声处理、超滤、透析、各种色谱(例如分子筛色谱(凝胶过滤)、吸附色谱、离子交换色谱和亲和色谱)、HPLC, 以及这些方法的组合, 并且可以使用本领域已知的合适方法从培养基或宿主细胞中回收变体。

[0118] 在另一个实施方案中, 可以不回收在培养步骤中由宿主细胞表达的变体。在该实施方案中, 表达变体的宿主细胞本身可以用作变体的来源。

[0119] 本申请的又一方面提供了包含本申请的丝氨酸蛋白酶变体和表达丝氨酸蛋白酶变体的微生物中至少一者的饲料组合物。

[0120] 包含在本申请变体的饲料组合物中的丝氨酸蛋白酶变体可以以这样的方式包含在饲料组合物中: 表达丝氨酸蛋白酶变体的微生物包含在饲料组合物中, 或者可以是表达丝氨酸蛋白酶变体的微生物中分离并纯化的形式, 但是本申请不限于此。

[0121] 本文所用术语“饲料组合物”是指任何天然或人造食物、膳食等的任何制备物, 或用于动物食用、摄入和消化或适合的膳食组分, 并且饲料可以以本领域已知的各种形式制备。

[0122] 饲料组合物可以是饲料添加剂。

[0123] 饲料的类型没有特别限制, 并且可以使用本领域中常用的饲料。饲料的非限制性

实例可以包括：蔬菜饲料，例如谷物、根/果实、食品加工副产物、藻类、纤维、药物副产物、油和脂肪、淀粉、瓜类、以及谷物副产物；和动物饲料，例如蛋白质，无机材料，油和脂肪，矿物质，单细胞蛋白质，动物浮游生物或食品。这些饲料可以单独使用或至少两种组合使用。

[0124] 本申请的饲料组合物可以进一步包括选自以下的至少一种：有机酸，例如柠檬酸、富马酸、己二酸、乳酸和苹果酸；磷酸盐，例如磷酸钠，磷酸钾，酸式焦磷酸盐和多磷酸盐（聚合磷酸盐）；以及天然抗氧化剂，例如多酚，儿茶素， α -生育酚，迷迭香提取物，维生素C，绿茶提取物，甘草提取物，壳聚糖，鞣酸和植酸。

[0125] 本申请的饲料组合物可以进一步包括选自以下的至少一种：辅助物质，例如氨基酸、矿物质、维生素、抗生素、抗菌物质、抗氧化剂、抗真菌剂和不同益生菌形式的微生物制剂；谷物，（例如，粉末或粉碎的小麦、燕麦、大麦、玉米和大米）；植物蛋白饲料，包括油菜、菜豆、向日葵为主要成分；动物蛋白饲料，例如血粉，肉粉，骨粉和鱼粉；糖和乳制品（例如，由各种类型的干奶粉和乳清粉形成的干组分；脂质（例如，活性成分，诸如通过加热液化的任何动物脂肪和植物油）；以及添加剂，例如营养补充剂，消化和吸收促进剂，生长促进剂和预防剂。

[0126] 本申请的饲料组合物可以是干燥或液体制剂的形式，并且可以进一步包括用于饲料的赋形剂。用于饲料的赋形剂可以是，例如，沸石，玉米粉，米糠等，但不限于此。

[0127] 除了丝氨酸蛋白酶变体之外，本申请的饲料组合物还可以包括酶制剂。例如，饲料组合物可进一步包括至少一种选自以下的酶：脂质降解酶，如脂肪酶，将植酸降解成磷酸盐和磷酸肌醇的植酸酶，催化淀粉、糖原等中含有的 α -1,4-糖苷键水解的淀粉酶，催化有机磷酸酯水解的磷酸酶，催化麦芽糖生成两个葡萄糖分子的麦芽酶，以及催化蔗糖水解为葡萄糖-果糖混合物的转化酶。然而，本申请的饲料组合物不限于此。

[0128] 本申请的饲料组合物可以单独施用于动物，或与包含在可食用载体中的其它饲料添加剂组合施用于动物。此外，饲料组合物可以作为饲料添加剂或追加饲料容易地施用，直接施用到家畜饲料中或与饲料分开施用，或在单独的口服制剂中施用或与其它成分组合施用。此外，每日剂量可通过如本领域中通常已知的每日一次剂量或每日多次剂量来使用。

[0129] 施用本申请的饲料组合物的动物的实例可以包括家畜，例如肉牛、奶牛、小牛、猪、仔猪、绵羊、山羊、马、兔、狗和猫；和家禽，例如小鸡、母鸡、家鸡、公鸡、鸭、鹅、火鸡、鹌鹑和小禽，但不限于此。

[0130] 包括在本申请的饲料组合物中的丝氨酸蛋白酶变体的量没有特别限制，并且可以根据目的适当地调节。在一个实施方案中，可以以适当的量包含丝氨酸蛋白酶变体，用于降解蛋白质源物质，同时在家畜的消化道中长期存有，如本申请所属领域中通常已知的，但丝氨酸蛋白酶变体的量不限于此。

[0131] 本申请的又一方面提供了包含本申请的丝氨酸蛋白酶变体和表达丝氨酸蛋白酶变体的微生物中的至少一种的食物组合物。丝氨酸蛋白酶变体可用于液体或固体食物组合物中。此外，食品可以是粉剂、丸剂、饮料、茶或普通食品的添加剂。

[0132] 在一个实施方案中，食物可以是一组需要蛋白酶的食物，例如乳制品，用于改善肠运动和体重减轻的健康功能食物，和用于预防高血压的健康功能食物。

[0133] 在另一个实施方案中，丝氨酸蛋白酶变体可以作为食品增溶剂、食品软化剂和肉质改良剂包含在各种食品组合物中。在各种其它实施方案中，丝氨酸蛋白酶变体可以在麸

质网络分解的步骤中加入到焙烤混合物中。或者,丝氨酸蛋白酶变体可用于催化食物蛋白(例如乳蛋白)的水解。或者,为了提供、制备调味剂,降低苦味,改变乳化性,产生生物活性肽,降低蛋白质中引起变态反应的抗原,丝氨酸蛋白酶变体可以包含在各种食品组合物中。然而,这仅仅是说明性的实施方案,丝氨酸蛋白酶变体的用途不限于此。

[0134] 本申请的食物组合物中包含的丝氨酸蛋白酶变体的量可以由本领域技术人员根据目的适当地调节。

[0135] 本申请的又一方面提供包含本申请的丝氨酸蛋白酶变体和表达丝氨酸蛋白酶变体的微生物中的至少一种的除垢剂组合物。

[0136] 本申请的除垢剂组合物可以是第一和第二水性除垢剂组合物、非水性液体除垢剂组合物、流延固体、粗粒、颗粒、压缩片剂、凝胶、糊剂或浆料的形式。除垢剂组合物可用于除去顽固的食物污渍、食物残渣膜和其它少量的食物组合物。

[0137] 根据本申请的除垢剂组合物可以以用于清洁硬表面的除垢剂组合物,用于清洁织物的除垢剂组合物,用于洗碗的除垢剂组合物,用于口腔清洁的除垢剂组合物,用于清洁义齿的除垢剂或隐形眼镜清洁溶液的形式提供。然而,本申请的除垢剂组合物不限于此。

[0138] 本申请的又一方面提供包含本申请的丝氨酸蛋白酶变体和表达丝氨酸蛋白酶变体的微生物中的至少一种的药物组合物。

[0139] 本申请的药物组合物可用作消化酶的药物组合物,以改善消化疾病、消化障碍和消化道手术后异常,直接施用于血凝块以溶解纤维蛋白的溶栓或抗血栓组合物,用作体内防御系统以去除炎性物质或坏死组织的抗炎药物,或用于减轻手术或伤口后的水肿的抗炎药物。

[0140] 根据使用的方法或目的,药物组合物可以进一步包括药学上可接受的或营养学可接受的载体、赋形剂、稀释剂或辅助成分。载体、赋形剂或稀释剂可包括选自乳糖、右旋糖、蔗糖、山梨糖醇、甘露糖醇、木糖醇、赤藓糖醇、麦芽糖醇、淀粉、阿拉伯胶、藻酸盐、明胶、磷酸钙、硅酸钙、纤维素、甲基纤维素、微晶纤维素、聚乙烯吡咯烷酮、水、羟基苯甲酸甲酯、羟基苯甲酸丙酯、滑石、硬脂酸镁、矿物油、糊精、碳酸钙、丙二醇、液体石蜡和盐水中的至少一种,但不限于此。

[0141] 除了上述用途之外,本申请的丝氨酸蛋白酶变体或表达丝氨酸蛋白酶变体的微生物还可以用于生产化妆品、加工皮革、制备药物、制备诊断剂、处理废物、制备用于学术研究的化学品的目的。然而,这仅仅是示例性地描述了目的,而丝氨酸蛋白酶变体也可以用于本领域已知的变性、降解或去除蛋白类物质的任何其它目的。

实施例

[0142] 在下文中,将参考以下实施例和实验例更详细地描述本申请。然而,这些实施例和实验例仅用于说明目的,而不旨在限制本申请的范围。

[0143] 实施例1.源自*Thermobifida fusca*的丝氨酸蛋白酶变体的筛选

[0144] 实施例1-1:源自*Thermobifida fusca*的丝氨酸蛋白酶文库的制备

[0145] 通过易错PCR将随机突变引入编码对应于源自*Thermobifida fusca*的丝氨酸蛋白酶的成熟区的氨基酸(SEQ ID NO:31)的基因中。使用DiversifyTM PCR Random Mutagenesis Kit(Clontech,货号630703)实施易错PCR,并且使用的PCR条件如下表1中所

述。证实突变以6.2个突变/kb的频率被引入。

[0146] 表1

组成	量	注释
模板 DNA	0.5 ng	SEQ ID NO: 1
正向引物	5 μ M	SEQ ID NO: 11
反向引物	5 μ M	SEQ ID NO: 12
10 x Titanium Taq 缓冲液	5 μ L	
MnSO ₄	640 μ M	
dGTP	40 μ M	
50 x Diversify dNTP 混合物	1 μ L	
50 x dNTP 混合物	1 μ L	
Titanium Taq 聚合物	1 μ L	
总计	50 μ L	用蒸馏水调至 50 μ L

[0147]

[0148]

PCR 条件		
I. 94°C	30 秒钟	重复 50 个循环的 II 和 III
II. 94°C	30 秒钟	
III. 68°C	1 分钟	
IV. 68°C	1 分钟	
V. 4°C	∞	

[0149] 使用下表2中所示的引物扩增,使用In-FusionR HD克隆试剂盒(Clontech),将在上述过程中获得的PCR片段连接到载体上,并转化到DH5 α 细胞中以获得菌落。纯化所得菌落中的质粒,以获得大小为约 5×10^4 的文库。

[0150] 表2

[0151]

模板DNA (pBE-S-TAP)	SEQ ID NO:1
正向引物	SEQ ID NO:13
反向引物	SEQ ID NO:14

[0152] 实施例1-2:源自Thermobifida fusca的丝氨酸蛋白酶文库的筛选

[0153] 用实施例1-1中制备的蛋白酶文库转化容易释放蛋白质的枯草芽孢杆菌LB700菌株并进行筛选。筛选通过两阶段方法进行。在第一阶段,将用文库转化的枯草芽孢杆菌菌株接种在2%脱脂奶平板上,并根据晕圈大小选择所希望的菌落。根据Groningen方法进行枯草芽孢杆菌的转化,筛选中使用的脱脂奶的组成如下表3所示。

[0154] 表3

	组成	量	注释
[0155]	M9 少盐培养基	11.28g	BD, 货号 248510
	琼脂	20 g	
	脱脂奶	20 g	BD, 货号 232100
	50%葡萄糖	20 g	
[0156]	1M MgSO ₄	2 mM	
	1M CaCl ₂	0.1 mM	
	卡那霉素	50μg/mL	

[0157] (每1L)

[0158] 第二阶段涉及通过偶氮酪蛋白显色来重新选择在第一阶段中选择的菌落的方法。将含有卡那霉素抗生素的脑心灌注 (BHI, bd, 货号53286) 液体培养基添加到96深孔板中, 并且将在第一阶段中选择的菌落接种到其中, 随后在37℃下孵育20至24小时。孵育后, 通过离心获得包含酶的上清液, 并将上清液与等量的作为底物的2% (w/v) 偶氮酪蛋白混合, 然后在37℃反应1小时。通过向酶反应溶液中加入3倍体积的10% 三氯乙酸 (TCA) 终止反应, 并通过离心除去凝结的蛋白质。通过将产物与等量的NaOH混合进行显色反应, 然后在440nm处测量吸光度, 以比较显色程度。通过该方法, 选择与野生型丝氨酸蛋白酶相比吸光度增加150%或更多的菌落。

[0159] 实施例2. 选择的变体的制备和活性评价

[0160] 实施例2-1: 变体制备

[0161] 作为分析来自筛选的变体序列的结果, 证实SEQ ID NO:31的第12位氨基酸 (苯丙氨酸, Phe) 和第116位氨基酸 (天冬酰胺, Asn) 分别被酪氨酸 (Tyr) 和天冬氨酸 (Asp) 取代。在图1中, 显示了基于SEQ ID NO:2的氨基酸序列的突变的位置。通过定点诱变将两个选择的突变 (F12和N116) 以单一突变的形式重新引入到pBE-S-TAP质粒中, 将具有双突变和单一突变的菌株的活性分别与野生型菌株的活性进行比较。用于制备变体的引物如下表4所示。

[0162] 表4

[0163]	TAP_F12Y_F	SEQ ID NO: 15
	TAP_F12Y_R	SEQ ID NO: 16
	TAP_N116D_F	SEQ ID NO: 17
[0164]	TAP_N116D_R	SEQ ID NO: 18

[0165] 实施例2-2: 活性评价

[0166] 在用制备的质粒转化枯草芽孢杆菌LB700菌株后, 使用N-琥珀酰-Ala-Ala-Pro-Phe-p-硝基苯胺 (Sigma, 货号S7388, 下文称为SUC-AAPF-pNA) 肽作为底物进行转化株的活性评价。将转化的枯草芽孢杆菌菌株接种到含有卡那霉素抗生素的脑心灌注 (BHI, bd, 货号53286) 液体培养基中, 并在37℃下培养20至24小时, 将除细胞外的一部分培养液与25mM

Tris-HCl (pH7.5) 缓冲液和1mM Suc-AAPF-pNA混合,然后在37℃下反应30分钟。在410nm测量反应溶液的吸光度。由文献中已知的酶产生的对硝基苯胺的消光系数在410nm处为 $8800\text{M}^{-1}\text{cm}^{-1}$,并基于此计算酶的单位(Barrett,A.J.,Cathepsin G.Methods Enzymol.,80,Pt.C,561-565,(1981))。测量的活性示于下表5中。

[0167] 表5

	酶活性(单位/mL)
野生型	16.3
F12Y	34.0
F12YN116D	63.8

[0169] 作为测量的结果,证实F12Y和F12YN116D变体的活性在pH 7.5和37℃的条件下相比野生型分别增加约2.1倍和3.9倍。

[0170] 实施例2-3:热稳定性的评价

[0171] 进行下面描述的实验以证实引入突变对热稳定性的影响。

[0172] 具体而言,在将实施例2-2中使用的样品分别置于室温、70℃、80℃和90℃下5分钟之后,测量样品的酶活性,用于活性评价。测量的活性示于下表6中。

[0173] 表6

	酶活性(单位/mL)			
热处理条件	室 温 , 5min	70℃, 5min	80℃ , 5min	90℃ , 5min
野生型	16.3	15.9	10.6	0.3
F12Y	34.0	34.7	22.8	0.1
F12YN116D	63.8	63.3	41.7	0.1

[0175] 作为测量的结果证实,即使在80℃下,F12Y和F12YN116D变体显示出比野生型菌株高约2倍和4倍的酶活性。由此证实本申请的丝氨酸蛋白酶变体在高温下也保持高活性,因此丝氨酸蛋白酶变体可以有效地用于工业中。

[0176] 实施例3. 饱和诱变文库的制备和筛选

[0177] 实施例3-1.F12和N116残基的饱和诱变文库制备

[0178] 为了确认用酪氨酸和天冬氨酸以外的残基取代F12和N116残基(即,先前选择的变体)对活性的影响,制备了关于这两个残基的饱和诱变文库。

[0179] 分别使用pBE-S-TAP质粒作为模板和引物对SEQ ID NO:11和12以及SEQ ID NO:13和14获得两个PCR片段。使用In-Fusion HD克隆试剂盒将各片段连接,然后转化到DH5α细胞中,从而获得菌落。纯化所得菌落中的质粒,以获得大小为约 4×10^3 的文库。

[0180] 表7

模板DNA (pBE-S-TAP)	SEQ ID NO:1
饱和诱变_F_1	SEQ ID NO:19
饱和诱变_R_1	SEQ ID NO:20
饱和诱变_F_2	SEQ ID NO:21

饱和诱变_R_2	SEQ ID NO:22
----------	--------------

[0182] 实施例3-2:饱和诱变文库的筛选和活性评价

[0183] 以与实施例1-2相同的方式对实施例3-1中制备的饱和诱变文库进行筛选。通过筛选、序列分析、使用Suc-AAPF-pNA作为底物对这些变体进行活性评价,选择与F12YN116D变体相比具有相同或增加的活性的变体。

[0184] 表8

[0185]		酶活性(单位/mL)
	野生型	23.52
	F12YN116D	65
	F12YN116S	77.35
	F12SN116D	61.75
	F12SN116T	117.65
	F12AN116G	94.9
	F12A	94.25
	F12R	58.5

[0186] 结果发现,F12和N116变体的活性增加,即使它们的残基各自被除了实施例2中证实的酪氨酸和天冬氨酸以外的不同氨基酸(例如,F12S、F12A、F12R、N116S、N116T、N116G)取代。

[0187] 实施例3-3:F12S变体的制备和活性评价

[0188] 通过定点诱变将突变(F12S)以单一突变形式重新引入pBE-S-TAP质粒后,将变体的活性与野生型的活性进行比较。

[0189] 使用Suc-AAPF-pNA作为底物对变体(F12S)进行活性评价。

[0190] [表9]

	酶活性(单位/mL)
[0191]	野生型 23.0

[0192]	F12S 34.5
--------	-----------

[0193] 作为测量的结果,证实F12S变体的活性增加了约1.5倍。

[0194] 实施例4.源自Thermobifida fusca的丝氨酸蛋白酶样蛋白的第12和116位残基的影响确认

[0195] 实施例4-1:野生型和变体的制备

[0196] 为了研究对应于SEQ ID NO:31的第12和116位的氨基酸残基是否影响与SEQ ID NO:31具有序列同源性的其它丝氨酸蛋白酶的活性的增加,分别用酪氨酸(Y)和天冬氨酸(D)取代序列同源性为87.2%、81.8%、81.3%、73.8%、69.9%和66.7%的丝氨酸蛋白酶的第12和116位残基,然后将丝氨酸蛋白酶的活性与野生型的活性进行比较。各丝氨酸蛋白酶的来源和序列信息如表10所示。

[0197] [表10]

来源	SEQ ID NO:	与 SEQ ID NO: 31 的同源性(%)	与 SEQ ID NO: 54 的同源性(%)
<i>Thermobifida fusca</i>	SEQ ID NO: 31	-	66.7
<i>Thermobifida cellulosilytica</i>	SEQ ID NO: 49	87.2	65.1
[0198] <i>Thermobifida halotolerans</i>	SEQ ID NO: 50	81.8	61.3
<i>Actinorugispora endophytica</i>	SEQ ID NO: 51	81.3	64
<i>Spinactinospora alkalitolerans</i>	SEQ ID NO: 52	73.8	75.4
堆肥拟诺卡氏菌	SEQ ID NO: 53	69.9	86.6
<i>Nocardiopsis potens</i>	SEQ ID NO: 54	66.7	-

[0199] 用酪氨酸(Y)和天冬氨酸(D)取代各种丝氨酸蛋白酶的12和116位残基,以制备SEQ ID NO:55至66所示的变体。

[0200] 实施例4-2:活性评价

[0201] 将制备的质粒在枯草芽孢杆菌LB700菌株中转化和表达,然后通过实施例2-2中所述的活性测量方法相同的方法进行活性评价。测量的活性如表11所示。

[0202] [表11]

[0203]	来源	变体	活性(U/mL)
	<i>Thermobifida cellulosilytica</i>	野生型	13.6
		F12Y	16.3
		F12YN116D	29.1
	<i>Thermobifida halotolerans</i>	野生型	4.0
		F12Y	6.3
	<i>Actinorugispora endophytica</i>	野生型	10.4
		F12Y	14.0
	<i>Spinactinospora alkalitolerans</i>	野生型	2.0
		P12Y	2.6
	堆肥拟诺卡氏菌	野生型	12.1
		P12Y	26.5
	<i>Nocardiopsis potens</i>	野生型	33.1
[0204]		P12Y	86.3

[0205] 作为测量的结果,证实了当突变被引入第12位残基时,与源自*Thermobifida fusca*的丝氨酸蛋白酶具有序列同源性的六种蛋白的活性也增加,如同源自*Thermobifida fusca*的丝氨酸蛋白酶。根据上述事实,证实了第12和116位残基是展现丝氨酸蛋白酶活性的重要残基,并且酶活性可以通过用不同的氨基酸取代这些残基来提高,正如通过SEQ ID NO:31所证实的。因此,本申请的具有增加的酶活性的丝氨酸蛋白酶变体可有效地用于工业中。

[0206] 综上所述,本申请所属领域的技术人员将能够理解,本申请可以以其它具体形式来实现,而不改变本申请的技术构思或本质特征。在这方面,本文公开的示例性实施方案仅用于说明目的,而不应被解释为限制本申请的范围。相反,本申请旨在不仅涵盖示例性实施方案,而且涵盖可包括在如所附权利要求书所定义的本申请的实质和范围内的各种替代、修改、等同方案和其它实施方案。

序列表

<110> CJ 第一制糖株式会社 (CJ CheilJedang Corporation)

<120> 新的丝氨酸蛋白酶变体

<130> OPA22007

<150> KR 10-2021-0032885

<151> 2021-03-12

<160> 72

<170> KoPatentIn 3.0

<210> 1

<211> 6922

<212> DNA

<213> 人工序列 (Artificial Sequence)

<220>

<221>

<222>

<223> pBE-S-TFP

<400> 1

```
actagtgttc ttttctgtat gaaaatagtt atttcgagtc tctacgaaa tagcgagaga 60
tgatatacct aaatagagat aaaatcatct caaaaaaatg ggtctactaa aatattattc 120
catctattac aataaattca cagaatagtc ttttaagtaa gtctactctg aacttaagca 180
aaaggagagg gacgcgtgtg agaagcaaaa aattgtggat cagcttggtt tttgcgttaa 240
cgттаатстт tacgatggcg ttcagcaaca tgtctgcgca ggctgcggcc ggtgcacata 300
tgcaagagct ggсgttgaaa cgggacctcg gcctctctga cgcagaagta gccgaactcc 360
gggctgctga ggсggaagcg gtcgagctcg aggaggagct ccgcgattca ttagggtcag 420
acttcggcgг tgtatatctg gatgctgaca ccaccgaaat tacggtcgcg gtaaccgacc 480
cggcagcggt aagtcgtgtc gacgcggatg atgtcacagt tgatgttgtc gatttcgggg 540
aaacagcttt gaatgatttt gtggcttcat taaatgccat tgccgacacg gcagacccta 600
aagtcactgg atggтатаcc gatctcgaaa gtgatgccgt agtcattacg accttgсgtg 660
gcgggactcc tgctgccgag gaacttgctg agagagcgгг tctcgacgaa agagccgttc 720
ggattgtgga agaagatgaa gaaccacaga gcttgгctgc aattattggt ggaaaccсgt 780
actatttcгг aaattacaga tgcagtatcg ggttttagtgt ccgtcagggc tctcaaacгг 840
gattcgсgac cgcaggccac tgcggatcca cggggacггг tgtgtcttct cttcaggaa 900
cagttgcagg aagttatttc ccgggtcггг atatggгctg ggtgcгgatt acatcagcag 960
atactgtaac accactгgta aatcгgtata atgggggaaс tgttacгgtc actggгtcac 1020
aagaagctgc caccгgatcc tccgtttgtc gctctгgagc aacaacгггс tgгcгctггг 1080
gaactatcca atcaaaaaac caaacгgttc gctatгcaga agggactgtt actгgtttta 1140
caagaactac агсctgtгct gaaggtгггг attctгgagg gccatггctc acaggtagcc 1200
aggcгcaagg ggttacaagc ggсggaacag гсgattгcag aagtгgaggг attacctttt 1260
```

```

tccaaccaat caatccattg cttagctatt tcggccttca attagtgacc ggctgaaagc 1320
ttgtcgacct gcagctctaga catcaccatc atcaccacta atgcggtagt ttatcacagt 1380
taaattgcta acgcagtcag gcaccgtgta tgaaatctaa caatgcgctc atcgctcatcc 1440
tcggcaccgt caccctggat gctgtaggca taggcttggt tatgccggtg ctgccgggcc 1500
tattttacttt tttgcattct acaaactgca taactattat gtaaactcgct ctttttttagg 1560
tggcacaaaat gtgaggcatt ttcgctcttt ccggcaacca cttccaagta aagtataaca 1620
cactatacttt tatattcata aagtgtgtgc tctgcgaggc tgtcggcagt gccgaccaa 1680
accataaaac ctttaagacc tttctttttt ttacgagaaa aaagaaacaa aaaaacctgc 1740
cctctgccac ctcagcaaaag gggggttttg ctctcgtgct cgttttaaaaa tcagcaaggg 1800
acaggtagta ttttttgaga agatcactca aaaaatctcc acctttaaac ctttgccaat 1860
ttttattttg tccgttttgt ctagcttacc gaaagccaga ctcagcaaga ataaaatttt 1920
tattgtcttt cggttttcta gtgtaacgga caaaaccact caaaataaaa aagatacaag 1980
agaggctctt cgtatctttt attcagcaat cgcgcccgat tgctgaacag attaataata 2040
gatttttagct ttttatttgt tgaaaaaagc taatcaaatt gttgtcggga tcaattactg 2100
caaagtctcg ttcattccac cactgatctt ttaatgatgt attgggggtgc aaaatgccca 2160
aaggcttaat atgttgatat aattcatcaa ttccctctac ttcaatgcgg caactagcag 2220
taccagcaat aaacgactcc gcacctgtac aaaccggtga atcattacta cgagagcgcc 2280
agccttcac acttgccctc catagatgaa tccgaacctc attacacatt agaactgcga 2340
atccatcttc atggtgaacc aaagtgaac ctagttttatc gcaataaaaa cctatactct 2400
ttttaatatc cccgactggc aatgccggga tagactgtaa cattctcacg cataaaatcc 2460
cctttcatct tctaattgaa atctattacc ttattattaa ttcaattcgc tcataattaa 2520
tcctttttct tattacgcaa aatggcccgga ttaagcaca ccctttattc cgttaatgcg 2580
ccatgacagc catgataaatt actaatacta ggagaagtta ataaatacga gcaaaaaggcc 2640
agcaaaaagg caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 2700
cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 2760
tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgacct 2820
tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 2880
gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc 2940
acgaaccccc cgttcagccc gaccgctgcg cttatccgg taactatcgt cttgagtcca 3000
accggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 3060
cgaggatatg aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta 3120
gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 3180
gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 3240
agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 3300
ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 3360
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 3420
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 3480
tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 3540
gggagggtct accatctggc cccagtgtcg caatgatacc gcgagacca cgctcaccgg 3600

```

```

ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg 3660
caactttatc cgctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt 3720
cgccagttaa tagtttgccg aacgttggtg ccattgctac aggcacgtg gtgtcacgct 3780
cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 3840
cccccatggt gtgcaaaaaa gcggttagct cttcgggtcc tccgatcggt gtcagaagta 3900
agttggccgc agtggttatca ctcatggta tggcagcact gcataattct cttactgtca 3960
tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 4020
agtgatatgcg gcgaccgagt tgctcttgcc cggcgtaaat acgggataat accgcgccac 4080
atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 4140
ggatcttacc gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt 4200
cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg 4260
caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 4320
attattgaag catttatcag gggtattgtc tcatgagcgg atacatattt gaatgtattt 4380
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtcca cctgacgtct 4440
aagaaaccat tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc 4500
gtctcgcgcg tttcgggtgat gacggtgaaa acctctgaca gtaaccaaca tgattaacaa 4560
ttattagagg tcatcggtca aaatgggtatg cgttttgaca catccactat atatccgtgt 4620
cgttctgtcc actcctgaat cccattccag aaattctcta gcgattccag aagtttctca 4680
gagtcggaaa gttgaccaga cattacgaac tggcacagat ggtcataacc tgaaggaaga 4740
tctgattgct taactgcttc agttaagacc gaagcgtcg tcgtataaca gatgcgatga 4800
tgcagaccaa tcaacatggc acctgccatt gctacctgta cagtcaagga tggtagaaat 4860
gttgtcggtc cttgcacacg aatattacgc catttgctg catattcaaa cagctcttct 4920
acgataaggg cacaaatcgc atcgtggaac gtttgggctt ctaccgattt agcagtttga 4980
tacactttct ctaagtatcc acctgaatca taaatcggca aaatagagaa aaattgacca 5040
tgtgtaagcg gccaatctga ttccacctga gatgcataat ctagtagaat ctcttcgcta 5100
tcaaaattca cttccacctt ccactcacg gttgtccatt catggctgaa ctctgcttcc 5160
tctgttgaca tgacacacat catctcaata tccgaatagg gccatcagt ctgacgacca 5220
agagagccat aaacaccaat agccttaaca tcatcccat atttatccaa tattcgttcc 5280
ttaatttcat gaacaatctt cattctttct tctctagtea ttattattgg tccattcact 5340
attctcattc ctttttcaga taattttaga tttgcttttc taaataagaa tatttgaga 5400
gcaccgttct tattcagcta ttaataactc gtcttcttaa gcaccttca atccttttaa 5460
taacaattat agcatctaatt cttcaacaaa ctggcccgtt tgttgaacta ctctttaata 5520
aaataatttt tccgttccca attccacatt gcaataatag aaaatccatc ttcacggct 5580
ttttcgtcat catctgtatg aatcaaatcg cttcttctg tgtcatcaag gtttaatttt 5640
ttatgtattt cttttaacaa accaccatag gagattaacc ttttacggtg taaaccttcc 5700
tccaaatcag acaaacgttt caaattcttt tcttcatcat cggtcataaa atccgtatcc 5760
tttacaggat attttgagcgt ttcgtcaatt gccgattgta tatccgattt atatttattt 5820
ttcggtcgaa tcatttgaac ttttacattt ggatcatagt ctaatttcat tgccttttcc 5880
caaaattgaa tccattgttt ttgattcacg tagttttctg tattcttaaa ataagttggt 5940

```

tccacacata ccaatacatg catgtgctga ttataagaat tatctttatt atttattgtc 6000
 acttccgttg cacgcataaa accaacaaga tttttattaa tttttttata ttgcatcatt 6060
 cggcgaaatc cttgagccat atctgacaaa ctcttattta attcttcgcc atcataaaca 6120
 tttttaactg ttaatgtgag aaacaaccaa cgaactgttg gcttttggtt aataacttca 6180
 gcaacaacct tttgtgactg aatgccatgt ttcattgctc tcctccagtt gcacattgga 6240
 caaagcctgg atttacaaaa ccacactcga tacaactttc tttcgccctgt ttcacgattt 6300
 tgtttatact ctaatatattc agcacaatct tttactcttt cagccttttt aaattcaaga 6360
 atatgcagaa gttcaaagta atcaacatta gcgattttct tttctctcca tggctctact 6420
 tttccacttt ttgtcttgct cactaaaacc cttgattttt catctgaata aatgctacta 6480
 ttaggacaca taatattaaa agaaaccccc atctatttag ttatttggtt agtcacttat 6540
 aactttaaca gatgggggtt ttctgtgcaa ccaattttta gggttttcaa tacttttaaa 6600
 cacatacata ccaacacttc aacgcacett tcagcaacta aaataaaaat gacgttattt 6660
 ctatatgtat caagataaga aagaacaagt tcaaaacat caaaaaaga caccttttca 6720
 ggtgcttttt ttattttata aactcattcc ctgatctcga cttcgttctt tttttacctc 6780
 tcggttatga gttagttaa attcgttctt tttaggttct aaatcgtgtt tttcttgaa 6840
 ttgtgctggt ttatccttta ccttgtctac aaaccctta aaaacgtttt taaagcctt 6900
 taagccgtct gtacgttcct aa 6922

<210> 2

<211> 338

<212> PRT

<213> 未知(Unknown)

<220>

<221>

<222>

<223> Thermobifida fusca

<400> 2

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

100	105	110
Ala Val Val Ile Thr Thr Leu Arg Gly Gly Thr Pro Ala Ala Glu Glu		
115	120	125
Leu Ala Glu Arg Ala Gly Leu Asp Glu Arg Ala Val Arg Ile Val Glu		
130	135	140
Glu Asp Glu Glu Pro Gln Ser Leu Ala Ala Ile Ile Gly Gly Asn Pro		
145	150	155
Tyr Tyr Phe Gly Asn Tyr Arg Cys Ser Ile Gly Phe Ser Val Arg Gln		
165	170	175
Gly Ser Gln Thr Gly Phe Ala Thr Ala Gly His Cys Gly Ser Thr Gly		
180	185	190
Thr Arg Val Ser Ser Pro Ser Gly Thr Val Ala Gly Ser Tyr Phe Pro		
195	200	205
Gly Arg Asp Met Gly Trp Val Arg Ile Thr Ser Ala Asp Thr Val Thr		
210	215	220
Pro Leu Val Asn Arg Tyr Asn Gly Gly Thr Val Thr Val Thr Gly Ser		
225	230	235
Gln Glu Ala Ala Thr Gly Ser Ser Val Cys Arg Ser Gly Ala Thr Thr		
245	250	255
Gly Trp Arg Cys Gly Thr Ile Gln Ser Lys Asn Gln Thr Val Arg Tyr		
260	265	270
Ala Glu Gly Thr Val Thr Gly Leu Thr Arg Thr Thr Ala Cys Ala Glu		
275	280	285
Gly Gly Asp Ser Gly Gly Pro Trp Leu Thr Gly Ser Gln Ala Gln Gly		
290	295	300
Val Thr Ser Gly Gly Thr Gly Asp Cys Arg Ser Gly Gly Ile Thr Phe		
305	310	315
Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu Gln Leu Val		
325	330	335

Thr Gly

<210> 3

<211> 338

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163Y)

<400> 3

Met	Gln	Glu	Leu	Ala	Leu	Lys	Arg	Asp	Leu	Gly	Leu	Ser	Asp	Ala	Glu
1				5					10					15	
Val	Ala	Glu	Leu	Arg	Ala	Ala	Glu	Ala	Glu	Ala	Val	Glu	Leu	Glu	Glu
				20					25					30	
Glu	Leu	Arg	Asp	Ser	Leu	Gly	Ser	Asp	Phe	Gly	Gly	Val	Tyr	Leu	Asp
				35					40					45	
Ala	Asp	Thr	Thr	Glu	Ile	Thr	Val	Ala	Val	Thr	Asp	Pro	Ala	Ala	Val
				50					55					60	
Ser	Arg	Val	Asp	Ala	Asp	Asp	Val	Thr	Val	Asp	Val	Val	Asp	Phe	Gly
65					70					75					80
Glu	Thr	Ala	Leu	Asn	Asp	Phe	Val	Ala	Ser	Leu	Asn	Ala	Ile	Ala	Asp
				85						90					95
Thr	Ala	Asp	Pro	Lys	Val	Thr	Gly	Trp	Tyr	Thr	Asp	Leu	Glu	Ser	Asp
				100						105					110
Ala	Val	Val	Ile	Thr	Thr	Leu	Arg	Gly	Gly	Thr	Pro	Ala	Ala	Glu	Glu
				115						120					125
Leu	Ala	Glu	Arg	Ala	Gly	Leu	Asp	Glu	Arg	Ala	Val	Arg	Ile	Val	Glu
				130						135					140
Glu	Asp	Glu	Glu	Pro	Gln	Ser	Leu	Ala	Ala	Ile	Ile	Gly	Gly	Asn	Pro
145					150					155					160
Tyr	Tyr	Tyr	Gly	Asn	Tyr	Arg	Cys	Ser	Ile	Gly	Phe	Ser	Val	Arg	Gln
				165						170					175
Gly	Ser	Gln	Thr	Gly	Phe	Ala	Thr	Ala	Gly	His	Cys	Gly	Ser	Thr	Gly
				180						185					190
Thr	Arg	Val	Ser	Ser	Pro	Ser	Gly	Thr	Val	Ala	Gly	Ser	Tyr	Phe	Pro
				195						200					205
Gly	Arg	Asp	Met	Gly	Trp	Val	Arg	Ile	Thr	Ser	Ala	Asp	Thr	Val	Thr
				210						215					220
Pro	Leu	Val	Asn	Arg	Tyr	Asn	Gly	Gly	Thr	Val	Thr	Val	Thr	Gly	Ser
225					230					235					240
Gln	Glu	Ala	Ala	Thr	Gly	Ser	Ser	Val	Cys	Arg	Ser	Gly	Ala	Thr	Thr
				245						250					255
Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln	Ser	Lys	Asn	Gln	Thr	Val	Arg	Tyr
				260						265					270
Ala	Glu	Gly	Thr	Val	Thr	Gly	Leu	Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu
				275						280					285
Gly	Gly	Asp	Ser	Gly	Gly	Pro	Trp	Leu	Thr	Gly	Ser	Gln	Ala	Gln	Gly
				290						295					300
Val	Thr	Ser	Gly	Gly	Thr	Gly	Asp	Cys	Arg	Ser	Gly	Gly	Ile	Thr	Phe

305	310	315	320
Phe Gln Pro Ile Asn	Pro Leu Leu Ser Tyr	Phe Gly Leu Gln Leu	Val
	325	330	335
Thr Gly			
<210> 4			
<211> 338			
<212> PRT			
<213> 人工序列 (Artificial Sequence)			
<220>			
<221>			
<222>			
<223> Thermobifida fusca (F163Y_N267D)			
<400> 4			
Met Gln Glu Leu Ala Leu Lys Arg Asp Leu Gly Leu Ser Asp Ala Glu			
1	5	10	15
Val Ala Glu Leu Arg Ala Ala Glu Ala Glu Ala Val Glu Leu Glu Glu			
	20	25	30
Glu Leu Arg Asp Ser Leu Gly Ser Asp Phe Gly Gly Val Tyr Leu Asp			
	35	40	45
Ala Asp Thr Thr Glu Ile Thr Val Ala Val Thr Asp Pro Ala Ala Val			
	50	55	60
Ser Arg Val Asp Ala Asp Asp Val Thr Val Asp Val Val Asp Phe Gly			
65	70	75	80
Glu Thr Ala Leu Asn Asp Phe Val Ala Ser Leu Asn Ala Ile Ala Asp			
	85	90	95
Thr Ala Asp Pro Lys Val Thr Gly Trp Tyr Thr Asp Leu Glu Ser Asp			
	100	105	110
Ala Val Val Ile Thr Thr Leu Arg Gly Gly Thr Pro Ala Ala Glu Glu			
	115	120	125
Leu Ala Glu Arg Ala Gly Leu Asp Glu Arg Ala Val Arg Ile Val Glu			
	130	135	140
Glu Asp Glu Glu Pro Gln Ser Leu Ala Ala Ile Ile Gly Gly Asn Pro			
145	150	155	160
Tyr Tyr Tyr Gly Asn Tyr Arg Cys Ser Ile Gly Phe Ser Val Arg Gln			
	165	170	175
Gly Ser Gln Thr Gly Phe Ala Thr Ala Gly His Cys Gly Ser Thr Gly			
	180	185	190
Thr Arg Val Ser Ser Pro Ser Gly Thr Val Ala Gly Ser Tyr Phe Pro			
	195	200	205

Gly Arg Asp Met Gly Trp Val Arg Ile Thr Ser Ala Asp Thr Val Thr
 210 215 220
 Pro Leu Val Asn Arg Tyr Asn Gly Gly Thr Val Thr Val Thr Gly Ser
 225 230 235 240
 Gln Glu Ala Ala Thr Gly Ser Ser Val Cys Arg Ser Gly Ala Thr Thr
 245 250 255
 Gly Trp Arg Cys Gly Thr Ile Gln Ser Lys Asp Gln Thr Val Arg Tyr
 260 265 270
 Ala Glu Gly Thr Val Thr Gly Leu Thr Arg Thr Thr Ala Cys Ala Glu
 275 280 285
 Gly Gly Asp Ser Gly Gly Pro Trp Leu Thr Gly Ser Gln Ala Gln Gly
 290 295 300
 Val Thr Ser Gly Gly Thr Gly Asp Cys Arg Ser Gly Gly Ile Thr Phe
 305 310 315 320
 Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu Gln Leu Val
 325 330 335

Thr Gly

<210> 5

<211> 338

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163Y_N267S)

<400> 5

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

100	105	110
Ala Val Val Ile Thr Thr Leu Arg Gly Gly Thr Pro Ala Ala Glu Glu		
115	120	125
Leu Ala Glu Arg Ala Gly Leu Asp Glu Arg Ala Val Arg Ile Val Glu		
130	135	140
Glu Asp Glu Glu Pro Gln Ser Leu Ala Ala Ile Ile Gly Gly Asn Pro		
145	150	155
Tyr Tyr Tyr Gly Asn Tyr Arg Cys Ser Ile Gly Phe Ser Val Arg Gln		
165	170	175
Gly Ser Gln Thr Gly Phe Ala Thr Ala Gly His Cys Gly Ser Thr Gly		
180	185	190
Thr Arg Val Ser Ser Pro Ser Gly Thr Val Ala Gly Ser Tyr Phe Pro		
195	200	205
Gly Arg Asp Met Gly Trp Val Arg Ile Thr Ser Ala Asp Thr Val Thr		
210	215	220
Pro Leu Val Asn Arg Tyr Asn Gly Gly Thr Val Thr Val Thr Gly Ser		
225	230	235
Gln Glu Ala Ala Thr Gly Ser Ser Val Cys Arg Ser Gly Ala Thr Thr		
245	250	255
Gly Trp Arg Cys Gly Thr Ile Gln Ser Lys Ser Gln Thr Val Arg Tyr		
260	265	270
Ala Glu Gly Thr Val Thr Gly Leu Thr Arg Thr Thr Ala Cys Ala Glu		
275	280	285
Gly Gly Asp Ser Gly Gly Pro Trp Leu Thr Gly Ser Gln Ala Gln Gly		
290	295	300
Val Thr Ser Gly Gly Thr Gly Asp Cys Arg Ser Gly Gly Ile Thr Phe		
305	310	315
Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu Gln Leu Val		
325	330	335

Thr Gly

<210> 6

<211> 338

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163S_N267D)

<400> 6

Met	Gln	Glu	Leu	Ala	Leu	Lys	Arg	Asp	Leu	Gly	Leu	Ser	Asp	Ala	Glu
1				5					10					15	
Val	Ala	Glu	Leu	Arg	Ala	Ala	Glu	Ala	Glu	Ala	Val	Glu	Leu	Glu	Glu
			20					25					30		
Glu	Leu	Arg	Asp	Ser	Leu	Gly	Ser	Asp	Phe	Gly	Gly	Val	Tyr	Leu	Asp
		35					40					45			
Ala	Asp	Thr	Thr	Glu	Ile	Thr	Val	Ala	Val	Thr	Asp	Pro	Ala	Ala	Val
	50					55					60				
Ser	Arg	Val	Asp	Ala	Asp	Asp	Val	Thr	Val	Asp	Val	Val	Asp	Phe	Gly
65				70						75				80	
Glu	Thr	Ala	Leu	Asn	Asp	Phe	Val	Ala	Ser	Leu	Asn	Ala	Ile	Ala	Asp
			85					90					95		
Thr	Ala	Asp	Pro	Lys	Val	Thr	Gly	Trp	Tyr	Thr	Asp	Leu	Glu	Ser	Asp
		100						105					110		
Ala	Val	Val	Ile	Thr	Thr	Leu	Arg	Gly	Gly	Thr	Pro	Ala	Ala	Glu	Glu
		115					120						125		
Leu	Ala	Glu	Arg	Ala	Gly	Leu	Asp	Glu	Arg	Ala	Val	Arg	Ile	Val	Glu
	130					135					140				
Glu	Asp	Glu	Glu	Pro	Gln	Ser	Leu	Ala	Ala	Ile	Ile	Gly	Gly	Asn	Pro
145				150						155				160	
Tyr	Tyr	Ser	Gly	Asn	Tyr	Arg	Cys	Ser	Ile	Gly	Phe	Ser	Val	Arg	Gln
			165					170					175		
Gly	Ser	Gln	Thr	Gly	Phe	Ala	Thr	Ala	Gly	His	Cys	Gly	Ser	Thr	Gly
		180						185					190		
Thr	Arg	Val	Ser	Ser	Pro	Ser	Gly	Thr	Val	Ala	Gly	Ser	Tyr	Phe	Pro
	195						200					205			
Gly	Arg	Asp	Met	Gly	Trp	Val	Arg	Ile	Thr	Ser	Ala	Asp	Thr	Val	Thr
	210				215						220				
Pro	Leu	Val	Asn	Arg	Tyr	Asn	Gly	Gly	Thr	Val	Thr	Val	Thr	Gly	Ser
225				230						235				240	
Gln	Glu	Ala	Ala	Thr	Gly	Ser	Ser	Val	Cys	Arg	Ser	Gly	Ala	Thr	Thr
			245					250					255		
Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln	Ser	Lys	Asp	Gln	Thr	Val	Arg	Tyr
		260					265						270		
Ala	Glu	Gly	Thr	Val	Thr	Gly	Leu	Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu
	275					280						285			
Gly	Gly	Asp	Ser	Gly	Gly	Pro	Trp	Leu	Thr	Gly	Ser	Gln	Ala	Gln	Gly
	290					295					300				
Val	Thr	Ser	Gly	Gly	Thr	Gly	Asp	Cys	Arg	Ser	Gly	Gly	Ile	Thr	Phe

305	310	315	320
Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu Gln Leu Val			
	325	330	335
Thr Gly			
<210> 7			
<211> 338			
<212> PRT			
<213> 人工序列(Artificial Sequence)			
<220>			
<221>			
<222>			
<223> Thermobifida fusca (F163S_N267T)			
<400> 7			
Met Gln Glu Leu Ala Leu Lys Arg Asp Leu Gly Leu Ser Asp Ala Glu			
1 5 10 15			
Val Ala Glu Leu Arg Ala Ala Glu Ala Glu Ala Val Glu Leu Glu Glu			
20 25 30			
Glu Leu Arg Asp Ser Leu Gly Ser Asp Phe Gly Gly Val Tyr Leu Asp			
35 40 45			
Ala Asp Thr Thr Glu Ile Thr Val Ala Val Thr Asp Pro Ala Ala Val			
50 55 60			
Ser Arg Val Asp Ala Asp Asp Val Thr Val Asp Val Val Asp Phe Gly			
65 70 75 80			
Glu Thr Ala Leu Asn Asp Phe Val Ala Ser Leu Asn Ala Ile Ala Asp			
85 90 95			
Thr Ala Asp Pro Lys Val Thr Gly Trp Tyr Thr Asp Leu Glu Ser Asp			
100 105 110			
Ala Val Val Ile Thr Thr Leu Arg Gly Gly Thr Pro Ala Ala Glu Glu			
115 120 125			
Leu Ala Glu Arg Ala Gly Leu Asp Glu Arg Ala Val Arg Ile Val Glu			
130 135 140			
Glu Asp Glu Glu Pro Gln Ser Leu Ala Ala Ile Ile Gly Gly Asn Pro			
145 150 155 160			
Tyr Tyr Ser Gly Asn Tyr Arg Cys Ser Ile Gly Phe Ser Val Arg Gln			
165 170 175			
Gly Ser Gln Thr Gly Phe Ala Thr Ala Gly His Cys Gly Ser Thr Gly			
180 185 190			
Thr Arg Val Ser Ser Pro Ser Gly Thr Val Ala Gly Ser Tyr Phe Pro			
195 200 205			

Gly Arg Asp Met Gly Trp Val Arg Ile Thr Ser Ala Asp Thr Val Thr
 210 215 220
 Pro Leu Val Asn Arg Tyr Asn Gly Gly Thr Val Thr Val Thr Gly Ser
 225 230 235 240
 Gln Glu Ala Ala Thr Gly Ser Ser Val Cys Arg Ser Gly Ala Thr Thr
 245 250 255
 Gly Trp Arg Cys Gly Thr Ile Gln Ser Lys Thr Gln Thr Val Arg Tyr
 260 265 270
 Ala Glu Gly Thr Val Thr Gly Leu Thr Arg Thr Thr Ala Cys Ala Glu
 275 280 285
 Gly Gly Asp Ser Gly Gly Pro Trp Leu Thr Gly Ser Gln Ala Gln Gly
 290 295 300
 Val Thr Ser Gly Gly Thr Gly Asp Cys Arg Ser Gly Gly Ile Thr Phe
 305 310 315 320
 Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu Gln Leu Val
 325 330 335

Thr Gly

<210> 8

<211> 338

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163A_N267G)

<400> 8

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

100	105	110
Ala Val Val Ile Thr Thr Leu Arg Gly Gly Thr Pro Ala Ala Glu Glu		
115	120	125
Leu Ala Glu Arg Ala Gly Leu Asp Glu Arg Ala Val Arg Ile Val Glu		
130	135	140
Glu Asp Glu Glu Pro Gln Ser Leu Ala Ala Ile Ile Gly Gly Asn Pro		
145	150	155
Tyr Tyr Ala Gly Asn Tyr Arg Cys Ser Ile Gly Phe Ser Val Arg Gln		
165	170	175
Gly Ser Gln Thr Gly Phe Ala Thr Ala Gly His Cys Gly Ser Thr Gly		
180	185	190
Thr Arg Val Ser Ser Pro Ser Gly Thr Val Ala Gly Ser Tyr Phe Pro		
195	200	205
Gly Arg Asp Met Gly Trp Val Arg Ile Thr Ser Ala Asp Thr Val Thr		
210	215	220
Pro Leu Val Asn Arg Tyr Asn Gly Gly Thr Val Thr Val Thr Gly Ser		
225	230	235
Gln Glu Ala Ala Thr Gly Ser Ser Val Cys Arg Ser Gly Ala Thr Thr		
245	250	255
Gly Trp Arg Cys Gly Thr Ile Gln Ser Lys Gly Gln Thr Val Arg Tyr		
260	265	270
Ala Glu Gly Thr Val Thr Gly Leu Thr Arg Thr Thr Ala Cys Ala Glu		
275	280	285
Gly Gly Asp Ser Gly Gly Pro Trp Leu Thr Gly Ser Gln Ala Gln Gly		
290	295	300
Val Thr Ser Gly Gly Thr Gly Asp Cys Arg Ser Gly Gly Ile Thr Phe		
305	310	315
Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu Gln Leu Val		
325	330	335

Thr Gly

<210> 9

<211> 338

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163A)

<400> 9

Met	Gln	Glu	Leu	Ala	Leu	Lys	Arg	Asp	Leu	Gly	Leu	Ser	Asp	Ala	Glu
1				5					10					15	
Val	Ala	Glu	Leu	Arg	Ala	Ala	Glu	Ala	Glu	Ala	Val	Glu	Leu	Glu	Glu
				20					25					30	
Glu	Leu	Arg	Asp	Ser	Leu	Gly	Ser	Asp	Phe	Gly	Gly	Val	Tyr	Leu	Asp
				35					40					45	
Ala	Asp	Thr	Thr	Glu	Ile	Thr	Val	Ala	Val	Thr	Asp	Pro	Ala	Ala	Val
				50					55					60	
Ser	Arg	Val	Asp	Ala	Asp	Asp	Val	Thr	Val	Asp	Val	Val	Asp	Phe	Gly
65					70					75					80
Glu	Thr	Ala	Leu	Asn	Asp	Phe	Val	Ala	Ser	Leu	Asn	Ala	Ile	Ala	Asp
				85						90					95
Thr	Ala	Asp	Pro	Lys	Val	Thr	Gly	Trp	Tyr	Thr	Asp	Leu	Glu	Ser	Asp
				100						105					110
Ala	Val	Val	Ile	Thr	Thr	Leu	Arg	Gly	Gly	Thr	Pro	Ala	Ala	Glu	Glu
				115						120					125
Leu	Ala	Glu	Arg	Ala	Gly	Leu	Asp	Glu	Arg	Ala	Val	Arg	Ile	Val	Glu
				130						135					140
Glu	Asp	Glu	Glu	Pro	Gln	Ser	Leu	Ala	Ala	Ile	Ile	Gly	Gly	Asn	Pro
145					150					155					160
Tyr	Tyr	Ala	Gly	Asn	Tyr	Arg	Cys	Ser	Ile	Gly	Phe	Ser	Val	Arg	Gln
				165						170					175
Gly	Ser	Gln	Thr	Gly	Phe	Ala	Thr	Ala	Gly	His	Cys	Gly	Ser	Thr	Gly
				180						185					190
Thr	Arg	Val	Ser	Ser	Pro	Ser	Gly	Thr	Val	Ala	Gly	Ser	Tyr	Phe	Pro
				195						200					205
Gly	Arg	Asp	Met	Gly	Trp	Val	Arg	Ile	Thr	Ser	Ala	Asp	Thr	Val	Thr
				210						215					220
Pro	Leu	Val	Asn	Arg	Tyr	Asn	Gly	Gly	Thr	Val	Thr	Val	Thr	Gly	Ser
225					230					235					240
Gln	Glu	Ala	Ala	Thr	Gly	Ser	Ser	Val	Cys	Arg	Ser	Gly	Ala	Thr	Thr
				245						250					255
Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln	Ser	Lys	Asn	Gln	Thr	Val	Arg	Tyr
				260						265					270
Ala	Glu	Gly	Thr	Val	Thr	Gly	Leu	Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu
				275						280					285
Gly	Gly	Asp	Ser	Gly	Gly	Pro	Trp	Leu	Thr	Gly	Ser	Gln	Ala	Gln	Gly
				290						295					300
Val	Thr	Ser	Gly	Gly	Thr	Gly	Asp	Cys	Arg	Ser	Gly	Gly	Ile	Thr	Phe

305	310	315	320
Phe Gln Pro Ile Asn	Pro Leu Leu Ser Tyr	Phe Gly Leu Gln Leu	Val
	325	330	335
Thr Gly			
<210> 10			
<211> 338			
<212> PRT			
<213> 人工序列 (Artificial Sequence)			
<220>			
<221>			
<222>			
<223> Thermobifida fusca (F163R)			
<400> 10			
Met Gln Glu Leu Ala Leu Lys Arg Asp Leu Gly Leu Ser Asp Ala Glu			
1	5	10	15
Val Ala Glu Leu Arg Ala Ala Glu Ala Glu Ala Val Glu Leu Glu Glu			
	20	25	30
Glu Leu Arg Asp Ser Leu Gly Ser Asp Phe Gly Gly Val Tyr Leu Asp			
	35	40	45
Ala Asp Thr Thr Glu Ile Thr Val Ala Val Thr Asp Pro Ala Ala Val			
	50	55	60
Ser Arg Val Asp Ala Asp Asp Val Thr Val Asp Val Val Asp Phe Gly			
65	70	75	80
Glu Thr Ala Leu Asn Asp Phe Val Ala Ser Leu Asn Ala Ile Ala Asp			
	85	90	95
Thr Ala Asp Pro Lys Val Thr Gly Trp Tyr Thr Asp Leu Glu Ser Asp			
	100	105	110
Ala Val Val Ile Thr Thr Leu Arg Gly Gly Thr Pro Ala Ala Glu Glu			
	115	120	125
Leu Ala Glu Arg Ala Gly Leu Asp Glu Arg Ala Val Arg Ile Val Glu			
	130	135	140
Glu Asp Glu Glu Pro Gln Ser Leu Ala Ala Ile Ile Gly Gly Asn Pro			
145	150	155	160
Tyr Tyr Arg Gly Asn Tyr Arg Cys Ser Ile Gly Phe Ser Val Arg Gln			
	165	170	175
Gly Ser Gln Thr Gly Phe Ala Thr Ala Gly His Cys Gly Ser Thr Gly			
	180	185	190
Thr Arg Val Ser Ser Pro Ser Gly Thr Val Ala Gly Ser Tyr Phe Pro			
	195	200	205

Gly	Arg	Asp	Met	Gly	Trp	Val	Arg	Ile	Thr	Ser	Ala	Asp	Thr	Val	Thr
210							215				220				
Pro	Leu	Val	Asn	Arg	Tyr	Asn	Gly	Gly	Thr	Val	Thr	Val	Thr	Gly	Ser
225						230				235				240	
Gln	Glu	Ala	Ala	Thr	Gly	Ser	Ser	Val	Cys	Arg	Ser	Gly	Ala	Thr	Thr
				245					250					255	
Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln	Ser	Lys	Asn	Gln	Thr	Val	Arg	Tyr
			260					265						270	
Ala	Glu	Gly	Thr	Val	Thr	Gly	Leu	Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu
			275				280						285		
Gly	Gly	Asp	Ser	Gly	Gly	Pro	Trp	Leu	Thr	Gly	Ser	Gln	Ala	Gln	Gly
			290			295					300				
Val	Thr	Ser	Gly	Gly	Thr	Gly	Asp	Cys	Arg	Ser	Gly	Gly	Ile	Thr	Phe
305						310				315				320	
Phe	Gln	Pro	Ile	Asn	Pro	Leu	Leu	Ser	Tyr	Phe	Gly	Leu	Gln	Leu	Val
				325					330					335	

Thr Gly

<210> 11

<211> 47

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> TAP_成熟区_F

<400> 11

aagaagatga agaaccacag agcttggtg caattattgg tggaaac 47

<210> 12

<211> 44

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> TAP_成熟区_R

<400> 12

gactgcaggt cgacaagctt tcagccggtc actaattgaa ggcc 44

<210> 13

<211> 23

<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<221>
<222>
<223> pBE-S-载体-F
<400> 13
aagcttggtcg acctgcagtc tag 23
<210> 14
<211> 24
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<221>
<222>
<223> pBE-S-载体-R
<400> 14
ctgtggttct tcattcttcc ccac 24
<210> 15
<211> 37
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<221>
<222>
<223> TAP_F12Y_F
<400> 15
gtggaaaccc gtactattac ggaaattaca gatgcag 37
<210> 16
<211> 37
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<221>
<222>
<223> TAP_F12Y_R
<400> 16
ctgcatctgt aatttccgta atagtacggg tttccac 37
<210> 17

<211> 38
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> TAP_N116D_F
<400> 17
ggaactatcc aatcaaaaaga ccaaacgggtt cgctatgc 38
<210> 18
<211> 38
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> TAP_N116D_R
<400> 18
gcatagcgaa ccgtttggtc ttttgattgg atagttcc 38
<210> 19
<211> 37
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> 饱和诱变_F_1
<400> 19
gtggaaaccc gtactatnnn ggaaattaca gatgcag 37
<210> 20
<211> 38
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> 饱和诱变_R_1
<400> 20
gcatagcgaa ccgtttggnn ttttgattgg atagttcc 38

<210> 21
 <211> 38
 <212> DNA
 <213> 人工序列(Artificial Sequence)
 <220>
 <221>
 <222>
 <223> 饱和诱变_F_2
 <400> 21
 ggaactatcc aatcaaaaann ncaaacggtt cgctatgc 38
 <210> 22
 <211> 37
 <212> DNA
 <213> 人工序列(Artificial Sequence)
 <220>
 <221>
 <222>
 <223> 饱和诱变_R_2
 <400> 22
 ctgcatctgt aatttccnnn atagtacggg tttccac 37
 <210> 23
 <211> 1017
 <212> DNA
 <213> 人工序列(Artificial Sequence)
 <220>
 <221>
 <222>
 <223> Thermobifida fusca (F12Y)
 <400> 23
 atgcaagagc tggcgttgaa acgggacctc ggctctcttg acgcagaagt agccgaactc 60
 cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
 gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
 ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
 gaaacagctt tgaatgattt tgtggettca ttaaatgcc a ttgccgacac ggcagaccct 300
 aaagtcactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
 ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
 cggattgtgg aagaagatga agaaccacag agcttggttg caattattgg tggaaacccg 480
 tactattacg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
 ggattcgcga ccgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600

acagttgcag gaagttatTTT cccgggtcgc gatatgggct ggggtgcggat tacatcagca 660
gatactgtaa caccactcgt aaatcgggtat aatgggggaa ctgttacgggt cactgggtca 720
caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
ggaactatcc aatcaaaaaa ccaaacgggt cgctatgcag aagggactgt tactggttta 840
acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattaccttt 960
ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

<210> 24

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163Y_N267D)

<400> 24

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
gaaacagctt tgaatgattt tgtggcttca ttaaattgcca ttgccgacac ggcagacctt 300
aaagtcactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
cggattgtgg aagaagatga agaaccacag agcttggctg caattattgg tggaaacccg 480
tactattacg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
ggattcgca cgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600
acagttgcag gaagttatTTT cccgggtcgc gatatgggct ggggtgcggat tacatcagca 660
gatactgtaa caccactcgt aaatcgggtat aatgggggaa ctgttacgggt cactgggtca 720
caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
ggaactatcc aatcaaaaaga ccaaacgggt cgctatgcag aagggactgt tactggttta 840
acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattaccttt 960
ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

<210> 25

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163Y_N267S)

<400> 25

```

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
gaaacagctt tgaatgattt tgttgcttca ttaaattgcca ttgccgacac ggcagaccct 300
aaagtcaactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
cggattgtgg aagaagatga agaaccacag agcttggttg caattattgg tggaaacccg 480
tactattacg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
ggattcgcga ccgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600
acagttgcag gaagttattht cccgggtcgc gatatgggct ggggtgcggat tacatcagca 660
gatactgtaa caccactcgt aaatcgggtat aatgggggaa ctgttacggg cactgggtca 720
caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
ggaactatcc aatcaaaaag tcaaacggtt cgctatgcag aagggactgt tactggttta 840
acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattacctt 960
ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

```

<210> 26

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163S_N267D)

<400> 26

```

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
gaaacagctt tgaatgattt tgttgcttca ttaaattgcca ttgccgacac ggcagaccct 300
aaagtcaactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
cggattgtgg aagaagatga agaaccacag agcttggttg caattattgg tggaaacccg 480
tactatagtg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
ggattcgcga ccgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600

```

acagttgcag gaagttatTTT CCCGGGTCGC gatatgggct gggTgcggat tacatcagca 660
gatactgtaa caccactcgt aaatcggtat aatgggggaa ctgttacggT cactgggtca 720
caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
ggaactatcc aatcaaaaaga tcaaacggtt cgctatgcag aagggactgt tactggttta 840
acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattaccttt 960
ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

<210> 27

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163S_N267T)

<400> 27

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attagggTca 120
gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
gaaacagctt tgaatgattt tgtggcttca ttaaattgcca ttgccgacac ggcagacctt 300
aaagtcactg gatggTatac cgatctcgaa agtgatgccg tagtcattac gaccttgctg 360
ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
cggattgtgg aagaagatga agaaccacag agcttggctg caattattgg tggaaacccg 480
tactatagtg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
ggattcgca cgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600
acagttgcag gaagttatTTT CCCGGGTCGC gatatgggct gggTgcggat tacatcagca 660
gatactgtaa caccactcgt aaatcggtat aatgggggaa ctgttacggT cactgggtca 720
caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
ggaactatcc aatcaaaaaac tcaaacggtt cgctatgcag aagggactgt tactggttta 840
acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattaccttt 960
ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

<210> 28

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163A_N267G)

<400> 28

```

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
gaaacagctt tgaatgattt tgtggcttca ttaaatacca ttgccgacac ggcagaccct 300
aaagtcaactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
cggattgtgg aagaagatga agaaccacag agcttggttg caattattgg tggaaacccg 480
tactatgctg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
ggattcgcga ccgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600
acagttgcag gaagttatth cccgggtcgc gatatgggct ggggtgcggat tacatcagca 660
gatactgtaa caccactcgt aaatcggtat aatgggggaa ctgttacggt cactgggtca 720
caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
ggaactatcc aatcaaaaagg tcaaacggtt cgctatgcag aagggactgt tactggttta 840
acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattacctt 960
ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

```

<210> 29

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163A)

<400> 29

```

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
gaaacagctt tgaatgattt tgtggcttca ttaaatacca ttgccgacac ggcagaccct 300
aaagtcaactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
cggattgtgg aagaagatga agaaccacag agcttggttg caattattgg tggaaacccg 480
tactatgctg gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
ggattcgcga ccgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600

```

acagttgcag gaagttatTTT cccgggtcgc gatatgggct ggggtgcggat tacatcagca 660
 gatactgtaa caccactcgt aaatcggtat aatgggggaa ctgttacgggt cactgggtca 720
 caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
 ggaactatcc aatcaaaaaa tcaaacgggt cgctatgcag aagggactgt tactggttta 840
 acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
 caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattaccttt 960
 ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

<210> 30

<211> 1017

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca (F163R)

<400> 30

atgcaagagc tggcgttgaa acgggacctc ggcctctctg acgcagaagt agccgaactc 60
 cgggctgctg aggcggaagc ggtcgagctc gaggaggagc tccgcgattc attaggggtca 120
 gacttcggcg gtgtatatct ggatgctgac accaccgaaa ttacggtcgc ggtaaccgac 180
 ccggcagcgg taagtcgtgt cgacgcggat gatgtcacag ttgatgttgt cgatttcggg 240
 gaaacagctt tgaatgattt tgtggcttca ttaaatacca ttgccgacac ggcagaccct 300
 aaagtcactg gatgggtatac cgatctcgaa agtgatgccg tagtcattac gaccttgcgt 360
 ggcgggactc ctgctgccga ggaacttgct gagagagcgg gtctcgacga aagagccgtt 420
 cggattgtgg aagaagatga agaaccacag agcttggctg caattattgg tggaaacccg 480
 tactatagag gaaattacag atgcagtatc gggtttagtg tccgtcaggg ctctcaaacg 540
 ggattcgca cgcaggcca ctgcggatcc acggggacgc gtgtgtcttc tccttcagga 600
 acagttgcag gaagttatTTT cccgggtcgc gatatgggct ggggtgcggat tacatcagca 660
 gatactgtaa caccactcgt aaatcggtat aatgggggaa ctgttacgggt cactgggtca 720
 caagaagctg ccaccggatc ctccgtttgt cgctctggag caacaacggg ctggcgctgc 780
 ggaactatcc aatcaaaaaa tcaaacgggt cgctatgcag aagggactgt tactggttta 840
 acaagaacta cagcctgtgc tgaaggtggg gattctggag ggccatggct cacaggtagc 900
 caggcgcaag gggttacaag cggcggaaca ggcgattgca gaagtggagg gattaccttt 960
 ttccaaccaa tcaatccatt gcttagctat ttcggccttc aattagtgac cggctga 1017

<210> 31

<211> 187

<212> PRT

<213> 未知(Unknown)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区

<400> 31

```

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

```

<210> 32

<211> 187

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12Y)

<400> 32

```

Leu Ala Ala Ile Ile Gly Gly Asn Pro Tyr Tyr Tyr Gly Asn Tyr Arg
1           5           10           15
Cys Ser Ile Gly Phe Ser Val Arg Gln Gly Ser Gln Thr Gly Phe Ala

```

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

<210> 33

<211> 187

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12Y_N116D)

<400> 33

Leu Ala Ala Ile Ile Gly Gly Asn Pro	Tyr Tyr Tyr Gly Asn Tyr Arg
1	5 10 15
Cys Ser Ile Gly Phe Ser Val Arg	Gln Gly Ser Gln Thr Gly Phe Ala
20	25 30
Thr Ala Gly His Cys Gly Ser Thr	Gly Thr Arg Val Ser Ser Pro Ser
35	40 45
Gly Thr Val Ala Gly Ser Tyr Phe	Pro Gly Arg Asp Met Gly Trp Val
50	55 60
Arg Ile Thr Ser Ala Asp Thr Val	Thr Pro Leu Val Asn Arg Tyr Asn

65	70	75	80
Gly Gly Thr Val Thr Val Thr Gly Ser Gln Glu Ala Ala Thr Gly Ser			
	85	90	95
Ser Val Cys Arg Ser Gly Ala Thr Thr Gly Trp Arg Cys Gly Thr Ile			
	100	105	110
Gln Ser Lys Asp Gln Thr Val Arg Tyr Ala Glu Gly Thr Val Thr Gly			
	115	120	125
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Pro			
	130	135	140
Trp Leu Thr Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Thr Gly			
145	150	155	160
Asp Cys Arg Ser Gly Gly Ile Thr Phe Phe Gln Pro Ile Asn Pro Leu			
	165	170	175
Leu Ser Tyr Phe Gly Leu Gln Leu Val Thr Gly			
	180	185	

<210> 34

<211> 187

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12Y_N116S)

<400> 34

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

115	120	125
Leu Thr Arg Thr Thr Ala Cys	Ala Glu Gly Gly Asp Ser Gly Gly Pro	
130	135	140
Trp Leu Thr Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Thr Gly		
145	150	155
Asp Cys Arg Ser Gly Gly Ile Thr Phe Phe Gln Pro Ile Asn Pro Leu		
165	170	175
Leu Ser Tyr Phe Gly Leu Gln Leu Val Thr Gly		
180	185	
<210> 35		
<211> 187		
<212> PRT		
<213> 人工序列(Artificial Sequence)		
<220>		
<221>		
<222>		
<223> Thermobifida fusca_成熟区 (F12S_N116D)		
<400> 35		
Leu Ala Ala Ile Ile Gly Gly Asn Pro Tyr Tyr Ser Gly Asn Tyr Arg		
1	5	10
Cys Ser Ile Gly Phe Ser Val Arg Gln Gly Ser Gln Thr Gly Phe Ala		
20	25	30
Thr Ala Gly His Cys Gly Ser Thr Gly Thr Arg Val Ser Ser Pro Ser		
35	40	45
Gly Thr Val Ala Gly Ser Tyr Phe Pro Gly Arg Asp Met Gly Trp Val		
50	55	60
Arg Ile Thr Ser Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn		
65	70	75
Gly Gly Thr Val Thr Val Thr Gly Ser Gln Glu Ala Ala Thr Gly Ser		
85	90	95
Ser Val Cys Arg Ser Gly Ala Thr Thr Gly Trp Arg Cys Gly Thr Ile		
100	105	110
Gln Ser Lys Asp Gln Thr Val Arg Tyr Ala Glu Gly Thr Val Thr Gly		
115	120	125
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Pro		
130	135	140
Trp Leu Thr Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Thr Gly		
145	150	155
Asp Cys Arg Ser Gly Gly Ile Thr Phe Phe Gln Pro Ile Asn Pro Leu		

	165	170	175
Leu Ser Tyr Phe Gly Leu Gln Leu Val Thr Gly			
	180	185	
<210> 36			
<211> 187			
<212> PRT			
<213> 人工序列(Artificial Sequence)			
<220>			
<221>			
<222>			
<223> Thermobifida fusca_成熟区 (F12S_N116T)			
<400> 36			
Leu Ala Ala Ile Ile Gly Gly Asn Pro Tyr Tyr Ser Gly Asn Tyr Arg			
1 5 10 15			
Cys Ser Ile Gly Phe Ser Val Arg Gln Gly Ser Gln Thr Gly Phe Ala			
20 25 30			
Thr Ala Gly His Cys Gly Ser Thr Gly Thr Arg Val Ser Ser Pro Ser			
35 40 45			
Gly Thr Val Ala Gly Ser Tyr Phe Pro Gly Arg Asp Met Gly Trp Val			
50 55 60			
Arg Ile Thr Ser Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn			
65 70 75 80			
Gly Gly Thr Val Thr Val Thr Gly Ser Gln Glu Ala Ala Thr Gly Ser			
85 90 95			
Ser Val Cys Arg Ser Gly Ala Thr Thr Gly Trp Arg Cys Gly Thr Ile			
100 105 110			
Gln Ser Lys Thr Gln Thr Val Arg Tyr Ala Glu Gly Thr Val Thr Gly			
115 120 125			
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Pro			
130 135 140			
Trp Leu Thr Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Thr Gly			
145 150 155 160			
Asp Cys Arg Ser Gly Gly Ile Thr Phe Phe Gln Pro Ile Asn Pro Leu			
165 170 175			
Leu Ser Tyr Phe Gly Leu Gln Leu Val Thr Gly			
180 185			
<210> 37			
<211> 187			
<212> PRT			

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12A_N116G)

<400> 37

```

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

```

<210> 38

<211> 187

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12A)

<400> 38

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

<210> 39

<211> 187

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12R)

<400> 39

Leu	Ala	Ala	Ile	Ile	Gly	Gly	Asn	Pro	Tyr	Tyr	Arg	Gly	Asn	Tyr	Arg
1				5					10					15	
Cys	Ser	Ile	Gly	Phe	Ser	Val	Arg	Gln	Gly	Ser	Gln	Thr	Gly	Phe	Ala
			20					25					30		
Thr	Ala	Gly	His	Cys	Gly	Ser	Thr	Gly	Thr	Arg	Val	Ser	Ser	Pro	Ser
		35					40					45			

Gly Thr Val Ala Gly Ser Tyr Phe Pro Gly Arg Asp Met Gly Trp Val		
50	55	60
Arg Ile Thr Ser Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn		
65	70	75 80
Gly Gly Thr Val Thr Val Thr Gly Ser Gln Glu Ala Ala Thr Gly Ser		
85	90	95
Ser Val Cys Arg Ser Gly Ala Thr Thr Gly Trp Arg Cys Gly Thr Ile		
100	105	110
Gln Ser Lys Asn Gln Thr Val Arg Tyr Ala Glu Gly Thr Val Thr Gly		
115	120	125
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Pro		
130	135	140
Trp Leu Thr Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Thr Gly		
145	150	155 160
Asp Cys Arg Ser Gly Gly Ile Thr Phe Phe Gln Pro Ile Asn Pro Leu		
165	170	175
Leu Ser Tyr Phe Gly Leu Gln Leu Val Thr Gly		
180	185	

<210> 40

<211> 368

<212> PRT

<213> 未知(Unknown)

<220>

<221>

<222>

<223> Thermobifida fusca

<400> 40

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

Val	Asp	Ala	Asp	Asp	Val	Thr	Val	Asp	Val	Val	Asp	Phe	Gly	Glu	Thr	
100					105					110						
Ala	Leu	Asn	Asp	Phe	Val	Ala	Ser	Leu	Asn	Ala	Ile	Ala	Asp	Thr	Ala	
115					120					125						
Asp	Pro	Lys	Val	Thr	Gly	Trp	Tyr	Thr	Asp	Leu	Glu	Ser	Asp	Ala	Val	
130					135					140						
Val	Ile	Thr	Thr	Leu	Arg	Gly	Gly	Thr	Pro	Ala	Ala	Glu	Glu	Leu	Ala	
145					150					155					160	
Glu	Arg	Ala	Gly	Leu	Asp	Glu	Arg	Ala	Val	Arg	Ile	Val	Glu	Glu	Asp	
165					170					175						
Glu	Glu	Pro	Gln	Ser	Leu	Ala	Ala	Ile	Ile	Gly	Gly	Asn	Pro	Tyr	Tyr	
180					185					190						
Phe	Gly	Asn	Tyr	Arg	Cys	Ser	Ile	Gly	Phe	Ser	Val	Arg	Gln	Gly	Ser	
195					200					205						
Gln	Thr	Gly	Phe	Ala	Thr	Ala	Gly	His	Cys	Gly	Ser	Thr	Gly	Thr	Arg	
210					215					220						
Val	Ser	Ser	Pro	Ser	Gly	Thr	Val	Ala	Gly	Ser	Tyr	Phe	Pro	Gly	Arg	
225					230					235					240	
Asp	Met	Gly	Trp	Val	Arg	Ile	Thr	Ser	Ala	Asp	Thr	Val	Thr	Pro	Leu	
245					250					255						
Val	Asn	Arg	Tyr	Asn	Gly	Gly	Thr	Val	Thr	Val	Thr	Gly	Ser	Gln	Glu	
260					265					270						
Ala	Ala	Thr	Gly	Ser	Ser	Val	Cys	Arg	Ser	Gly	Ala	Thr	Thr	Gly	Trp	
275					280					285						
Arg	Cys	Gly	Thr	Ile	Gln	Ser	Lys	Asn	Gln	Thr	Val	Arg	Tyr	Ala	Glu	
290					295					300						
Gly	Thr	Val	Thr	Gly	Leu	Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu	Gly	Gly	
305					310					315					320	
Asp	Ser	Gly	Gly	Pro	Trp	Leu	Thr	Gly	Ser	Gln	Ala	Gln	Gly	Val	Thr	
325					330					335						
Ser	Gly	Gly	Thr	Gly	Asp	Cys	Arg	Ser	Gly	Gly	Ile	Thr	Phe	Phe	Gln	
340					345					350						
Pro	Ile	Asn	Pro	Leu	Leu	Ser	Tyr	Phe	Gly	Leu	Gln	Leu	Val	Thr	Gly	
355					360					365						

<210> 41

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12Y)

<400> 41

```
ttggctgcaa ttattggtgg aaacccgtac tattacggaa attacagatg cagtatcggg 60
ttagtggtcc gtcagggtc tcaaacggga ttgcgaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcgat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacgggtcac tgggtcacaa gaagctgcca ccggtcctc cgtttgctgc 300
tctggagcaa caacgggtg gcgctgcgga actatccaat caaaaacca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catgggtcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat taccttttcc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564
```

<210> 42

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12Y_N116D)

<400> 42

```
ttggctgcaa ttattggtgg aaacccgtac tattacggaa attacagatg cagtatcggg 60
ttagtggtcc gtcagggtc tcaaacggga ttgcgaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcgat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacgggtcac tgggtcacaa gaagctgcca ccggtcctc cgtttgctgc 300
tctggagcaa caacgggtg gcgctgcgga actatccaat caaaagacca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catgggtcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat taccttttcc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564
```

<210> 43

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12Y_N116S)

<400> 43

```
ttggctgcaa ttattggtgg aaaccggtac tattacggaa attacagatg cagtatcggg 60
tttagtgtcc gtcagggctc tcaaacggga ttcgcgaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacggtcac tgggtcacia gaagctgcc a cggatcctc cgtttgctgc 300
tctggagcaa caacgggctg gcgctgcgga actatccaat caaaaagtca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catggctcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat tacctttttc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564
```

<210> 44

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12S_N116D)

<400> 44

```
ttggctgcaa ttattggtgg aaaccggtac tatagtggaa attacagatg cagtatcggg 60
tttagtgtcc gtcagggctc tcaaacggga ttcgcgaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacggtcac tgggtcacia gaagctgcc a cggatcctc cgtttgctgc 300
tctggagcaa caacgggctg gcgctgcgga actatccaat caaaagatca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catggctcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat tacctttttc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564
```

<210> 45

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12S_N116T)

<400> 45

```
ttggctgcaa ttattggtgg aaacccgtac tatagtggaa attacagatg cagtatcggg 60
tttagtgtcc gtcagggctc tcaaacggga ttcgcgaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcgat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacgggtcac tgggtcacia gaagctgcca ccggtacctc cgtttgctgc 300
tctggagcaa caacgggctg gcgctgcgga actatccaat caaaaactca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catgggtcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat tacctttttc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564
```

<210> 46

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12A_N116G)

<400> 46

```
ttggctgcaa ttattggtgg aaacccgtac tatgctggaa attacagatg cagtatcggg 60
tttagtgtcc gtcagggctc tcaaacggga ttcgcgaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcgat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacgggtcac tgggtcacia gaagctgcca ccggtacctc cgtttgctgc 300
tctggagcaa caacgggctg gcgctgcgga actatccaat caaaaggtca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catgggtcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat tacctttttc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564
```

<210> 47

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12A)

<400> 47

ttggctgcaa ttattggtgg aaacccgtac tatgctggaa attacagatg cagtatcggg 60
tttagtgtcc gtcagggctc tcaaacggga ttgcgcaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcgat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacggtcac tgggtcacia gaagctgcc a cggatcctc cgtttgtcgc 300
tctggagcaa caacgggctg gcgctgcgga actatccaat caaaaaatca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catggctcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat tacctttttc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564

<210> 48

<211> 564

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida fusca_成熟区 (F12R)

<400> 48

ttggctgcaa ttattggtgg aaacccgtac tatagaggaa attacagatg cagtatcggg 60
tttagtgtcc gtcagggctc tcaaacggga ttgcgcaccg caggccactg cggatccacg 120
gggacgcgtg tgtcttctcc ttcaggaaca gttgcaggaa gttatttccc gggtcgcgat 180
atgggctggg tgcggattac atcagcagat actgtaacac cactcgtaaa tcggtataat 240
gggggaactg ttacggtcac tgggtcacia gaagctgcc a cggatcctc cgtttgtcgc 300
tctggagcaa caacgggctg gcgctgcgga actatccaat caaaaaatca aacggttcgc 360
tatgcagaag ggactgttac tggtttaaca agaactacag cctgtgctga aggtggggat 420
tctggagggc catggctcac aggtagccag gcgcaagggg ttacaagcgg cggaacaggc 480
gattgcagaa gtggagggat tacctttttc caaccaatca atccattgct tagctatttc 540
ggccttcaat tagtgaccgg ctga 564

<210> 49

<211> 187

<212> PRT

<213> 未知(Unknown)

<220>

<221>

<222>

<223> Thermobifida cellulosilytica

<400> 49

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

<210> 50

<211> 187

<212> PRT

<213> 未知 (Unknown)

<220>

<221>

<222>

<223> Thermobifida halotolerans

<400> 50

Phe	Thr	Asp	Ile	Ile	Gly	Gly	Asn	Pro	Tyr	Tyr	Phe	Asp	Gly	Tyr	Arg
1				5					10					15	
Cys	Ser	Ile	Gly	Phe	Ser	Val	Arg	Arg	Gly	Ser	Glu	Ser	Gly	Phe	Ala
			20					25					30		
Thr	Ala	Gly	His	Cys	Gly	Glu	Glu	Gly	Thr	Glu	Thr	Ser	Asp	Pro	Glu
		35					40					45			

Gly Thr Val Ala Gly Ala Tyr Phe Pro Gly Arg Asp Met Gly Trp Val		
50	55	60
Arg Ile Thr Asp Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn		
65	70	75 80
Gly Glu Asn Val Thr Val Ala Gly Ser Arg Glu Ala Ala Thr Gly Ser		
85	90	95
Ser Val Cys Arg Ser Gly Ser Thr Thr Gly Trp Arg Cys Gly Thr Ile		
100	105	110
Arg Ser Lys Asn Gln Thr Val Arg Tyr Ile Glu Gly Thr Val Thr Gly		
115	120	125
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Pro		
130	135	140
Trp Leu Thr Gly Ser Gln Gly Gln Gly Val Thr Ser Gly Gly Ser Gly		
145	150	155 160
Asn Cys Thr Leu Gly Gly Val Thr Tyr Phe Gln Pro Leu Asn Pro Leu		
165	170	175
Leu Ser His Phe Asp Leu Asp Leu Val Thr Gly		
180	185	

<210> 51

<211> 187

<212> PRT

<213> 未知(Unknown)

<220>

<221>

<222>

<223> Actinorugispora endophytica

<400> 51

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

Ser	Val	Cys	Arg	Ser	Gly	Ser	Thr	Thr	Gly	Trp	His	Cys	Gly	Thr	Ile
			100						105					110	
Gln	Ser	Lys	Asn	Gln	Thr	Val	Asn	Tyr	Ala	Glu	Gly	Ser	Val	Ala	Gly
			115						120					125	
Leu	Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu	Gly	Gly	Asp	Ser	Gly	Gly	Ser
			130						135					140	
Trp	Leu	Thr	Gly	Thr	Gln	Ala	Gln	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly
			145						150					155	
Asn	Cys	Thr	Trp	Gly	Gly	Thr	Thr	Tyr	Phe	Gln	Pro	Ile	Asn	Pro	Leu
									165					170	
Leu	Ser	Tyr	Phe	Asn	Leu	Thr	Leu	Val	Thr	Gly					
									180					185	

<210> 52

<211> 183

<212> PRT

<213> 未知 (Unknown)

<220>

<221>

<222>

<223> Spinactinospora alkalitolerans

<400> 52

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

<210> 54

<211> 186

<212> PRT

<213> 未知 (Unknown)

<220>

<221>

<222>

<223> Nocardiosis potens

<400> 54

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

<210> 55

<211> 187

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida cellulosilytica_F12Y

<400> 55

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

<210> 56

<211> 187

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida cellulosilytica_F12YN116D

<400> 56

Phe Ala Asp Val Ile Gly Gly Asn Pro Tyr Tyr Tyr Gly Gly Tyr Arg
 1 5 10 15
 Cys Ser Ile Gly Phe Ser Val Arg Lys Gly Ser Asp Thr Gly Phe Ala

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

<210> 57

<211> 187

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida halotolerans_F12Y

<400> 57

Phe Thr Asp Ile Ile Gly Gly Asn Pro Tyr Tyr Tyr Asp Gly Tyr Arg		
1	5	10
Cys Ser Ile Gly Phe Ser Val Arg Arg Gly Ser Glu Ser Gly Phe Ala		
20	25	30
Thr Ala Gly His Cys Gly Glu Glu Gly Thr Glu Thr Ser Asp Pro Glu		
35	40	45
Gly Thr Val Ala Gly Ala Tyr Phe Pro Gly Arg Asp Met Gly Trp Val		
50	55	60
Arg Ile Thr Asp Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn		

65	70	75	80
Gly Glu Asn Val Thr	Val Ala Gly Ser Arg	Glu Ala Ala Thr	Gly Ser
	85	90	95
Ser Val Cys Arg Ser	Gly Ser Thr Thr	Gly Trp Arg Cys	Gly Thr Ile
	100	105	110
Arg Ser Lys Asn Gln Thr	Val Arg Tyr Ile	Glu Gly Thr Val	Thr Gly
	115	120	125
Leu Thr Arg Thr Thr	Ala Cys Ala Glu Gly	Gly Asp Ser Gly	Gly Pro
	130	135	140
Trp Leu Thr Gly Ser	Gln Gly Gln Gly	Val Thr Ser Gly	Gly Ser Gly
145	150	155	160
Asn Cys Thr Leu Gly	Gly Val Thr Tyr	Phe Gln Pro Leu	Asn Pro Leu
	165	170	175
Leu Ser His Phe Asp	Leu Asp Leu Val	Thr Gly	
	180	185	

<210> 58

<211> 187

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Thermobifida halotolerans_F12YN116D

<400> 58

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

115	120	125
Leu Thr Arg Thr Thr Ala Cys	Ala Glu Gly Gly Asp Ser Gly Gly Pro	
130	135	140
Trp Leu Thr Gly Ser Gln Gly Gln Gly Val Thr Ser Gly Gly Ser Gly		
145	150	155
Asn Cys Thr Leu Gly Gly Val Thr Tyr Phe Gln Pro Leu Asn Pro Leu		
165	170	175
Leu Ser His Phe Asp Leu Asp Leu Val Thr Gly		
180	185	
<210> 59		
<211> 187		
<212> PRT		
<213> 人工序列(Artificial Sequence)		
<220>		
<221>		
<222>		
<223> Actinorugispora endophytica_F12Y		
<400> 59		
Leu Ala Asn Val Ile Gly Gly Asn Ala Tyr Tyr Tyr Gly Gly Tyr Arg		
1	5	10
Cys Ser Val Gly Phe Ser Val Arg His Ser Ser Gly Pro Gly Phe Ala		
20	25	30
Thr Ala Gly His Cys Gly Asp Val Gly Thr Arg Thr Thr Ser Pro Thr		
35	40	45
Gly Thr Ile Ala Gly Ser Tyr Phe Pro Gly Arg Asp Met Gly Trp Val		
50	55	60
Arg Ile Thr Ser Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn		
65	70	75
Gly Ser Tyr Ile Thr Val Thr Gly Ser Ser Glu Ala Ala Asn Gly Ser		
85	90	95
Ser Val Cys Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile		
100	105	110
Gln Ser Lys Asn Gln Thr Val Asn Tyr Ala Glu Gly Ser Val Ala Gly		
115	120	125
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Ser		
130	135	140
Trp Leu Thr Gly Thr Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly		
145	150	155
Asn Cys Thr Trp Gly Gly Thr Thr Tyr Phe Gln Pro Ile Asn Pro Leu		

	165	170	175
Leu Ser Tyr Phe Asn Leu Thr Leu Val Thr Gly			
	180	185	
<210> 60			
<211> 187			
<212> PRT			
<213> 人工序列(Artificial Sequence)			
<220>			
<221>			
<222>			
<223> Actinorugispora endophytica_F12YN116D			
<400> 60			
Leu Ala Asn Val Ile Gly Gly Asn Ala Tyr Tyr Tyr Gly Gly Tyr Arg			
1 5 10 15			
Cys Ser Val Gly Phe Ser Val Arg His Ser Ser Gly Pro Gly Phe Ala			
20 25 30			
Thr Ala Gly His Cys Gly Asp Val Gly Thr Arg Thr Thr Ser Pro Thr			
35 40 45			
Gly Thr Ile Ala Gly Ser Tyr Phe Pro Gly Arg Asp Met Gly Trp Val			
50 55 60			
Arg Ile Thr Ser Ala Asp Thr Val Thr Pro Leu Val Asn Arg Tyr Asn			
65 70 75 80			
Gly Ser Tyr Ile Thr Val Thr Gly Ser Ser Glu Ala Ala Asn Gly Ser			
85 90 95			
Ser Val Cys Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile			
100 105 110			
Gln Ser Lys Asp Gln Thr Val Asn Tyr Ala Glu Gly Ser Val Ala Gly			
115 120 125			
Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser Gly Gly Ser			
130 135 140			
Trp Leu Thr Gly Thr Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly			
145 150 155 160			
Asn Cys Thr Trp Gly Gly Thr Thr Tyr Phe Gln Pro Ile Asn Pro Leu			
165 170 175			
Leu Ser Tyr Phe Asn Leu Thr Leu Val Thr Gly			
180 185			
<210> 61			
<211> 183			
<212> PRT			

<400> 62

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

<210> 63

<211> 186

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Nocardiosis composta_F12Y

<400> 63

Phe	Gly	Asp	Ile	Val	Gly	Gly	Asn	Ala	Tyr	Tyr	Tyr	Gly	Gly	Ser	Arg
1				5					10					15	
Cys	Ser	Ile	Gly	Phe	Ser	Val	Gln	Gly	Gly	Phe	Ala	Thr	Ala	Gly	His
			20					25					30		
Cys	Gly	Ser	Gln	Gly	Thr	Arg	Val	Thr	Gly	Gly	Ala	Gly	Glu	Ser	Gly
			35				40					45			

Thr	Val	Ala	Gly	Ser	Ile	Phe	Pro	Gly	Arg	Asp	Met	Gly	Trp	Val	Arg
50						55					60				
Val	Asn	Ser	Gly	Trp	Asn	Pro	Ser	Pro	Tyr	Val	Asn	Asn	Tyr	Ser	Gly
65					70				75					80	
Gly	Arg	Val	Leu	Val	Thr	Gly	Ser	Gln	Glu	Ala	Ser	Val	Gly	Ala	Ser
				85				90					95		
Val	Cys	Arg	Ser	Gly	Ser	Thr	Thr	Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln
			100					105					110		
Ala	Lys	Asn	Gln	Thr	Val	Arg	Tyr	Pro	Glu	Gly	Thr	Val	Asn	Gly	Leu
		115					120					125			
Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Trp
	130					135					140				
Leu	Ser	Gly	Asn	Gln	Ala	Gln	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn
145				150					155					160	
Cys	Ser	Ser	Gly	Gly	Thr	Thr	Phe	Phe	Gln	Pro	Leu	Asn	Pro	Ile	Leu
			165						170					175	
Ser	Gln	Trp	Gly	Leu	Thr	Leu	Thr	Thr	Gly						
			180					185							

<210> 64

<211> 186

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Nocardiosis composta_F12YN116D

<400> 64

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

Val	Cys	Arg	Ser	Gly	Ser	Thr	Thr	Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln	
				100					105					110		
Ala	Lys	Asp	Gln	Thr	Val	Arg	Tyr	Pro	Glu	Gly	Thr	Val	Asn	Gly	Leu	
				115					120					125		
Thr	Arg	Thr	Thr	Ala	Cys	Ala	Glu	Pro	Gly	Asp	Ser	Gly	Gly	Ser	Trp	
				130					135					140		
Leu	Ser	Gly	Asn	Gln	Ala	Gln	Gly	Val	Thr	Ser	Gly	Gly	Ser	Gly	Asn	
				145					150					155		
Cys	Ser	Ser	Gly	Gly	Thr	Thr	Phe	Phe	Gln	Pro	Leu	Asn	Pro	Ile	Leu	
				165					170					175		
Ser	Gln	Trp	Gly	Leu	Thr	Leu	Thr	Thr	Gly							
				180					185							

<210> 65

<211> 186

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> Nocardiosis potens_F12Y

<400> 65

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

<210> 67
 <211> 374
 <212> PRT
 <213> 未知 (Unknown)
 <220>
 <221>
 <222>
 <223> Thermobifida cellulosilytica
 <400> 67
 Met Asn Arg Pro Ser Thr Arg Arg Thr Val Arg Ala Leu Leu Thr Ala
 1 5 10 15
 Ala Leu Ala Ala Thr Ala Leu Thr Ala Pro Ala Ala Pro Ser Leu Ala
 20 25 30
 Gln Glu Ala Ser Gln Glu Ala Ala Leu Thr Arg Asp Leu Asp Leu Thr
 35 40 45
 Gly Thr Glu Val Ala Leu Leu Arg Ala Ala Glu Ser Glu Ala Met Asp
 50 55 60
 Arg Glu Glu Glu Leu Ser Ala Val Leu Gly Ser Asp Phe Gly Gly Val
 65 70 75 80
 Tyr Leu Ala Pro Glu Thr Gly Glu Val Thr Val Ala Val Thr Asp Pro
 85 90 95
 Ala Ala Val Pro Val Val Glu Gln Ser Gly Ala Thr Ala Gln Val Val
 100 105 110
 Thr Phe Gly Glu Thr Ala Leu Asn Asp Phe Val Asp Ser Leu Asn Ala
 115 120 125
 Val Ala Asp Arg Ala Asp Glu Gln Ile Thr Gly Trp Tyr Thr Asp Leu
 130 135 140
 Ala Ala Asp Thr Val Val Ile Thr Ala Phe Pro Gly Gly Ser Ala Ala
 145 150 155 160
 Ala Glu Glu Leu Ala Ala Leu Ala Gly Val Asp Glu Arg Ala Val Arg
 165 170 175
 Val Thr Glu Ser Ala Ala Arg Pro Gln Leu Phe Ala Asp Val Ile Gly
 180 185 190
 Gly Asn Pro Tyr Tyr Phe Gly Gly Tyr Arg Cys Ser Ile Gly Phe Ser
 195 200 205
 Val Arg Lys Gly Ser Asp Thr Gly Phe Ala Thr Ala Gly His Cys Gly
 210 215 220
 Glu Thr Gly Thr Leu Thr Arg Ser Pro Glu Gly Val Val Ala Gly Ser
 225 230 235 240

Tyr Phe Pro Gly Arg Asp Met Gly Trp Val Arg Leu Thr Gly Ala Asp
 245 250 255
 Thr Val Thr Pro Leu Val Asn Arg Tyr Asp Gly Gly Thr Val Thr Val
 260 265 270
 Thr Gly Ser Gln Glu Ala Val Thr Gly Ser Ser Val Cys Arg Ser Gly
 275 280 285
 Ser Thr Thr Gly Trp Arg Cys Gly Ile Ile Gln Ser Lys Asn Gln Thr
 290 295 300
 Val Arg Tyr Ala Glu Gly Thr Val Thr Gly Leu Thr Arg Thr Thr Ala
 305 310 315 320
 Cys Ala Glu Ala Gly Asp Ser Gly Gly Pro Trp Leu Thr Gly Ser Gln
 325 330 335
 Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly Gly
 340 345 350
 Ile Thr Tyr Phe Gln Pro Ile Asn Pro Leu Leu Ser Tyr Phe Gly Leu
 355 360 365
 Glu Leu Val Thr Gly Ala
 370

<210> 68

<211> 354

<212> PRT

<213> 未知(Unknown)

<220>

<221>

<222>

<223> Thermobifida halotolerans

<400> 68

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

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

Gly Ala

<210> 69

<211> 383

<212> PRT

<213> 未知 (Unknown)

<220>

<221>

<222>

<223> Actinorugispora endophytica

<400> 69

Met	Lys	Arg	Ser	Ser	Val	Phe	Arg	Ala	Leu	Gly	Gly	Thr	Ile	Leu	Thr
1				5					10					15	
Ala	Gly	Leu	Val	Ile	Thr	Ala	Ala	Pro	Phe	Ala	Ser	Ala	Ala	Pro	Val
			20					25						30	
His	Thr	Glu	Thr	Thr	Pro	Thr	Ala	Ala	Glu	Ala	Gly	Asp	Gln	Leu	Ser
		35					40					45			
Ala	Leu	Lys	Arg	Asp	Leu	Gly	Leu	Ser	Thr	Ala	Glu	Val	Glu	Glu	Leu
		50				55					60				
Gln	Ala	Ala	Glu	Ala	Glu	Ala	Met	Asp	Val	Glu	Glu	Gly	Leu	Arg	Glu
65					70					75					80
Thr	Leu	Gly	Ser	Asp	Phe	Gly	Gly	Ala	His	Phe	Asp	Ile	Asp	Ser	Gly
				85					90					95	
Glu	Leu	Thr	Val	Ser	Val	Thr	Asp	Ala	Ala	Ala	Val	Ser	Thr	Val	Glu
			100					105						110	
Ala	Ala	Gly	Ala	Asn	Ala	Glu	Val	Val	Asp	Phe	Gly	Glu	Pro	Ala	Leu
			115					120				125			
Asp	Ala	Ile	Val	Glu	Asp	Leu	Asn	Thr	Val	Ala	Glu	Glu	Ala	Asp	Asp
		130					135					140			
Ser	Val	Thr	Gly	Trp	Tyr	Val	Asp	Thr	Ala	Asp	Asp	Ser	Val	Val	Ile
145					150					155					160
Thr	Val	Leu	Glu	Gly	Asp	Thr	Glu	Ala	Ala	Glu	Ala	Leu	Val	Ala	Glu
				165					170					175	
Ala	Asp	Val	Asp	Gly	Lys	Ala	Val	Arg	Val	Glu	Glu	Thr	Thr	Glu	Gln
			180					185						190	
Pro	Lys	Leu	Leu	Ala	Asn	Val	Ile	Gly	Gly	Asn	Ala	Tyr	Tyr	Phe	Gly
		195					200					205			
Gly	Tyr	Arg	Cys	Ser	Val	Gly	Phe	Ser	Val	Arg	His	Ser	Ser	Gly	Pro
		210				215					220				
Gly	Phe	Ala	Thr	Ala	Gly	His	Cys	Gly	Asp	Val	Gly	Thr	Arg	Thr	Thr
225					230					235					240
Ser	Pro	Thr	Gly	Thr	Ile	Ala	Gly	Ser	Tyr	Phe	Pro	Gly	Arg	Asp	Met
				245					250					255	
Gly	Trp	Val	Arg	Ile	Thr	Ser	Ala	Asp	Thr	Val	Thr	Pro	Leu	Val	Asn
			260					265					270		
Arg	Tyr	Asn	Gly	Ser	Tyr	Ile	Thr	Val	Thr	Gly	Ser	Ser	Glu	Ala	Ala
		275						280					285		

Asn Gly Ser Ser Val Cys Arg Ser Gly Ser Thr Thr Gly Trp His Cys																	
290							295								300		
Gly Thr Ile Gln Ser Lys Asn Gln Thr Val Asn Tyr Ala Glu Gly Ser																	
305						310					315					320	
Val Ala Gly Leu Thr Arg Thr Thr Ala Cys Ala Glu Gly Gly Asp Ser																	
					325					330						335	
Gly Gly Ser Trp Leu Thr Gly Thr Gln Ala Gln Gly Val Thr Ser Gly																	
					340					345						350	
Gly Ser Gly Asn Cys Thr Trp Gly Gly Thr Thr Tyr Phe Gln Pro Ile																	
					355					360						365	
Asn Pro Leu Leu Ser Tyr Phe Asn Leu Thr Leu Val Thr Gly Ala																	
					370					375						380	
<210> 70																	
<211> 375																	
<212> PRT																	
<213> 未知 (Unknown)																	
<220>																	
<221>																	
<222>																	
<223> Spinactinospora alkalitolerans																	
<400> 70																	
Met Arg Lys Ser Pro Ile Ile Arg Ala Val Gly Gly Ala Ala Ile Thr																	
1					5					10					15		
Phe Gly Leu Val Ile Ala Ala Ala Pro Phe Ala Ser Ala Asp Ser Gly																	
					20					25					30		
Ser Glu Thr Thr Ala Gly Ser Val Gly Gln Leu Gly Ala Met Gln Arg																	
					35					40					45		
Asp Leu Gly Leu Ser Ala Thr Glu Ala Thr Ala Leu Leu Asp Gln Glu																	
					50					55					60		
Glu Gln Ala Arg Thr Leu Glu Gly Glu Leu Arg Glu Thr Leu Gly Gly																	
65						70					75					80	
Asp Phe Gly Gly Ala Val Phe Asp Ile Glu Ser Gly Glu Leu Thr Val																	
					85					90						95	
Ser Val Thr Asp Glu Asp Ala Val Asp Glu Val Arg Glu Ala Gly Ala																	
					100					105						110	
Glu Ala Glu Val Val Thr Tyr Gly Glu Gln Arg Leu Asp Ala Ile Val																	
					115					120						125	
Asp Asp Leu Asn Ala Thr Glu Asp Thr Ala Asp Glu Ser Val Thr Gly																	
					130					135						140	

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

<210> 71
 <211> 376
 <212> PRT
 <213> 未知 (Unknown)
 <220>
 <221>
 <222>
 <223> Nocardiosis composta
 <400> 71

Met	Arg	Lys	Ser	Pro	Tyr	Ile	Pro	Leu	Leu	Gly	Ala	Ser	Val	Leu	Ala
1				5					10					15	
Leu	Gly	Met	Ile	Ala	Ala	Ser	Pro	Thr	Ala	Ala	Ser	Ala	Asp	Glu	Ala
			20					25					30		
Thr	Asp	Ser	Ser	Pro	Ala	Arg	Ala	Leu	Ala	Ser	Gly	Leu	Asp	Met	Ser
		35					40					45			
Thr	Ala	Gln	Ala	Ala	Glu	Leu	Leu	Asp	Ala	Glu	Ala	Gln	Ala	Arg	Thr
	50					55					60				
Ala	Glu	Gln	Glu	Ala	Arg	Glu	Leu	Ala	Gly	Ala	Ser	Phe	Ala	Gly	Ala
65					70				75					80	
Val	Phe	Asp	Ala	Asp	Thr	Arg	Lys	Leu	Thr	Val	Ser	Val	Thr	Asp	Ala
			85					90					95		
Ala	Ala	Ala	Glu	Ala	Val	Gln	Ala	Thr	Gly	Ala	Glu	Thr	Arg	Val	Val
			100					105					110		
Glu	Ala	Ser	Ala	Asp	Glu	Leu	Asp	Ala	Ala	Val	Ala	Asp	Leu	Asn	Ala
		115					120					125			
Glu	Glu	Arg	Gly	Leu	Gly	Ser	Glu	Ile	Asp	Gly	Val	Thr	Gly	Trp	Tyr
	130					135					140				
Val	Asp	Gln	Ala	Ala	Asn	Glu	Leu	Val	Val	Thr	Val	Leu	Asp	Gly	Glu
145					150				155					160	
Thr	Glu	Ala	Ala	Glu	Thr	Leu	Leu	Asp	Glu	Ala	Gly	Val	Asp	Ser	Val
			165					170					175		
Pro	Val	Arg	Val	Asp	Gln	Gly	Ala	Glu	Gln	Pro	Glu	Thr	Phe	Gly	Asp
		180						185					190		
Ile	Val	Gly	Gly	Asn	Ala	Tyr	Tyr	Pro	Gly	Gly	Ser	Arg	Cys	Ser	Ile
	195						200					205			
Gly	Phe	Ser	Val	Gln	Gly	Gly	Phe	Ala	Thr	Ala	Gly	His	Cys	Gly	Ser
	210					215					220				
Gln	Gly	Thr	Arg	Val	Thr	Gly	Gly	Ala	Gly	Glu	Ser	Gly	Thr	Val	Ala
225					230				235					240	
Gly	Ser	Ile	Phe	Pro	Gly	Arg	Asp	Met	Gly	Trp	Val	Arg	Val	Asn	Ser
			245					250					255		
Gly	Trp	Asn	Pro	Ser	Pro	Tyr	Val	Asn	Asn	Tyr	Ser	Gly	Gly	Arg	Val
		260					265					270			
Leu	Val	Thr	Gly	Ser	Gln	Glu	Ala	Ser	Val	Gly	Ala	Ser	Val	Cys	Arg
	275						280				285				
Ser	Gly	Ser	Thr	Thr	Gly	Trp	Arg	Cys	Gly	Thr	Ile	Gln	Ala	Lys	Asn
	290					295					300				
Gln	Thr	Val	Arg	Tyr	Pro	Glu	Gly	Thr	Val	Asn	Gly	Leu	Thr	Arg	Thr

305	310	315	320
Thr Ala Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Trp Leu Ser Gly			
	325	330	335
Asn Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ser			
	340	345	350
Gly Gly Thr Thr Phe Phe Gln Pro Leu Asn Pro Ile Leu Ser Gln Trp			
	355	360	365
Gly Leu Thr Leu Thr Thr Gly Ala			
	370	375	
<210> 72			
<211> 376			
<212> PRT			
<213> 未知 (Unknown)			
<220>			
<221>			
<222>			
<223> Nocardiosis potens			
<400> 72			
Met Arg Lys Ser Pro Tyr Ile Ser Phe Leu Gly Ala Ser Ala Leu Ala			
1	5	10	15
Leu Gly Met Ile Ala Ala Ser Pro Ala Ala Ala Ser Ala Asp Glu Ala			
	20	25	30
Ala Asp Thr Ser Pro Ala Glu Ala Leu Ala Ser Gly Leu Asp Met Ser			
	35	40	45
Ala Ser Gln Ala Ala Asp Leu Leu Asp Ala Glu Ala Glu Ala Arg Gly			
	50	55	60
Thr Glu Ala Glu Ala Arg Glu Ala Ala Gly Gly Ser Phe Ala Gly Ala			
65	70	75	80
Val Phe Asp Ala Glu Ser Gln Val Leu Thr Val Ser Val Thr Asp Ala			
	85	90	95
Ala Ala Ala Glu Ala Val Glu Ala Thr Gly Ala Glu Thr Arg Val Val			
	100	105	110
Glu Ala Ser Glu Asp Glu Leu Asp Ser Ala Val Ser Asp Leu Asn Ala			
	115	120	125
Glu Glu Ser Ser Leu Gly Ser Ala Ile Glu Gly Val Thr Gly Trp Tyr			
	130	135	140
Val Asp Pro Ala Ala Asn Glu Val Val Val Thr Val Leu Asp Gly Glu			
145	150	155	160
Thr Ala Ala Ala Glu Thr Leu Leu Asp Glu Ala Gly Val Asp Gly Val			

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

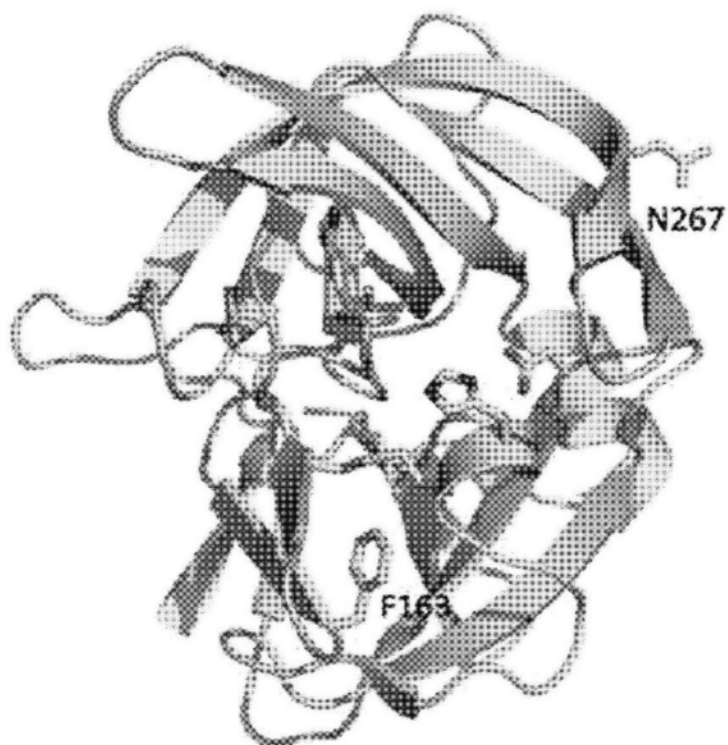


图1