[19]中华人民共和国国家知识产权局

[51]Int.Cl[7]
C12N 15/57
C12N 9/54
C12N 1/21

[12]发明专利说明书

[21] ZL 专利号 94193785.2

[45]授权公告日 2000 年 9 月 20 日

[11]授权公告号 CN 1056649C

[22]申请日 1994.10.13 [24]颁证日 2000.6.30
[21]申请号 94193785.2
[30]优先权
[32]1993.10.14 [33]US [31]08/137,240
[86]国际申请 PCT/US94/11562 1994.10.13
[87]国际公布 WO95/10615 英 1995.4.20
[85]进入国家阶段日期 1996.4.15
[73]专利权人 吉恩康国际股份有限公司
地址 美国纽约州
[72]发明人 T·P·格雷卡 R·R·博特
L·J·威尔逊
[56]参考文献
EP405901 1991. 1. 2 C11D3/386
审查员 曾繁辉

[74]专利代理机构 上海专利商标事务所
代理人 林蕴和

权利要求书 4 页 说明书 53 页 附图页数 20 页

[54]发明名称 枯草杆菌蛋白酶变异体

[57]摘要

公开了衍生自天然或重组的非人体羰基水解酶的新变异体。一般,这种变异羰 基水解酶是通过体外修饰编码天然或重组羰基水解酶的前体 DNA 序列,使其 仅编码前体氨基酸序列中的多个取代的氨基酸残基。与前体水解酶相比,变异 羰基水解酶具有不同的性质如不同的蛋白水解活性,稳定性等。被取代的氨基 酸残基对应于解淀粉芽孢杆菌枯草杆菌蛋白酶残基 + 76 位置的残基结合一个或 多个选自下组的残基:+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265 和/或 +274。



ISSN1008-4274

权 利 要 求 书

1.一种枯草杆菌蛋白酶变异体，其特征在于，它具有衍生自前体枯草杆菌蛋白酶的、在自然界没有发现的氨基酸序列，它在相当于下列位置的多个氨基酸残基处具有不同的氨基酸取代:

76/99,76/101,76/103,76/104,76/107,76/123,

76/99/101,76/99/103,76/99/104,76/101/103,76/101/104,

76/101/104,76/104/107,76/104/123,76/107/123,76/99/101/103,

76/99/101/104,76/99/103/104,76/101/103/104,76/103/104/123,

76/104/107/123,76/99/101/103/104,76/99/103/104/123,

76/99/101/103/104/123,76/103/104/126,76/103/104/135,

76/103/104/197,76/103/104/222,76/103/104/260,

76/103/104/265,76/103/104/126/265,27/76/104/123/274,

27/76/104/109/123/274,27/76/104/123/218/274,27/76/104/123,

27/76/104/107/123,27/76/104/109/123,

27/76/104/109/123/218/274,27/76/104/123/197,

27/76/104/123/204,27/76/104/123/206,27/76/104/123/216,

27/76/104/123/218,27/76/104/123/260,27/76/104/123/195/197,

27/76/104/123/195/218,27/76/104/123/197/218,

27/76/104/123/204/218,27/76/104/123/206/218,

27/76/104/123/218/260,27/76/104/123/195/197/218,

76/103/104/217,76/103/104/156,76/103/104/166,

76/103/104/105,76/101/103/104,76/103/104/128,

76/103/104/210,76/103/104/107,76/103/104/204,76/217,

和 76/103/104/156/166,

其中，上述位置的取代氨基酸如下:

氨基酸残基位置	取代氨基酸
+76	D,H
+99	D,T,N,G

+101	R,G,D,K,L,A,E
+103	A,Q,T,D,E,Y,K,G,R
+104	I,Y,S,L,A,T,G,F,M,W,D,V,N
+107	V,L,Y,G,F,T,S,A,N
+123	S,T,I
+27	K
+105	A,D
+109	S,K,R
+126	A,I,V,F
+128	G,L
+135	I,A,S,F,V
+156	E,D,Q
+166	D,G,E,K,N,A,F,I,V,L
+195	E
+197	E
+204	A,G,C
+206	L
+210	I,S,C
+216	V
+217	H,I,Y,C,A,G,F,S,N,E,K
+218	S
+222	A,Q,S,C,I,K
+260	P,A,S,N,G
+265	N,A,G,S
+274	A,S 。

2.如权利要求 1 所述的枯草杆菌蛋白酶变异体，其特征在于取代组合是在相当于下列的位置上进行:

76/99,76/101,76/103,76/104,76/107,76/123,

76/99/101,76/99/103,76/99/104,76/101/103,76/101/104,

76/103/104,76/104/107,76/104/123,76/107/123,76/99/101/103,
76/99/101/104,76/99/103/104,76/101/103/104,76/103/104/123,
76/104/107/123,76/99/101/103/104,76/99/103/104/123;
76/99/101/103/104/123;76/103/104/128;76/103/104/260;
76/103/104/265;76/103/104/197;76/103/104/105;
76/103/104/135;76/103/104/126;76/103/104/107;
76/103/104/210;76/103/104/126/265 和/或 76/103/104/222 。

3.如权利要求 2 所述的枯草杆菌蛋白酶变异体，其特征在于，取代组合是在相当于下列的位置上进行：76/99, 76/104, 76/99/104, 76/103/104, 76/104/107, 76/101/103/104, 76/99/103/104 和 76/101/104 。

4.如权利要求 2 所述的枯草杆菌蛋白酶变异体，其特征在于，枯草杆菌蛋白酶变异体含有：

N76D/S99D;N76D/S101R;
N76D/S103A;N76D/V104I;N76D/I107V;N76D/N123S;
N76D/S99D/S101R;N76D/S99D/S103A;N76D/S99D/V104I;
N76D/S101R/S103A;N76D/S101R/V104I;N76D/S103A/V104I;
N76D/V104I/I107V;N76D/V104Y/I107V;N76D/V104I/N123S;
N76D/I107V/N123S;N76D/S99D/S101R/S103A;
N76D/S99D/S101R/V104I;N76D/S99D/S103A/V104I;
N76D/S101R/S103A/V104I;N76D/S103A/V104I/N123S;
N76D/V104I/I107V/N123S;N76D/S99D/S101R/S103A/V104I;
N76D/S99D/S103A/V104I/N123S;
N76D/S99D/S101R/S103A/V104I/N123S;N76D/S103A/V104I/S128G;
N76D/S103A/V104I/T260P;N76D/S103A/V104I/S265N;
N76D/S103A/V104I/D197E;N76D/S103A/V104I/S105A;
N76D/S103A/V104I/L135I;N76D/S103A/V104I/L126F;
N76D/S103A/V104T/L107T;N76D/S103A/V104I/P210I;
N76D/S103A/V104I/L126F/S265N 和 N76D/S103A/V104I/M222A 。

5.如权利要求 4 所述的枯草杆菌蛋白酶变异体，其特征在于，枯草杆菌蛋白酶变异体含有：N76D/S99D，N76D/V104I，N76D/S99D/V104I，N76D/S103A/V104I，N76D/V104I/I107V，N76D/V104Y/I107V，N76D/S101R/S103A/V104I，N76D/S99D/S101R/S103A/V104I 和 N76D/S101R/V104I。

6.如权利要求 2 所述的枯草杆菌蛋白酶变异体，其特征在于，它衍生自杆菌属枯草杆菌蛋白酶。

7.如权利要求 6 所述的枯草杆菌蛋白酶变异体，其特征在于，它衍生自透镜状芽孢杆菌枯草杆菌蛋白酶。

8.编码权利要求 1 所述的枯草杆菌蛋白酶变异体的 DNA。

说　明　书

枯草杆菌蛋白酶变异体

相关申请

本申请是1993年10月14日申请的美国申请No. 08/137,240(未审定)的部分续展申请,原案申请在此全部引用作为参考。

发明领域

本发明涉及新的羰基水解酶变异体,其氨基酸序列中相应于前体羰基水解酶的多个氨基酸残基,尤其是那些对应于或相当于解淀粉芽孢杆菌(Bacillus amyloliquefaciens)枯草杆菌蛋白酶+76位置的残基并结合一个或多个选自下组的残基:+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265和/或+274,被不同的氨基酸所取代。一般,这种突变/变异羰基水解酶的获得是通过体外修饰编码天然或重组羰基水解酶的前体DNA序列,使其仅编码前体氨基酸序列中取代的多个氨基酸残基,或同时还编码前体氨基酸序列中的其它取代、插入或缺失。

发明背景

丝氨酸蛋白酶是羰基水解酶的一个亚类。它们包括一组不同的、具有广谱特异性和生物学功能的酶。参见Stroud, R. Sci. Amer., 131:74—88。尽管它们功能各异,但是丝氨酸蛋白酶的的催化机制已由对至少两类遗传学上截然不同家族酶的研究所提出:枯草杆菌蛋

白酶和相关的哺乳动物胰凝乳蛋白酶以及同源的细菌丝氨酸蛋白酶(如胰蛋白酶和 S. gresius 胰蛋白酶)。这两个家族的丝氨酸蛋白酶具有非常类似的催化机制。参见 Kraut, J. (1977), Ann. Rev. Biochem., 46:331－358。此外,尽管这两类酶家族的一级结构不相关,但是它们的三级结构都形成由丝氨酸、组氨酸和天冬氨酸组成的保守性催化三氨基酸结构。

枯草杆菌蛋白酶是一种丝氨酸内切蛋白酶(分子量 27,500),各种不同杆菌属种和其他微生物都大量分泌该酶。已经确定了至少 4 种不同杆菌属种的枯草杆菌蛋白酶的蛋白质序列。参见 Markland, F. S., et al. (1983), Honne－Seyler's Z. Physiol. Chem., 364: 1537－1540。已经报道了分辨率为 2.5 埃的解淀粉芽孢杆菌枯草杆菌蛋白酶的三维晶体结构。参见 Wright, C.S., et al. (1969) Nature, 221: 235 － 242; Drenth, J., et al. (1972), Eur. J. Biochem., 26:177－181。这些研究表明,尽管枯草杆菌蛋白酶与哺乳动物丝氨酸蛋白酶在遗传学上是不相关的,但是具有类似的活性部位结构。含有共价相连的肽抑制剂的枯草杆菌蛋白酶(Robertus, J.D., et al. (1972), Biochemistry, 11:2439－2449)或产物复合物(Robertus, J.D., et al. (1976), J. Biol Chem., 251:1097－1103)的 X 射线晶体结构提供了有关枯草杆菌蛋白酶活性部位和假设的底物结合裂隙的信息。此外,已报道了大量枯草杆菌蛋白酶动力学和化学修饰方面的研究(Philipp, M., et al. (1983), Mol. Cell. Biochem.., 51: 5 － 32; Svendsen, B. (1976), Carlsberg Res. Comm., 41:237－291; Markland, F.S. Id.), 而且至少有一则报道说枯草杆菌残基 222 处的甲硫氨酸侧链被过氧化氢转变成甲硫氨酸亚砜(Stauffer, D.C., et al. (1965), J. Biol. Chem., 244:5333－5338), 并且残基 221 处的丝氨酸侧链用化学修饰法转变成半胱氨酸(Polgar, et al. (1981), Biochimica et Biophsica Acta, 667:

351－354.)。

美国专利 4,760,025(RE 34,606)公开了对枯草杆菌蛋白酶氨基酸残基的修饰，这些残基对应于解淀粉芽孢杆菌枯草杆菌蛋白酶位置中的酪氨酸－1、天冬氨酸＋32、天冬酰胺＋155、酪氨酸＋104、甲硫氨酸＋222、甘氨酸＋166、组氨酸＋64、甘氨酸＋169、苯丙氨酸＋189、丝氨酸＋33、丝氨酸＋221、酪氨酸＋217、谷氨酸＋156 和丙氨酸＋152。美国专利 5,182,204 公开了解淀粉芽孢杆菌枯草杆菌蛋白酶中氨基酸残基＋224 及其他枯草杆菌蛋白酶中相应位置的修饰情况，其中是通过取代、插入或缺失而进行修饰而且可以与美国专利 4,760,025(RE 34,606)所述的残基修饰结合以形成有用的枯草杆菌蛋白酶突变体或变异体。美国专利 5,155,033 公开了类似的突变型枯草杆菌蛋白酶，其中它在相应于解淀粉芽孢杆菌枯草杆菌蛋白酶＋225 位有修饰。美国专利 5,185,285 和 5,204,015 公开了在＋123 和/或＋274 位有修饰的突变型枯草杆菌蛋白酶。这些专利的公开内容在此引用作为参考，同时引用作为参考的还有美国专利申请 07/898,382。该申请公开了枯草杆菌蛋白酶中许多氨基酸残基的修饰情况，具体包括＋99、＋101、＋103、＋107、＋126、＋128、＋135、＋197 和＋204。所有这些专利/申请都是共同拥有的。美国专利 4,914,031 公开了某些枯草杆菌蛋白酶类似物，包括在＋76 位被修饰的枯草杆菌蛋白酶。该专利的公开内容也在此引用作为参考。但是，本文指明的特定残基和/或此处提出权利要求的具体组合并没有在这些文献中提及。

因此，本发明的一个目的是提供一种羰基水解酶(最好是枯草杆菌蛋白酶)变异体，它含有 DNA 编码的前体羰基水解酶中多个取代的氨基酸残基，这些取代对应于或相当于解淀粉芽孢杆菌枯草杆菌蛋白酶残基＋76 位置的残基并结合一个或多个选自下组的残基：＋99、＋101、＋103、＋104、＋107、＋123、＋27、＋105、＋109、＋126、＋

128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265 和/或+274，被不同的氨基酸所取代。与衍生出该变异体氨基酸的羰基水解酶前体比较，这些变异体通常至少有一种不同于前体相应特性的性质。

本发明的另一目的是提供编码该羰基水解酶变异体的 DNA 序列，以及含有该突变 DNA 序列的表达载体。

此外，本发明的另一目的是提供用该载体转化的宿主细胞，以及能够表达该 DNA 从而能在胞内或胞外产生羰基水解酶变异体的宿主细胞。

列出上述文献仅仅是因为它们是在本申请的申请日之前公开的，并不意味着本发明的发明人不能凭借先发明或基于更早申请的优先权而位于这些公开文献之前。

发明概述

本发明包括非天然产生的羰基水解酶变异体，与衍生出该变异体氨基酸的前体羰基水解酶比较，它们具有不同的蛋白水解活性、稳定性、底物特异性、pH 分布图和/或工作性能。前体羰基水解酶可以是天然产生的羰基水解酶或重组的羰基水解酶。具体地，这种羰基水解酶变异体具有自然界没有发现的氨基酸序列，它是通过用不同的氨基酸置换前体羰基水解酶的多个氨基酸衍生而得。前体酶的这些多个氨基酸残基对应于+76 位并结合一个或多个选自下组的残基：+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265 和/或+274，其中编号位置对应于天然解淀粉芽孢杆菌枯草杆菌蛋白酶或相应于其他羰基水解酶或枯草杆菌蛋白酶如透镜状芽孢杆菌枯草杆菌蛋白酶的氨基酸残基。本发明的羰基水解酶变异体含有氨基酸残基+76 处的置换并结

合一个或多个其他修饰。较佳地，本发明的变异酶含有下列组合氨基酸残基的取代、缺失或插入：76/99；76/101；76/103；76/104；76/107；76/123；76/99/101；76/99/103；76/99/104；76/101/103；76/101/104；76/103/104；76/104/107；76/104/123；76/107/123；76/99/101/103；76/99/101/104；76/99/103/104；76/101/103/104；76/103/104/123；76/104/107/123；76/99/101/103/104；76/99/103/104/123；76/99/101/103/104/123；76/103/104/128；76/103/104/260；76/103/104/265；76/103/104/197；76/103/104/105；76/103/104/135；76/103/104/126；76/103/104/107；76/103/104/210；76/103/104/126/265 和/或 76/103/104/222。最佳地，本发明的变异酶含有下列残基组合的氨基酸残基取代、缺失或插入：解淀粉芽孢杆菌枯草杆菌蛋白酶的 76/99；76/104；76/99/104；76/103/104；76/104/107；76/101/103/104；76/99/101/103/104 和 76/101/104。

本发明还包括编码这种羰基水解酶或枯草杆菌蛋白酶变异体的变异 DNA 序列。这些变异 DNA 序列衍生自编码天然存在的或重组前体酶的前体 DNA 序列。变异 DNA 序列是通过修饰前体 DNA 序列使其编码一个或多个具体取代的氨基酸残基而得到，其中这些被前体 DNA 序列编码的氨基酸残基对应于解淀粉芽孢杆菌中 76、99、101、103、104、107、123、27、105、109、126、128、135、156、166、195、197、204、206、210、216、217、218、222、260、265 和/或 274 或其任何组合。尽管确定的被修饰氨基酸残基与可应用于解淀粉芽孢杆菌的编号相同(这种编号方法已成为确定所有枯草杆菌蛋白酶中残基位置的常规方法。)，但是用于本发明的优选前体 DNA 序列是如图 6 所示的透镜状芽孢杆菌的 DNA 序列(SEQ ID No. 11)。

本发明的变异 DNA 序列编码 76 位插入或取代的氨基酸残基并结合一个或多个其他修饰。较佳地，变异 DNA 序列编码下列组合

的取代或插入氨基酸残基:76/99; 76/101; 76/103; 76/104; 76/107; 76/123; 76/99/101; 76/99/103; 76/99/104; 76/101/103; 76/101/104; 76/103/104; 76/104/107; 76/104/123; 76/107/123; 76/99/101/103; 76/99/101/104; 76/99/103/104; 76/101/103/104; 76/103/104/123; 76/104/107/123; 76/99/101/103/104; 76/99/103/104/123; 76/99/101/103/104/123; 76/103/104/128; 76/103/104/260; 76/103/104/265; 76/103/104/197; 76/103/104/105; 76/103/104/135; 76/103/104/126; 76/103/104/107; 76/103/104/210; 76/103/104/126/265 和/或 76/103/104/222。最佳地,变异 DNA 序列编码下列组合修饰的残基:76/99; 76/104; 76/99/104; 76/103/104; 76/104/107; 76/101/103/104; 76/99/101/103/104 和 76/101/104。这些重组 DNA 序列编码的羰基水解酶变异体具有新的氨基酸序列,而且与前体羰基水解酶 DNA 序列编码的酶相比,通常至少有一种基本上不同于前体酶相应特性的性质。这些性质包括蛋白水解活性、底物特异性、稳定性、改变的 pH 分布图和/或增强的工作性能。

本发明包括在指定的氨基酸残基位置处的 19 种天然 L—氨基酸的任何取代。这种取代可以在任何前体枯草杆菌蛋白酶上进行(原核的、真核的、哺乳动物的等)。较佳地,取代是在各确定的氨基酸残基位置进行,包括(但并不限于)这些位置:76 位的取代包括 D、H、E、G、F、K、P 和 N;99 位的取代包括 D、T、N、Q、G 和 S;101 位的取代包括 G、D、K、L、A、E、S 和 R;103 位的取代包括 Q、T、D、E、Y、K、G、R、S 和 A;76 位的取代包括 D、H、E、G、F、K、P 和 N;104 位的取代包括所有 19 种天然氨基酸;107 位的取代包括 V、L、M、Y、G、E、F、T、S、A、N 和 I;123 位的取代包括 N、T、I、G、A、C 和 S;27 位的取代包括 K、N、C、V 和 T;105 位的取代包括 A、D、G、R 和 N;107 位的取代包括 A、L、V、Y、G、G、T、S 和 A;109 位的取代包括 S、K、

R、A、N 和 D;126 位的取代包括 A、F、I、V 和 G;128 位的取代包括 G、L 和 A;135 位的取代包括 A、F、I、S 和 V;156 位的取代包括 D、E、A、G、Q 和 K;166 位的取代包括所有 19 种天然氨基酸;195 位的取代包括 E;197 位的取代包括 E;204 位的取代包括 A、G、C、S 和 D;206 位的取代包括 L、Y、N、D 和 E;210 位的取代包括 L、I、S、C 和 F;216 位的取代包括 V、E、T 和 K;217 位的取代包括所有 19 种天然氨基酸;218 位的取代包括 S、A、G、T 和 V;222 位的取代包括所有 19 种天然氨基酸;260 位的取代包括 P、N、G、A、S、C、K 和 D;265 位的取代包括 N、G、A、S、C、K、Y 和 H;和 274 位的取代包括 A 和 S。在各位置上特别优选的取代氨基酸列于表 I。尽管在表 I 中列出了具体的氨基酸,但是应理解,所确定的残基可用任何氨基酸进行取代。

表 I

氨基酸残基	优选的取代/插入氨基酸
+76	D,H
+99	D,T,N,G
+101	R,G,D,K,L,A,E
+103	A,Q,T,D,E,Y,K,G,R
+104	I,Y,S,L,A,T,G,F,M,W,D,V,N
+107	V,L,Y,G,F,T,S,A,N
+123	S,T,I
+27	K
+105	A,D
+109	S,K,R
+126	A,I,V,F
+128	G,L
+135	I,A,S

+156	E,D,Q
+166	D,G,E,K,N,A,F,I,V,L
+195	E
+197	E
+204	A,G,C
+206	L
+210	I,S,C
+216	V
+217	H,I,Y,C,A,G,F,S,N,E,K
+218	S
+222	A,Q,S,C,I,K
+260	P,A,S,N,G
+265	N,A,G,S
+274	A,S

此外，本发明包括含有该变异体羰基水解酶 DNA 序列的表达载体及用该载体转化的、能够产生该变异体的宿主细胞。本发明还涉及含有本发明羰基水解酶变异体的去垢剂组合物。

附图简述

图 1A－C 是解淀粉芽孢杆菌枯草杆菌蛋白酶的 DNA 和氨基酸序列以及该基因(SEQ ID No. 6)的部分限制酶图谱。

图 2 显示了解淀粉芽孢杆菌(BPN′)和透镜状芽孢杆菌(野生型)的枯草杆菌蛋白酶中保守氨基酸残基。

图 3A 和 3B 显示了 4 种枯草杆菌蛋白酶的氨基酸序列。最上面一行是解淀粉芽孢杆菌枯草杆菌蛋白酶(有时也称为枯草杆菌蛋白酶 BPN′)的氨基酸序列(SEQ ID No. 7)。第二行是枯草杆菌(Bacil-

lus subtilis)的枯草杆菌蛋白酶氨基酸序列(SEQ ID No. 8)。第三行是地衣芽孢杆菌(Bacillus licheniformis)的枯草杆菌蛋白酶氨基酸序列(SEQ ID No. 9)。第四行是透镜状芽孢杆菌的枯草杆菌蛋白酶(在 PCT WO89/06276 中也称为枯草杆菌蛋白酶 309)氨基酸序列(SEQ ID No. 10)。符号"*"表示与枯草杆菌 BPN′相比,缺失特定的氨基酸残基。

图 4 是质粒 GGA274 的构建图。

图 5 是质粒 GGT274 的构建图,该质粒是本申请中所用的某些表达质粒的中间质粒。

图 6A 和 6B 是透镜状芽孢杆菌的枯草杆菌蛋白酶的 DNA 和氨基酸序列(SEQ ID No. 11)。成熟的枯草杆菌蛋白酶由从对应于 Ala 的密码子 GCG(334－336)开始的密码子所编码。

图 7A 和 7B 是本发明的一个优选例子(N76D/S103A/V104I)的 DNA 和氨基酸序列(SEQ ID No. 12)。该图中的 DNA 用所述的方法进行修饰,从而编码 76 位的天冬氨酸、103 位的丙氨酸和 104 位的异亮氨酸。成熟的枯草杆菌蛋白酶变异体由从对应于 Ala 的密码子 GCG(334－336)开始的密码子所编码。

图 8 是载体 pBCDAICAT 的构建图。

图 9 是载体 pUCCATFNA 的构建图。

图 10 显示了在液态去垢剂配方中,与野生型相比,优选的突变酶的稳定性。

发明详述

已经发现,对透镜状芽孢杆菌中相当于解淀粉芽孢杆菌枯草杆菌蛋白酶＋76 位的氨基酸残基进行体外突变,可以产生稳定性与前体枯草杆菌蛋白酶不同(例如改变的自我蛋白水解稳定性)的枯草杆菌蛋白酶变异体(见表 IV 和 VI)。

还发现，对相当于解淀粉芽孢杆菌枯草杆菌蛋白酶＋99、＋101、＋103、＋104、＋107、＋123、＋27、＋105、＋109、＋126、＋128、＋135、＋156、＋166、＋195、＋197、＋204、＋206、＋210、＋216、＋217、＋218、＋222、＋260、＋265 和/或＋274 位的残基进行体外突变，无论单独或互相组合的突变，以及以任何与＋76 位突变组合一起进行，则产生的枯草杆菌蛋白酶变异体具有不同的蛋白水解活性、不同的热稳定性、不同的 pH 分布图、不同的底物特异性和/或不同的工作性能。

羰基水解酶是水解含有 $\overset{O}{\overset{\|}{C}}-X$ 键的化合物的酶，其中，X 为氧或氮。它们包括天然存在的羰基水解酶和重组羰基水解酶。天然存在的羰基水解酶主要包括水解酶，例如肽水解酶如枯草杆菌蛋白酶或金属蛋白酶。肽水解酶包括 α-氨酰基肽水解酶、肽基(peptidyl)氨基酸水解酶、酰氨基水解酶、丝氨酸羧肽酶、金属羧肽酶、巯基蛋白酶、羧基蛋白酶和金属蛋白酶。丝氨酸蛋白酶、金属蛋白酶、巯基蛋白酶和酸性蛋白酶被包括在内，还包括内切和外切蛋白酶。

“重组羰基水解酶”指这样的羰基水解酶，其中编码天然羰基水解酶的 DNA 序列被修饰以产生编码羰基水解酶氨基酸序列中一个或多个氨基酸取代、插入或缺失的突变型 DNA 序列。合适的修饰方法公开于本文、美国专利 4,760,025(RE 34,606)、美国专利 5,204,015 和美国专利 5,185,258 中，这些文献在此引用作为参考。

枯草杆菌蛋白酶是细菌或真菌的羰基水解酶，它们一般切除蛋白质或肽的肽键。如本文所用，“枯草杆菌蛋白酶”指天然存在的枯草杆菌蛋白酶或重组的枯草杆菌蛋白酶。已知各种微生物会产生和分泌一系列天然存在的枯草杆菌蛋白酶。这一系列中的各成员的氨基酸序列并不完全同源。但是，该系列中的枯草杆菌蛋白酶具有相同或相似类型的蛋白水解活性。这类丝氨酸蛋白酶具有共同的确定催化

的三元结构的氨基酸序列，该催化性三元结构使其与胰凝乳蛋白酶相关类的丝氨酸蛋白酶区分开。枯草杆菌蛋白酶和胰凝乳蛋白酶相关的丝氨酸蛋白酶两者都具有由天冬氨酸、组氨酸和丝氨酸组成的催化性三元结构。在枯草杆菌蛋白酶相关的蛋白酶中，这些氨基酸的相对次序从氨基端到羧基端依次为天冬氨酸-组氨酸-丝氨酸。但是，在胰凝乳蛋白酶相关的蛋白酶中，相对次序为组氨酸－天冬氨酸－丝氨酸。因此，本文中枯草杆菌蛋白酶指具有枯草杆菌蛋白酶相关的蛋白酶的催化性三元结构的丝氨酸蛋白酶。具体例子包括(但不限于)图 3 所示的枯草杆菌蛋白酶。

“重组枯草杆菌蛋白酶”指这样的枯草杆菌蛋白酶，其中编码枯草杆菌蛋白酶的 DNA 序列被修饰从而产生编码天然存在枯草杆菌蛋白酶氨基酸序列中一个或多个氨基酸的取代、缺失或插入的变异(或突变)DNA 序列。用于产生这种修饰的合适方法，以及可以与此处公开的方法结合使用的方法，包括在下列文献中公开的方法：美国专利 4,760,025(RE 34,606)、美国专利 5,204,015 和美国专利 5,185,258。

“非人体羰基水解酶”及其编码 DNA 可以从众多原核和真核有机体中获得。原核有机体的合适例子包括革兰氏阴性有机体如大肠杆菌或假单胞菌(Pseudomonas)，革兰氏阳性细菌如微球菌(Micrococuus)或杆菌(Bacillus)。可获得羰基水解酶及其基因的真核有机体的例子包括酵母如酿酒酵母(Saccharomyces cerevisiae)，真菌如曲霉属(Aspergillus sp.)和人以外的哺乳动物如牛，从中可以获得编码羰基水解酶凝乳酶的基因。与枯草杆菌蛋白酶相同，从各种相关物种中可以获得一系列羰基水解酶，它们的氨基酸序列在该系列中的各成员之间并不完全同源，但是却具有相同或相似类型的生物活性。因此，此处所用的非人体羰基水解酶具有一个功能上的定义，即与原核和真核来源直接或间接相关的羰基水解酶。

"羰基水解酶变异体"的氨基酸序列衍生自"前体羰基水解酶"的氨基酸序列。前体羰基水解酶(如枯草杆菌蛋白酶)包括天然存在的羰基水解酶(枯草杆菌蛋白酶)和重组羰基水解酶(枯草杆菌蛋白酶)。羰基水解酶变异体的氨基酸序列是通过前体氨基酸序列中一个或多个氨基酸的取代、缺失或插入而从前体蛋白酶氨基酸序列衍生而得。这种修饰是对编码前体羰基水解酶(枯草杆菌蛋白酶)的"前体DNA序列"进行的,而不是对前体羰基水解酶(枯草杆菌蛋白酶)本身进行修饰。合适的对前体DNA序列进行操作的方法包括此处公开的方法以及本领域技术人员已知的方法(例如参见,EP 0 328299,WO89/06279和那些在此引用的美国专利和申请。)。

对应于解淀粉芽孢杆菌枯草杆菌蛋白酶+76位并结合一个或多个选自下组位置:+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265和/或+274的特定残基被用于突变。较佳地,修饰的残基选自下列组合:76/99;76/101;76/103;76/104;76/107;76/123;76/99/101;76/99/103;76/99/104;76/101/103;76/101/104;76/103/104;76/104/107;76/104/123;76/107/123;76/99/101/103;76/99/101/104;76/99/103/104;76/101/103/104;76/103/104/123;76/104/107/123;76/99/101/103/104;76/99/103/104/123;76/99/101/103/104/123;76/103/104/128;76/103/104/260;76/103/104/265;76/103/104/197;76/103/104/105;76/103/104/135;76/103/104/126;76/103/104/107;76/103/104/210;76/103/104/126/265和/或76/103/104/222;最佳地是:76/99;76/104;76/99/104;76/103/104;76/104/107;76/101/103/104;76/99/101/103/104和76/101/104。这些氨基酸位置的编号对应于图1中给出的成熟解淀粉芽孢杆菌枯草杆菌蛋白酶的编号。但是,本发明并不局限于

这个特定枯草杆菌蛋白酶的突变，还延伸包括在"相当于"解淀粉芽孢杆菌枯草杆菌蛋白酶中各确定残基的位置处具有氨基酸残基的前体羰基水解酶的突变形式。在本发明的一个优选例子中，前体枯草杆菌蛋白酶是透镜状芽孢杆菌枯草杆菌蛋白酶，而取代、缺失或插入是在透镜状芽孢杆菌中相当于上述位置的氨基酸残基处进行的。

如果前体羰基水解酶的残基(氨基酸)与解淀粉芽孢杆菌枯草杆菌蛋白酶的特定残基或残基部分是同源的(即在一级或三级结构中位置对应)或类似的(即具有相同或相似的结合、反应或化学作用等功能)，那么它与解淀粉芽孢杆菌枯草杆菌蛋白酶的残基是相当的。

为了确立与一级结构的同源性，前体羰基水解酶的氨基酸序列与解淀粉芽孢杆菌枯草杆菌蛋白酶的一级序列直接进行比较，尤其与一系列在已知序列的枯草杆菌蛋白酶中已知的不变残基进行比较。图2显示了解淀粉芽孢杆菌和透镜状芽孢杆菌枯草杆菌蛋白酶之间的保守残基。将保守残基排列好之后，其中允许必要的插入和缺失以维持排列(即避免通过任意的缺失和插入而导致消除保守残基)，便可确定与解淀粉芽孢杆菌枯草杆菌蛋白酶一级序列中特定氨基酸相当的残基。保守残基的排列最好100%地保留这种残基。但是，大于75%或少至50%的排列也已足以确定相当的残基。催化性三元结构即Asp32/His64/Ser221的保守性应保留。

例如，在图3中，来自解淀粉芽孢杆菌、枯草杆菌、地衣芽孢杆菌(carlsbergensis)和透镜状芽孢杆菌的枯草杆菌蛋白酶的氨基酸序列被排列在一起以提供序列间最大程度的同源性。这些序列间的比较表明，在各序列中都含有一些保守残基。这些保守残基(BPN′和透镜状芽孢杆菌之间)在图2中指明。

因此，这些保守残基可用于确定解淀粉芽孢杆菌枯草杆菌蛋白酶在其他羰基水解酶如来自透镜状芽孢杆菌的枯草杆菌蛋白酶(PCT出版物No. WO89/06279，1989年7月13日出版)(此处优选

的枯草杆菌蛋白前体酶)或称为 PB92 的枯草杆菌蛋白酶(EP 0328 299)(它与优选的透镜状芽孢杆菌枯草杆菌蛋白酶高度同源)中对应的相当氨基酸残基。在图 3A 和 3B 中排列了某些这种枯草杆菌蛋白酶和解淀粉芽孢杆菌枯草杆菌蛋白酶的氨基酸序列,以显示保守残基的同源性最大。从中可以看出,与解淀粉芽孢杆菌枯草杆菌蛋白酶相比,透镜状芽孢杆菌的序列中有一些缺失。因此,例如在解淀粉芽孢杆菌枯草杆菌蛋白酶中 Val165 的相当氨基酸在透镜状芽孢杆菌和地衣芽孢杆菌中为异亮氨酸。

因此,例如在解淀粉芽孢杆菌和透镜状芽孢杆菌枯草杆菌蛋白酶中+76 位的氨基酸都是天冬酰胺(N)。但是,在本发明的优选枯草杆菌蛋白酶变异体中,相当于解淀粉芽孢杆菌枯草杆菌蛋白酶中+76 的氨基酸可以用天冬氨酸(D)取代。出于阐述目的,在表 II 中列出了此处指明的被取代的某些氨基酸残基和各位置的最佳取代之间的比较情况。

表 II

	+76	+99	+101	+103	+104	+107	+123
液化性淀粉杆菌(野生型)	N	D	S	Q	Y	I	N
透镜状芽孢杆菌(野生型)	N	S	S	S	V	I	N
最佳取代	D	D	R	A	I/Y	V	S

对于三级结构已经用 X 射线晶体法确定的前体羰基水解酶,也可以通过比较三级结构的同源性而确定相当的残基。相当的残基的定义是,在排列之后,前体羰基水解酶和解淀粉芽孢杆菌枯草杆菌蛋白酶的一个特定氨基酸残基的两个或多个主链原子的原子坐标(N 对 N,CA 对 CA,C 对 C 和 O 对 O)在 0.13nm 之内,更佳地在 0.

1nm 之内的残基。在最佳模型被取向和定位以给出该羰基水解酶的非氢蛋白质原子与解淀粉芽孢杆菌枯草杆菌蛋白酶的原子坐标间最大重迭后，便完成了排列。最佳模型是在最高分辨率下使试验衍射数据的 R 因子最低的晶体模型。

$$R\text{因子}=\frac{\sum_h|Fo(h)|-|Fc(h)|}{\sum_h|Fo(h)|}$$

与解淀粉芽孢杆菌枯草杆菌蛋白酶的特定残基在功能上类似的相当残基被定义为前体羰基水解酶的这种氨基酸，它们采用某种构象使得它们或者改变、修饰或对蛋白质结构、底物结合或催化等起作用，并且按照解淀粉芽孢杆菌枯草杆菌蛋白酶的特定残基所限度或归属的方式进行。此外，它们是前体羰基水解酶（已获得 X 晶体衍射图）中这样的氨基酸，它们占据一个类似位置使得残基的至少两个侧链原子的原子坐标在对应的解淀粉芽孢杆菌枯草杆菌蛋白酶的侧链原子的 0.13nm 之内，尽管给定的残基的主侧链原子没有满足基于占据同源位置的相当残基标准。解淀粉芽孢杆菌枯草杆菌蛋白酶的三维结构的坐标显示在 EPO 出版物 No. 0 251 446（对应于美国专利申请 08/212,291，其公开内容在此引用作为参考）中，可如上所述用其确定三级结构水平的相当残基。

指定被取代、插入或缺失的残基中，一些是保守残基，而另一些不是保守残基。当残基不是保守性的时，一个或多个氨基酸的置换被限定于这样的取代，它们导致产生具有不同于自然界中发现的氨基酸序列的变异体。当残基是保守的时，这种置换不应产生天然存在的序列。本发明的羰基水解酶变异体包括羰基水解酶变异体的成熟形式，以及这种水解酶变异体的原（pro-）或前原（prepro-）形式。前原形式是优选的，因为这便于羰基水解酶变异体的表达、分泌和成熟化。

"原序列"指与成熟形式的羰基水解酶的N末端部分相连的氨基酸序列，但它被去除时便出现成熟的羰基水解酶。许多在自然界发现的蛋白水解酶是转译时的原酶产物，并且以这种形式表达而不经过翻译后加工。优选的产生羰基水解酶变异体、尤其是枯草杆菌蛋白酶变异体的原序列是假设的解淀粉芽孢杆菌枯草杆菌蛋白酶的原序列，尽管也可以使用其它枯草杆菌蛋白酶原序列。在实施例1和2中，使用透镜状芽孢杆菌(ATCC 21536)的枯草杆菌蛋白酶的原序列。

"信号序列"或"前序列"指与羰基水解酶的N末端部分相连或与原水解酶的N末端部分相连的任何氨基酸序列，它参与成熟或原形式的水解酶的分泌。信号序列的定义是一种功能性的序列，包括由枯草杆菌蛋白酶基因或其他可分泌羰基水解酶的N末端编码的所有氨基酸序列，这些序列参与实现自然条件下枯草杆菌蛋白酶或其他羰基水解酶的分泌。本发明采用这种序列以实现此处定义的羰基水解酶变异体的分泌。优选的用于实施例的信号序列包括解淀粉芽孢杆菌枯草杆菌蛋白酶信号序列的前7个氨基酸残基，它与透镜状芽孢杆菌(ATCC 21536)的枯草杆菌蛋白酶信号序列的其余部分融合在一起。

羰基水解酶变异体的"前原"形式由具有可操作地连于水解酶氨基端原序列的水解酶成熟形式和可操作地连于原序列氨基端的"前"或"信号"序列组成。

"表达载体"指这样的DNA构建物，它含有可操作地连于合适的调控序列的DNA序列，并能够在合适的宿主中实现该DNA的表达。这种调控序列包括实现转录的启动子、任选的控制转录的操纵子序列、编码合适的mRNA核糖体结合位点的序列和控制转录和翻译终止的序列。载体可以是质粒、噬菌体颗粒或就是潜在的基因组插入子。一旦转入合适宿主，载体会独立于基因组而进行复制并发挥功

能，或者在某些情况下整合入基因组。在本说明书中，“质粒”和“载体”有时可互换使用，因为质粒是目前最常用的载体形式。但是，本发明包括其他形式的起相当功能的，以及本发明中已知或将要已知的表达载体。

用于本发明的“宿主细胞”一般是原核或真核宿主，最好是用美国专利 4,760,025(RE 34,606)公开方法进行操作从而使其不能分泌具有酶活性的内切蛋白酶。优选的表达枯草杆菌蛋白酶的宿主细胞是杆菌属菌株 BG2036，它缺乏具有酶活性的中性蛋白酶和碱性蛋白酶(枯草杆菌蛋白酶)。菌株 BG2036 的构建在美国专利 5,264,366 中有详细描述。其他表达枯草杆菌蛋白酶的宿主细胞包括枯草杆菌 I168(在美国专利 4,760,025(RE 34,606)和美国专利 5,264,366 中有描述，这些专利的内容在此引用作为参考。)以及任何合适的杆菌属菌株如地前芽孢杆菌、透镜状芽孢杆菌等。

用重组 DNA 技术构建的载体转化或转染宿主细胞。这种转化的宿主细胞能够复制编码羰基水解酶变异体的载体，或者能够表达所需的羰基水解酶变异体。当载体编码前一或前原一形式的羰基水解酶变异体时，这种变异体(表达时)通常是以宿主细胞分泌到宿主细胞培养基中。

用“可操作地相连”描述两个 DNA 区域的关系时，就表明它们在功能上是相互相关的。例如，原序列与肽可操作地相连，如果它作为信号序列参与成熟形式的蛋白质分泌，其中最可能涉及信号序列的切除。启动子与编码序列是可操作地相连的，如果它控制序列的转录；核糖体结合位点与编码序列是可操作地相连的，如果它被置于允许翻译的位置。

编码天然存在前体羰基水解酶的基因可用本领域技术人员已知的通用方法获得。这些方法通常包括合成具有假定的编码感兴趣的水解酶区域的标记探针，从表达水解酶的有机体制备基因组文库，以

及通过与探针的杂交在文库中筛选感兴趣的基因。对阳性杂交克隆作图谱分析和测序。用于实施例的透镜状芽孢杆菌基因是按美国专利5,185,258中实施例1所述的方法进行克隆的，该专利的公开内容在此引用作为参考。用于实施例5的BPN′基因是按RE 34,606中实施例1所述的方法进行克隆的，该专利的公开内容在此引用作为参考。

克隆的羰基水解酶接着被用于转化宿主细胞以表达水解酶。然后水解酶基因被连入高拷贝数的质粒。该质粒在宿主中复制，只要它含有质粒复制所需的、众所周知的元件：可操作地连于感兴趣基因的启动子(如果能被宿主识别(即转录)，那么也可以使用基因本身的同源启动子。)、转录终止和聚腺苷酸化区域(在某些真核宿主细胞中，对于宿主中从水解酶基因转录下来的mRNA稳定性是必须的。)，该区域可以是外源的，或由水解酶基因的内源终止区域提供，同时更佳地有选择基因如抗生素抗性基因(它能够通过在含抗生素的培养基中生长而连续培养维持被质粒感染的宿主细胞)。高拷贝数质粒还可含有在宿主中复制的起始点，从而能够不受染色体的限制而在细胞质中产生大量的质粒。但是，将多拷贝的水解酶基因整合入宿主基因组中也在本发明范围之内。这对于特别适合同源重组的原核或真核有机体而言很方便。

用于实施例的基因是天然透镜状芽孢杆菌基因和天然解淀粉芽孢杆菌基因。或者，可产生编码天然存在或突变前体羰基水解酶(枯草杆菌蛋白酶)的合成基因。在这种方法中，应确定前体羰基水解酶(枯草杆菌蛋白酶)的DNA和/或氨基酸序列。随后合成多个重迭的合成单链DNA片段，然后通过杂交和连接而形成编码前体水解酶的合成DNA。合成基因构建物的一个例子是美国专利5,204,015实施例3中所述的构建物，该专利的公开内容在此引用作为参考。

一旦天然存在或合成的前体羰基水解酶基因被克隆，便可以进

行一系列的修饰，从而在合成天然存在前体羰基水解酶之外，更广泛地使用基因。这种修饰包括如美国专利 4,760,025(RE 34,606)和 EPO 出版物 No. 0 251 446 中所述的那样产生重组羰基水解酶，以及如此处所述的那样产生羰基水解酶变异体。

下列盒式诱变方法可用于协助构建和鉴别本发明的羰基水解酶变异体，尽管也可使用其他方法如定点诱变。首先，获得天然存在的编码水解酶的基因，进行全长或部分测序。然后，扫描序列，在编码的酶中找出需要产生一个或多个氨基酸突变(缺失、插入或取代)的位置。评估该位置两侧的序列是否存在限制性位点，以便用寡核苷酸混合物(oligonucleotide pool)置换某个基因短片段，这些寡核苷酸混合物被表达时会编码不同的突变型。这种限制性位点较佳地是水解酶基因中的唯一位点，以便置换基因片段。然而，在水解酶基因中并不过多的任何方便的限制性位点都可使用，只要限制性消化产生的基因片段能够按适当的次序重新装配在一起。如果从选定的位置起在方便的距离(10—15 个核苷酸)内没有限制性位点，那么限制性位点可通过取代基因中一些核苷酸而产生，其中取代时不会在最终的构建中改变阅读框或编码的氨基酸。可按常规方法，通过 M13 引物延伸而使基因发生突变从而改变其序列成为所需序列。确定合适的两侧区域以及衡量为了具有两个方便的限制性位点序列而所需改变等工作是常规性的，因为有遗传密码子的兼并性，基因限制性内切酶图谱和大量不同的限制性内切酶。应注意，如果存在一个方便的侧向限制性位点，那么仅需对不含有限制性位点的侧向区域使用上述方法。

一旦天然存在 DNA 或合成 DNA 被克隆，那么在待突变位置两侧的限制性位点可用同族限制性酶进行消化，然后将多种末端互补寡核苷酸盒连入基因。该方法简化了诱变过程，因为可合成所有的寡核苷酸使其具有相同的限制性位点，从而不需要合成接头来产生限

制性位点。

如此处所用，蛋白水解活性被定义为每毫克活性酶水解肽键的速率。有许多熟知的程序可测量蛋白水解活性(K. M. Kalisz, "Microbial Proteinases," Advances in Biochemical Engineering/Biotechnology, A. Fiechter ed., 1988)。除了具有改变的蛋白水解活性之外，本发明的变异酶还可有其他改变性能如 K_m、K_{cat}、K_{cat}/K_m 比率和/或改变的底物特异性和/或改变的 pH 活性分布图。可以对酶进行改造使其适应预计会存在的特定底物，例如在制备肽中存在的底物，或者适应水解过程例如冲洗使用。

本发明的一方面目的是为了获得与前体羰基水解酶相比，蛋白水解活性有所改变的变异羰基水解酶，因为增加这种活性(数值上更大)可以使酶更有效地作用于靶底物。此外，与前体相比，热稳定性和/或底物特异性改变的变异酶也是感兴趣的。较佳地，被突变的羰基水解酶是枯草杆菌蛋白酶。在实施例中给出了用于从枯草杆菌蛋白酶型羰基水解酶获得这种效果的具体氨基酸，它们相当于解淀粉芽孢杆菌枯草杆菌蛋白酶残基+76、+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265 和/或+274，或其组合。在某些情况下，更低的蛋白水解活性是有利的，例如蛋白水解活性的下降在需要羰基水解酶合成活性的场合(如用于合成肽)是有用的。人们可能希望降低会破坏合成产物的蛋白水解活性。相反，在某些情况下，与前体相比增加变异酶的蛋白水解活性是有利的。此外，增加或减少(改变)变异体的稳定性，无论是碱性稳定性或热稳定性，都可能是有利的。K_m，K_{cat} 或 K_{cat}/K_m 的增加或降低对用来确定这些动力学参数的底物是特异的。

在本发明的另一方面，已确定枯草杆菌蛋白酶中相当于+76 的残基与大量其他修饰的组合对于调节酶的整体稳定性和/或蛋白水

解活性是至关重要的。因此，如实施例中所述，在优选例子中，在透镜状芽孢杆菌枯草杆菌蛋白酶中相当于+76位上的天冬酰胺(N)可以用天冬氨酸(D)取代并结合一个或多个下列氨基酸残基的修饰：+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265和/或+274，从而产生稳定性更好和/或活性更好的突变酶。

本发明的最佳例子在实施例中给出。这些例子包括下列取代残基的特定组合：N76D/S99；N76D/V104I；N76D/S99D/V104I；N76D/S103A/V104I；N76D/V104I/I107V；N76D/V104Y/I107V和N76D/S101R/S103A/V104I。所有在本发明中提出权利要求的突变组合都在实施例中有描述。这些取代最好是在透镜状芽孢杆菌(重组的或天然的)枯草杆菌蛋白酶上进行，尽管也可在任何杆菌属枯草杆菌蛋白酶上进行取代。

根据这种以及其他变异枯草杆菌蛋白酶的结果，似乎羰基水解酶(较佳地是枯草杆菌蛋白酶)中相当于解淀粉芽孢杆菌枯草杆菌蛋白酶残基+76、+99、+101、+103、+104、+107、+123、+27、+105、+109、+126、+128、+135、+156、+166、+195、+197、+204、+206、+210、+216、+217、+218、+222、+260、+265和/或+274的残基，对于酶的蛋白水解活性、工作性能和/或稳定性以及变异酶的清洗或冲洗性能是至关重要的。

本发明的许多羰基水解酶变异体，尤其是枯草杆菌蛋白酶，可用于配制各种去垢剂组合物。有大量已知化合物是能用于含有本发明的羰基水解酶突变型的组合物的合适表面活性剂。其中包括非离子型、阴离子型、阳离子型或两性离子型洗涤剂，比如参见Barry J. Anderson的US 4,404,128及Jiri Flora等人的US 4,261,868。一种合适的去垢剂配方是在美国专利5,204,015(前面引用作为参考)

实施例 7 中描述的配方。对于可用作清洗组合物的各种不同配方，本领域是很熟悉的。除了典型的清洗组合物，很清楚本发明的枯草杆菌变异体也可用于任何使用天然或野生型枯草杆菌蛋白酶的场合。因此，这些变异体可以用于例如条状或液体肥皂、洗碗剂、隐形眼镜清洗液或产品、肽水解、废物处理、织物应用以及蛋白质生产中用作切割融合的酶。在去垢剂组合物中，本发明的变异体可有增强的性能（与前体相比）。如本文所用，在去垢剂中的增强性能的定义是，在标准洗涤过程后，用通常的评估方法所确定的某些酶敏感污迹如草或血的清洗效果提高。

本发明的枯草杆菌蛋白酶可配制成熟知的粉末状或液体去垢剂，其浓度为约 0.01—5 重量%（较佳地为 0.1—0.5%）时 pH 为 6.5—12.0。这些去垢清洗组合物还可含有其他酶如蛋白酶、淀粉酶、纤维素酶、脂酶或内糖苷酶以及助洗剂和稳定剂。

在常规清洗组合物中加入本发明的枯草杆菌蛋白酶并不会形成任何特殊的使用限制。换言之，对去垢剂合适的任何温度和 pH 也适用于本发明的组合物，只要 pH 在上述范围内而温度低于所述的枯草杆菌蛋白酶变性温度。此外，本发明的枯草杆菌蛋白酶可用于不含去垢剂的清洗组合物，同样可仅含有该酶或者同时含有助洗剂和稳定剂。

下面通过实施例描述本发明，这些实施例不用于限制本发明范围。

实施例 1

构建在枯草杆菌中表达 GG36 基因

为了表达透镜状芽孢杆菌的枯草杆菌蛋白酶基因而进行的克隆和构建，基本上是按照美国专利 5,185,258 所述的方法进行。质粒 GGA274（见图 4）进一步按图 5 所示的方式进行修饰。在构建

GGA274 中引入的 PstI 位点通过下述的寡核苷酸定点诱变而切除，其中寡核苷酸序列如下：

5′-GAAGCTGCAA CTCGTTAAA-3′(SEQ ID NO:1)

带下划线的“A”残基消除了限制性酶 PstI 的识别序列，并将 274 位的相应氨基酸残基从 Ala 变为 Thr。274 位的苏氨酸是最初在克隆的透镜状芽孢杆菌枯草杆菌蛋白酶基因序列中发现的野生型残基。从质粒 GGA274 或其衍生物(图 5 中的 GGT274)中通过 EcoRI 和 BamHI 消化而切下编码枯草杆菌蛋白酶的 DNA 片段。将 DNA 片段亚克隆至细菌噬菌体 M13 类型载体如 MP19 以便诱变。诱变后，进行 EcoRI 和 HindIII 消化，随后克隆，从而将突变的枯草杆菌蛋白酶基因重新连入诸如 GGA274 之类的表达载体以便表达和回收突变枯草杆菌蛋白酶蛋白质。

实施例 2

寡核苷酸定点诱变

寡核苷酸定点诱变按照 Zoller，M. et al. (1983)，Methods Enzymol.，100:486－500 所述的方法进行。一个例子是，用人工合成的序列

5′-GCTGCTCTAG ACAATTCG-3′ (SEQ ID NO:2)

寡核苷酸将76位的氨基酸残基从天冬酰胺(N)变为天冬氨酸(D)，即 N76D。带下划线的“G”和“C”残基表示从野生型基因序列改变而来。CA 保留了＋75位的亮氨酸，同时改变了氨基酸序列从而引入 XbaI 限制性酶的识别位点(TCTAGA)，而GAC 的改变使＋76天冬酰胺变为天冬氨酸。

对于99、101、103和104位的诱变，可使用不同的寡核苷酸，这取决于所需的突变组合情况。例如，序列

5′-GTATTAGGGG CGGACGGTCG AGGCGCCATC AGCTCGATT-3′

(SEQ ID NO:3)

寡核苷酸可用于在单个枯草杆菌蛋白酶分子中同时产生下列改变：S99D;S101R;S103A 和 V104I。类似地，序列

5′-TCAGGTTCGG TCTCGAGCGT TGCCCAAGGA TTG-3′(SEQ ID NO:4)

和5′-CACGTTGCTA GCTTGAGTTT AG-3′(SEQ ID NO:5)

可分别用于产生 I107V 和 N123S。同样，带下划线的残基表示从野生型序列改变而来，以用于产生所需的氨基酸序列或限制性酶识别序列的改变。

实施例3

枯草杆菌蛋白酶变异体的蛋白水解活性

按照实施例2的方法，制得表 III 中列出的变异体。每种枯草杆菌蛋白酶变异体的蛋白水解活性都列在表 III 中。用 P. Bonneau, et al. (1991) J. Am. Chem. Soc., Vol. 113, No. 3, p. 1030中所述的方法，通过水解合成肽底物琥珀酰-L-Ala-L-Ala-L-Pro-L-Phe-对硝基酰苯胺而测量动力学参数 K_{cat}，K_M和 K_{cat}/K_M。简言之，将一少量枯草杆菌蛋白酶变异体原液加入1厘米的小试管中，试管中含有溶解于0.1M Tris-HClpH 8.6缓冲液中的底物并保温于25℃。反应后，用光谱分光光度法测量在410nm 处反应产物对硝基酰苯胺的吸收值。用非线性回归算法获得动力学参数，使各反应的反应速度和产物浓度与 Michaelis-Menten 方程式相符。

表 III

对透镜状芽孢杆菌枯草杆菌蛋白酶和突变蛋白测得的

动力学参数 K_{cat}，K_M和 K_{cat}/K_M

酶	$K_{cat}(s^{-1})$	$K_M(M)$	$K_{cat}/K_M(s^{-1}M^{-1})$
透镜状芽孢杆菌枯草杆菌蛋白酶	170	0.00078	2.18×10^5
N76D	219	0.0008	2.74×10^5
M76D/S99D	88	0.00061	1.44×10^5
N76D/S103A	400	0.0014	2.86×10^5

N76D/V104I	459	0.0011	4.17×10^5
N76D/I107V	219	0.0011	1.99×10^5
N76D/N123S	115	0.0018	6.40×10^4
N76D/S99D/S101R	146	0.00038	3.84×10^5
N76D/S99D/S103A	157	0.0012	1.31×10^5
N76D/S99D/V104I	247	0.00097	2.55×10^5
N76D/S101R.S103A	405	0.00069	5.90×10^5
N76D/S101R/V104I	540	0.00049	1.10×10^6
N76D/S102A/V104I	832	0.0016	5.20×10^5
N76D/V104I/I107V	497	0.00045	1.10×10^6
N76D/V104Y/I107V	330	0.00017	1.90×10^6
N76D/V104I/N123S	251	0.0026	9.65×10^4
N76D/I107V/N123S	147	0.0035	4.20×10^4
N76D/S99D/S101R/S103A	242	0.00074	3.27×10^5
N76D/S99D/S101R/V104I	403	0.00072	5.60×10^5
N76D/SS99D/S103A/V104I	420	0.0016	2.62×10^5
N76D/S101R/S103A/V104I	731	0.00065	1.12×10^6
N76D/S103A/V104I/N123S	321	0.0026	1.23×10^5
N76D/V104I/I107V/N123S	231	0.003	7.70×10^4
N76D/S99D/S101R/S103A/V104I	624	0.00098	6.37×10^5
N76D/S99D/S103A/V104I/N123S	194	0.0043	4.51×10^4
N76D/S99D/S101R/S103A/V104I/N123S	311	0.0023	1.35×10^5

表 III 中列出的数据表明，所有测试的枯草杆菌蛋白酶变异体都保留了蛋白水解活性。此外，仔细分析数据揭示，相当于解淀粉芽孢杆菌中76、99、101、103、104、107和123位的氨基酸残基的各种取代组合，会明显改变透镜状芽孢杆菌枯草杆菌蛋白酶的蛋白水解活性。

实施例4

枯草杆菌蛋白酶变异体的热稳定性

对于透镜状芽孢杆菌枯草杆菌蛋白酶及用实施例2方法制得的变异体，观察比较其热稳定性，结果列于表IV。将纯化的酶(15μg/ml于0.1M 甘氨酸，0.01吐温(% Tween－80) pH 10.0中，含有或不含50mM $CaCl_2$)一份份地装入各小试管中，在10℃保温5分钟，10－60℃1分钟，60℃20分钟。然后将试管置于冰上10分钟。分析各试管中样品的酶活性，即将其加入1厘米的小试管中，试管中含有溶解于0.1M Tris-HClpH 8.6缓冲液中的底物琥珀酰-L-Ala-L-Ala-L-Pro-L-Phe-对硝基酰苯胺并保温于25℃。在最初线性反应速度之后，用光谱分光光度法测量在410nm 处反应产物对硝基酰苯胺的吸收值，以时间作为函数。数据为加热前活性的百分比。列于表 IV 中的数据表明，在加入50mM $CaCl_2$的测试条件下，绝大多数变异体(26个中的24个)的热稳定性与透镜状芽孢杆菌枯草杆菌蛋白酶的热稳定性相近。在不加入50mM $CaCl_2$的测试条件下，绝大多数变异体(26个中的19个)比透镜状芽孢杆菌枯草杆菌蛋白酶明显地更具热稳定性。此外，在不加入50mM $CaCl_2$的测试条件下，变异体 N76D/S99D、N76D/V104I、N76D/ S99D/V104I、N76D/S103A/V104I、N76D/V104I/I107V、 N76D/V104Y/I107V 和 N76D/S101R/S103A/V104I 比单个取代的N76D 变异体明显地更稳定。

表 IV

对透镜状芽孢杆菌枯草杆菌蛋白酶和突变蛋白测得的热稳定性(pH 10，60℃，+/－ 50mM $CaCl_2$)

酶	最初活性的残留百分比	
	$-CaCl_2$	$+CaCl_2$
透镜状芽孢杆菌枯草杆菌蛋白酶	2	96
N76D	34	97
N76D/S99D	49	98

N76D/S103A	26	92
N76D/V104I	58	98
N76D/I107V	32	96
N76D/N123S	0	97
N76D/S99D/S101R	30	100
N76D/S99D/S103A	36	100
N76D/S99D/V104I	48	97
N76D/S101R/S103A	26	100
N76D/S101R/V104I	38	100
N76D/S103A/V104I	58	100
N76D/V104I/I107V	60	97
N76D/V104Y/I107V	48	74
N76D/V104I/N123S	0	98
N76D/I107V/N123S	16	100
N76D/S99D/S101R/S103A	38	100
N76D/S99D/S101R/V104I	33	100
N76D/S99D/S103A/V104I	38	98
N76D/S101R/S103A/V104I	40	99
N76D/S103A/V104I/N123S	1	98
N76D/V104I/I107V/N123S	3	99
N76D/S99D/S101R/S103A/V104I	36	99
N76D/S99D/S103A/V104I/N123S	2	95
N76D/S99D/S101R/S103A/V104I/N123S	0	100

实施例5

用噬菌粒(phagemid)产生的单链DNA模板进行寡核苷酸定点诱变

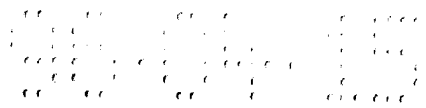
A. 构建透镜状芽孢杆菌变异体

诱变方案与上面实施例2所述的方法基本相同。用噬菌粒方法产

生单链DNA模板。为了构建用于产生单链DNA模板的噬菌粒，我们首先构建载体pBCDAICAT。载体构建流程图示于图8。首先，来自pC194质粒(Horinouchi，S. and Weisblum，B.，J. Bacteriol.，150:8－15，1982)的编码CAT基因的ClaI－ClaI片段被克隆入pUC19(New England BioLabs，Beverly，MA)多接头区域的AccI位点中，形成质粒pUCCHL。来自GG36DAI编码DNA 5′端的0.6KB EcoRI－DraI片段被克隆入pBSKS质粒(Stratagene，Inc.，San Diego，CA)的EcoRI和EcoRV位点，形成质粒pBC2SK5。质粒pBC2SK5的单个EcoRI位点的去除是通过EcoRI消化、随后用T4 DNA聚合酶催化填平再重新连入形成不含EcoRI位点质粒pBC2SK－5R。被克隆入pBC2SK－5R的EcoRI－DraI片段以PstI－HindIII片段形式分离出，然后克隆入pUCCHL的PstI－HindIII位点(pUC19的多接头部分)形成质粒pUCCHL5R。GG36DAI基因的编码序列以EcoRI－BamHI片段形式切下，克隆入pUCCHL5R的EcoRI－BamHI位点，形成pUCCAT。pUCCAT的大EcoRI－HindIII片段接着被克隆入BS2KS+的EcoRI和HindIII位点，形成质粒pBCDAICAT。

为了产生单链DNA，按照Russel，M.，Kidd，S. and Kelley，M.R.，GENE45:333－338，1986的方法，用噬菌体R408(得自Stratagene，Inc.，San Diego，CA)感染含有pBCDAICAT的大肠杆菌。一旦形成单链DNA模板，就进行实施例2所述的标准诱变方法。某些突变型的构建在下面详细给出以便阐述。

为了构建透镜状芽孢杆菌(GG36)N76D/S103A/V104I/L217H，用编码GG36 N76D/S103A/V104I的EcoRI－BamHI片段构建pUCCAT(见图8)，形成质粒pBCDAICAT。一旦形成单链DNA模板，就按照上述方案，用如下序列的诱变引物形成L217H



* * * * * * x ClaI

5′-TAT GCC AGC CAC AAC GGT ACT TCG ATG GCT-3′ (SEQ ID NO:13)

与前面相同，带下划线的残基表示被改变的核苷酸，而 x ClaI 表示原存在的 ClaI 位点在诱变后被消除。诱变方案描述于实施例2。在诱变后，首先筛选质粒 DNA 是否消除了 ClaI 位点，那些缺少 ClaI 位点的克隆通过 DNA 序列分析以确定217位氨基酸残基从 L 变为 H 的 DNA 序列。

B. BPN′变异体的构建及其在枯草杆菌中的表达

解淀粉芽孢杆菌(BPN′)N76D/Q103A/Y104I/Y217L 构建按类似的方式进行，除了两个连续步骤。首先将 N76D 引入 BPN′ Y217L 中形成 BPN′ N76D/Y217L，然后进行第二次诱变将 BPN′ N76D/Y217L 变为 BPN′ N76D/Q103A/Y104I/Y217L。为了产生用于第一次诱变的单链 DNA 模板，用编码 BPN′ Y217L 枯草杆菌蛋白酶的 EcoRI—BamHI 片段(衍生自 Wells，J.，et al.，PNAS，84，5167，1087中所述的 Y217L 质粒)构建质粒 pUCCATFNA(见图9)。用含有 BPN′ Y217L 的 pUCCATFNA 质粒构建质粒 pBCFNACAT(图9)。如上产生单链 DNA。为了产生 BPN′ N76D/Y217L，用如下序列的寡核苷酸引物形成 N76D 变化

* * * * * * x XbaI

5′-C ACA GTT GCG GCT CTA GAT AAC TCA ATC GGT G-3′

(SEQ ID NO:14)

然后用含有 BPN′ N76D/Y217L 的质粒 pBCFNACAT(N76D 诱变后的质粒 pBCFNACAT)制备单链 DNA，并用另一如下序列的寡核苷酸引物形成 BPN′ N76D/Q103A/Y104I/Y217L

* * * * * * x PvuII

5′-GCT GAC GGT TCC GGC GCT ATT AGT TGG ATC ATT-3′

(SEQ ID NO:15)

所有克隆、制备单链 DNA、诱变和筛选突变型的步骤都按上述方法

进行。

BPN′基因及其变异体的表达是通过将质粒DNA(含有不同变异BPN′基因的pBCFNACAT)直接整合入RE 34,606中所述的缺少蛋白酶的枯草杆菌菌株。

按实施例2和5所述的方法制备各种变异体。对这些变异都测其动力学数据和稳定性数据。动力学数据是用实施例3的方法获得的，列于表V。稳定性数据的获得如下所述。结果列于表VI。

热稳定性测试程序

纯化的酶是在0.1M甘氨酸pH 10.0,0.01% Tween－80中进行缓冲液交换(buffer－exchanged)，即将酶上柱于Sephadex G－25柱，该柱已用缓冲液平衡过，然后用相同缓冲液将酶从柱中洗脱出来。

向含有0.1M甘氨酸，0.01% Tween－80 pH 10.0并保温于60℃的试管中，加入缓冲液交换过的酶，使最终酶浓度为15μg/ml。

按不同时间从60℃保温中取出一份样品，立刻测试其酶活性，即将其加入1厘米的小试管中，试管中含有溶解于0.1M Tris－HCl pH 8.6缓冲液中的1.2mM人工合成肽底物琥珀酰－L－Ala－L－Ala－L－Pro－L－Phe－对硝基酰苯胺并保温于25℃。在最初线性反应速度之后，用光谱分光光度法测量在410nm处反应产物对硝基酰苯胺的吸收值，以时间作为函数。

以60℃保温时间作为函数，根据反应速度的一级坐标图确定半衰期(即50%酶失去活性所需的时间)。

表VI中的数据是在相同条件下测得的透镜状芽孢杆菌枯草杆菌蛋白酶(GG36)半衰期的百分比。

表 V

酶	K_{cat} (s^{-1})	K_M (mM)	K_{cat}/K_M ($s^{-1}M^{-1}$)
透明状芽孢杆菌枯草杆菌蛋白酶	170	0.78	2.20E+05
N76D/S103G/V104I*	380	1.4	2.70E+05
N76D/S103A/V104F	730	0.33	2.20E+06
N76D/S103A/V104N	790	2.8	2.80E+05
N76D/S103A/V104S	170	0.83	2.00E+05
N76D/S103A/V104T	370	1.9	2.00E+05
N76D/S103A/V104W	880	0.31	2.80E+06
N76D/S103A/V104Y	690	0.5	1.40E+06
K27R/N76D/V104Y/N123S	500	1.2	4.20E+05
N76D/S101G/S103A/V104I*	620	1.3	4.80E+05
N76D/S103A/V104I/S105A*	550	1.3	4.20E+05
N76D/S103A/V104I/S105G*	440	1.7	2.60E+05
N76D/S103A/V104T/I107A*	120	5.7	2.10E+04
N76D/S103A/V104T/I107L*	310	3.2	9.70E+04
N76D/S103A/V104I/L126A	90	2.2	4.10E+04
N76D/S103A/V104I/L126F	180	1.9	9.50E+04
N76D/S103A/V104I/126I	100	2.4	4.20E+04
N76D/S103A/V104I/L126V	64	3.2	2.00E+04
N76D/S103A/V104I/S128G*	560	1.7	3.30E+05
N76D/S103A/V104I/S128L*	430	3.8	1.10E+05
N76D/S103A/V104I/L135A	140	0.76	1.80E+05
N76D/S103A/V104I/L135F	390	0.69	5.70E+05
N76D/S103A/V104I/L135I	110	0.73	1.50E+05
N76D/S103A/V104I/L135V	140	0.86	1.60E+05
N76D/S103A/V104I/S156E*	170	2.6	6.50E+04
N76D/S103A/V104I/S166D*	160	3.5	4.60E+04

N76D/S103A/V104I/D197E	510	1.4	3.60E+05
N76D/S103A/V104I/N204A*	530	1.1	4.80E+05
N76D/S103A/V104I/N204G*	580	1.4	4.10E+05
N76D/S103A/V104I/N204C*	370	1.3	2.90E+05
N76/S103A/V104I/P201I*	500	1.2	4.20E+05
N76D/S103A/V104I/L217H*	80	0.63	1.30E+05
N76D/S103A/V104I/M222A	70	3.1	2.30E+04
N76D/S103A/V104I/M222S	80	3.1	2.60E+04
N76D/S103A/V104I/T260P	660	1.5	4.40E+05
N76D/S103A/V104I/S265N	590	1.3	4.50E+05
K27R/N76D/V104Y/I107V/N123S	220	1.4	1.60E+05
K27R/N76D/V104Y/N123S/D197E	430	1.1	3.90E+05
K27R/N76D/V104Y/N123S/N204C	400	1.1	3.60E+05
K27R/N76D/V104Y/N123S/Q206L	440	1.2	3.70E+05
K27R/N76D/V104Y/N123S/S216V	440	1.2	3.70E+05
K27R/N76D/V104Y/N123S/N218S	760	0.98	7.80E+05
K27R/N76D/V104Y/N123S/T260P	410	1.2	3.40E+05
K27R/N76D/V104Y/N123S/T274A	390	1	3.90E+05
N76D/S103A/V104I/I126F/S265N	170	2.1	8.10E+04
N76D/S103A/V104I/S156E/S166D*	40	6.3	6.40E+03
K27R/N76D/V104Y/N123S/G195E/G197E	410	0.98	4.20E+05
K27R/N76D/V104Y/N123S/G195E/N218S	540	0.66	8.20E+05
K27R/N76D/V104Y/N123S/D197E/N218S	770	0.79	9.80E+05
K27R/N76D/V104Y/N123S/N204C/N218S	610	0.99	6.20E+05
K27R/N76D/V104Y/N123S/Q206L/N218S	580	0.78	7.40E+05
K27R/N76D/V104Y/N123S/N218S/T260P	660	1	6.60E+05
K27R/N76D/V104Y/N123S/N218S/T274A	590	0.89	6.60E+05
K27R/N76D/V104Y/Q109S/N123S/N218S/T274A	520	1	5.20E+05
K27R/N76D/V104Y/N123S/G195E/D197E/N218S	460	0.65	7.10E+05

液化性淀粉杆菌枯草杆菌蛋白酶(BPN')	50	0.14	3.60E+05
BPN'-N76D/Y217L *	380	0.46	8.30E+05

* 这些突变型是按实施例5方法制备，其他所有的是按实施例2方法制备。

表 VI

酶	热稳定性（天然酶半衰期的百分比）
透镜状芽孢杆菌枯草杆菌蛋白酶	100
N76D	590
N76D/S99D	840
N76D/S103A	390
N76D/V104I	660
N76D/I107V	710
N76D/N123S	70
N76D/S99D/S101R	610
N76D/S99D/S103A	590
N76D/S99D/V104I	910
N76D/S101R/S103A	930
N76D/S101R/V104I	500
N76D/S103A/V104I	460
N76D/S103G/V104I*	370
N76D/S103A/V104F	480
N76D/S103A/V104N	230
N76D/S103A/V104S	230
N76D/S103A/V104T	370
N76D/S103A/V104W	280
N76D/S103A/V104Y	400
N76D/V104I/I107V	940

N76D/V104Y/I107V	820
N76D/V104I/N123S	80
N76D/I107V/N123S	150
K27R/N76D/V104Y/N123S	100
N76D/S99D/S101R/S103A	570
N76D/S99D/S101R/V104I	1000
N76D/S99D/S103A/V104I	680
N76D/S101G/S103A/V104I*	390
N76D/S101R/S103A/V104I	470
N76D/S103A/V104I/S105A*	360
N76D/S103A/V104I/S105D*	370
N76D/S103A/V104T/I107A*	270
N76D/S103A/V104T/I107L*	230
N76D/S103A/V104I/N123S	110
N76D/V104I/I107V/N123S	220
N76D/S103A/V104I/L126A	270
N76D/S103A/V104I/L126F	950
N76D/S103A/V104I/L126I	410
N76D/S103A/V104I/L126V	320
N76D/S103A/V104I/S128G*	640
N76D/S103A/V104I/S128L*	760
N76D/S103A/V104I/L135A	230
N76D/S103A/V104I/L135F	200
N76D/S103A/V104I/L135I	510
N76D/S103A/V104I/L135V	500
N76D/S103A/V104I/S156E*	120
N76D/S103A/V104I/S166D*	590

N76D/S103A/V104I/D197E	460
N76D/S103A/V104I/N204A*	230
N76D/S103A/V104I/N204G*	240
N76D/S103A/V104I/N204c*	500
N76D/S103A/V104I/P210I*	1370
N76D/S103A/V104I/L217H*	60
N76D/S103A/V104I/M222A	520
N76D/S103A/V104I/M222S	490
N76D/S103A/V104I/T260P	490
N76D/S103A/V104I/S265N	360
K27R/N76D/V104Y/I107V/N123S	210
K27R/N76D/V104Y/N123S/D197E	120
K27R/N76D/V104Y/N123S/N204C	110
K27R/N76D/V104Y/N123S/Q206L	380
K27R/N76D/V104Y/N123S/S216V	140
K27R/N76D/V104Y/N123S/N218S	270
K27R/N76D/V104Y/N123S/T260P	40
K27R/N76D/V104Y/N123S/T274A	60
N76D/S99D/S101R/S103A/V104I	590
N76D/S99D/S103A/V104I/N123S	110
N76D/S103A/V104I/L126F/S265N	810
N76D/S103A/V104I/S156E/S166D*	220
K27R/N76D/V104Y/N123S/G195E/G197E	90
K27R/N76D/V104Y/N123S/G195E/N218S	250
K27R/N76D/V104Y/N123S/D197E/N218S	270
K27R/N76D/V104Y/N123S/C204C/N218S	460
K27R/N76D/V104Y/N123S/Q206L/N218S	1400

K27R/N76D/V104Y/N123S/N218S/T260P	310
K27R/N76D/V104Y/N123S/N218S/T274A	180
N76D/S99D/S101R/S103A/V104I/N123S	90
K27R/N76D/V104Y/Q109S/N123S/N218S/T274A	230
K27R/N76D/V104Y/N123S/G195E/D197E/N218S	240
液化性淀粉杆菌枯草杆菌蛋白酶(BPN′)	100
BPN′−N76D/Y217L *	420

*这些突变型是按实施例5方法制备,其他所有的是按实施例2方法制备。

实施例6

洗涤性能测试

在前面实施例中描述过的变异体洗涤性能的评估是通过从EMPA116布块(Testfabrics, Inc., Middlesex, NJ 07030)上去除污迹(布上的血/牛奶/碳黑)情况进行的。

6块被切成3×4.5英寸、边缘为锯齿状的EMPA116布块,被置于7243S型Terg−O−Tometer(United States Testing Co., Inc., Hoboken, NJ)的各个盆中,其中含有1000毫升水、硬度为15gpg (Ca^{++}:Mg^{++}::3:1::w:w),7克去垢剂和适当的酶。去垢剂基料是wfk−Testgewebe GmbH(Adlerstrasse 42, Postfach 13 07 62, D−47759 Krefeld, Germany)的WFK1去垢剂。

组份	终配方百分比
沸石A	25%
硫酸钠	25%
苏打粉	10%
线性烷基苯磺酸盐	8.8%
醇乙氧化物(7−8EO)	4.5%
钠皂	3%
硅酸钠(SiO_2:Na_2O::3.3:1)	3%

向该基本去垢剂中，加入下列组份：

组份	终配方百分比
过硼酸钠一水合物	13%
共聚物(Sokalan CP5)	4%
TAED (Mykon ATC Green)	3%
酶	0.5%
增白剂(Tinopal AMS—GX)	0.2%

过硼酸钠一水合物是从 Degussa Corporation，Ridgefield—Park，NJ 07660获得的。Sokalan CP5是从BASF Corporation，Parsippany，NJ 07054获得的。Mykon ATC Green(TAED，四乙酰基乙二胺)是从 Warwick International，Limited，Mostyn，Holywell，Clwyd CH8 9HE，England 获得。Tinopal AMSGX 是从 Ciba—Geigy Corporation，Greensboro，NC 27419获得的。

6块 EMPA116布块在去垢剂中用酶在60℃洗30分钟，然后在1000毫升水中漂洗5分钟二次。加入的酶终浓度对标准曲线为0.05—1ppm，常规分析为0.25ppm。干燥并压平布块，用 Minolta Chroma Meter，Model CR — 200 (Minolta Corporation，Ramsey，NJ 07446)L＊a＊b 刻度上的 L 值测量布块的反射系数。性能以透镜状芽孢杆菌(GG36)蛋白酶性能的百分比表示，用透镜状芽孢杆菌(GG36)蛋白酶的数量除以得到同样污迹去除效果所需的变异体蛋白酶的数量，然后乘以100。数据列于表 VII。

表 VII

酶	清洗性能
透镜状芽孢杆菌枯草杆菌蛋白酶	100
N76D	310
N76D/S103A	230
N76D/V104I	130
N76D/I107V	160
N76D/S99D/S101R	370
N76D/S99D/S103A	290
N76D/S101R/S103A	130
N76D/S101R/V104I	300
N76D/S103A/V104I	320
N76D/S103G/V104I	160
N76D/S103A/V104F	210
N76D/S103A/V104N	110
N76D/S103A/V104T	170
N76D/V104I/I107V	210
N76D/S99D/S101R/S103A	220
N76D/S99D/S101R/V104I	140
N76D/S101G/S103A/V104I	170
N76D/S101R/S103A/V104I	150
N76D/S103A/V104I/S105A	170
N76D/S103A/V104T/I107A	120
N76D/S103A/V104T/I107L	110
N76D/S103A/V104I/L126F	110
N76D/S103A/V104I/S128G	280
N76D/S103A/V104I/L135I	160

N76D/S103A/V104I/L135V	160
N76D/S103A/V104I/D197E	170
N76D/S103A/V104I/N204A	160
N76D/S103A/V104I/N204G	150
N76D/S103A/V104I/P210I	470
N76D/S103A/V104I/M222A	100
N76D/S103A/V104I/T260P	280
N76D/S103A/V104I/S265N	190

实施例7

在液体去垢剂配方中蛋白酶的稳定性

对透镜状芽孢杆菌枯草杆菌蛋白酶及其变异酶N76D/S103A/V104I，根据此处所述的程序比较它们在液体去垢剂配方中的蛋白酶稳定性(导致失活)。用于研究的去垢剂配方是市场上出售的瓶装Tide Ultra液体洗衣去垢剂(美国的Procter& Gamble Company制造)。为了使原来的酶失活，必须对去垢剂制剂进行热处理。这通过将去垢剂在96℃保温4.5小时而实现。然后在室温下，将透镜状芽孢杆菌枯草杆菌蛋白酶和N76D/S103A/V104I变异体的浓制剂(20克酶/升)加入热处理过的Tide Ultra中，使去垢剂制剂中的终浓度为0.3克酶/升。将带有加入酶的、热处理去垢剂在50℃水浴中保温。在0，24，46，76和112小时时从保温试管中取出各份样品，测试其酶活性，即将其加入1厘米的小试管中，试管中含有溶解于0.1M Tris－HCl pH 8.6缓冲液中的底物琥珀酰－L－Ala－L－Ala－L－Pro－L－Phe－对硝基酰苯胺并保温于25℃。在最初线性反应速度之后，用光谱分光光度法测量在410nm处反应产物对硝基酰苯胺的吸收值，以时间作为函数。如图10所示，观察到N76D/S103A/V104I变异体的抗失活稳定性远大于天然透镜状芽孢杆菌酶。在特定的测试条件下，

估计在 Tide Ultra 去垢剂制剂中两种酶的失活半衰期是，透镜状芽孢杆菌枯草杆菌蛋白酶为45小时，而 N76D/S103A/V104I 变异体是125小时。

在整个申请中，通用的单字符和三字符表示各种氨基酸。这些字符可参见 Dale，J. W.（1989），Molecular Genetics of Bacteria，John Wiley & Sons，Ltd.，Appendix B.

尽管上面描述了本发明的优选实施例，很明显本发明相关领域的熟练技术人员在完整地领会本发明之后，可对本发明进行各种改变或各种相应改动，这些都在所附权利要求书所限定的范围之内。

序列表

(1) 一般信息

(i)申请人：Genencor International, Inc.

(ii) 发明名称：枯草杆菌蛋白酶变异体

(iii)序列数目：15

(iv) 通信地址：

(A) 地址：Genencor International, Inc

(B) 街道：180 Kimball Way

(C) 城市：So. San Francisco

(D) 州：CA

(E) 国家：USA

(F) 邮编：94080

(v) 计算机可读形式：

(A) 记录介质类型：软盘

(B) 计算机：IBM PC 兼容型

(C) 操作系统：PC－DOS/MS－DOS

(D) 软件：PatentIn Release #1.0，版本#1.25

(vi) 本申请资料：

(A) 申请号：PCT/US94/11562

(B) 申请日：13－OCT－1994

(C) 分类：

(viii) 律师/代理人信息：

(A) 姓名：Horn, Margaret A.

(B) 登记号：33,401

(C) 参考/案卷号：GC235－2 PCT

(ix) 通讯信息：

(A) 电话：(415) 742－7536

(B) 传真：(415) 742－7217

(2) SEQ ID NO:1的信息：

(i) 序列特征：

(A) 长度：19 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:1：

GAAGCTGCAA CTCGTTAAA

(2) SEQ ID NO:2的信息：

(i) 序列特征：

(A) 长度：18 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:2：

GCTGCTCTAG ACAATTCG

18

(2) SEQ ID NO:3的信息：

(i) 序列特征：

(A) 长度：39 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:3：

GTATTAGGGG CGGACGGTCG AGGCGCCATC AGCTCGATT 39

(2) SEQ ID NO:4的信息：

(i) 序列特征：

(A) 长度：33 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:4：

TCAGGTTCGG TCTCGAGCGT TGCCCAAGGA TTG 33

(2) SEQ ID NO:5的信息：

(i) 序列特征：

(A) 长度：22 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:5：

CACGTTGCTA GCTTGAGTTT AG 22

(2) SEQ ID NO:6的信息：

(i) 序列特征：

(A) 长度：1497 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:6：

GGTCTACTAA AATATTATTC CATACTATAC AATTAATACA CAGAATAATC TGTCTATTGG 60

TTATTCTGCA AATGAAAAAA AGGAGAGGAT AAAGAGTGAG AGGCAAAAAA GTATGGATCA 120

GTTTGCTGTT TGCTTTAGCG TTAATCTTTA CGATGGCGTT CGGCAGCACA TCCTCTGCCC 180

AGGCGGCAGG GAAATCAAAC GGGGAAAAGA AATATATTGT CGGGTTTAAA CAGACAATGA 240

GCACGATGAG CGCCGCTAAG AAGAAAGATG TCATTTCTGA AAAAGGCGGG AAAGTGCAAA 300

AGCAATTCAA ATATGTAGAC GCAGCTTCAG TCACATTAAA CGAAAAAGCT GTAAAAGAAT 360

TGAAAAAAGA CCCGAGCGTC GCTTACGTTG AAGAAGATCA CGTAGCACAT GCGTACGCGC 420

AGTCCGTGCC TTACGGCGTA TCACAAATTA AAGCCCCTGC TCTGCACTCT CAAGGCTACA 480

CTGGATCAAA TGTTAAAGTA GCGGTTATCG ACAGCGGTAT CGATTCTTCT CATCCTGATT 540

TAAAGGTAGC AAGCGGAGCC AGCATGGTTC CTTCTGAAAC AAATCCTTTC CAAGACAACA 600

ACTCTCACGG AACTCACGTT GCCGGCACAG TTGCGGCTCT TAATAACTCA ATCGGTGTAT 660

TAGGCGTTGC GCCAAGCGCA TCACTTTACG CTGTAAAAGT TCTCGGTGCT GACGGTTCCG 720

GCCAATACAG CTGGATCATT AACGGAATCG AGTGGGCGAT CGCAAACAAT ATGGACGTTA 780

TTAACATGAG CCTCGGCGGA CCTTCTGGTT CTGCTGCTTT AAAAGCGGCA GTTGATAAAG 840

CCGTTGCATC CGGCGTCGTA GTCGTTGCGG CAGCCGGTAA CGAAGGCACT TCCGGCAGCT 900

CAAGCACAGT GGGCTACCCT GGTAAATACC CTTCTGTCAT TGCAGTAGGC GCTGTTGACA 960

GCAGCAACCA AAGAGCATCT TTCTCAAGCG TAGGACCTGA GCTTGATGTC ATGGCACCTG 1020

GCGTATCTAT CCAAAGCACG CTTCCTGGAA ACAAATACGG GGCGTACAAC GGTACGTCAA 1080

TGGCATCTCC GCACGTTGCC GGAGCGGCTG CTTTGATTCT TTCTAAGCAC CCGAACTGGA 1140

CAAACACTCA AGTCCGCAGC AGTTTAGAAA ACACCACTAC AAAACTTGGT GATTCTTTGT 1200

ACTATGGAAA AGGGCTGATC AACGTACAAG CGGCAGCTCA GTAAAACATA AAAAACCGGC 1260

CTTGGCCCCG CCGGTTTTTT ATTATTTTTC TTCCTCCGCA TGTTCAATCC GCTCCATAAT 1320

CGACGGATGG CTCCCTCTGA AAATTTTAAC GAGAAACGGC GGGTTGACCC GGCTCAGTCC 1380

CGTAACGGCC AACTCCTGAA ACGTCTCAAT CGCCGCTTCC CGGTTTCCGG TCAGCTCAAT 1440

GCCATAACGG TCGGCGGCGT TTTCCTGATA CCGGGAGACG GCATTCGTAA TCGGATC 1497

(2) SEQ ID NO:7的信息：

(i) 序列特征：

(A) 长度：275 氨基酸

(B) 类型：氨基酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：蛋白质

(xi) 序列描述：SEQ ID NO:7：

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

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

(2) SEQ ID NO:8的信息：

(i) 序列特征：

(A) 长度：275 氨基酸

(B) 类型：氨基酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：蛋白质

(xi) 序列描述：SEQ ID NO:8：

Ala Gln Ser Val Pro Tyr Gly Ile Ser Gln Ile Lys Ala Pro Ala Leu
1 5 10 15
His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp
20 25 30
Ser Gly Ile Asp Ser Ser His Pro Asp Leu Asn Val Arg Gly Gly Ala

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

(2) SEQ ID NO:9的信息：

(i)序列特征：

(A) 长度：274 氨基酸

(B) 类型：氨基酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：蛋白质

(xi) 序列描述：SEQ ID NO:9：

Ala Gln Thr Val Pro Tyr Gly Ile Pro Leu Ile Lys Ala Asp Lys Val
1 5 10 15

Gln Ala Gln Gly Phe Lys Gly Ala Asn Val Lys Val Ala Val Leu Asp
20 25 30

Thr Gly Ile Gln Ala Ser His Pro Asp Leu Asn Val Val Gly Gly Ala
35 40 45

Ser Phe Val Ala Gly Glu Ala Tyr Asn Thr Asp Gly Asn Gly His Gly
50 55 60

Thr His Val Ala Gly Thr Val Ala Ala Leu Asp Asn Thr Thr Gly Val
65 70 75 80

Leu Gly Val Ala Pro Ser Val Ser Leu Tyr Ala Val Lys Val Leu Asn
85 90 95

Ser Ser Gly Ser Gly Ser Tyr Ser Gly Ile Val Ser Gly Ile Glu Trp
100 105 110

Ala Thr Thr Asn Gly Met Asp Val Ile Asn Met Ser Leu Gly Gly Ala
115 120 125

Ser Gly Ser Thr Ala Met Lys Gln Ala Val Asp Asn Ala Tyr Ala Arg
130 135 140

Gly Val Val Val Val Ala Ala Ala Gly Asn Ser Gly Asn Ser Gly Ser
145 150 155 160

Thr Asn Thr Ile Gly Tyr Pro Ala Lys Tyr Asp Ser Val Ile Ala Val
165 170 175

Gly Ala Val Asp Ser Asn Ser Asn Arg Ala Ser Phe Ser Ser Val Gly
180 185 190
Ala Glu Leu Glu Val Met Ala Pro Gly Ala Gly Val Tyr Ser Thr Tyr
195 200 205
Pro Thr Asn Thr Tyr Ala Thr Leu Asn Gly Thr Ser Met Ala Ser Pro
210 215 220
His Val Ala Gly Ala Ala Ala Leu Ile Leu Ser Lys His Pro Asn Leu
225 230 235 240
Ser Ala Ser Gln Val Arg Asn Arg Leu Ser Ser Thr Ala Thr Tyr Leu
245 250 255
Gly Ser Ser Phe Tyr Tyr Gly Lys Gly Leu Ile Asn Val Glu Ala Ala
260 265 270
Ala Gln

(2) SEQ ID NO:10的信息：

(i) 序列特征：

(A) 长度：269 氨基酸

(B) 类型：氨基酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：蛋白质

(xi) 序列描述：SEQ ID NO:10：

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

His Val Ala Gly Thr Ile Ala Ala Leu Asn Asn Ser Ile Gly Val Leu
65 70 75 80
Gly Val Ala Pro Ser Ala Glu Leu Tyr Ala Val Lys Val Leu Gly Ala
85 90 95
Ser Gly Ser Gly Ser Val Ser Ser Ile Ala Gln Gly Leu Glu Trp Ala
100 105 110
Gly Asn Asn Gly Met His Val Ala Asn Leu Ser Leu Gly Ser Pro Ser
115 120 125
Pro Ser Ala Thr Leu Glu Gln Ala Val Asn Ser Ala Thr Ser Arg Gly
130 135 140
Val Leu Val Val Ala Ala Ser Gly Asn Ser Gly Ala Gly Ser Ile Ser
145 150 155 160
Tyr Pro Ala Arg Tyr Ala Asn Ala Met Ala Val Gly Ala Thr Asp Gln
165 170 175
Asn Asn Asn Arg Ala Ser Phe Ser Gln Tyr Gly Ala Gly Leu Asp Ile
180 185 190
Val Ala Pro Gly Val Asn Val Gln Ser Thr Tyr Pro Gly Ser Thr Tyr
195 200 205
Ala Ser Leu Asn Gly Thr Ser Met Ala Thr Pro His Val Ala Gly Ala
210 215 220
Ala Ala Leu Val Lys Gln Lys Asn Pro Ser Trp Ser Asn Val Gln Ile
225 230 235 240
Arg Asn His Leu Lys Asn Thr Ala Thr Ser Leu Gly Ser Thr Asn Leu
245 250 255
Tyr Gly Ser Gly Leu Val Asn Ala Glu Ala Ala Thr Arg
260 265

(2) SEQ ID NO:11的信息：

(i) 序列特征：

(A) 长度：1140 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA (基因组)

(xi) 序列描述：SEQ ID NO:11：

ATGAAGAAAC CGTTGGGGAA AATTGTCGCA AGCACCGCAC TACTCATTTC TGTTGCTTTT 60

AGTTCATCGA TCGCATCGGC TGCTGAAGAA GCAAAAGAAA AATATTTAAT TGGCTTTAAT 120

GAGCAGGAAG CTGTCAGTGA GTTTGTAGAA CAAGTAGAGG CAAATGACGA GGTCGCCATT 180

CTCTCTGAGG AAGAGGAAGT CGAAATTGAA TTGCTTCATG AATTTGAAAC GATTCCTGTT 240

TTATCCGTTG AGTTAAGCCC AGAAGATGTG GACGCGCTTG AACTCGATCC AGCGATTTCT 300

TATATTGAAG AGGATGCAGA AGTAACGACA ATGGCGCAAT CAGTGCCATG GGGAATTAGC 360

CGTGTGCAAG CCCCAGCTGC CCATAACCGT GGATTGACAG GTTCTGGTGT AAAAGTTGCT 420

GTCCTCGATA CAGGTATTTC CACTCATCCA GACTTAAATA TTCGTGGTGG CGCTAGCTTT 480

GTACCAGGGG AACCATCCAC TCAAGATGGG AATGGGCATG GCACGCATGT GGCCGGGACG 540

ATTGCTGCTT TAAACAATTC GATTGGCGTT CTTGGCGTAG CGCCGAGCGC GGAACTATAC 600

GCTGTTAAAG TATTAGGGGC GAGCGGTTCA GGTTCGGTCA GCTCGATTGC CCAAGGATTG 660

GAATGGGCAG GGAACAATGG CATGCACGTT GCTAATTTGA GTTTAGGAAG CCCTTCGCCA 720

AGTGCCACAC TTGAGCAAGC TGTTAATAGC GCGACTTCTA GAGGCGTTCT TGTTGTAGCG 780

GCATCTGGGA ATTCAGGTGC AGGCTCAATC AGCTATCCGG CCCGTTATGC GAACGCAATG 840

GCAGTCGGAG CTACTGACCA AAACAACAAC CGCGCCAGCT TTTCACAGTA TGGCGCAGGG 900

CTTGACATTG TCGCACCAGG TGTAAACGTG CAGAGCACAT ACCCAGGTTC AACGTATGCC 960

AGCTTAAACG GTACATCGAT GGCTACTCCT CATGTTGCAG GTGCAGCAGC CCTTGTTAAA 1020

CAAAAGAACC CATCTTGGTC CAATGTACAA ATCCGCAATC ATCTAAAGAA TACGGCAACG 1080

AGCTTAGGAA GCACGAACTT GTATGGAAGC GGACTTGTCA ATGCAGAAGC GGCAACACGC 1140

(2)SEQIDNO:12的信息：

(i) 序列特征：

(A) 长度：1140 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA（基因组）

(xi) 序列描述：SEQ ID NO:12：

ATGAAGAAAC CGTTGGGGAA AATTGTCGCA AGCACCGCAC TACTCATTTC TGTTGCTTTT 60

AGTTCATCGA TCGCATCGGC TGCTGAAGAA GCAAAAGAAA AATATTTAAT TGGCTTTAAT 120

GAGCAGGAAG CTGTCAGTGA GTTTGTAGAA CAAGTAGAGG CAAATGACGA GGTCGCCATT 180

CTCTCTGAGG AAGAGGAAGT CGAAATTGAA TTGCTTCATG AATTTGAAAC GATTCCTGTT 240

TTATCCGTTG AGTTAAGCCC AGAAGATGTG GACGCGCTTG AACTCGATCC AGCGATTTCT 300

TATATTGAAG AGGATGCAGA AGTAACGACA ATGGCGCAAT CAGTGCCATG GGGAATTAGC 360

CGTGTGCAAG CCCCAGCTGC CCATAACCGT GGATTGACAG GTTCTGGTGT AAAAGTTGCT 420

GTCCTCGATA CAGGTATTTC CACTCATCCA GACTTAAATA TTCGTGGTGG CGCTAGCTTT 480

GTACCAGGGG AACCATCCAC TCAAGATGGG AATGGGCATG GCACGCATGT GGCCGGGACG 540

ATTGCTGCTT TAGACAACTC GATTGGCGTT CTTGGCGTAG CGCCGAGCGC GGAACTATAC 600

GCTGTTAAAG TATTAGGGGC GAGCGGTTCA GGCGCCATCA GCTCGATTGC CCAAGGATTG 660

GAATGGGCAG GGAACAATGG CATGCACGTT GCTAATTTGA GTTTAGGAAG CCCTTCGCCA 720

AGTGCCACAC TTGAGCAAGC TGTTAATAGC GCGACTTCTA GAGGCGTTCT TGTTGTAGCG 780

GCATCTGGGA ATTCAGGTGC AGGCTCAATC AGCTATCCGG CCCGTTATGC GAACGCAATG 840

GCAGTCGGAG CTACTGACCA AAACAACAAC CGCGCCAGCT TTTCACAGTA TGGCGCAGGG 900

CTTGACATTG TCGCACCAGG TGTAAACGTG CAGAGCACAT ACCCAGGTTC AACGTATGCC 960

AGCTTAAACG GTACATCGAT GGCTACTCCT CATGTTGCAG GTGCAGCAGC CCTTGTTAAA 1020

CAAAAGAACC CATCTTGGTC CAATGTACAA ATCCGCAATC ATCTAAAGAA TACGGCAACG 1080

AGCTTAGGAA GCACGAACTT GTATGGAAGC GGACTTGTCA ATGCAGAAGC GGCAACACGC 1140

(2)SEQ ID NO:13的信息：

(i) 序列特征：

(A) 长度：30 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA（基因组）

(xi) 序列描述：SEQ ID NO:13：

TATGCCAGCC ACAACGGTAC TTCGATGGCT 30

(2) SEQ ID NO:14的信息：

(i) 序列特征：

(A) 长度：31 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA（基因组）

(xi) 序列描述：SEQ ID NO:14：

CACAGTTGCG GCTCTAGATA ACTCAATCGG T 31

(2) SEQ ID NO:15的信息：

(i) 序列特征：

(A) 长度：33 碱基对

(B) 类型：核酸

(C) 股性：单链

(D) 拓扑结构：线性

(ii) 分子类型：DNA（基因组）

(xi) 序列描述：SEQ ID NO:15：

GCTGACGGTT CCGGCGCTAT TAGTTGGATC ATT 33

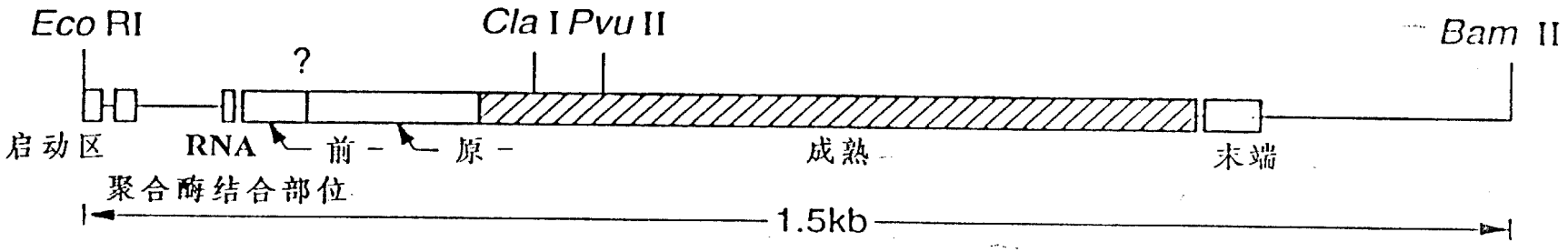


图 1A

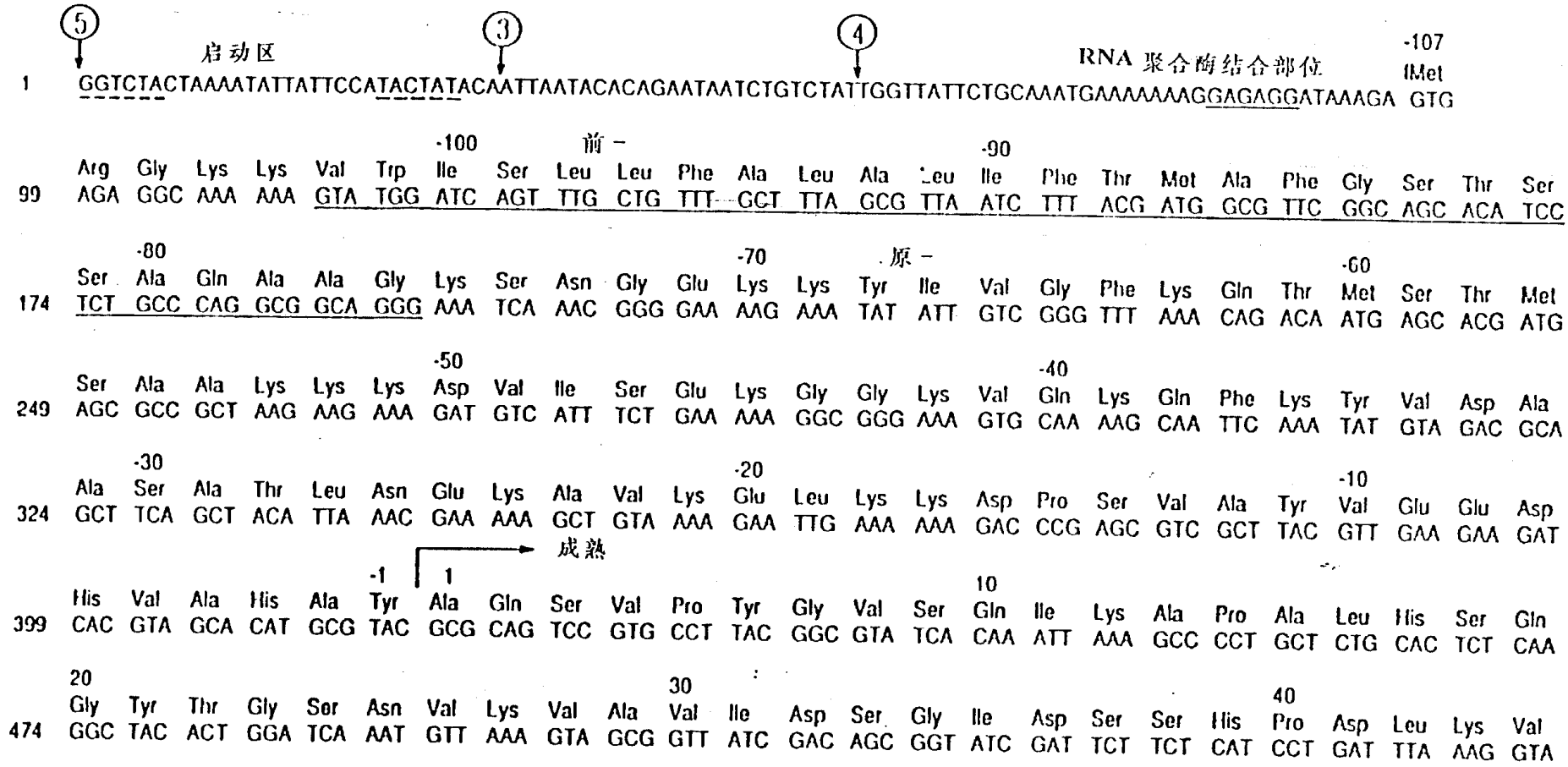

⑤
启动区
③
④
RNA 聚合酶结合部位
-107
fMet
1 GGTCTACTAAAATATTATTCCATACTATACAATTAATACACAGAATAATCTGTCTATTGGTTATTCTGCAAATGAAAAAAAGGAGAGGATAAAGA GTG
-100
前 -
-90
99 Arg Gly Lys Lys Val Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu Ile Phe Thr Met Ala Phe Gly Ser Thr Ser
AGA GGC AAA AAA GTA TGG ATC AGT TTG CTG TTT GCT TTA GCG TTA ATC TTT ACG ATG GCG TTC GGC AGC ACA TCC
-80
-70
原 -
-60
174 Ser Ala Gln Ala Ala Gly Lys Ser Asn Gly Glu Lys Lys Tyr Ile Val Gly Phe Lys Gln Thr Met Ser Thr Met
TCT GCC CAG GCG GCA GGG AAA TCA AAC GGG GAA AAG AAA TAT ATT GTC GGG TTT AAA CAG ACA ATG AGC ACG ATG
-50
-40
249 Ser Ala Ala Lys Lys Lys Asp Val Ile Ser Glu Lys Gly Gly Lys Val Gln Lys Gln Phe Lys Tyr Val Asp Ala
AGC GCC GCT AAG AAG AAA GAT GTC ATT TCT GAA AAA GGC GGG AAA GTG CAA AAG CAA TTC AAA TAT GTA GAC GCA
-30
-20
-10
324 Ala Ser Ala Thr Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala Tyr Val Glu Glu Asp
GCT TCA GCT ACA TTA AAC GAA AAA GCT GTA AAA GAA TTG AAA AAA GAC CCG AGC GTC GCT TAC GTT GAA GAA GAT
成熟
-1
1
10
399 His Val Ala His Ala Tyr Ala Gln Ser Val Pro Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln
CAC GTA GCA CAT GCG TAC GCG CAG TCC GTG CCT TAC GGC GTA TCA CAA ATT AAA GCC CCT GCT CTG CAC TCT CAA
20
30
40
474 Gly Tyr Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser Ser His Pro Asp Leu Lys Val
GGC TAC ACT GGA TCA AAT GTT AAA GTA GCG GTT ATC GAC AGC GGT ATC GAT TCT TCT CAT CCT GAT TTA AAG GTA

图 1B—1

						50						Pro	Asn			60	Asp							
	Ala	Gly	Gly	Ala	Ser	Met	Val	Pro	Ser	Glu	Thr	Asn	Pro	Phe	Gln	Asp	Asn	Asn	Ser	His	Gly	Thr	His	Val	Ala
549	GCA	GGC	GGA	GCC	AGC	ATG	GTT	CCT	TCT	GAA	ACA	AAT	CCT	TTC	CAA	GAC	AAC	AAC	TCT	CAC	GGA	ACT	CAC	GTT	GCC
		70									80								Ser	Ala	90				
	Gly	Thr	Val	Ala	Ala	Leu	Asn	Asn	Ser	Ile	Gly	Val	Leu	Gly	Val	Ala	Pro	Ser	Ala	Ser	Leu	Tyr	Ala	Val	Lys
624	GGC	ACA	GTT	GCG	GCT	CTT	AAT	AAC	TCA	ATC	GGT	GTA	TTA	GGC	GTT	GCG	CCA	AGC	GCA	TCA	CTT	TAC	GCT	GTA	AAA
				Asp	Ala	100										110									
	Val	Leu	Gly	Ala	Asp	Gly	Ser	Gly	Gln	Tyr	Ser	Trp	Ile	Ile	Asn	Gly	Ile	Glu	Trp	Ala	Ile	Ala	Asn	Asn	Met
699	GTT	CTC	GGT	GCT	GAC	GGT	TCC	GGC	CAA	TAC	AGC	TGG	ATC	ATT	AAC	GGA	ATC	GAG	TGG	GCG	ATC	GCA	AAC	AAT	ATG
	120										130										140				
	Asp	Val	Ile	Asn	Met	Ser	Leu	Gly	Gly	Pro	Ser	Gly	Ser	Ala	Ala	Leu	Lys	Ala	Ala	Val	Asp	Lys	Ala	Val	Ala
774	GAC	GTT	ATT	AAC	ATG	AGC	CTC	GGC	GGA	CCT	TCT	GGT	TCT	GCT	GCT	TTA	AAA	GCG	GCA	GTT	GAT	AAA	GCC	GTT	GCA
						150								Ser	Thr	160									
	Ser	Gly	Val	Val	Val	Val	Ala	Ala	Ala	Gly	Asn	Glu	Gly	Thr	Ser	Gly	Ser	Ser	Ser	Thr	Val	Gly	Tyr	Pro	Gly
849	TCC	GGC	GTC	GTA	GTC	GTT	GCG	GCA	GCC	GGT	AAC	GAA	GGC	ACT	TCC	GGC	AGC	TCA	AGC	ACA	GTG	GGC	TAC	CCT	GGT
	170										180										190				
	Lys	Tyr	Pro	Ser	Val	Ile	Ala	Val	Gly	Ala	Val	Asp	Ser	Ser	Asn	Gln	Arg	Ala	Ser	Phe	Ser	Ser	Val	Gly	Pro
924	AAA	TAC	CCT	TCT	GTC	ATT	GCA	GTA	GGC	GCT	GTT	GAC	AGC	AGC	AAC	CAA	AGA	GCA	TCT	TTC	TCA	AGC	GTA	GGA	CCT
						200										210									
	Glu	Leu	Asp	Val	Met	Ala	Pro	Gly	Val	Ser	Ile	Gln	Ser	Thr	Leu	Pro	Gly	Asn	Lys	Tyr	Gly	Ala	Tyr	Asn	Gly
999	GAG	CTT	GAT	GTC	ATG	GCA	CCT	GGC	GTA	TCT	ATC	CAA	AGC	ACG	CTT	CCT	GGA	AAC	AAA	TAC	GGG	GCG	TAC	AAC	GGT
	220										230										240				
	Thr	Ser	Met	Ala	Ser	Pro	His	Val	Ala	Gly	Ala	Ala	Ala	Leu	Ile	Leu	Ser	Lys	His	Pro	Asn	Trp	Thr	Asn	Thr
1074	ACG	TCA	ATG	GCA	TCT	CCG	CAC	GTT	GCC	GGA	GCG	GCT	GCT	TTG	ATT	CTT	TCT	AAG	CAC	CCG	AAC	TGG	ACA	AAC	ACT

图 1B—2

```
                                 250  Gln                                                  260
     Gln  Val  Arg  Ser  Ser  Leu  Glu  Asn  Thr  Thr  Thr  Lys  Leu  Gly  Asp  Ser  Phe  Tyr  Tyr  Gly  Lys  Gly  Leu  Ile  Asn
1149 CAA  GTC  CGC  AGC  AGT  TTA  GAA  AAC  ACC  ACT  ACA  AAA  CTT  GGT  GAT  TCT  TTC  TAC  TAT  GGA  AAA  GGG  CTG  ATC  AAC

     270                      275
     Val  Gln  Ala  Ala  Ala  Gln  OC                          末端
1224 GTA  CAG  GCG  GCA  GCT  CAG  TAA  AACATAAAAAACCGGCCTTGGCCCCGCCGGTTTTTTATTTTTCTTCCTCCGCATGTTCAATCCGCTCC

1316 ATAATCGACGGATGGCTCCCTCTGAAAATTTTAACGAGAAACGGCGGGTTGACCCGGCTCAGTCCCGTAACGGCCAAGTCCTGAAACGTCTCAATCGCCG

1416 CTTCCCGGTTTCCGGTCAGCTCAATGCCGTAACGGTCGGCGGCGTTTTCCTGATACCGGGAGACGGCATTCGTAATCGGATC
```

图 1B—3

解淀粉芽孢杆菌枯草杆菌蛋白酶中的保守残基

```
1                   10                  20
A Q S V P . G . . . . . A P A . H . . G

21                  30                  40
. T G S . V K V A V . D . G . . . . H P

41                  50                  60
D L . . . G G A S . V P . . . . . Q D

61                  70                  80
. N . H G T H V A G T . A A L N N S I G

81                  90                  100
V L G V A P S A . L Y A V K V L G A . G

101                 110                 120
S G . . S . L . . G . E W A . N . . . .

121                 130                 140
V . N . S L G . P S . S . . . . A . .

141                 150                 160
. . . . . G V . V V A A . G N . G . . .

161                 170                 180
. . . . . Y P . . Y . . . . A V G A .

181                 190                 200
D . . N . . A S F S . . G . . L D . . A

201                 210                 220
P G V . . Q S T . P G . . Y . . . N G T

221                 230                 240
S M A . P H V A G A A A L . . . K . . .

241                 250                 260
W . . . Q . R . . L . N T . . . L G . .

261                 270
. . Y G . G L . N . . A A . .
```

图 2

枯草杆菌蛋白酶序列的比较:
解淀粉芽孢杆菌
枯草杆菌
地衣芽孢杆菌
嗜碱芽孢杆菌

图 3A

```
01                  10                  20                  30
A Q S V P Y G V S Q I K A P A L H S Q G Y T G S N V K V A V I D S G I D S S H P
A Q S V P Y G I S Q I K A P A L H S Q G Y T G S N V K V A V I D S G I D S S H P
A Q T V P Y G I P L I K A D K V Q A Q G F K G A N V K V A V L D T G I Q A S H P
A Q S V P W G I S R V Q A P A A H N R G L T G S G V K V A V L D T G I S T * H P

41                  50                  60                  70
D L K V A G G A S M V P S E T N P F Q D N N S H G T H V A G T V A A L N N S I G
D L N V R G G A S F V P S E T N P Y Q D G S S H G T H V A G T I A A L N N S I G
D L N V V G G A S F V A G E A Y N * T D G N G H G T H V A G T V A A L D N T T G
D L N I R G G A S F V P G E * P S T Q D G N G H G T H V A G T I A A L N N S I G

81                  90                  100                 110
V L G V A P S A S L Y A V K V L G A D G S G Q Y S W I I N G I E W A I A N N M D
V L G V S P S A S L Y A V K V L D S T G S G Q Y S W I I N G I E W A I S N N M D
V L G V A P S V S L Y A V K V L N S S G S G S Y S G I V S G I E W A T T N G M D
V L G V A P S A E L Y A V K V L G A S G S G S V S S I A Q G L E W A G N N G M H

121                 130                 140                 150
V I N M S L G G P S G S A A L K A A V D K A V A S G V V V V A A A G N E G T S G
V I N M S L G G P T G S T A L K T V V D K A V S S G I V V A A A A G N E G S S G
V I N M S L G G A S G S T A M K Q A V D N A Y A R G V V V V A A A G N S G N S G
V A N L S L G S P S P S A T L E Q A V N S A T S R G V L V V A A S G N S G A G S
```

```
161                 170                 180                 190
S S S T V G Y P G K Y P S V I A V G A V D S S N Q R A S F S S V G P E L D V M A
S T S T V G Y P A K Y P S T I A V G A V N S S N Q R A S F S S A G S E L D V M A
S T N T I G Y P A K Y D S V I A V G A V D S N S N R A S F S S V G A E L E V M A
* * * * I S Y P A R Y A N A M A V G A T D Q N N N R A S F S Q Y G A G L D I V A

201                 210                 220                 230
P G V S I Q S T L P G N K Y G A Y N G T S M A S P H V A G A A A L I L S K H P N
P G V S I Q S T L P G G T Y G A Y N G T S M A T P H V A G A A A L I L S K H P T
P G A G V Y S T Y P T N T Y A T L N G T S M A S P H V A G A A A L I L S K H P N
P G V N V Q S T Y P G S T Y A S L N G T S M A T P H V A G A A A L V K Q K N P S

241                 250                 260                 270
W T N T Q V R S S L E N T T T K L G D S F Y Y G K G L I N V Q A A A A Q
W T N A Q V R D R L E S T A T Y L G N S F Y Y G K G L I N V Q A A A A Q
L S A S Q V R N R L S S T A T Y L G S S F Y Y G K G L I N V E A A A A Q
W S N V Q I R N H L K N T A T S L G S T N L Y G S G L V N A E A A T R
```

图 3B

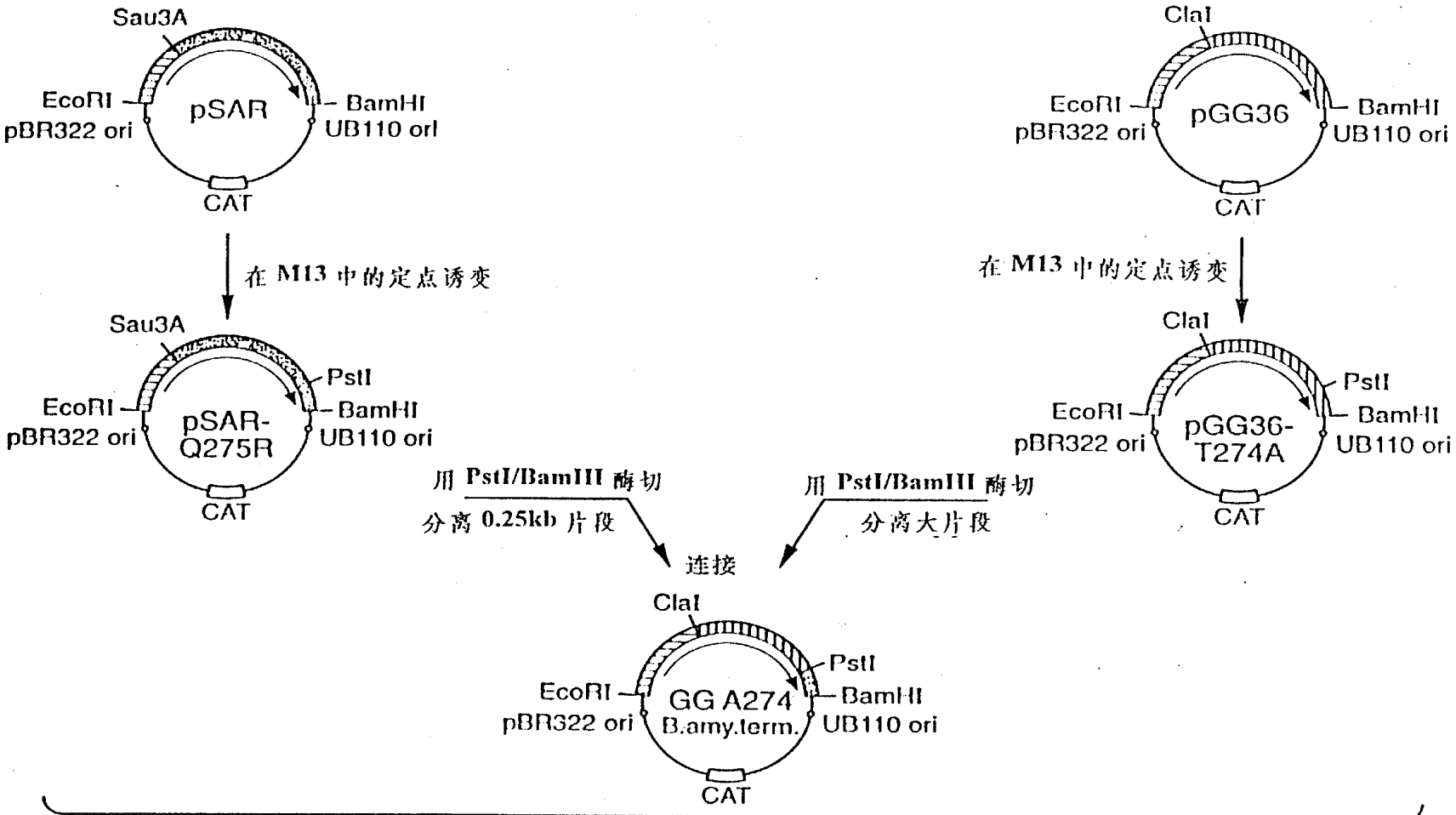

Sau3A
EcoRI
pBR322 ori
pSAR
BamHI
UB110 ori
CAT
在 M13 中的定点诱变
Sau3A
PstI
EcoRI
pBR322 ori
pSAR-
Q275R
BamHI
UB110 ori
CAT
ClaI
EcoRI
pBR322 ori
pGG36
BamHI
UB110 ori
CAT
在 M13 中的定点诱变
ClaI
PstI
EcoRI
pBR322 ori
pGG36-
T274A
BamHI
UB110 ori
CAT
用 PstI/BamHI 酶切
分离 0.25kb 片段
用 PstI/BamHI 酶切
分离大片段
连接
ClaI
PstI
EcoRI
pBR322 ori
GG A274
B.amy.term.
BamHI
UB110 ori
CAT

图 4

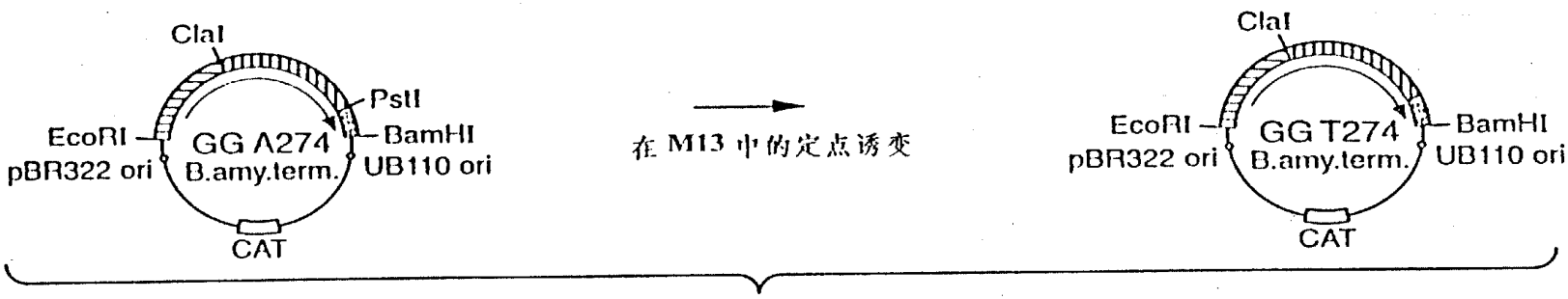

ClaI
PstI
EcoRI
GG A274
BamHI
pBR322 ori
B.amy.term.
UB110 ori
CAT
在 M13 中的定点诱变
ClaI
EcoRI
GG T274
BamHI
pBR322 ori
B.amy.term.
UB110 ori
CAT

图 5

```
        10                  30                  50
ATGAAGAAACCGTTGGGGAAAATTGTCGCAAGCACCGCACTACTCATTTCTGTTGCTTTT
MetLysLysProLeuGlyLysIleValAlaSerThrAlaLeuLeuIleSerValAlaPhe

        70                  90                  110
AGTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAAT
SerSerSerIleAlaSerAlaAlaGluGluAlaLysGluLysTyrLeuIleGlyPheAsn

        130                 150                 170
GAGCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATT
GluGlnGluAlaValSerGluPheValGluGlnValGluAlaAsnAspGluValAlaIle

        190                 210                 230
CTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTCCTGTT
LeuSerGluGluGluGluValGluIleGluLeuLeuHisGluPheGluThrIleProVal

        250                 270                 290
TTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACTCGATCCAGCGATTTCT
LeuSerValGluLeuSerProGluAspValAspAlaLeuGluLeuAspProAlaIleSer

        310                 330                 350
TATATTGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCAGTGCCATGGGGAATTAGC
TyrIleGluGluAspAlaGluValThrThrMetAlaGlnSerValProTrpGlyIleSer

        370                 390                 410
CGTGTGCAAGCCCCAGCTGCCCATAACCGTGGATTGACAGGTTCTGGTGTAAAAGTTGCT
ArgValGlnAlaProAlaAlaHisAsnArgGlyLeuThrGlySerGlyValLysValAla
```

图 6A

```
        430                   450                   470
GTCCTCGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTGGCGCTAGCTTT
ValLeuAspThrGlyIleSerThrHisProAspLeuAsnIleArgGlyGlyAlaSerPhe

        490                   510                   530
GTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCCGGGACG
ValProGlyGluProSerThrGlnAspGlyAsnGlyHisGlyThrHisValAlaGlyThr

        550                   570                   590
ATTGCTGCTTTAAACAATTCGATTGGCGTTCTTGGCGTAGCGCCGAGCGCGGAACTATAC
IleAlaAlaLeuAsnAsnSerIleGlyValLeuGlyValAlaProSerAlaGluLeuTyr

        610                   630                   650
GCTGTTAAAGTATTAGGGGCGAGCGGTTCAGGTTCGGTCAGCTCGATTGCCCAAGGATTG
AlaValLysValLeuGlyAlaSerGlySerGlySerValSerSerIleAlaGlnGlyLeu

        670                   690                   710
GAATGGGCAGGGAACAATGGCATGCACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCA
GluTrpAlaGlyAsnAsnGlyMetHisValAlaAsnLeuSerLeuGlySerProSerPro

        730                   750                   770
AGTGCCACACTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTTGTAGCG
SerAlaThrLeuGluGlnAlaValAsnSerAlaThrSerArgGlyValLeuValValAla

        790                   810                   830
GCATCTGGGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCCGTTATGCGAACGCAATG
AlaSerGlyAsnSerGlyAlaGlySerIleSerTyrProAlaArgTyrAlaAsnAlaMet
```

图 6B

```
        850                       870                       890
GCAGTCGGAGCTACTGACCAAAACAACAACCGCGCCAGCTTTTCACAGTATGGCGCAGGG
AlaValGlyAlaThrAspGlnAsnAsnAsnArgAlaSerPheSerGlnTyrGlyAlaGly

        910                       930                       950
CTTGACATTGTCGCACCAGGTGTAAACGTGCAGAGCACATACCCAGGTTCAACGTATGCC
LeuAspIleValAlaProGlyValAsnValGlnSerThrTyrProGlySerThrTyrAla

        970                       990                      1010
AGCTTAAACGGTACATCGATGGCTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAA
SerLeuAsnGlyThrSerMetAlaThrProHisValAlaGlyAlaAlaAlaLeuValLys

       1030                      1050                      1070
CAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCAATCATCTAAAGAATACGGCAACG
GlnLysAsnProSerTrpSerAsnValGlnIleArgAsnHisLeuLysAsnThrAlaThr

       1090                      1110                      1130
AGCTTAGGAAGCACGAACTTGTATGGAAGCGGACTTGTCAATGCAGAAGCGGCAACACGC
SerLeuGlySerThrAsnLeuTyrGlySerGlyLeuValAsnAlaGluAlaAlaThrArg
```

图 6C

```
          10                  30                  50
ATGAAGAAACCGTTGGGGAAAATTGTCGCAAGCACCGCACTACTCATTTCTGTTGCTTTT
MetLysLysProLeuGlyLysIleValAlaSerThrAlaLeuLeuIleSerValAlaPhe

          70                  90                  110
AGTTCATCGATCGCATCGGCTGCTGAAGAAGCAAAAGAAAAATATTTAATTGGCTTTAAT
SerSerSerIleAlaSerAlaAlaGluGluAlaLysGluLysTyrLeuIleGlyPheAsn

          130                 150                 170
GAGCAGGAAGCTGTCAGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATT
GluGlnGluAlaValSerGluPheValGluGlnValGluAlaAsnAspGluValAlaIle

          190                 210                 230
CTCTCTGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAATTTGAAACGATTCCTGTT
LeuSerGluGluGluGluValGluIleGluLeuLeuHisGluPheGluThrIleProVal

          250                 270                 290
TTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACTCGATCCAGCGATTTCT
LeuSerValGluLeuSerProGluAspValAspAlaLeuGluLeuAspProAlaIleSer

          310                 330                 350
TATATTGAAGAGGATGCAGAAGTAACGACAATGGCGCAATCAGTGCCATGGGGAATTAGC
TyrIleGluGluAspAlaGluValThrThrMetAlaGlnSerValProTrpGlyIleSer

          370                 390                 410
CGTGTGCAAGCCCCAGCTGCCCATAACCGTGGATTGACAGGTTCTGGTGTAAAAGTTGCT
ArgValGlnAlaProAlaAlaHisAsnArgGlyLeuThrGlySerGlyValLysValAla
```

图 7A

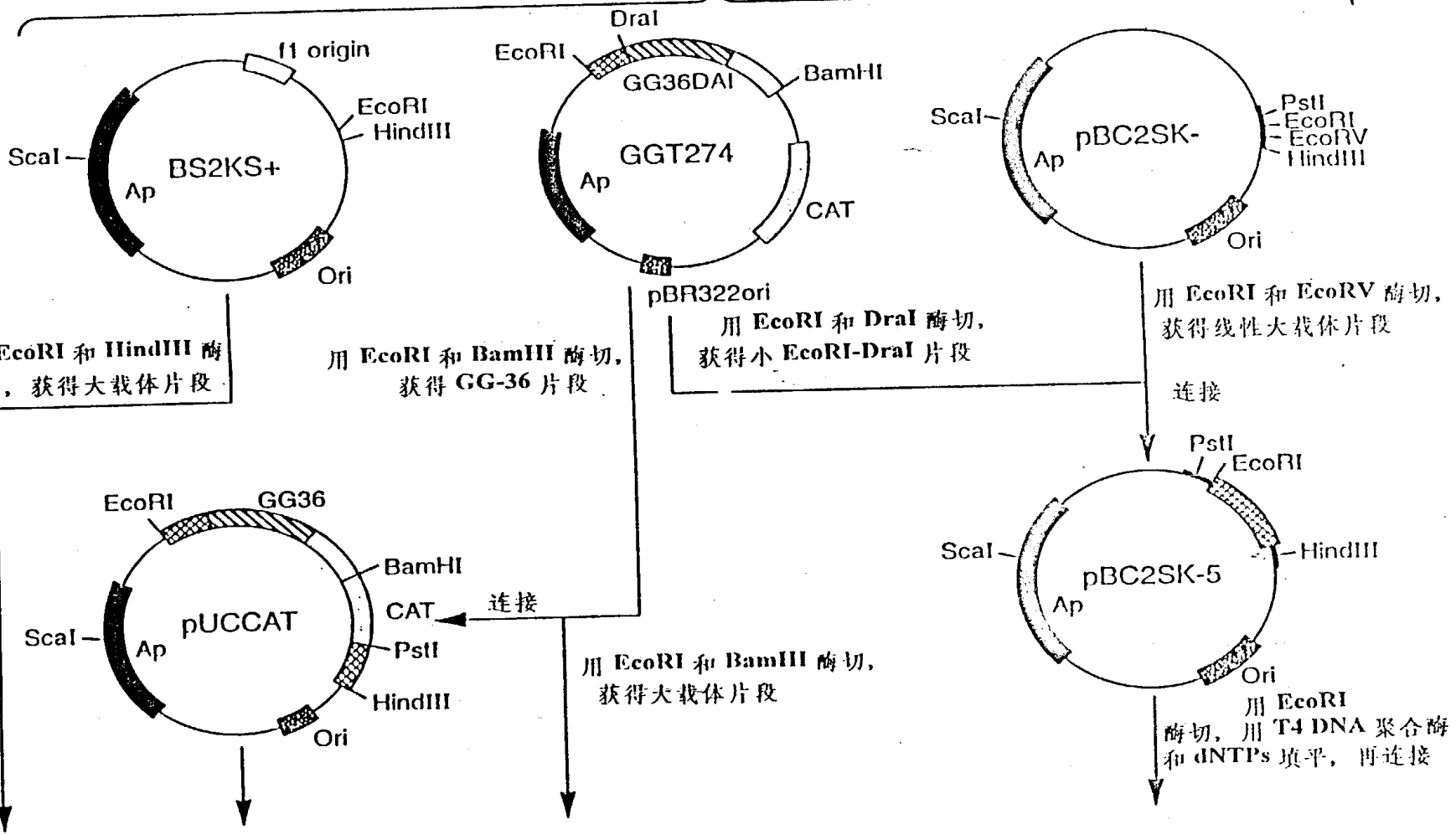
```
        430                       450                       470
GTCCTCGATACAGGTATTTCCACTCATCCAGACTTAAATATTCGTGGTGGCGCTAGCTTT
ValLeuAspThrGlyIleSerThrHisProAspLeuAsnIleArgGlyGlyAlaSerPhe

        490                       510                       530
GTACCAGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCCGGGACG
ValProGlyGluProSerThrGlnAspGlyAsnGlyHisGlyThrHisValAlaGlyThr

        550                       570                       590
ATTGCTGCTTTAGACAACTCGATTGGCGTTCTTGGCGTAGCGCCGAGCGCGGAACTATAC
IleAlaAlaLeuAspAsnSerIleGlyValLeuGlyValAlaProSerAlaGluLeuTyr

        610                       630                       650
GCTGTTAAAGTATTAGGGGCGAGCGGTTCAGGCGCCATCAGCTCGATTGCCCAAGGATTG
AlaValLysValLeuGlyAlaSerGlySerGlyAlaIleSerSerIleAlaGlnGlyLeu

        670                       690                       710
GAATGGGCAGGGAACAATGGCATGCACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCA
GluTrpAlaGlyAsnAsnGlyMetHisValAlaAsnLeuSerLeuGlySerProSerPro

        730                       750                       770
AGTGCCACACTTGAGCAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTTGTAGCG
SerAlaThrLeuGluGlnAlaValAsnSerAlaThrSerArgGlyValLeuValValAla

        790                       810                       830
GCATCTGGGAATTCAGGTGCAGGCTCAATCAGCTATCCGGCCCGTTATGCGAACGCAATG
AlaSerGlyAsnSerGlyAlaGlySerIleSerTyrProAlaArgTyrAlaAsnAlaMet
```

图 7B

```
       850                       870                       890
GCAGTCGGAGCTACTGACCAAAACAACAACCGCGCCAGCTTTTCACAGTATGGCGCAGGG
AlaValGlyAlaThrAspGlnAsnAsnAsnArgAlaSerPheSerGlnTyrGlyAlaGly

       910                       930                       950
CTTGACATTGTCGCACCAGGTGTAAACGTGCAGAGCACATACCCAGGTTCAACGTATGCC
LeuAspIleValAlaProGlyValAsnValGlnSerThrTyrProGlySerThrTyrAla

       970                       990                      1010
AGCTTAAACGGTACATCGATGGCTACTCCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAA
SerLeuAsnGlyThrSerMetAlaThrProHisValAlaGlyAlaAlaAlaLeuValLys

      1030                      1050                      1070
CAAAAGAACCCATCTTGGTCCAATGTACAAATCCGCAATCATCTAAAGAATACGGCAACG
GlnLysAsnProSerTrpSerAsnValGlnIleArgAsnHisLeuLysAsnThrAlaThr

      1090                      1110                      1130
AGCTTAGGAAGCACGAACTTGTATGGAAGCGGACTTGTCAATGCAGAAGCGGCAACACGC
SerLeuGlySerThrAsnLeuTyrGlySerGlyLeuValAsnAlaGluAlaAlaThrArg
```

图 7C

图 8A


f1 origin
EcoRI
HindIII
ScaI
BS2KS+
Ap
Ori
DraI
EcoRI
GG36DAI
BamHI
GGT274
Ap
CAT
pBR322ori
ScaI
pBC2SK-
PstI
EcoRI
EcoRV
HindIII
Ap
Ori
用 EcoRI 和 HindIII 酶切，获得大载体片段
用 EcoRI 和 BamHI 酶切，获得 GG-36 片段
用 EcoRI 和 DraI 酶切，获得小 EcoRI-DraI 片段
用 EcoRI 和 EcoRV 酶切，获得线性大载体片段
连接
EcoRI
GG36
BamHI
CAT
ScaI
pUCCAT
Ap
PstI
HindIII
Ori
连接
用 EcoRI 和 BamHI 酶切，获得大载体片段
PstI
EcoRI
ScaI
pBC2SK-5
HindIII
Ap
Ori
用 EcoRI 酶切，用 T4 DNA 聚合酶和 dNTPs 填平，再连接

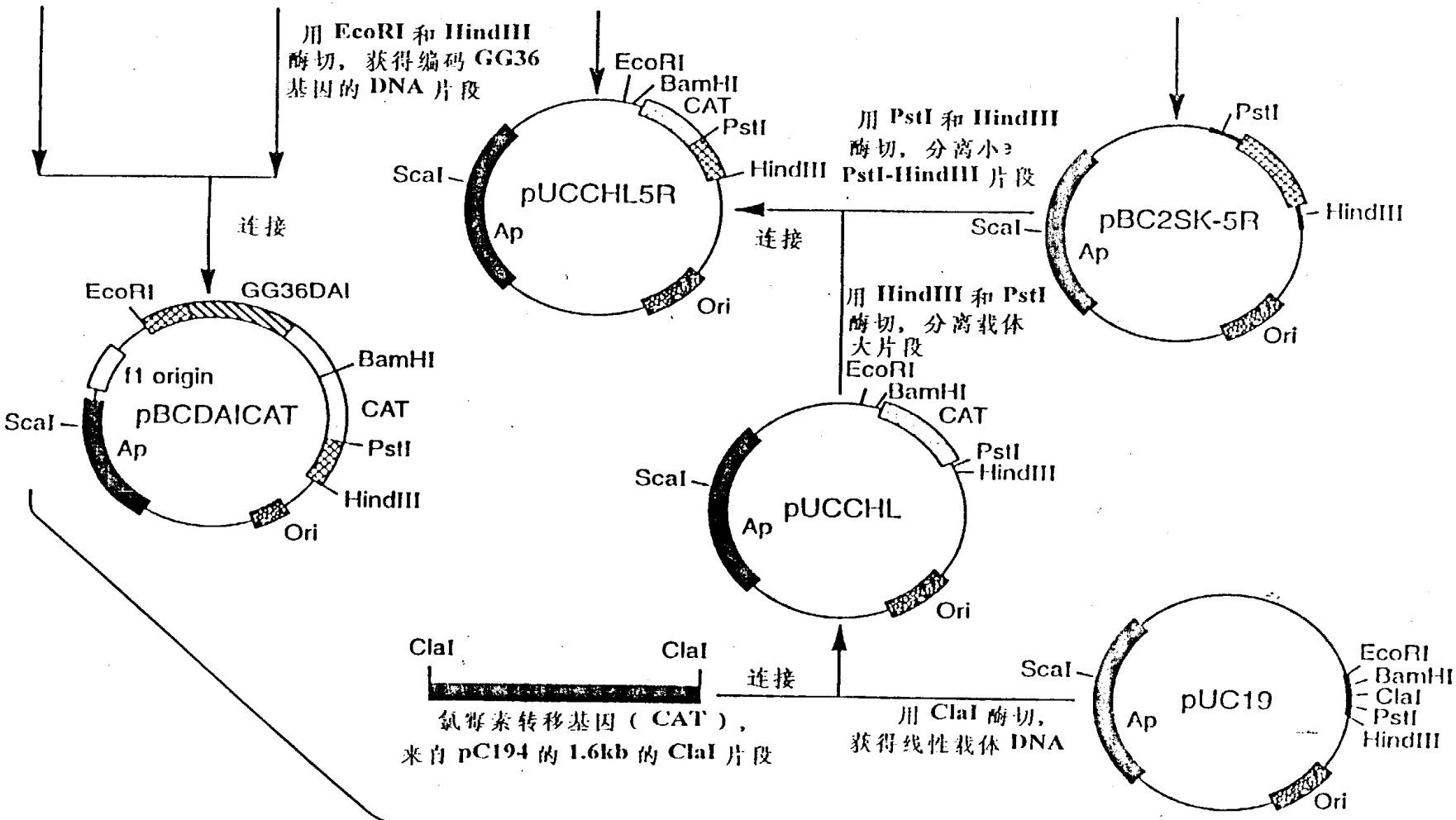
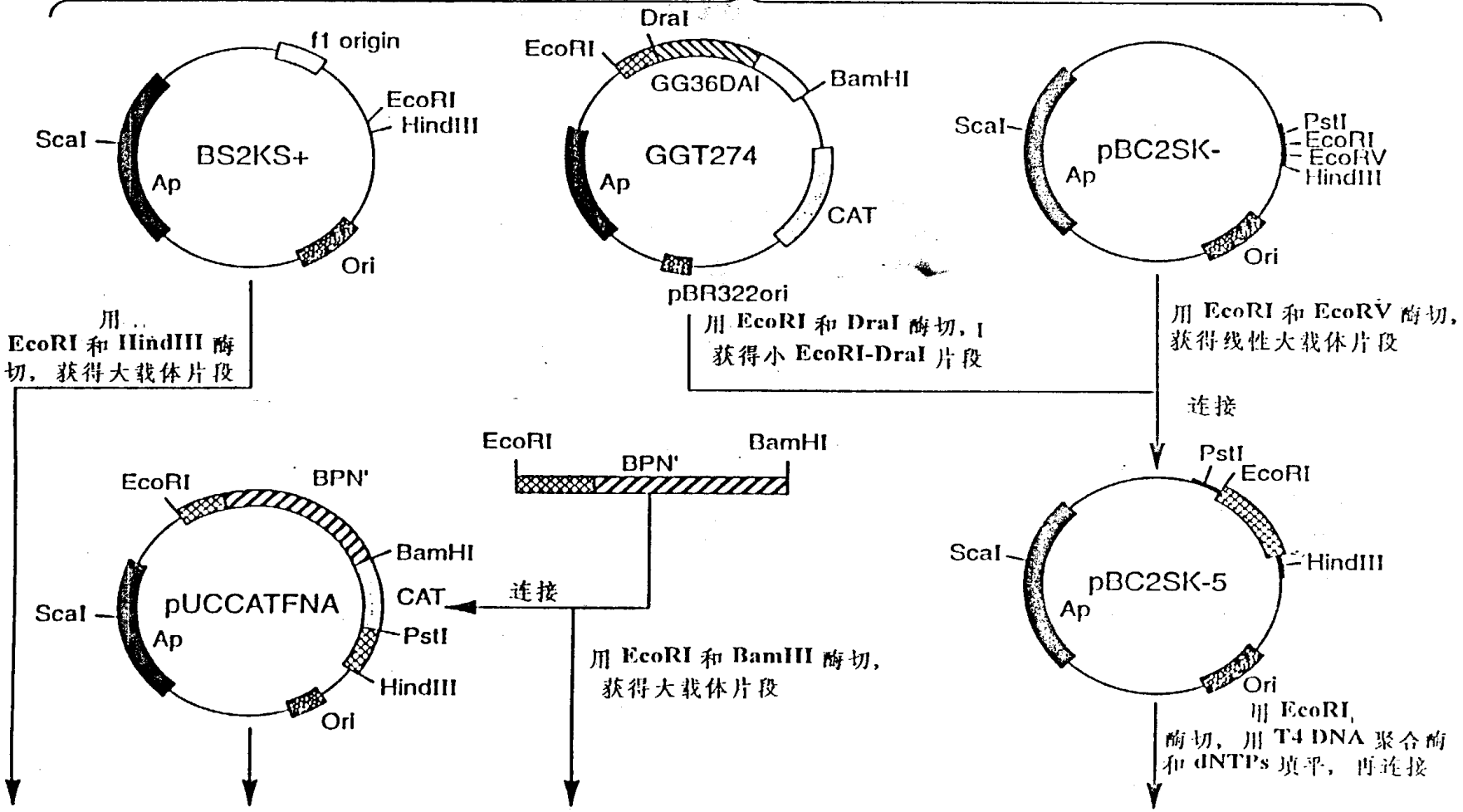

用 EcoRI 和 HindIII
酶切，获得编码 GG36
基因的 DNA 片段
连接
EcoRI
GG36DAI
f1 origin
BamHI
CAT
PstI
HindIII
Ori
ScaI
Ap
pBCDAICAT
EcoRI
BamHI
CAT
PstI
HindIII
ScaI
Ap
pUCCHL5R
Ori
连接
用 PstI 和 HindIII
酶切，分离小3
PstI-HindIII 片段
PstI
HindIII
ScaI
Ap
pBC2SK-5R
Ori
用 HindIII 和 PstI
酶切，分离载体
大片段
EcoRI
BamHI
CAT
PstI
HindIII
ScaI
Ap
pUCCHL
Ori
ClaI
ClaI
氯霉素转移基因（CAT），
来自 pC194 的 1.6kb 的 ClaI 片段
连接
用 ClaI 酶切，
获得线性载体 DNA
ScaI
Ap
pUC19
EcoRI
BamHI
ClaI
PstI
HindIII
Ori

图 8B

图 9A


f1 origin
EcoRI
HindIII
ScaI
BS2KS+
Ap
Ori
DraI
EcoRI
GG36DAI
BamHI
GGT274
Ap
CAT
pBR322ori
ScaI
pBC2SK-
Ap
PstI
EcoRI
EcoRV
HindIII
Ori
用 EcoRI 和 HindIII 酶切，获得大载体片段
用 EcoRI 和 DraI 酶切，获得小 EcoRI-DraI 片段
用 EcoRI 和 EcoRV 酶切，获得线性大载体片段
连接
EcoRI
BPN'
BamHI
EcoRI
BPN'
BamHI
ScaI
pUCCATFNA
Ap
CAT
PstI
HindIII
Ori
连接
用 EcoRI 和 BamHI 酶切，获得大载体片段
PstI
EcoRI
ScaI
pBC2SK-5
Ap
HindIII
Ori
用 EcoRI，酶切，用 T4 DNA 聚合酶和 dNTPs 填平，再连接

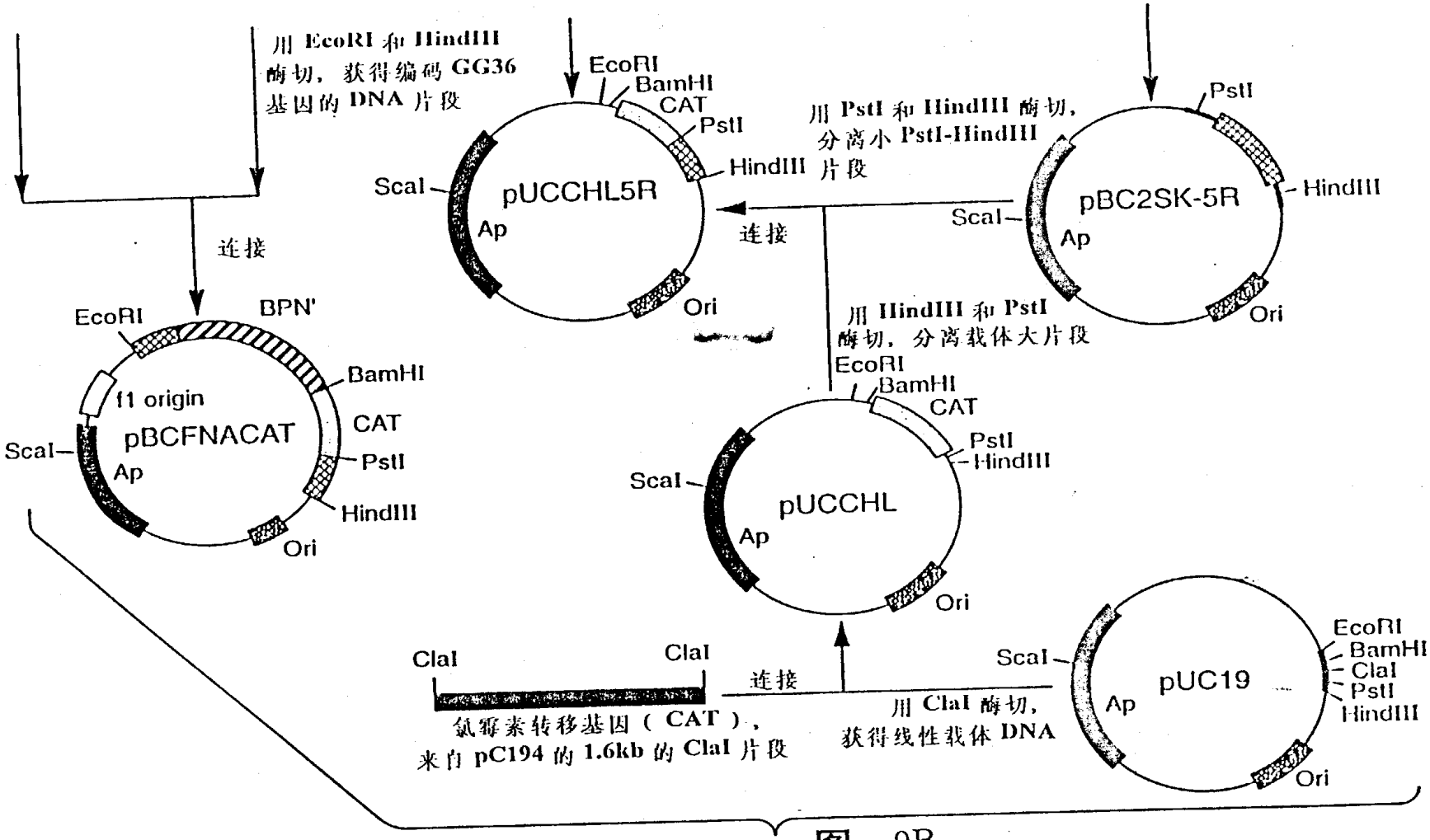

用 EcoRI 和 HindIII
酶切，获得编码 GG36
基因的 DNA 片段
连接
EcoRI
BPN'
BamHI
f1 origin
CAT
pBCFNACAT
ScaI
Ap
PstI
HindIII
Ori
EcoRI
BamHI
CAT
PstI
HindIII
ScaI
pUCCHL5R
Ap
Ori
用 PstI 和 HindIII 酶切，
分离小 PstI-HindIII
片段
连接
PstI
HindIII
ScaI
pBC2SK-5R
Ap
Ori
用 HindIII 和 PstI
酶切，分离载体大片段
EcoRI
BamHI
CAT
PstI
HindIII
ScaI
pUCCHL
Ap
Ori
ClaI
ClaI
氯霉素转移基因（CAT），
来自 pC194 的 1.6kb 的 ClaI 片段
连接
用 ClaI 酶切，
获得线性载体 DNA
ScaI
pUC19
Ap
EcoRI
BamHI
ClaI
PstI
HindIII
Ori

图 9B

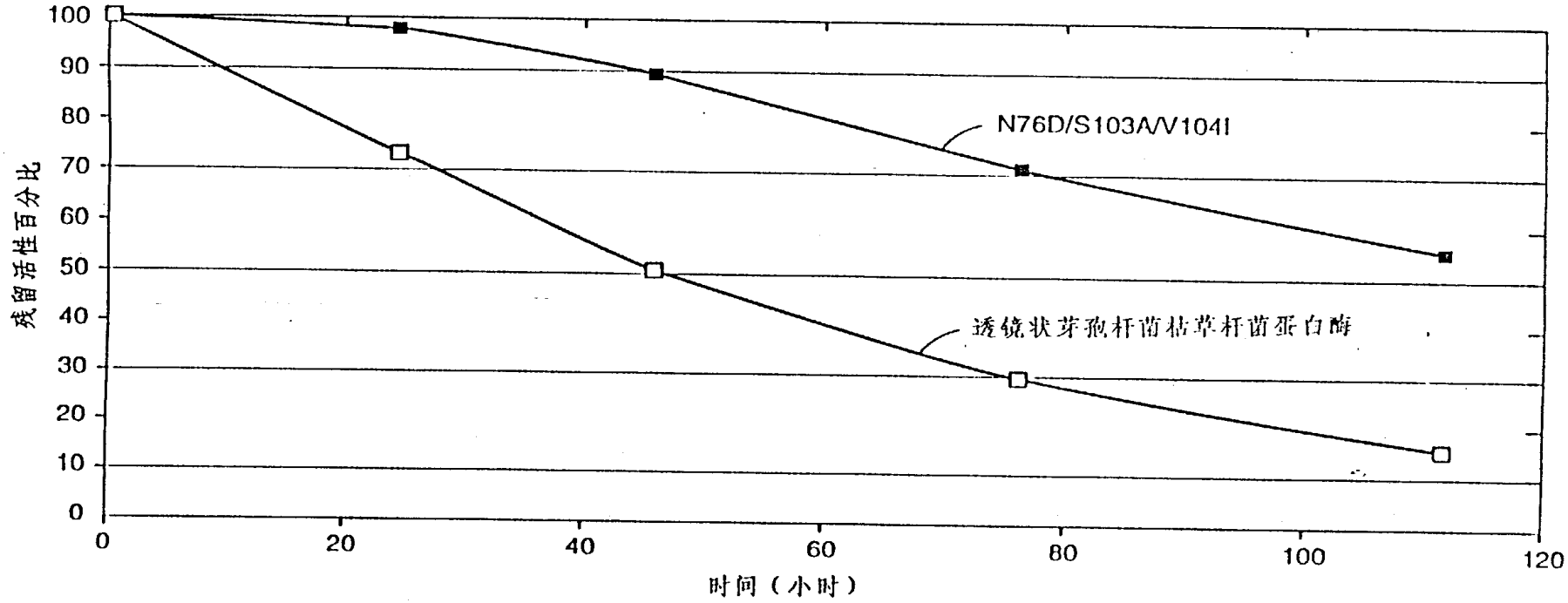

100
90
80
70
60
50
40
30
20
10
0
残留活性百分比
0
20
40
60
80
100
120
时间（小时）
N76D/S103A/V104I
透镜状芽孢杆菌枯草杆菌蛋白酶

图 10