



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 00805245. X

[45] 授权公告日 2004 年 12 月 8 日

[11] 授权公告号 CN 1179041C

[22] 申请日 2000.1.21 [21] 申请号 00805245. X
[30] 优先权
[32] 1999. 1.22 [33] DK [31] PA199900092
[32] 1999. 9.21 [33] DK [31] PA199901340
[86] 国际申请 PCT/DK2000/000025 2000.1.21
[87] 国际公布 WO2000/043503 英 2000.7.27
[85] 进入国家阶段日期 2001.9.19
[71] 专利权人 诺维信公司
地址 丹麦鲍斯韦
[72] 发明人 马丁·莱曼 索伦·F·拉森
审查员 陈 敏

[74] 专利代理机构 北京市柳沈律师事务所
代理人 巫肖南

权利要求书 2 页 说明书 144 页 附图 56 页

[54] 发明名称 改进的植酸酶

[57] 摘要

本发明涉及改进的植酸酶，优选是增加了热稳定性的植酸酶，以及产生植酸酶的方法。具体的，将稳定的氨基酸突变导入到同源蛋白中，或使植酸酶的活性位点部分或全部被取代。本发明还公开了相应的 DNA 序列及其制备方法，以及生产改进的植酸酶的方法，及其应用。本发明还公开了烟曲霉植酸酶和共有植酸酶的特异性突变体。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种植酸酶，其含有与 SEQ ID NO:26 中氨基酸-26 到+441 的序列具有至少 93.80 % 一致性的氨基酸序列，所述一致性用 GCG 软件包 9.1 版所提供的 GAP 程序以下述设置确定：长度权重为 0，缺口权重为 3.0。
2. 一种植酸酶，其由与 SEQ ID NO:25 之核苷酸 12-1412 具有至少 95.88 % 一致性的 DNA 序列编码，所述一致性用 GCG 软件包 9.1 版所提供的 GAP 程序以下述设置确定：GAP 产生罚分为 50，GAP 延伸罚分为 3。
3. 一种植酸酶，其含有选自下组的氨基酸序列
- 10 (i) SEQ ID NO:26 的序列-26 至+441， +1 至+441，或+1 至+438，或
- (ii) 由 SEQ ID NO:25 的核苷酸 12-1412 或 90-1412 编码的氨基酸序列。
4. 一种植酸酶，其包含选自下组的氨基酸序列
- (i) SEQ ID NO:26 的变体 K94A+V158I+A396S，或
- (ii) (i) 的变体 Q50T， K91A，或 Q50T+K91A，或
- 15 (iii) 序列(i)和(ii)之任一的序列+1 至 441，或
- (iv) SEQ ID NO:31 的序列-26 至+441，或+1 至+441；或
- (v) 由 SEQ ID NO:30 的核苷酸 1-1401，或 79-1401 编码的氨基酸序列。
5. 一种植酸酶，其包含选自下组的氨基酸序列
- (i) SEQ ID NO:29 的变体 T50Q+A91K，或
- 20 (ii) (i) 的变体 Q50T， K91A，或 Q50T+K91A，或
- (iii) 序列(i)和(ii)之任一的序列+1 至 441，或
- (iv) SEQ ID NO:29 的序列-26 至+441，或+1 至+441；或
- (v) 由 SEQ ID NO:28 的核苷酸 1-1401，或 79-1401 编码的氨基酸序列。
6. 一种植酸酶，其含有 SEQ ID NO:27 的氨基酸序列。
- 25 7. 一种 DNA 序列，其含有编码权利要求 1-6 之任一植酸酶的 DNA 序列。
8. 一种含有编码植酸酶之 DNA 序列的 DNA 序列，所述编码植酸酶的 DNA 序列与 SEQ ID NO:25 中核苷酸 12-1412 具有至少 95.88 % 一致性，所述一致性用 GCG 软件包 9.1 版所提供的 GAP 程序以下述设置确定：GAP 产生罚分为 50，GAP 延伸罚分为 3。
- 30 9. 一种含有编码植酸酶之 DNA 序列的 DNA 序列，所述植酸酶含有与

SEQ ID NO:26 之氨基酸-26 至+441 具有至少 93.80% 一致性的氨基酸序列，所述一致性用 GCG 软件包 9.1 版所提供的 GAP 程序以下述设置确定：GAP 产生罚分为 50，GAP 延伸罚分为 3。

- 5 10. 一种含有编码植酸酶之 DNA 序列的 DNA 序列，其中所述编码植酸酶的 DNA 序列含有：
- (i)SEQ ID NO:25 的核苷酸 12-1412，或 90-1412；或
 - (ii)SEQ ID NO:30 的核苷酸 1-1401，或 79-1401；或
 - (iii)SEQ ID NO:28 的核苷酸 1-1401，或 79-1401。
11. 一种载体，其含有权利要求 7-10 之任一的 DNA 序列。
- 10 12. 一种微生物宿主细胞，其含有权利要求 7-10 之任一的 DNA 序列，或权利要求 11 的载体。
13. 一种生产植酸酶的方法，该方法包括在允许植酸酶产生的条件下培养权利要求 12 的宿主细胞，并从培养基中回收植酸酶。
14. 一种食品，其含有权利要求 1-6 之任一的植酸酶。
- 15 15. 一种饲料，其含有权利要求 1-6 之任一的植酸酶。
16. 一种药物组合物，其含有权利要求 1-6 之任一的植酸酶。

改进的植酸酶

- 5 植酸酶是一种将植酸盐(肌醇六磷酸)水解为肌醇和无机磷酸的酶。已知其是一种有价值的饲料添加剂。

本发明涉及改进的植酸酶，即具有改良特性(如改良的活性特性)的植酸酶，其参照为例如其衍生源或已知的植酸酶。改良的活性特性是指在植酸酶活性的至少一个相关方面进行改良，例如(非-排它列举)：pH 稳定性，
10 温度稳定性，pH 范围，温度范围，比活性(特别是相关于 pH 和温度的)，底物特异性，底物裂解模式，底物结合，位点特异性，磷酸从谷物中释放的速度和水平，反应速度，植酸盐降解速率)，释放的磷酸所达到的终水平。

改良的活性特性的例子为改良的比活性(如增加，如在 pH3，4，5，或 6 时增加)；改良的 pH 或温度范围；和/或改良(如增加)的热稳定性，例如当
15 使用差示扫描量热(DSC)测量时增加的解链温度。

本发明还涉及一种制备修饰型蛋白的方法，其中第一步是通过下述步骤 a)，b)和 c)从多个高度同源序列中测定出一个共有序列：

- a)通过本领域已知的任何标准比对程序对至少 3 个，优选至少 4 个氨基酸序列进行比对；
20 b)在氨基酸序列比对的每一位点，评价氨基酸的进化相似性，而共有残基通过本领域公知的标准程序选择，其中为了预测(calculation)共有残基，至少应保证在某指定残基仅出现在比对序列中的两个序列上时，所述程序能确定一个共有残基。然而，如果在所比对的氨基酸序列中有一亚组序列，其彼此间的相似性远高于与其它序列之间的相似性，则仅当其共有序列可
25 以相同于 EP 897985 所述的方式而确定时，或给该亚组中每一序列分配一个由 1 除以该亚组序列数所得的权重时，该亚组可用所述预测表示；

c)如果该程序未能在指定位置鉴定出共有序列，则可选择任何氨基酸，优选在此位置出现频率最高的氨基酸。

- 在本发明第二方面，使一个同源序列与该共有序列比对，并使该同源
30 序列中一或多个非共有残基被相应共有残基取代。

优选的，该同源氨基酸序列仅在所述程序于中度严谨条件下可明确确定

义共有残基的位置上进行氨基酸残基的取代，而在未能于中度严谨条件下确定优选的共有氨基酸的所有位置上，该同源蛋白的氨基酸保持不变。

在发明第三方面，测定了目的蛋白的活性中心，包括共有序列中以及同源蛋白的序列中参与形成活性中心的所有氨基酸残基；随后，将同源蛋白的部分或全部趋异氨基酸残基插入共有序列的骨架中。

在此方法的一个实施方案中，用于比对特定位置之氨基酸的进化相似性的程序为程序"PRETTY"。

蛋白质活性中心可利用蛋白质三维结构分析测定。

同源蛋白如酶家族，限定的蛋白家族的实例为植酸酶家族，例如真菌来源的。

例如，植酸酶的氨基酸序列可通过引入至少一个选自下列的突变或取代被改变：

	E58A	F54Y
	D69K	I73V
15	D197N	K94A
	T214L	R101A
	E222T	N153K
	E267D	V158I
	R291I	A203G
20	R329H	S205G
	S364T	V217A
	A379K	A227V
	G404A	V234L
		P238A
25		Q277E
		A287H
		A292Q
		V366I
		A396S
30		E415Q
		G437A

R451E

为了解释这些缩写，举一个例子，突变 E58A 被解释如下：当从该数字减去 26，可以得到在如图 1 所示共有植酸酶序列或另一被比对的植酸酶序列中的位置或残基号(对应于在图 1 所示序列中加入一段 26 个氨基酸的信号序列)。例如，在 E58A 中，数字 58 是指位置号 32(58-26=32)。该数字前的字母，即 E，代表该植酸酶中被修饰的氨基酸，其被所述数字之后的氨基酸，即 A，取代。

上述氨基酸取代，单独和/或联合，对氨基酸的稳定性有积极影响。

下组突变(即含有选自下组的至少一个突变的植酸酶)也是兴趣所在：

10 E58A, D69K, D197N, T214L, E222T, E267D, R291I, R329H, S364T, A379K, G404A;

F54Y, I73V, K94A, R101A, N153K, V158I, A203G, S205G, V217A, A227V, V234L, P238A, Q277E, A287H, A292Q, V366I, A396S, E415Q, G437A, R451E;

15 E58A, D69K, D197N, F54Y, I73V, K94A;

T214L, E222T, E267D, R101A, N153K, V158I;

R291I, R329H, S364T, A203G, S205G, V217A;

A379K, G404A, A227V, V234L, P238A, Q277E;

A287H, A292Q, V366I, A396S, E415Q, G437A, R451E;

20 T214L, E222T, S364T, V158I, A203G, G404A, A227V, P238A, A396S, G437A, R451E。

宿主细胞的实例为植物细胞，动物细胞，和微生物细胞，例如原核或真核细胞，如细菌，真菌或酵母细胞。真菌宿主的实例为曲霉属菌株，酵母宿主的实例为糖酵母属菌株，汉逊氏酵母属(*Hansenula*)菌株。

25 本发明也涉及通过上述任一方法可获得的或已获得的修饰蛋白。

本发明也涉及植酸酶例如(但不限于)共有植酸酶-1 的变体或突变蛋白，其中，在图 2 的氨基酸序列中，已进行至少一个下述取代：Q50L, Q50T, Q50G, Q50T-Y51N, Q50L-Y51N 或 Q50T-K91A。

在上述第三方面，如上所述从同源序列中确定一个共有序列；第二步，
30 在所述共有序列及单个同源蛋白的序列中，测定蛋白活性中心，其包含参与形成活性中心的所有氨基酸残基。所述单个同源蛋白可具有优选的特性

如高比活性或酶活性的不同 pH 依赖性。第三步，将参与形成该同源蛋白活性中心的部分或全部氨基酸残基插入所述共有序列的骨架中。结果产生嵌合蛋白，其具有来自单个蛋白的活性中心以及共有序列的骨架。

5 蛋白的活性中心可通过例如任何的蛋白三维结构分析来检测，例如通过基于已知蛋白的已知 3D-结构的同源模拟。

本发明也提供了通过此方法可获得的或已获得的共有蛋白，特别是含有至少一个示于图 2-6, 10 或 21 的氨基酸序列的蛋白，或其变体或突变蛋白。此类变体的实例见图 7-9。

10 此类变体或突变蛋白可如欧洲专利申请 EP0897010 所述定义并制备，例如 Q50L, Q50T, Q50G, Q50L-Y51N, 或 Q50T-Y51N。这些突变如上述定义，或参见图 2。当参见图 2 时，不要求减去所述 26 个氨基酸的信号肽(例如，在图 2 中氨基酸序列第 50 位的“Q50L,”中，氨基酸 Q 被氨基酸 L 取代)。

含有本发明植酸酶的食品，饲料，或药物组合物是本发明的另一方面。

15 在此上下文及本发明的方法中，“此类指定蛋白质家族的至少 3 个，优选至少 4 个氨基酸序列”是指 3, 4, 5, 6 至 12, 20, 50 个，或甚至更多的序列可用于比对和比较以产生共有蛋白的氨基酸序列。“指定的蛋白家族的序列”是指被折叠成三维结构的序列，其中 α -螺旋， β -折叠和 β -转角在同样的位置，使此结构被本领域技术人员称作可高度重叠。此外，这些序

20 列定性了显示同类生物学活性的蛋白质，例如，特定酶类如植酸酶。这类蛋白的三维结构足以允许模拟此家族其它同源蛋白的结构。实施例 1 举例说明如何操作。本发明中“进化相似性”指根据结构相似性对氨基酸分类的方案，其允许一个氨基酸被另一个氨基酸取代且对整个结构影响最小，这可通过本领域公知的程序如“PRETTY”程序来进行。短语“此程序提供的相似程度...被设置到较小的严谨数”在本发明中指，可将确定本发明所用

25 程序中相似性程度的参数值设定为，使该程序能针对完整氨基酸序列的最多数位置确定一个共有氨基酸，例如在程序 PRETTY 中，可将 THRESHOLD 的值选为 2 或 3，PLURALITY 的值为 2。此外，“1 除以这些序列的数目所得的权重”在本发明中指，能限定与用于确定共有序列之其它序列具有较

30 高相似程度的一组序列的序列仅为所述确定提供一个系数，该系数等于 1 除以该组所有序列的数目。

如上所述, 如果所述程序无法选出共有氨基酸, 则将最常出现的氨基酸选为共有氨基酸; 如果后一种情况不可能, 本领域技术人员将从用于比较的所有序列中选出本领域已知其特性可改进目标蛋白热稳定性的氨基酸, 方式如 Janecek, S. (1993), 生物化学方法(Process Biochem.) 28, 435-445; 5 Fersht, A.R.和 Serrano, L.(1993), 结构生物学最新观点(Curr. Opin. Struct. Biol.) 3, 75-83; Alber, T.(1989), 生物化学年鉴 58, 765-798; Matthews, B. W. (1987), 生物化学 26, 6885-6888; 或 Matthews, B. W. (1991), 结构生物学最新观点 1, 17-21。

10 酶的稳定性与多种工业应用相关。因此, 已有大量成功或不太成功的尝试, 其通过合理或随机的方法用来改进酶的稳定性, 优选热稳定性。

在此我们提供了另一种改进蛋白热稳定性的方式。

本发明提供了一种制备共有蛋白的方法, 包括对所谓共有蛋白的几乎所有位置的氨基酸残基进行预测的方法, 以及从该序列合成能在原核或真核表达系统中表达的完整基因的方法。

15 本发明的 DNA 序列可从编码目的蛋白如植酸酶的基因组或 cDNA 序列开始构建。例如, 可通过体外诱变构建[见, 例如 Sambrook 等, 分子克隆, 冷泉港实验室出版社, 纽约]。广泛使用的“定点诱变”最初由 Hurchinson 和 Edgell[病毒学杂志 8, 181(1971)]所述, 其包括使携有所需核苷酸取代的合成寡核苷酸与单链 DNA 序列中待导入突变的靶区域退火[见 Smith, 遗传学 20 年鉴 19, 423 (1985), 改良方法见 Stanssen 等, 核酸研究 17, 4441-4454 (1989) 中文献 2-6。对指定 DNA 序列的诱变还可通过聚合酶链式反应(PCR)进行。DNA 作为起始材料可通过本领域公知的如 Sambrook 等(分子克隆)所述的方法从各个菌株中分离。

25 菌株信息参见如 EP 684313 或下述任何保藏单位。黑曲霉(*Aspergillus niger*)[ATCC 9142], 嗜热毁丝霉(*Myceliophthora thermophila*)[ATCC 48102], 嗜热踝节菌(*Talaromyces thermophilus*)[ATCC 20186]与烟曲霉(*Aspergillus fumigatus*)[ATCC 34625]已根据布达佩斯条约保藏在美国典型培养物保藏中心, 保藏号分别为: ATCC 74337, ATCC 74340, ATCC 74338 和 ATCC 74339。但应理解, 编码本发明共有蛋白的 DNA 可以如 EP 747483 或 EP 30 897985, 或实施例所述, 通过本领域公知的方法合成。

序列信息参见如 EP684313, 或序列数据库如 Genbank(Intelligenetics,

加利福尼亚, 美国), 欧洲生物信息协会(European Bioinformatics Institute) (HinstonHall, 剑桥, 英国), NBRF(Georgetown 大学, 医疗中心, 华盛顿特区, 美国)和 Vecbase(威斯康星大学, 生物技术中心, Madison, 威斯康星, 美国)。

5 本发明的方法可例如用于改进植酸酶的热稳定性。

一旦获得本发明的全长 DNA 序列, 可通过本领域公知以及 Sambrook 等(同上(s.a.))所述的方法将它们整合至载体, 从而在适当宿主系统中过表达所编码的多肽。但本领域技术人员公知, DNA 序列本身可用于转化本发明的适当宿主系统以获得所编码多肽的过表达。适当宿主系统有真菌, 如黑
10 曲霉[ATCC 9142]或无花果曲霉(*Aspergillus ficuum*)[NRRL 3135]等曲霉属或如 *Trichoderma reesei* 等木霉属; 或酵母, 如酿酒酵母等糖酵母属, 或巴斯德毕赤酵母(*Pichia pastoris*)等毕赤酵母属, 或多形汉逊氏酵母(DSM5215)等汉逊氏酵母多形体; 或植物, 如 Pen 等, 生物/技术 11, 811-814 (1994)所述。本领域技术人员公知此类微生物可来自保藏机构, 如美国典型培养物
15 保藏中心(ATCC), 真菌菌株保藏中心(CBS)或德意志微生物和细胞培养物保藏中心(DSM)或“Industrial Property” [(1991) 1, 29-40 页]所列的任何其它保藏机构。可使用的细菌有大肠杆菌; 枯草芽孢杆菌(*Bacillus subtilis*)等芽孢杆菌属; 或变青链霉菌(*Streptomyces lividans*)等链霉菌属(参见如, Anne 和 Mallaert, FEMS 微生物通讯 114, 121 (1993)。优选大肠杆菌菌株, 如
20 大肠杆菌 K12 菌株, 如 M15[Villarejo, 细菌学杂志 120, 466-474 (1974)所述 DZ 291], HB 101[ATCC 33694]或大肠杆菌 SG13009[Gottesman 等, 细菌学 148, 265-273 (1981)]。

可用于真菌中表达的载体是本领域公知的且述于 EP 420358, 或 Cullen 等[生物/技术 5, 369-376(1987)], Ward[分子工业真菌学, 丝状真菌的系统
25 和应用, Marcel Dekker, 纽约(1991)], Upshall 等[生物/技术 5, 1301-1304(1987)], Gwynne 等[生物/技术 5, 71-79 (1987)], 或 Punt 等[生物技术杂志 17, 19-34(1991)]; 用于酵母中表达的载体见 Sreekrishna 等[基础微生物学杂志 28, 265-278(1988), 生物化学 28, 4117-4125(1989)], Hitzemann 等[自然 293, 717-722 (1981)]或 EP 183070, EP 183071, EP 248227, 或 EP
30 263311。用于大肠杆菌中表达的合适载体见 Sambrook 等[同上], Fiers 等[第 8 次国际生物技术专题讨论会论文集, 法国微生物协会(Soc. Franc. de

Microbiol.), 巴黎(Durand 等编), 680-697 页(1988)], Bujard 等[酶学方法 155, 416-433(1987)], 或 Stuber 等[免疫学方法, Lefkovits 和 Pernis 编, Academic Press, Inc., 卷 IV, 121-152 (1990)]。用于芽孢杆菌中表达的载体是本领域公知的且述于如 EP 207459, EP 405370, 美国国家科学院院报, 81, 439 (1984) 5 或 Yansura 和 Henner, 酶学方法 185, 199-228 (1990)。用于多形汉逊氏酵母中表达的载体是本领域公知的且述于如 Gellissen 等, 生物技术 9, 291-295(1991)。

已携有调控元件如启动子的载体, 或本发明的 DNA 序列可通过改造而含有这种元件。可使用的合适启动子元件是本领域公知的, 例如 10 *Trichoderma reesei* 的 cbhl-启动子[Haarki 等, 生物技术 7, 596-600 (1989)]或 pkil-启动子[Schindler 等, 基因 130, 271-275(1993)]; 用于米曲霉的 amy-启动子[Christensen 等, Abstr. 19th Lunteren Lectures on Molecular Genetics F23 (1987), Christensen 等, 生物/技术 6, 1419-1422(1988), Tada 等, 分子普通遗传学 229, 301 (1991)]; 用于黑曲霉的 glaA-启动子[Cullen 等, 15 生物/技术 5, 369-376(1987), Gwynne 等, 生物/技术 5, 713-719(1987), Ward, 分子工业真菌学, 丝状真菌的系统和应用, Marcel Dekker, 纽约, 83-106 (1991)], alcA-启动子[Gwynne 等, 生物技术 5, 718-719(1987)], sucl-启动子[Boddy 等, 现代遗传学 24, 60-66 (1993)], aphA-启动子[MacRae 等, 基因 71, 339-348(1988), MacRae 等, 基因 132, 193-198(1993)], tpiA-启动 20 子[McKnight 等, 细胞 46, 143-147(1986), Upshall 等, 生物/技术 5, 1301-1304(1987)], gpdA-启动子[Punt 等, 基因 69, 49-57(1988), Punt 等, 生物技术杂志 17, 19-37(1991)]以及 pkiA-启动子[de Graaff 等, 现代遗传学 22, 21-27(1992)]。可用于酵母中表达的适当启动子是本领域已知的, 例如 pho5-启动子[Vogel 等, 分子细胞生物学, 2050-2057 (1989); Rudolf 和 25 Hinnen, 美国国家科学院院报(Proc. Nati. Acad. Sci.) 84, 1340-1344(1987)]或用于酿酒酵母中表达的 gap-启动子; 用于在巴斯德毕赤氏酵母中表达的 aoxl-启动子[Koutz 等, 酵母 5, 167-177(1989); Sreekrishna 等, 基础微生物学杂志 28, 265-278(1988)]; 或用于在多形汉逊氏酵母中表达的 FMD 启动子[Hollenberg 等, EPA 0299108]或 MOX-启动子[Ledeboer 等, 核酸研究 30 13, 3063-3082(1985)]。

因此含有本发明 DNA 序列的载体, 优选使所述 DNA 序列在细菌或真

菌或酵母中表达的载体，以及所述转化的细菌或真菌或酵母，均是本发明的一部分。

5 本发明还提供一种允许蛋白，尤其本发明植酸酶，高表达的系统，例如重组汉逊氏酵母菌株。为此，可根据用于表达的生物如本案中的酵母(见实施例 1)的密码子频率表来选择此类蛋白的密码子。随意地，信号序列的密码子可如实施例 1 中具体案例所述选择；这意味着密码子频率表是依据编码指定蛋白家族的氨基酸序列的 DNA 序列中所用的密码子来制备。然后从用于表达的宿主细胞的密码子频率表中选择用于信号序列之 DNA 序列的密码子，其中利用了在一个表中出现频率相当的密码子。

10 一旦这些 DNA 序列被适当培养基中的适当宿主细胞表达后，所编码的蛋白可通过本领域公知的蛋白纯化方法或就植酸酶而言 EP 420358 中的方法，在被分泌到培养基中的情况下从培养基中分离，或在蛋白存在于胞内的情况下从宿主细胞中分离。因此，本发明多肽的制备方法，所述方法产生的多肽，或由本发明 DNA 序列编码的多肽，都是本发明的一部分，其中所述方法包括将上述转化的细菌或宿主细胞在适当条件下培养，和从中回收所述多肽。

本发明的多肽一经获得，可通过本领域公知的任何试验，按其具有农业用途的特性进行鉴定。

20 通常，本发明的多肽可用于任何具体领域，例如植酸酶可用于将肌醇聚磷酸如植酸盐转化为肌醇和无机磷酸。

此外，本发明的多肽可在药用组合物或复合食品或饲料的制备方法中使用，其中将所述组合物的成分与至少一种本发明的多肽混合。因此，含有至少一种本发明多肽的复合食品或饲料或药用组合物也是本发明的一部分。本领域技术人员熟知其制备方法。这类药用组合物或复合食品或饲料可进一步含有通常用于此目的且本领域公知的添加剂或成分。

本发明进一步提供了减少动物肥料中植酸盐水平的方法，其中给动物喂食所述饲料组合物，其量可将饲料中所含植酸盐有效转化为低级肌醇磷酸和/或肌醇，和无机磷酸。

30 在上下文中，植酸酶是一种具有植酸酶活性的酶或多肽。植酸酶可为例如肌醇六磷酸的磷酸水解酶，例如(肌醇六磷酸 3-磷酸水解酶，EC 3.1.3.8)与(肌醇六磷酸 6-磷酸水解酶，EC 3.1.3.26)。

在一个实施方案中, 通过 SDS-PAGE 评估, 植酸酶被纯化, 即纯度至少 85%, 优选至少 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99%。将植酸酶分离。植酸酶活性可通过本领域公知的任何植酸酶试验来测定, 例如下述试验(见实施例 9)。试验温度可以是真实(actual)植酸酶的最适温度, 试验 pH 可以是真实植酸酶的最适 pH。

试验温度可在例如 20-90℃, 或 30-80℃, 或 35-75℃的范围内选择, 例如 37℃, 50℃, 60℃, 或 70℃。

试验 pH 可在例如 pH 2-9, 或 3-8, 或 3-6 的范围内选择, 例如试验的 pH 值可选择 3, 4, 5, 6, 或 7。

测定氨基酸序列的同源性(或多肽或氨基酸同源性)作为两个序列相同的程度。这可通过本领域已知的预测机程序适当测定, 例如 GCG 软件包所提供的 GAP 程序[威斯康星软件包程序手册, 遗传学预测机组, 575 Science Drive, Madison, 威斯康星 53711, 美国], 也参见 Needleman, S.B. 和 Wunsch, C.D., (1970), 分子生物学杂志, 48, 443-453]。在 9.1 版本中, 为了进行多肽序列对比, 将长度权重设为 0, 缺口权重设为 3.0。

两个 DNA(核酸)序列之间的同一性或同源性的程度可通过本领域已知的预测机程序测定, 例如 GCG 软件包所提供的 GAP 程序[威斯康星软件包程序手册, 遗传学预测机组, 575 Science Drive, Madison, 威斯康星 53711, 美国]来测定, 也参见 Needleman, S.B. 和 Wunsch, C.D., (1970), 分子生物学杂志, 48, 443-453]。在 9.1 版本中, 为了进行 DNA 序列对比, 使用 GAP 并使用下列设置: 缺口产生罚分为 50, 缺口延伸罚分为 3。

用于测定指定 DNA 或 RNA 序列是否与特定核苷酸和寡核苷酸探针杂交的适当实验条件包括, 将含有所述 DNA 或 RNA 片段、用于检测杂交的膜在 5xSSC(氯化钠/柠檬酸钠; (J.Sambrook, E.F. Fritsch, 和 T. Maniatis, 1989, 分子克隆, 实验室手册, 第二版, 冷泉港, 纽约)中预浸渍 10 min, 并使该膜在 5x SSC, 5x Denhardt's 溶液, 0.5 % SDS 和 100 μ g/ml 超声处理的变性鲑精 DNA (Sambrook 等 1989)中预杂交, 然后在含 10 ng/ml 随机引发的 32 P-dCTP-标记探针(比活性 $> 1 \times 10^9$ cpm/ μ g)的相同溶液中于约 45℃杂交 12 小时(Feinberg, A.P.和 Vogelstein, B.(1983)生物化学年鉴 132:6-13)。

然后将此膜在 2 x SSC, 0.5 % SDS 中洗涤两次, 每次 30 分钟, 温度为至少 55℃(低严谨度), 至少 60℃ (中度严谨), 至少 65℃ (中/高严谨度),

至少 70℃(高严谨度), 或至少 75℃(非常高严谨度)。

在这些条件下与所述寡核苷酸探针杂交的分子可通过 x-线胶片检测。

具有改变的热稳定性的植酸酶, 或热稳定植酸酶, 均属于本发明。“热稳定”植酸酶是 T_m (解链温度)——经差示扫描量热(DSC)对植酸酶蛋白质的测定——至少为 65℃的一种植酸酶。DSC 可使用恒定加热率, 例如 10℃/min。在另外的实施方案中, T_m 至少为 66, 67, 68, 69, 70, 71, 72, 73, 74 或 75℃。或, T_m 等于或低于 150℃, 或等于或低于 145, 140, 135, 130, 125, 120, 115 或 110℃。因此, T_m 间距的实例为: 65-150℃, 66-150℃, -(等)-75-150℃; 65-145℃, 66-145℃, -(等)-75-145℃; 65-140℃, -(等)-75-140℃; -(等)-65-110℃, 66-110℃, -(等)-75-110℃。

T_m 的具体范围如下: 65 - 110℃; 70 - 110℃; 70 - 100℃; 75 - 95℃, 或 80 - 90℃。

在以下实施例 9 和 10 中, 描述了经 DSC 对 T_m 的测定, 且显示了多种植酸酶的 T_m 。

还显示了最适 pH, 因为——作为另一种方式——热稳定的植酸酶可被定义为一种具有至少 60℃的最适温度的植酸酶。优选该最适温度在 pH 5.0 或 5.5 时的底物植酸盐或植酸上测定。实施例 9 描述了一例植酸酶试验, 其包括对单位的定义。

在另外的实施方案中, 最适温度为至少 61, 62, 63, 64, 65, 66, 67, 68, 69 或 70℃。在一个具体实施方案中, 最适温度等于或低于 140℃, 或等于或低于 135, 130, 125, 120, 115, 110, 105 或 100℃。因此, 最适温度的间距可以为: 60-140℃, 61-140℃, -(等)-70-140℃; 60-135℃, 61-135℃, -(等)-70-135℃; 60-130℃, -(等)-70-130℃; -(等)-60-100℃, 61-100℃, -(等)-70-100℃。

在更具体地描述本发明之前, 对所附图表的简单解释如下。

图 1: 共有植酸酶-1 序列的图样。字母表示氨基酸残基的单字母代号。以下序列用于比对: 土曲霉(*Aspergillus terreus*)9A-1 的 phyA[Mitchell, D. B., Vogel, K., Weimann, B. J., Pasamontes, L. & van Loon, A. P. G. M. (1997)。组氨酸磷酸酶的植酸酶亚家族: 来自土曲霉和嗜热毁丝霉的两个新植酸酶的基因的分离, 微生物学 143, 245-252]; 始于氨基酸(aa) 27; SEQ ID NO:1];

土曲霉 cbs116.46 的 phyA [EP 897985]。一种在动物实验中表现优良的烟曲霉耐热植酸酶。植酸酶的优化和天然变异性。在：植酸盐和植酸酶的生物化学(Rasmussen, S.K; Raboy, V.; Dalb~ge, H. 和 Loewus, F.编; Kluwer Academic Publishers); 始于 aa 27; SEQ ID NO:2; 黑曲霉泡盛变种的 phyA(Piddington 等(1993)基因 133, 55-62; 始于 aa27; SEQ ID NO:3); 黑曲霉 T213 的 phyA(EP 897985); 始于 aa27; SEQ ID NO:4); 黑曲霉 NRRL3135 株的 phyA[van Hartingsveldt, W., van Zeijl, C. M. F., Harteveld, G. M., Gouka, R. J., Suykerbuyk, M. E. G., Luiten, R. G. M., van Paridon, P. A., Selten, G. C. M., Veenstra, A. E., van Gorcom, R. F. M.和 van den Hondel, C. A. M. J. J.(1993), 黑曲霉植酸酶编码基因(phyA)的克隆、鉴定和过表达。基因 127, 87-94; 始于 aa27; SEQ ID NO:5]; 烟曲霉 ATCC 13073 的 phyA (Pasamontes, L., Haiker, M., Wyss, M., Tessier, M.和 van Loon, A. P. G. M. (1997)烟曲霉耐热植酸酶的克隆、纯化和鉴定, 应用环境微生物学 63, 1696-1700; 始于 aa25; SEQ ID NO:6); 烟曲霉 ATCC 32722 的 phyA (EP 897985); 始于 aa27; SEQ ID NO:7); 烟曲霉 ATCC 58128 的 phyA (EP 897985); 始于 aa27; SEQ ID NO:8); 烟曲霉 ATCC 26906 的 phyA (EP 897985); 始于 aa27; SEQ ID NO:9); 烟曲霉 ATCC 32239 的 phyA (EP 897985); 始于 aa 30; SEQ ID NO:10; *Emericella nidulans* 的 phyA [Pasamontes, L., Haiker, M., HenriquezHuecas, M., Mitchell, D. B.和 van Loon, A. P. G. M. (1997a). *Emericella nidulans* 和嗜热真菌嗜热踝节菌的植酸酶的克隆, 生物化学和生物物理学学报 1353, 217-223; 始于 aa 25; SEQ ID NO:11]; 嗜热踝节菌的 phyA (Pasamontes 等, 1997a; 始于 aa 24; SEQ ID NO:12); 以及嗜热毁丝霉的 phyA (Mitchell 等, 1997; 始于 aa 19; SEQ ID NO:13)。比对利用程序 PILEUP 预测。缺口的位置经手工修正(refine)。比
 25 对中指定位置上的大写氨基酸残基属于确立共有残基的氨基酸的集合。最终构建的共有植酸酶(Fcp)的氨基酸序列(SEQ ID NO:14)以粗体显示在预测的共有序列的下方。所预测的共有序列中的缺口根据实施例 1 所述原则手工补齐。

图 2：共有植酸酶-1 基因(fcp)的 DNA 序列 (SEQ ID NO:15)和用于基
 30 因构建的引物序列。预测的氨基酸序列(图 1, SEQ ID NO:14)用程序 BACKTRANSLATE [Devereux, J., Haeberli, P. & Smithies, O. (1984), VAX

序列分析程序组，核酸研究，12，387-395]转化为 DNA 序列，还有高表达酵母基因的密码子频率表(GCG 程序包，9.0)。土曲霉 cbs 116.46 的植酸酶的信号肽被融合到 N-末端。图 2 所示氨基酸序列为 SEQ ID NO:16。粗体碱基表示用于产生基因的寡核苷酸序列。各个寡核苷酸的名字被交替注释在序列的上方或下方。加下划线的碱基代表基因的起始和终止密码子。斜体字表示的碱基代表被导入的两个 EcoRI 位点。

图 3：五种担子菌植酸酶的比对和共有序列。字母表示氨基酸残基的单字母代号。来自 *Paxillus involutus* 的植酸酶的氨基酸序列，phyA1 (始于 aa 21; SEQ ID NO:17); 以及 phyA2 (始于 aa 21, WO 98/28409; SEQ ID NO:18); 绒毛栓菌(始于 aa 24, WO 98/28409; SEQ ID NO:19); 平田头菇(始于 aa 19, WO98/28409; SEQ ID NO:20); 和 *Peniophora lycii*(始于 aa 21, WO 98/28409; SEQ ID NO:21)，从圆括号中的氨基酸残基开始，用于称作“Basidio”的相应共有序列(实施例 2; SEQ ID NO:22)的比对和预测。比对用程序 PILEPUP 来进行，缺口的位置经手工修正。共有序列用程序 PRETTY 预测。当两种 *P. involutus* 植酸酶的权重为 0.5 时，所有其它基因采用 1.0 的权重以便共有序列的预测。在所述程序不能测定共有残基之处，Basidio 序列含有一个破折号(dash)。比对中指定位置的大写氨基酸残基代表能确定共有残基的氨基酸集和。

图 4：共有植酸酶-10 氨基酸序列的图样。将 *Thermomyces lanuginosus* 植酸酶的序列[Berka, R. M., Rey, M. W., Brown, K. M., Byun, T.和 Klotz, A. V. (1998), 嗜热真菌 *Thermomyces lanuginosus* 的一个植酸酶基因的分子鉴定和表达，应用环境微生物学 64, 4423-4427; SEQ ID NO:23]和五种担子菌植酸酶的共有序列(SEQ ID NO:22)添加至图 1 的比对中，由程序 PRETTY 预测出一个改进的共有序列。此外，略去黑曲霉 T213 的氨基酸序列，给剩余两种黑曲霉植酸酶序列分配权重 0.5。详见实施例 2。

图 5：共有植酸酶-10 的 DNA 和氨基酸序列(分别为 SEQ ID NO:25, 和 SEQ ID NO:26)。氨基酸序列用单字母表示在相应 DNA 序列上方。用于装配基因的寡核苷酸的序列用粗体字母表示。不同于共有植酸酶-1 的寡核苷酸和氨基酸带下划线。fcpl0 基因用以下寡核苷酸装配：CP-1, CP-2, CP-3.10, CP-4.10, CP-5.10, CP-6, CP-7.10, CP-8.10, CP-9.10, CP-10.10, CP-11.10, CP-12.10, ap13.10, CP-14.10, CP-15.10, CP-16.10, CP-17.10, CP18.10,

CP19.10, CP-20.10, CP-21.10, 以及 CP-22.10。新合成的寡核苷酸进一步用数字 10 标记。植酸酶相对于共有植酸酶-1 含有以下 32 个取代: Y54F, E58A, D69K, D70G, A94K, N134Q, I158V, S187A, Q188N, D197N, S204A, T214L, D220E, L234V, A238P, D246H, T251N, Y259N, E267D,
 5 E277Q, A283D, R291I, A320V, R329H, S364T, I366V, A379K, S396A, G404A, Q415E, A437G, A463E。带下划线的突变显示了当作为共有植酸酶-1 中的单突变被测试时对共有植酸酶-1 的稳定效应。

图 6: 共有植酸酶-11(SEQ ID NO:27)的比对。与共有植酸酶-10 对照, 设计共有植酸酶-11 的氨基酸序列时, 所有担子菌植酸酶均作为独立序列使用, 每一序列分配一个 0.2 的权重。此外, 黑曲霉 T213 的氨基酸序列再次用于该比对中。

图 7: 共有植酸酶-1-thermo[8]-Q50T-K91A 的 DNA 和氨基酸序列(分别为 SEQ ID NO:28 和 SEQ ID NO:29)。氨基酸序列用单字母代码表示在 DNA 序列之上。被取代的氨基酸残基(相对于共有植酸酶-1)带下划线。基
 15 因的终止密码子用星号(*)标记。

图 8: 共有植酸酶-10-thermo[3]-Q50T-K91A 的 DNA 和氨基酸序列(分别为 SEQ ID NO:30 和 SEQ ID NO:31)。氨基酸序列用单字母代码表示在 DNA 序列之上。被取代的氨基酸残基(相对于共有植酸酶-10)带下划线。基
 20 因的终止密码子用星号(*)标记。

图 9: 烟曲霉 ATCC 13073 植酸酶 α -突变体 Q51T 的 DNA 和氨基酸序列(分别为 SEQ ID NO:32 和 SEQ ID NO:33)。氨基酸序列用单字母代码表示在 DNA 序列之上。被取代的氨基酸残基(相对于烟曲霉 ATCC13073 的植酸酶)带下划线。基因的终止密码子用星号(*)标记。

图 10: 共有植酸酶-7 的 DNA 和氨基酸序列(分别为 SEQ ID NO:34 和
 25 SEQ ID NO:35)。氨基酸序列用单字母代码表示在 DNA 序列之上。用于装配基因的寡核苷酸的序列用粗体字母表示。被取代的寡核苷酸和氨基酸(相对于共有植酸酶-1)带下划线, 相应的三联密码用小写字母表示。fcp7 基因由下列寡核苷酸装配: CP-1, CP-2, CP-3, CP-4.7, CP5.7, CP-6, CP-7, CP-8.7, CP-9, CP-10.7, CP-11.7, CP-12.7, CP13.7, CP-14.7, CP-15.7,
 30 CP-16, CP-17.7, CP-18.7, CP-19.7, CP-20, CP-21, 和 CP-22。新合成的寡核苷酸进一步用数字 7 标记。与原始的共有植酸酶-1 相比, 共有植酸酶-

7 含有以下 24 个取代：S89D, S92G, A94K, D164S, P201S, G203A, G205S, H212P, G224A, D226T, E255T, D256E, V258T, P265S, Q292H, G300K, Y305H, A314T, S364G, M365I, A397S, S398A, G404A, 和 A405S。

5 图 11：共有植酸酶-1 和共有植酸酶-10 的差示扫描量热(DSC)。将蛋白样品浓缩至约 50-60 mg/ml, 对 10 mM 乙酸钠, pH5.0 强力透析。以 10℃/min 的恒定加热速率加热到 95℃。共有植酸酶-10 (上图)的 DSC 产生 85.4℃的解链温度, 其比共有植酸酶-1(78.1℃, 下图)的解链温度高 7.3℃。

10 图 12：共有植酸酶-10-thermo[3]-Q50T 和共有植酸酶-10-thermo[3]-Q50T-K91A 的差示扫描量热(DSC)。将蛋白样品浓缩至约 50-60 mg/ml, 对 10 mM 乙酸钠, pH5.0 强力透析。以 10℃/min 的恒定加热速率加热到 95℃。共有植酸酶-10-thermo[3]-Q50T(上图)的 DSC 产生 88.6℃的解链温度, 而共有植酸酶-10-thermo[3]-Q50T-K91A 的解链温度被测定为 89.3℃。

15 图 13：共有植酸酶-1、共有植酸酶-10 和共有植酸酶-10-thermo[3]-Q50T 的最适温度比较。为了测定最适温度, 在 37-86℃之间的一系列温度下进行实施例 9 的植酸酶标准试验。用酿酒酵母的稀释上清进行检测。上清的其它组份对最适温度的测定没有影响：△, 共有植酸酶-1; ◇, 共有植酸酶-10; ■, 共有植酸酶-10-thermo[3]-Q50T。

20 图 14：共有植酸酶-10 及其突变体 thermo[3]-Q50T 和 thermo[3]-Q50T-K91A 的依赖 pH 的活性图谱和底物特异性。植酸酶活性通过在适当缓冲液中不同 pH 的标准试验测定(见实施例 9)。图 a)显示了共有植酸酶-10(□), 共有植酸酶-10-thermo[3]-Q50T(●)和共有植酸酶-10-thermo[3]-Q50T-K91A(△)的依赖 pH 的活性图谱, 图 b)显示了通过在标准试验中用所述化合物置换植酸测得的相应底物特异性; 空白栏, 共有植酸酶-10; 灰色栏, 共有植酸酶-10-thermo[3]-Q50T; 黑色栏, 共有植酸酶-10-thermo[3]-Q50T-K91A。数字对应下列底物：1.植酸；2.对硝基苯磷酸；3.苯基磷酸；4.果糖-1,6-二磷酸；5.果糖-6-磷酸；6.葡萄糖-6-磷酸；7.核糖-5-磷酸；8.DL-甘油-3-磷酸；9.甘油-2-磷酸；10.3-磷酸甘油酸；11.磷酸烯醇丙酮酸；12.AMP；13.ADP；14.ATP。

30 图 15：共有植酸酶-1-thermo[8]-Q50T 和共有植酸酶-1-thermo[8]-Q50T-K91A 的依赖 pH 的活性图谱和底物特异性。植酸酶活性通过在适当缓冲液中不同 pH 的标准试验测定(见实施例 9)。图 a)显示了 Q50T 变体(■)和

Q50T-K91A 变体(●)的依赖 pH 的活性图谱, 图 b)显示了通过在标准试验中用所述化合物置换植酸测得的相应底物特异性(空白栏, 共有植酸酶-1-thermo[8]-Q50T; 实心栏, 共有植酸酶-1-thermo[8]-Q50T-K91A)。底物列在图 14 的图说明中。

5 图 16: 共有植酸酶-1-thermo[8]-Q50T 和共有植酸酶-1-thermo[8]-Q50T-K91A 的差示扫描量热(DSC)。将蛋白样品浓缩至约 50-60 mg/ml, 对 10 mM 乙酸钠, pH5.0 强力透析。以 10℃/min 的恒定加热速率加热到 95℃。共有植酸酶-1-thermo[8]-Q50T(上图)的 DSC 产生 84.7℃的解链温度, 而共有植酸酶-1-thermo[8]-Q50T-K91A 的解链温度被测定为 85.7℃。

10 图 17: 共有植酸酶-1, 共有植酸酶-1 和共有植酸酶-1-thermo-[3]和共有植酸酶-1-thermo[8]的最适温度的比较。为了测定最适温度, 在 37-86℃之间的一系列温度下进行植酸酶标准试验。用来自转化的酿酒酵母菌株的纯化蛋白进行检测。○, 共有植酸酶-1; □, 共有植酸酶-1-thermo[3]; ▲, 共有植酸酶 1-thermo[8]。

15 图 18: 共有植酸酶-1, 共有植酸酶-7, 和黑曲霉 NRRL 3135 植酸酶的依赖 pH 的活性和底物特异性的比较。植酸酶活性通过在适当缓冲液中不同 pH 的标准试验测定(见实施例 9)。图 a)显示了植酸酶-1(■), 黑曲霉 NRRL 3135 植酸酶(○), 和共有植酸酶-7(▲)的依赖 pH 的活性图谱, 图表 b)显示了通过在标准试验中用所述化合物置换植酸测得的相应底物特异性(黑色栏, 黑曲霉 NRRL 3135 植酸酶; 空白栏, 共有植酸酶-1; 灰色栏, 共有植酸酶-7)。底物列在图 14 的图说明中。

图 19: 烟曲霉 ATCC 13073 的植酸酶及其含有氨基酸置换(F55Y, V100I, F114Y, A243L, S265P, 和 N294D)的稳定型 α -突变体的差示扫描量热(DSC)。

25 将蛋白样品浓缩至约 50-60 mg/ml, 对 10 mM 乙酸钠, pH5.0 强力透析。以 10℃/min 的恒定加热速率加热到 95℃。烟曲霉 ATCC 13073 植酸酶(上图)的 DSC 显示为 62.5℃的解链温度, 而所述 α -突变体的解链温度被测定为 67.0℃。

30 图 20: 烟曲霉 13073 的野生型植酸酶, 其 α -突变体, 以及其进一步稳定的 α -突变体(E59A-S154N-R329H-S364T-G404A)的最适温度比较。为了测定最适温度, 在 37-75℃之间的一系列温度下进行植酸酶标准试验。

用转化的酿酒酵母菌株的稀释上清进行检测。上清的其它组份对最适温度的测得没有影响。○, 烟曲霉 ATCC 13073 植酸酶; ▲, 烟曲霉 ATCC 13073 α -突变体; □, 烟曲霉 ATCC 13073 植酸酶 α -突变体-(E59A-S154N-R329H-S364T-G404A); ■, 烟曲霉 ATCC 13073 α -突变体(E59A-S154N-R329H-S364T-G404A)-Q51T-K92A。Q51T 和 K92A 分别对应于共有植酸酶-1 的取代 Q50T 和 K91A。

图 21: 共有植酸酶-12 的氨基酸序列(consphyl2; SEQ ID NO:36), 其含有从"basidio"共有序列转移到共有植酸酶-10-thermo[3]-Q50T-K91A 的多个活性位点残基(加下划线)。

10 图 22: 共有植酸酶-3-thermo[11]-Q50T 的 DNA 和氨基酸序列。氨基酸用单字母代码表示在相应 DNA 序列之下。

图 23: 共有植酸酶-3-thermo[11]-Q50T-K91A 的 DNA 和氨基酸序列。氨基酸用单字母代码表示在相应 DNA 序列之下。

15 图 24: 共有植酸酶-10-thermo[5]-Q50T 的 DNA 和氨基酸序列。氨基酸用单字母代码表示在相应 DNA 序列之下。

图 25: 共有植酸酶-10-thermo[5]-Q50T-K91A 的 DNA 和氨基酸序列。氨基酸用单字母代码表示在相应 DNA 序列之下。

本文所述的生产植酸酶的微生物菌株, 即 *Paxillus involutus* CBS 100231; *Peniophora lycii* CBS 686.96; 平田头菇 CBS 900.96; 以及绒毛栓菌 CBS100232 分别分离于来自丹麦; 丹麦; 丹麦; 和瑞典的天然样品(Uppsala 保藏中心。分别于 1992 年 11 月; 1993 年 10 月; 1995 年 6 月; 和 1995 年 11 月保藏。

实施例 1 共有植酸酶-1

25 共有植酸酶-1(真菌共有植酸酶, fcp)的氨基酸序列如 EP 897985 中实施例 1 和 2 所述设计并预测。下表 1 显示了用于设计共有植酸酶-1 的植酸酶氨基酸序列的起点和权重。进一步将共有植酸酶-1 序列如 EP 897985 的实施例 3 所述转化为 DNA 序列, 并如 EP897985 的实施例 4 所述构建和克隆共有植酸酶-1 基因。

30 表 1 植酸酶氨基酸序列的起点和权重

土曲霉 9A-1 的 phyA, aa 27, 权重 0.5(Mitchell 等, 1997)

- 土曲霉 cbs116.46 phyA, aa 27, 权重 0.5(EP 897985)
- 黑曲霉泡盛变种 phyA, aa 27, 权重 0.33[Piddington, C.S., Houston, C.S., Paloheimo M., Cantrell, M., Miettinen-Oinonen, A., Nevalainen, H., & Rambousek, J. (1993), 编码黑曲霉泡盛变种的植酸酶(phy)和 pH2.5 最适
- 5 型酸性磷酸酶(aph)的基因的克隆和测序, 基因 133, 55-62)
- 黑曲霉 T213 phyA (EP 897985), aa 27, 权重 0.33
- 黑曲霉 NRRL3135 株 phyA, aa 27, 权重 0.33 (van Hartingsveldt 等, 1993)
- 烟曲霉 ATCC 13073 phyA, aa 26, 权重 0.2 (Pasamontes 等, 1997)
- 烟曲霉 ATCC 32722 phyA, aa 26, 权重 0.2 (EP 897985)
- 10 烟曲霉 ATCC 58128 phyA, aa 26, 权重 0.2 (EP 897985)
- 烟曲霉 ATCC 26906 phyA, aa 26, 权重 0.2 (EP 897985)
- 烟曲霉 ATCC 32239 phyA, aa 30, 权重 0.2 (EP 897985)
- Emericella nidulans* phyA, aa 25, 权重 1.0(Pasamontes 等, 1997a)
- 嗜热踝节菌 ATCC 20186 phyA, aa 24, 权重 1.0 (Pasamontes 等, 1997a)
- 15 嗜热毁丝霉 phyA, aa 19, 权重 1.0(Mitchell 等, 1997)

实施例 2 一种改进型共有植酸酶(共有植酸酶-10)氨基酸序列的设计

- 用于设计共有植酸酶-10 的比对通过使用 GCG 序列分析软件包 9.0 版本(Devereux 等, 1984)的程序 PILEUP 以标准参数(缺口产生罚分 12, 缺口
- 20 延伸罚分 4)进行预测。缺口的定位使用文本编辑器(text editor)修正。
- 用以下序列对始于表 2 所述氨基酸(aa)的担子菌植酸酶进行比对:

表 2

5 种担子菌植酸酶用于预测相应共有氨基酸序列(basidio)的起点和权重

- Paxillus involutus* CBS 100231 phyA1, aa 21, 权重 0.5 (WO 98/28409)
- 25 *Paxillus involutus* CBS 100231 phyA2, aa 21, 权重 0.5 (WO 98/28409)
- 绒毛栓菌 CBS 100232 phyA, aa 24, 权重 1.0 (WO 98/28409)
- 平田头菇 CBS 900.96 phyA, aa 19, 权重 1.0 (WO 98/28409)
- Peniophora lycii* CBS 686.96 phyA, aa 21, 权重 1.0 (WO 98/28409)
- 该比对显示在图 3 中。

- 30 表 3 中列出了用于最终比对的基因。该比对中所用序列的第一氨基酸(aa)述于生物命名之后。

表3 用于设计共有植酸酶-10的植酸酶序列的起点和权重

	土曲霉 9A-1 phyA, aa 27, 权重 0.5 (Mitchell 等, 1997)
	土曲霉 cbs116.46 phyA, aa 27, 权重 0.5 (EP 897985)
	黑曲霉泡盛变种 phyA, aa27, 权重 0.5 (Piddington 等, 1993)
5	黑曲霉 NRRL3135 株 phyA, aa 27, 权重 0.5 (van Hartingsveldt 等, 1993)
	烟曲霉 ATCC 13073 phyA, aa 26, 权重 0.2 (Pasamontes 等, 1997)
	烟曲霉 ATCC 32722 phyA, aa 26, 权重 0.2 (EP 897985)
	烟曲霉 ATCC 58128 phyA, aa 26, 权重 0.2 (EP 897985)
	烟曲霉 ATCC 26906 phyA, aa 26, 权重 0.2 (EP 897985)
10	烟曲霉 ATCC 32239 phyA, aa 30, 权重 0.2 (EP 897985)
	<i>Emericella nidulans</i> phyA, aa 25, 权重 1.0 (Pasamontes 等, 1997a)
	嗜热踝节菌 ATCC 20186 phyA, aa 24, 权重 1.0 (Pasamontes 等, 1997a)
	嗜热毁丝霉 phyA, aa 19, 权重 1.0 (Mitchell 等, 1997)
	<i>Thermomyces lanuginosus</i> phyA, aa 36, 权重 1.0 (Berka 等, 1998)
15	五种担子菌植酸酶的共有序列, 权重 1.0 (Basidio, 图 3)
	相应比对显示于图 4。

共有植酸酶-10的氨基酸序列的预测

为了改进比对, 我们增加了来自 4 种不同担子菌之 5 种植酸酶的最初共有序列(称为 Basidio; 仍含有未确定的序列位置; 见图 3), 用于该最初共有植酸酶序列之预测的几乎所有植酸酶序列和一种新的来自子囊菌纲 *Thermomyces lanuginosus* 的植酸酶序列, 以便进行更大的比对。

我们设置复数性(plurality)为 2.0 和阈值为 3。所用权重列于表 3。比对及相应共有序列见表 4。新的共有植酸酶序列和最初的共有植酸酶-1 相比具有 32 个不同的氨基酸。程序 PRETTY 不能预测共有氨基酸残基之处按实施例 1 所述规则补齐。该程序所提示的残基无一被取代。

此外, 在另一预测中, 我们将所有担子菌植酸酶分别作为个氨基酸序列配以权重 0.2 来进行预测。相应的比对示于图 6。预测得到的共有氨基酸序列(共有植酸酶-11)与共有植酸酶-10 的序列有以下差异。字母 X 指该程序不能预测共有氨基酸; 圆括号中的氨基酸对应于最终包括在共有植酸酶-10 内的氨基酸。

D35X(第一个字母是共有植酸酶-10 的, 最后一个字母是共有植酸酶-11

的), X(K)69K, X(E)100E, A101R, Q134N, X(K)153N, X(H)190H, X(A)204S, X(E)220D, E222T, V227A, X(R)271R, H287A, X(D)288D, X(K)379K, X(I)389I, E390X, X(E)415E, X(A)416A, X(R)446L, E463A.编号如图 5。

我们也检查了改进型共有序列 10 和 11 所示单个氨基酸取代对最初共有植酸酶之稳定性的影响。该方法见实施例 3。

共有植酸酶-10 的氨基酸序列到 DNA 序列的转换

用土曲霉 cbs116.46 植酸酶的最初 26 个氨基酸残基作为信号肽, 将其融合至共有植酸酶-10 的 N-末端。所用方法进一步描述在实施例 1 中。

所得 fcp10 基因的序列见图 5。

10 共有植酸酶-10(fcp10)基因的构建和克隆

任选用正义链和反义链将预测的 fcp10DNA 序列分为 85bp 的寡核苷酸。每一寡核苷酸与相对链的前、后寡核苷酸重叠 20 bp。所有购自 Microsynth, Balgach(瑞士)的 PAGE-纯化型引物参见图 5。

PCR-反应

15 在三个 PCR 反应中, 将合成寡核苷酸被编入完整基因中。用来自 Boehringer Mannheim (Boehringer Mannheim, Mannheim, 德国)的 High Fidelity 试剂盒和来自 AMS 生物技术(欧洲)有限公司(Lugano, 瑞士)的热循环仪 “The Protokoll™” 实施 PCR。所用以下寡核苷酸浓度为 0.2 pMol/ml。

Mix 1.10: CP-1, CP-2, CP-3.10, CP-4.10, CP-5.10, CP-6,
20 CP-7.10, CP-8.10, CP-9.10, CP-10.10
Mix 2.10: CP-9.10, Cp-11.10, CP-12.10, CP-13.10, Cp-14.10,
CP-15.10, CP-16.10, CP-17.10, CP18.10, CP-19.10,
CP-20.10, CP-21.10, CP-22.10

新合成的寡核苷酸用数字 10 标记。共有植酸酶-10 与最初的共有植酸酶-1 相比含有下面的 32 个取代, 其在图 5 中被加下划线: Y54F, E58A, D69K, D70G, A94K, N134Q, I158V, S187A, Q188N, D197N, S204A, T214L, D220E, L234V, A238P, D246H, T251N, Y259N, E267D, E277Q, A283D, R291I, A320V, R329H, S364T, I366V, A379K, S396A, G404A, Q415E, A437G, A463E。

30 4 个短 PCR 引物用于寡核苷酸的合成:

CP-a: Eco RI

5'-TATATGAATTTCATGGGCGTGTTTCGTC-3' SEQ ID NO:37)

CP-b:

5'-TGAAAAGTTCATTGAAGGTTTC-3' (SEQ ID NO:38)

CP-c.10:

5'-TCTTCGAAAGCAGTACACAAAC-3' (SEQ ID NO:39)

CP-e: Eco RI

5'-TATATGAATTCTTAAGCGAAAC-3' (SEQ ID NO:40)

PCR 反应 a: 10 μ l Mix 1.10 (每一寡核苷酸 2.0pmol)

2 μ l 核苷酸 (每一寡核苷酸 10 mM)

2 μ l 引物 CP-a (10 pmol/ml)

2 μ l 引物 CP-c.10 (10 pmol/ml)

10.0 μ l PCR 缓冲液

0.75 μ l 聚合酶混合物 (2.6 U)

73.25 μ l H₂O

PCR 反应 b: 10 μ l Mix 2.10 (每一寡核苷酸 2.0 pmol)

2 μ l 核苷酸(每一核苷酸 10 mM)

2 μ l 引物 CP-b (10 pmol/ml)

2 μ l 引物 CP-e (10 pmol/ml)

10.0 μ l PCR 缓冲液

0.75 μ l 聚合酶混合物(2.6 U)

73.25 μ l H₂O

PCR 反应 a 和 b 的反应条件:

步骤 1 2 min-45°C

步骤 2 30 sec-72°C

步骤 3 30 sec-94°C

步骤 4 30 sec-52°C

步骤 5 1 min-72°C

将步骤 3 到 5 重复 40 次。

PCR 产物(670 和 905 bp)经琼脂糖凝胶电泳(0.9% 琼脂糖), 然后经凝胶
提取(QIAEX II 凝胶提取试剂盒, Qiagen, Hilden, 德国)纯化。用纯化的 DNA
片段进行 PCR 反应 c。

PCR 反应 c: 6 μ l PCR 反应 a 的产物 $\approx$ 50 ng)
6 μ l PCR 反应 b 的产物 $\approx$ 50 ng)
2 μ l 引物 CP-a (10 pmol/ml)
2 μ l 引物 CP-e (10 pmol/ml)
5 10.0 μ l PCR 缓冲液
0.75 μ l 聚合酶混合物 (2.6 U)
73.25 μ l H₂O

PCR 反应 c 的反应条件:

步骤 1 2 min-94 $^{\circ}$ C
10 步骤 2 30 sec-94 $^{\circ}$ C
步骤 3 30 sec-55 $^{\circ}$ C
步骤 4 1 min-55 $^{\circ}$ C

将步骤 2 到 4 重复 31 次。

所得 PCR 产物(1.4kb)如上述纯化, 用 EcoRI 降解, 并连接至 EcoRI-消
15 化并去磷酸化的 pBsk(-)-载体(Stratagene, La Jolla, CA, 美国)。用 1 μ l 连
接混合物转化大肠杆菌 XL-1 感受态细胞(Stratagene, La Jolla, CA, 美国)。
所有标准方法如 Sambrook 等(1987)所述。所构建的基因(fcp10)的 DNA 序
列通过本领域公知的测序方法检查。

20 实施例 3 通过导入共有植酸酶-10 和共有植酸酶-11 的氨基酸序列所 显示的单个突变增加共有植酸酶-1 的热稳定性

为了增加同源基因的热稳定性, 也可检验目的蛋白与预测的共有序列
之间每一差异氨基酸残基的稳定性效应, 并将所有稳定突变组合至目的蛋
白中。我们利用共有植酸酶-1 作为目的蛋白, 测试以单位点定点突变区别
25 于共有植酸酶-10 和/或-11 的 34 个氨基酸残基的蛋白稳定性效应。

为构建在黑曲霉, 酿酒酵母, 或多形汉逊氏酵母中表达的突变蛋白,
用含有共有植酸酶-1 基因的相应表达质粒作为模板进行定点诱变(见实施例
6-8)。突变用 Stratagene (La Jolla, CA, 美国)的"快速交换(quick exchange)TM
定点诱变试剂盒"按说明书指示并使用相应引物来导入。所有获得的突变及
30 相应引物概述于表 4。携有所需突变的质粒用本领域公知的 DNA 序列分析
来鉴定。

表4 用于共有植酸酶-1 定点诱变的引物

所取代的碱基用粗体来突出。导入的限制性位点在序列上方标出。当限制性位点被写在圆括号中时，所述位点通过导入突变而破坏。

5	突变	引物系列
		<i>Kpn</i> I
Q50T		5'-CACTTGTGGGGT ACCT ACTCTCCATACTTCTC-3' (SEQ ID NO: 41) 5'-GAGAAGTATGGAGAGTAGGT ACCC CACAAGTG-3'
Y54F		5'-GGTCAATACTCTCCATTCTTCTCTTTGGAAG-3' (SEQ ID NO: 42) 5'-CTTCCAAAGAGAAGAATGGAGAGTATTGACC-3'
E58A		5'-CATACTTCTCTTTGGCAGACGAATCTGC-3' (SEQ ID NO: 43) 5'-GCAGATTCGTCTG CCAA GAGAAGTATG-3'
		<i>Aat</i> II
D69K		5'-CTCCAGACGT CCCA AGGACTGTAGAGTTAC-3' (SEQ ID NO: 44) 5'-GTA ACT CTACAGTCCTTTGGGACGTCTGGAG-3'
		<i>Aat</i> II
D70G		5'-CTCCAGACGT CCCA GACGGCTGTAGAGTTAC-3' (SEQ ID NO: 45) 5'-GTA ACT CTACAGCCGTCTGGGACGTCTGGAG-3'
K91A		5'-GATAC CCA ACTTCTTCTGCGTCTAAGGCTTACTCTG-3' (SEQ ID NO: 46) 5'-CAGAGTAAGCCTTAGACGCAGAGAAGTTGGGTATC-3'
		<i>Sca</i> I
A94K		5'-CTTCTAAGTCTAAG AGT ACTCTGCTTTG-3' (SEQ ID NO: 47) 5'-CAAAGCAGAGT ACTT CTTAGACTTAGAAG-3'
A101R		5'-GCTTACTCTGCTTTGATTGAACGGATTCAA AGA ACGCTAC-3' (SEQ ID NO: 48) 5'-GTAGCGTTCTTTTGAATCCGTTCAATCAAAGCAGAGTAAGC-3'
N134Q		5'-CCATTCGGTGAACAGCAAATGGTTAACTC-3' (SEQ ID NO: 49) 5'-GAGTTAACCATTGCTGTT CA CCGAATGG-3'

<i>Nru I</i>	
K153N	5'-GATACAAGGCTCTCGCGAGAAACATTGTTC-3' (SEQ ID NO: 50) 5'-GGAACAATGTTTCTCGCGAGAGCCTTGTATC-3'
<i>Bss HI</i>	
I158V	5'-GATTGTTCCATTTCGTGCGCGCTTCTGGTTC-3' (SEQ ID NO: 51) 5'-GAACCAGAAGCGCGCACGAATGGAACAATC-3'
<i>Apa I</i>	
S187A	5'-GGCTGACCCAGGGGCCCAACCACACCAAGC-3' (SEQ ID NO: 53) 5'-GCTTGGTGTGGTTGGGCCCTGGGTCAGCC-3'
<i>Bcl I</i>	
D197N	5'-CTCCAGTTATTAACGTGATCATTCCAGAAGG-3' (SEQ ID NO: 52) 5'-CCTTCTGGAATGATCACGTTAATAACTGGAG-3'
<i>Nco I</i>	
T214L	5'-CACTTTGGACCATGGTCTTTGTACTGCTTTTCG-3' (SEQ ID NO: 54) 5'-CGAAAGCAGTACAAAGACCATGGTCCAAAGTG-3'
<i>Avr II</i>	
E222T	5'-GCTTTCGAAGACTCTACCCTAGGTGACGACGTTG-3' (SEQ ID NO: 55) 5'-CAACGTCGTCACCTAGGGTAGAGTCTTCGAAAGC-3'
V227A	5'-GGTGACGACGCTGAAGCTAACTTCAC-3' (SEQ ID NO: 56) 5'-GTGAAGTTAGCTTCAGCGTCGTCACC-3'
<i>Sac II</i>	
L234V	5'-CTAACTTCACCGCGGTGTTTCGCTCCAG-3' (SEQ ID NO: 57) 5'-CTGGAGCGAACACCGCGGTGAAGTTAG-3'
A238P	5'-GCTTTGTTTCGCTCCACCTATTAGAGCTAGATTGG-3' (SEQ ID NO: 58) 5'-CCAATCTAGCTCTAATAGGTGGAGCGAACAAAGC-3'
<i>Hpa I</i>	
T251N	5'-GCCAGGTGTAACTTGACTGACGAAG-3' (SEQ ID NO: 59) 5'-TTCGTCAGTCAAGTTAACACCTGGC-3'
<i>Aat II</i>	
Y259N	5'-GACGAAGACGTCGTTAACTTGATGGAC-3' (SEQ ID NO: 60) 5'-GTCCATCAAGTTAACGACGTCTTCGTC-3'
<i>Asp I</i>	
E267D	5'-GTCCATTCGACACTGTCGCTAGAACTTC-3' (SEQ ID NO: 61) 5'-GAAGTTCTAGCGACAGTGTGCAATGGAC-3'

- E277Q 5'-CTGACGCTACTCAGCTGTCTCCATTC-3' (SEQ ID NO: 62)
5'-GAATGGAGACAGCTGAGTAGCGTCAG-3'
- A283D 5'-GTCTCCATTCTGTGATTTGTTCACTCAC-3' (SEQ ID NO: 63)
5'-GTGAGTGAACAAATCACAGAATGGAGAC-3'
- Ksp I*
- H287A 5'-GCTTTGTTCAACCGCGGACGAATGGAG-3' (SEQ ID NO: 64)
5'-CTCCATTCGTCCGCGGTGAACAAAGC-3'
- Bam HI*
- R291I 5'-CACGACGAATGGATCCAATACGACTAC-3' (SEQ ID NO: 65)
5'-GTAGTCGTATTGGATCCATTCGTCGTG-3'
- Bsi WI*
- Q292A 5'-GACGAATGGAGAGCGTACGACTACTTG-3' (SEQ ID NO: 66)
5'-CAAGTAGTCGTACGCTCTCCATTCGTC-3'
- Hpa I*
- A320V 5'-GGTGTGTTGGTTTCGTTAAACGAATTGATTGC-3' (SEQ ID NO: 67)
5'-GCAATCAATTCGTTAAAGAAACCAACACC-3'
- (Bgl II)*
- R329H 5'-GCTAGATTGACTCACTCTCCAGTTCAAG-3' (SEQ ID NO: 68)
5'-CTTGAACTGGAGAGTGAGTCAATCTAGC-3'
- Eco RV*
- S364T 5'-CTCACGACAACACTATGATACTATTTTCTTC-3' (SEQ ID NO: 69)
5'-GAAGAAAATAGATATCATAGTGTGTCGTGAG-3'
- Nco I*
- I366V 5'-CGACAACTCCATGGTTTCTATTTTCTTCGC-3' (SEQ ID NO: 70)
5'-GCGAAGAAAATAGAAACCATGGAGTTGTGCG-3'
- Kpn I*
- A379K 5'-GTACAACGGTACCAAGCCATTGTCTAC-3' (SEQ ID NO: 71)
5'-GTAGACAATGGCTTGGTACCGTTGTAC-3'
- S396A 5'-CTGACGGTTACGCTGCTTCTTGGAC-3' (SEQ ID NO: 72)
5'-GTCCAAGAAGCAGCGTAACCGTCAG-3'
- G404A 5'-CTGTTCCATTCGCTGCTAGAGCTTAC-3' (SEQ ID NO: 73)
5'-GTAAGCTCTAGCAGCGAATGGAACAG-3'
- Q415E 5'-GATGCAATGTGAAGCTGAAAAGGAACC-3' (SEQ ID NO: 74)
5'-GGTTCCTTTTCAGCTTCACATTGCATC-3'

Sal I

A437G 5'-CACGGTTGTGGTGTCTGACAAGTTGGG-3' (SEQ ID NO: 75)
5'-CCCAACTTGTCTGACACCACAACCGTG-3'

Mun I

5 A463E 5'-GATCTGGTGGCAATTGGGAGGAATGTTTCG-3' (SEQ ID NO: 76)
5'-CGAAACATTCCTCCCAATTGCCACCAGATC-3'

以及,相应地,可用于其它突变。

如实施例 9 所述测定于酿酒酵母中表达的纯化植酸酶(实施例 7)的最适温度。表 5 显示所导入的每一突变对共有植酸酶-1 稳定性的影响。

10 表 5 单独氨基酸在共有植酸酶-1 中的取代对稳定性的影响

+或-分别指蛋白稳定性变动最多 1℃的正效应或负效应,++和--是指蛋白稳定性变动 1℃至 3℃的正效应或负效应;圆括号中的 10 或 11 指提示氨基酸取代的共有植酸酶序列。

稳定		中性		不稳定	
突变	效应	突变	效应	突变	效应
E58A (10)	+	D69A	±	Y54F (10)	-
D69K (11)	+	D70G (10)	±	V73I	-
D197N (10)	+	N134Q (10)	±	A94K (10)	-
T214L (10)	++	G186H	±	A101R (11)	-
E222T (11)	++	S187A (10)	±	K153N (11)	-
E267D (10)	+	T214V	±	I158V (10)	--
R291I	+	T251N (10)	±	G203A	--
R329H (10)	+	Y259N (10)	±	G205S	-
S364T (10)	++	A283D (10)	±	A217V	-
A379K (11)	+	A320V (10)	±	V227A (11)	--
G404A (10)	++	K445T	±	L234V (10)	-
		A463E (10)	±	A238P (10)	--
				E277Q (10)	-
				H287A (11)	-
				Q292A	-
				I366V (10)	-
				S396A (10)	--
				Q415E (11)	-
				A437G (10)	--
				E451R	--

15

本实施例用上述引物和技术将 8 种正突变(E58A, D197N, E267D, R291I, R329H, S364T, A379K, G404A)组合在共有植酸酶-1-thermo[8]中。此外,导入主要影响植酸酶催化特性的突变 Q50T 和/或 K91A(见专利申请

EP 897010 和 EP 897985, 以及实施例 9)。所得植酸酶(共有植酸酶-1-thermo[8]-Q50T-K91A)的 DNA 和氨基酸序列示于图 7。该共有植酸酶的最适温度和解链温度通过这种方式增加了 7℃(图 15, 16, 17)。

在另一个共有蛋白中, 我们在共有植酸酶-1-thermo[11]中组合了 11 个正突变(E58A, D69K, D197N, T214L, E222T, E267D, R291I, R329H, S364T, A379K, G404A)。此外, 导入突变 Q50T 和/或 K91A。通过这种方式使解链温度比共有植酸酶-1-thermo[8]又增加 3-4℃。

利用表 5 的结果, 我们通过回复突变 K94A, V158I, 和 A396S 进一步改善了共有植酸酶-10 的热稳定性, (A94K, I158V, 和 S396A)的回复显示对共有植酸酶-1 的稳定性的强烈负效应, 所得蛋白称为共有植酸酶-10-thermo[3]。SEQ ID NO:26 加上 3 个突变 K94A, V158I, 和 A396S。此外, 我们导入了主要影响共有植酸酶催化特性的突变 Q50T 和 K91A (参见专利申请 EP 897010 和 EP 897985, 以及实施例 9 和图 14 和 15)。所得 DNA 和氨基酸序列见图 8。所优化的植酸酶显示出比共有植酸酶-10 高 4℃的最适温度和解链温度(图 12 和 13)。此外该植酸酶液在 pH5.5 对 250U/mg 植酸底物的比活性大大增加(图 14)。

在另一种共有蛋白中, 两个另外的突变被导入到产生共有植酸酶-10 thermo[5]的共有植酸酶-10 thermo[3](E222T, G437A)中。此外, 导入突变 Q50T 和/或 K91A。通过这种方式使解链温度比共有植酸酶-10 thermo[3]又增加了 1-2℃。

实施例 4

烟曲霉 ATCC 13073 的植酸酶通过将其氨基酸残基用共有植酸酶-1 和/或共有植酸酶-10 的相关残基取代而稳定

在图 1 的比对中的 6 个氨基酸序列位置上, 仅有或几乎仅有烟曲霉 13073 植酸酶不含有相应的共有植酸酶氨基酸残基, 将这些非共有氨基酸残基用共有残基取代。以下氨基酸在烟曲霉 13073 植酸酶中被取代, 另含有 Q51(24)T 取代(其影响催化特性且对应于共有植酸酶中的 Q50T 取代)和土曲霉 cbs116.46 植酸酶的信号序列(参见欧洲专利申请 0897010, 和图 9): F55(28)Y, V100(73)I, F114(87)Y, A243(220)L, S265(242)P, N294(282)D。圆括号中的数字指图 1 中的编号。

在第二轮中，进一步将在共有植酸酶-10 中鉴定并检验为共有植酸酶-1(表 5)中单一突变的 11 个稳定氨基酸取代中的 4 个(E58A, R329H, S364T, G404A)导入烟曲霉 α -突变体中。此外，还导入氨基酸取代 S154N，其显示植酸酶的蛋白酶敏感性降低。

5 如实施例 3 所述导入突变(见表 6)，如实施例 6-8 所述使之表达。产生的烟曲霉 13073 植酸酶变体被称为 α 突变体(例如带有 Q24T, F28Y, V73I, F87Y, A220L, S242P, N282D 取代的烟曲霉 ATCC 13073 植酸酶)以及"优化的" α -突变体(例如进一步具有 E59A-S1S4N-R329H-S364T-G404A 取代的烟曲霉 α -突变体)。K92A 是一个额外的优选突变。

10 烟曲霉 13073 α -突变体的植酸酶的最适温度(60℃, 图 20)和解链温度(67.0℃, 图 9)比野生型植酸酶的(最适温度: 55℃, Tm: 60℃)提高 5-7℃。另 5 个氨基酸取代进一步使最适温度提高 3℃ (图 20)。

表 6 用于稳定烟曲霉 ATCC 13073 植酸酶的诱变引物

15	突变	引物
	F55Y	5'-CACGTACTCGCCATACTTTTCGCTCGAG-3' (SEQ ID NO: 77) 5'-CTCGAGCGAAAAGTATGGCGAGTACGTG-3'
	E58A	(Xho I) 5'-CCATACTTTTCGCTCGCGGACGAGCTGTCCGTG-3' (SEQ ID NO: 78) 5'-CACGGACAGCTCGTCCGCGAGCGAAAAGTAGG-3'
20	V100I	5'-GTATAAGAAGCTTATTACGGCGATCCAGGCC-3' (SEQ ID NO: 79) 5'-GGCCTGGATCGCCGTAATAAGCTTCTTATAC-3'
	F114Y	5'-CTTCAAGGGCAAGTACGCCTTTTTGAAGACG-3' (SEQ ID NO: 80) 5'-CGTCTTCAAAAAGGCGTACTTGCCCTTGAAG-3'
25	A243L	5'-CATCCGAGCTCGCCTCGAGAAGCATCTTC-3' (SEQ ID NO: 81) 5'-GAAGATGCTTCTCGAGGCGAGCTCGGATG-3'
	S265P	5'-CTAATGGA TGTGTCCGTTTGATACGGTAG-3' (SEQ ID NO: 82) 5'-CTACCGTATCAAACGGACACATGTCCATTAG-3'
30	N294D	5'-GTGGAAGAAGTACGACTACCTTCAGTC-3' (SEQ ID NO: 83) 5'-GACTGAAGGTAGTCGTA CTCTTCCAC-3'

(Mlu I)

R329H 5' -GCCCCGGTTGACGCAATTCGCCAGTGCAGG-3' (SEQ ID NO: 84)
 5' -CCTGCACTGGCGAATGCGTCAACCGGGC-3'

Nco I

5 S364T 5' -CACACGACAACACCATGGTTTCCATCTTC-3' (SEQ ID NO: 85)
 5' -GAAGATGGAAACCATGGTGTGTCGTGTG-3'

(Bss HI)

G404A 5' -GTGGTGCCTTTCGCCGCGCGAGCCTACTTC-3' (SEQ ID NO: 86)
 5' -GAAGTAGGCTCGCGCGCGAAAGGCACCAC-3'

实施例 5

10 将黑曲霉 NRRL3135 植酸酶的活性位点氨基酸残基导入共有植酸酶-1

我们使用黑曲霉 NRRL 3135 植酸酶的晶体结构来限定所有活性位点氨基酸残基(见实施例 1 和 EP 897010)。根据图 1 的比对, 我们将共有植酸酶-1 的下述活性位点残基以及不相同的邻近残基用黑曲霉植酸酶的那些取代:

15 S89D, S92G, A94K, D164S, P201S, G203A, G205S, H212P, G224A,
 D226T, E255T, D256E, V258T, P265S, Q292H, G300K, Y305H, A314T,
 S364G, M365I, A397S, S398A, G404A, 和 A405S。

如实施例 1 所述将新的共有植酸酶-7 蛋白序列反译为 DNA 序列(图 10)。相应的基因(fcp7)如实施例 1 所述用下述寡核苷酸混合物来产生:

Mix 1.7: CP-1, CP-2, CP-3, CP-4.7, CP-5.7, CP-6, CP-7, CP8.7,
 20 CP-9, CP-10.7

Mix 2.7: CP-9, CP-10.7, CP-11.7, CP-12.7, CP-13.7, CP-14.7, CP-15.7,
 CP-16, CP-17.7, CP-18.7, CP-19.7, CP-20, CP-21, CP-22.

上述寡核苷酸的 DNA 序列见图 10。新合成的寡核苷酸另用数字 7 标记。用如实施例 1 所述的相同 PCR 引物合成寡核苷酸后, 实施例 6-8 所述

25 将该基因克隆至表达载体中。

在多形汉逊氏酵母中表达和纯化后, 所检测的该酶的 pH-谱非常相似于黑曲霉植酸酶的(见图 18)。

实施例 6 共有植酸酶基因在多形汉逊氏酵母中的表达

30 用于转化多形汉逊氏酵母 RB11 的植酸酶表达载体[Gellissen, G., Hollenberg, C. P., Janowicz, Z. A.(1994), 甲基营养型酵母中的基因表达,

在 Smith, A.(编), 重组微生物中的基因表达。Dekker, 纽约, pp395-439] 通过将 pBsk-fcp 或其变体的 Eco RI 片段插入多形汉逊氏酵母表达载体 pFPMT121 的多克隆位点中而构建, 其中 pFPMT121 基于酿酒酵母的 ura3 选择标记, 多形汉逊氏酵母的甲酸脱氢酶(FMD)启动子元件和甲醇氧化酶(MO)终止子元件。将 Fcp 基因的 5' 末端融合至 FMD 启动子, 3' 末端融合至 MOX 终止子(Gellissen 等, 应用微生物学生物技术(Appl. Microbiol. Biotechnol.) 46, 46-54, 1996; EP 299108)。所产生的表达载体命名为 pFPMTfcp, pFPMTfcp10, 和 pFPMTfcp7。

使所构建的质粒在大肠杆菌中增殖。质粒 DNA 用现有技术中的常规方法纯化。利用如 Gellissen 等(1996)所述的感受态细胞制备方法和酵母转化方法将所述表达质粒转化至乳清苷-5'-磷酸脱羧酶(ura3)缺陷型多形汉逊氏酵母菌株 RB11 中。将每一种转化混合物涂布于含有 2% 葡萄糖和 1.8% 琼脂的 YNB 培养基(0.14% w/v Difco YNB 和 0.5% 硫酸铵)上, 37℃ 培养。4-5 天后挑出单个转化菌落, 如上述在液体培养基中 37℃ 培养 2 天。随后, 将此培养物的等分试样接种到装有含 2% 葡萄糖之 YNB 培养基的新鲜小瓶中。在选择培养基上进一步传 7 代后, 所述表达载体以多聚体形式整合到酵母基因组中。然后, 通过两个另外的在 3ml 非选择性液体培养基(YPD, 2% 葡萄糖, 10 g/l 酵母提取物, 和 20g/l 蛋白胨)中另两步培养, 获得有丝分裂稳定型转化子。为了获得遗传同源性重组菌株, 将最终的稳定培养物的等分试样涂布于选择平皿中。分离单菌落, 分析在含有 2% 甘油(其阻遏 FMD 启动子)而非葡萄糖的 YNB 中表达的植酸酶。共有植酸酶的纯化如实施 7 所述。

实施例 7

在酿酒酵母中表达共有植酸酶基因并从培养物上清中纯化植酸酶

共有植酸酶基因从相应蓝斑(Bluescript)-质粒(pBsk-fcp, pBSK-fcpl0, pBskfcp7)中分离, 将其连接至酿酒酵母表达载体 pYES2(Invitrogen, San Diego, CA, 美国)的表达盒的 Eco RI 位点, 或亚克隆至缩短的 GAPFL (甘油醛-3-磷酸脱氢酶)启动子和 pho5 终止子之间, 如 Janes 等, 现代遗传学 18, 97-103 所述。基因的正确方向通过 PCR 核对。酿酒酵母菌株, 例如 INVSc1 (Invitrogen, San Diego, CA, USA)如 Hinnen 等, 美国国家科学院院报 75,

1929-1933(1978)所述进行转化。挑出携有受控于 GAPFL 启动子之植酸酶基因的单菌落，在 5ml 选择培养基[SD-尿嘧啶；Sherman, J. P., Finck, G. R. 和 Hicks, J. B. (1986), 酵母遗传学方法实验室手册. 冷泉港大学]中 30℃ 强烈振荡(250 rpm)培养一天。再将预培养物加到 500 ml YPD 培养基(Sherman 等, 1986)中并在相同的条件下培养。按说明书对 gal1 启动子进行诱导。培养 4 天后，离心细胞培养物(7000 rpm, GS3 转头, 15 mm, 5℃)以除去细胞，上清通过 Amicon 8400 cell (PM30 膜；Grace AG, Wallizeller, 瑞士)和 ultrafree-15 离心过滤装置(Biomax-30K, Millipore, Bedford, MA, 美国)超滤浓缩。浓缩液(10 ml)上样至 40 ml Sephadex G25 Superfine 柱(Pharmacia Biotech, Freiburg, 德国)，用 10 mM 乙酸钠，pH 5.0 洗脱以脱盐。将脱盐样品加至 2M (NH₄)₂SO₄ 中，直接上样至 1ml 丁基 Sepharose 凝胶 4 Fast Flow 疏水作用层析柱(Pharmacia Biotech, Freiburg, 德国)，用在 10mM 乙酸钠，pH5.0 中从 2M 到 0M (NH₄)₂SO₄ 的线性梯度洗脱。植酸酶在临界点洗脱，浓缩并上样至 120 ml Sephacryl S-300 凝胶渗透层析柱(Pharmacia Biotech, Freiburg, 德国)。共有植酸酶-1, -7 和 -10 以均一对称峰形式洗脱，经 SDS-PAGE 证实纯度约 95%。

实施例 8 共有植酸酶基因在黑曲霉中的表达

以蓝斑质粒 pBsk-fcp, pBsk-fcp10, 和 pBskfcp7 为模板，在基因的起始密码子上游导入 Bsp HI-位点，终止密码子下游导入 Eco RV-位点。使用 Expand™ 高保真 PCR 试剂盒(Boehringer Mannheim, Mannheim, 德国)及以下引物：

引物 Asp-1：

Bsp HI

5'-TATA~~TCATGAGCGTGTTCGTCGTGCTACTGTTC~~-3' (SEQ ID NO:87)

用于克隆 fcp 和 fcp7 的引物 Asp-2：

Eco RV

3'-ACCCGACTTACAAAGCGAATTCTATAGATATAT-5' (SEQ ID NO:88)

用于克隆 fcp10 的引物 Asp-3：

Eco RV

3'-ACCCTTCTTACAAAGCGAATTCTATAGATATAT-5' (SEQ ID NO:89)

反应按厂家所述进行。经 PCR-扩增的 fcp-基因在起始密码子处有一个

新的 Bsp HI 位点, 其通过引物 Asp-1 导入, 导致第二个氨基酸残基甘氨酸被丝氨酸取代。随后, 将该 DNA-片段用 Bsp HI 和 Eco RV 消化并连接至黑曲霉葡萄糖淀粉酶启动子(glaA)下游的 Nco I 位点和构巢曲霉(*Aspergillus nidulans*)色氨酸 C 终止子(trpC)上游的 Eco RV 位点(Mullaney 等, 1985)。

- 5 在此克隆步骤后, 对该基因测序以检测可能由 PCR 导入的错误。所得表达质粒(基本对应于 EP 684313 之实施例 9 所述 PGLAC 载体)含粗糙脉孢霉(*Neurospora crassa*)的乳清苷-5'-磷酸脱羧酶基因(pyr4)作为筛选标记。黑曲霉的转化和共有植酸酶基因的表达如 EP 684313 所述。共有植酸酶如实施例 7 所述被纯化。

10

实施例 9 植酸酶活性和最适温度及 pH 的检测

此实施例涉及植酸酶活性和最适温度的检测。各种植酸酶都被检测。

如 EP 420358 以及 van Hartingsveldt 等(基因 127, 87-94, 1993)所述制备黑曲霉 NRRL 3135 的植酸酶。

- 15 烟曲霉 ATCC13073, 土曲霉 9A-1, 土曲霉 cbs116.46, *Emericella nidulans*, 嗜热毁丝霉, 及嗜热踝节菌的植酸酶如 EP-0897985 及其引用文献所述制备。

所测的其它植酸酶如本文所述制备。

- 共有植酸酶-1-thermo(8)是共有植酸酶-1 的一个变体, 其进一步包括 8 个突变, 如图 5 中下划线所示。共有植酸酶-1 的无信号肽形式示于图 1 (SEQ ID NO:14), 有信号肽形式示于图 2 (SEQ ID NO:16)。

- 植酸酶活性的测定基本上如 Mitchell 等(1997)所述。在含有 0.5 % 植酸(≈ 5 mM)的 200 mM 乙酸钠, pH 5.0 的试验混合物中进行活性测定。37℃培养 15min 后, 通过加入等体积的 15 % 三氯乙酸来终止反应。所释放的无机磷酸根通过使 100 μ l 试验混合物与 900 μ l H₂O 和 1 ml 0.6 M H₂SO₄, 2 % 抗坏血酸和 0.5 % 钼酸铵混合而定量测定。用磷酸钾的标准溶液作参照。1 个酶活单位定义为 37℃时每分钟释放 1 μ mol 磷酸根的酶量。蛋白浓度通过如 Pace 等人所述计算的 280 nm 处酶消光系数来测定[Pace N. C., Vajdos, F., Fee, L., Grimsley, G.和 Gray, T. (1995), 如何测量和预测蛋白质的摩尔吸收系数, 蛋白质科学, 4, 2411-2423]: 280 nm 处的 1 吸收单位(1 OD)相当于 1.101 mg/ml 共有植酸酶-1, 1.068 mg/ml 共有植酸酶-7, 以及 1.039 mg/ml
- 25
- 30

共有植酸酶-10。

在 pH-最适曲线中，纯化的酶用 10 mM 乙酸钠，pH 5.0 稀释。将稀释蛋白质的试样与等体积的 1% 植酸(约 10 mM)在一系列的不同缓冲液中混合以进行培养：0.4 M 甘氨酸/HCl，pH 2.5；0.4 M 乙酸盐/NaOH，pH 3.0，3.5，
5 4.0，4.5，5.0，5.5；0.4 M 咪唑/HCl，pH 6.0，6.5；0.4 M Tris/HCl pH 7.0，7.5，8.0，8.5，9.0。对照试验表明 pH 仅受混合步骤的影响。如上所述 37℃ 培养 15 分钟。

为了测定植酸酶的底物特异性，将试验混合物中的植酸用 5mM 浓度的各种磷酸盐化合物代替。此外，活性的测试如上所述。

10 为了测定最适温度，将酶(100 μ l)和底物溶液(100 μ l)在指定温度下预温 5min。通过将酶加入底物溶液中开始反应。保温 15 分钟后，用三氯乙酸终止反应，检测磷酸根的释放量。

共有植酸酶-1 的最适 pH 在 pH 6.0-6.5(70 U/mg)左右。导入 Q50T 突变使得最适 pH 渐变为 pH 6.0(130 U/mg)。导入 K91A 突变进一步使最适 pH
15 渐变至更酸性范围。Q50T 和 K91A 突变对共有植酸酶-10 和进一步稳定的共有植酸酶变体也有相当的效应(图 14 和 15)。

共有植酸酶-7(其构建为使黑曲霉 NRRL 3135 植酸酶的催化特性转移至共有植酸酶-1)具有与黑曲霉 NRRL 3135 植酸酶(见图 18)非常相似的 pH-谱。其底物特异性与黑曲霉 NRRL 3135 植酸酶的相似程度更甚于共有植酸
20 酶-1 的。

共有植酸酶-1 的最适温度(71℃)比用来预测共有序列的野生型植酸酶
的(45-55℃，表 7)高 16-26℃。改进型共有植酸酶-10 显示出其最适温度进一步增加到 80℃(图 13)。

当使用酿酒酵母菌株过度生产的上清时，发现共有植酸酶-1-thermo[8]
25 的最适温度在相同的范围(78℃)。经测定共有植酸酶-10-thermo[3]-Q50T-K91A 具有最高的最适温度为 82℃。

表 7 共有植酸酶以及来自烟曲霉，黑曲霉，*E. nidulans*，和嗜热毁丝霉的植酸酶的最适温度和 T_m-值。

最适温度的测定如实施例 9 所述。T_m-值通过实施例 10 所述差示扫描
30 量热来测定。

植酸酶	最适温度(℃)	T _m (℃)
黑曲霉 NRRL 3135	55	63.3
烟曲霉 ATCC 13073	55	62.5
土曲霉 9A-1	49	57.5
土曲霉 cbs116 .46	45	58.5
<i>Emericella nidulans</i>	45	55.7
嗜热毁丝霉	55	-
嗜热踝节菌	45	-
共有植酸酶-10-thermo[5]-Q50T-K91A	-	90.4
共有植酸酶-10-thermo[3]-Q50T-K91A	82	89.3
共有植酸酶-10-thermo[3]-Q50T	82	88.6
共有植酸酶-10	80	85.4
共有植酸酶-1-thermo[11]-Q50T-K91A	-	88.0
共有植酸酶-1-thermo[11]-Q50T	-	88.5
共有植酸酶-1-thermo[8]-Q50T-K91A	-	85.7
共有植酸酶-1-thermo[8]-Q50T	78	84.7
共有植酸酶-1-thermo[8]	81	-
共有植酸酶-1-thermo[3]	75	-
共有植酸酶-1-Q50T	-	78.9
共有植酸酶-1	71	78.1
烟曲霉 α -突变体 Q51T	60	67.0
烟曲霉 α -突变体, 加突变 E59A, S154N, R329H, S364T, G404A	63	-
烟曲霉"优化的" α 突变体, 加突变 K92A	63	-

实施例 10

通过差示扫描量热(DSC)测定解链温度

为了测定植酸酶的伸展温度, 如 Brugger 等, 1997[Brugger, R., Mascarello, F., Augem, S., van Loon, A. P. G. M.和 Wyss, M.(1997), 通过差示扫描量热研究植酸酶和 pH2.5 酸性植酸酶的热变性。植酸和植酸酶的生物化学(Rasmussen, S.K.; Raboy, V.; Dalbøge, H.和 Loewus, F.编; Kluwer Academic Publishers, Dordrecht, 荷兰]所述使用差示扫描量热法。用 50-60mg/ml 植酸酶均一溶液进行检验。以 10℃/min 的恒定加热速率加热到 90-95℃。

所测的解链点证实了最适温度检测结果(表 7)。迄今最稳定的共有植酸酶是共有植酸酶-10-thermo[3]-Q50T-K91A, 其在所选条件下显示出解链温

度 89.3℃。这比所用的野生型植酸酶的解链温度高 26.0 到 33.6℃。

实施例 11

将担子菌植酸酶活性位点转移到植酸酶-10-thermo[3]-Q50T-K91A

- 5 如上所述(实施例 5)，将衍生自担子菌植酸酶活性位点的突变导入共有植酸酶-10 中。制备下述 a) - e) 的 5 个构建体：

a) 称作共有植酸酶-12 的构建物，其含有担子菌共有序列中选定数目的活性位点残基。其氨基酸序列被显示于图 21(前 26 个氨基酸组成信号肽；不同于共有植酸酶-10-thermo[3]-Q50T-K91A 的位置带下划线)；

- 10 b) 将一簇突变(第 II 簇)转移到共有植酸酶-I 和-10 的序列中，即：S80Q, Y86F, S90G, K91A, S92A, K93T, A94R, Y95I；

c) 以类似的方式转移另一簇突变(第 III 簇)，即：T129V, E133A, Q134N, M136S, V137S, N138Q, S139A；

- 15 d) 以类似的方式转移另一簇突变(第 IV 簇)，即：A168D, E171T, K172N, F173W；

e) 最后，转移另一簇突变(第 V 簇)被，即：Q297G, S298D, G300D, Y305T。

这些构建物如实施例 6 到 8 所述被表达。

20 实施例 12 利用 GAP 进行植酸酶比对

使本文所述植酸酶—即氨基酸序列及相应 DNA 序列—相互比对。此外，对其它植酸酶也进行相应比对，它们是：

- EP 897985 所述的共有植酸酶-1；
- EP 420358 所述的黑曲霉(*ficuum*)NRRL 3135(黑曲霉 NRRL3135)的植
- 25 酸酶；
- EP 897010 所述的烟曲霉 ATCC 13073(烟曲霉 13073)；烟曲霉 ATCC 32239(烟曲霉 32239)；土曲霉 cbs116.46 (土曲霉 cbs)；*Emericella nidulans* (*E. nidulans*)；和嗜热踝节菌(*T. thermophilus*)的植酸酶；
- EP 684313 所述的嗜热毁丝霉和土曲霉 9-A1 的植酸酶；
- 30 - WO 9735017 (PCT/US97/04559)所述的 *Thermomyces lanuginosus*(*T. lanuginosus*)的植酸酶；

- WO98/28409 所述的平田头菇, *Paxillus involutus* 1 和 2 (*P. involutus* phyA1 和 phyA2); 以及绒毛栓菌的植酸酶;

- WO 98/28408 所述的 *Peniophora lycii*(*P. lycii*)的植酸酶。

用 GAP 程序及如上所述设置进行比对。

5 在多肽比对中, 除了与共有植酸酶-11 比对外, 还包括信号肽。

氨基酸序列比对的结果被显示于下面表 8 中。每一小格的第一个数字是氨基酸的相似性, 第二个数字是氨基酸的同源性。

在 DNA 序列比对中, 也总是包括信号肽。结果示于下表 9。

本发明包括以外实施方案(A)-(J)。

10 在这些实施方案中, 当测定另一种植酸酶在氨基酸水平的 % 同一性或 % 相似性时, 使用上述 GAP 等比对程序将其氨基酸序列与参照序列(例如在实施方案(A)中的共有植酸酶-10 氨基酸序列)相比对。同一性百分比和相似性百分比均通过此程序计算。其它植酸酶的氨基酸序列可包括或不包括信号肽。

15 当测定另一种编码植酸酶的 DNA 在 DNA 水平的 % 同一性时, 使用上述 GAP 等比对程序将该 DNA 序列与参照序列 [例如在实施方案(A)中的 SEQ ID NO:25 至核苷酸 12-1412(如图 5 所示的共有植酸酶-10(Fcpl0)的 DNA 序列]相比对。同一性百分比通过此程序计算。编码另一种植酸酶的 DNA 序列可以是包含内含子的基因组 DNA 序列, 或 cDNA 序列。其可以包括或不包括信号肽编码部分。

20 当测定杂交时, 所用探针是特异性 DNA 序列[例如在实施方案(A)中 SEQ ID NO:25 的核苷酸 12-1412 (如图 5 所示的共有植酸酶-10(Fcpl0)的 DNA 序列]。编码另一种植酸酶的 DNA 序列可以是含有植酸酶 DNA 序列的基因组 DNA; 纯化的基因组 DNA 序列(就植酸酶 DNA 序列而言为纯化的); 或其
25 可以是编码植酸酶的 cDNA 序列, 优选纯化型或扩增型, 例如 PCR-扩增型。编码植酸酶的 DNA, 无论什么类型, 可包括或不包括信号肽编码部分。合适的杂交条件如上所述。

术语 "DNA 序列" 包括在此例示的 DNA 序列的片段或部分, 只要其能够编码活性酶(例如植酸酶)。

30 术语"氨基酸序列"包括在此例示的氨基酸序列, 只要其具有酶活性 (例如表现植酸酶活性)。

(A)与共有植酸酶-10(CP10, Fcp10)相关的植酸酶及相应 DNA 序列

一种植酸酶，其含有与图 5 所示共有植酸酶-10(Fcp10)中 1-467 位的氨基酸序列有至少 93.80%；或至少 94, 94.5, 95, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 同一性的氨基酸序列。

- 5 一种植酸酶，含有与共有植酸酶-10 中 1-467 位的氨基酸序列有至少 95.09%；或至少 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 相似性的氨基酸序列。

一种植酸酶，其由与图 5 所示共有植酸酶-10(Fcp10)之 DNA 序列中核苷酸 12-1412 有至少 95.88%；或至少 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 同一性的 DNA 序列编码。

- 10 一种编码植酸酶的 DNA 序列，其与图 5 所示共有植酸酶-10 (Fcp10)之 DNA 序列的核苷酸 12-1412：(i)有至少 95.88%；或至少 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 相同；或(ii)在低度，或中度，中/高度，高度，或极高度严谨条件下可杂交。合适的阴性对照是编码共有植酸酶-1 的 DNA。合适的阳性对照是编码 CP10, CP10-thermo[3]-Q50T,K91A, CP1-thermo[8], CP1-thermo [83]Q50T,K91A 中任一个的 DNA。

一种编码植酸酶的 DNA 序列，所述酶含有与图 5 所示共有植酸酶-10(Fcp10)中 1-467 位的氨基酸序列有至少 93.80%；或至少 94, 94.5, 95, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 同一性的氨基酸序列。

- 20 (B)与共有植酸酶-10-thermo[3]-Q50T-K91A 相关的植酸酶及相应 DNA 序列

一种植酸酶，其含有与图 8 所示共有植酸酶-10-thermo[3]-Q50T-K91A 中 1-467 位的氨基酸序列有至少 93.37%；或至少 93.5, 94, 94.5, 95, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 同一性的氨基酸序列。

- 25 一种植酸酶，其含有与图 8 所示共有植酸酶-10-thermo[3]-Q50T-K91A 中 1-467 位的氨基酸序列有至少 94.66%；或至少 95.0, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 相似性的氨基酸序列。

一种植酸酶，其由与图 8 所示共有植酸酶-10-thermo[3]-Q50T-K91A 之 DNA 序列的核苷酸 12-1412 有至少 95.88%；或至少 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5% 同一性的 DNA 序列编码。

- 30 一种编码植酸酶的 DNA 序列，其与图 8 所示共有植酸酶-10-

thermo[3]-Q50T-K91A 之 DNA 序列的核苷酸 12-1412: (i)有至少 95.88 % ;
或至少 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5 % 相同; 或(ii)在低度, 或
中度, 中/高度, 高度, 或极高度严谨条件下可杂交。合适的阴性对照是编
码共有植酸酶-1 的 DNA。合适的阳性对照是编码 CPl0, CPl0-thermo[3]-
5 Q50T-K91A, CPl-thermo[8], 或 CPl-thermo [8]-Q50T-K91A 中任一种的
DNA 。

一种编码植酸酶的 DNA 序列, 所述酶含有与图 8 所示共有植酸酶-10-
thermo[3]-Q50T-K91A 中 1-467 位的氨基酸序列有至少 93.37 % ;或至少 93.5,
94, 94.5, 95, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5 % 同一性的
10 氨基酸序列。

(C)与共有植酸酶-1-thermo[8]相关的植酸酶及相应 DNA 序列

一种植酸酶, 其含有与共有植酸酶-1-thermo[8](示于图 7; 已添加回复
突变 T50Q 和 A91K)中 1-467 位的氨基酸序列有至少 98.30 % ; 或至少 98.5,
99, 99.5 % 同一性的氨基酸序列。

15 一种植酸酶, 其含有与共有植酸酶-1-thermo[8](示于图 7; 已添加回复
突变 T50Q 和 A91K)中 1-467 位的氨基酸序列有至少 98.51 % ; 或至少 99,
99.5 % 相似性的氨基酸序列。

一种植酸酶, 其由与共有植酸酶-1-thermo[8](示于图 7; 已添加回复突
变 T50Q 和 A91K)之 DNA 序列的核苷酸 1-1407 有至少 98.73 % ; 或至少 99,
20 99.5 % 同一性的 DNA 序列编码。

一种编码植酸酶的 DNA 序列, 其与共有植酸酶-1-thermo[8](示于图 7;
已添加回复突变 T50Q 和 A91K)之 DNA 序列的核苷酸 1-1407: (i)有至少
98.73 % ; 或至少 99, 99.5 % 相同; 或(ii)在低度, 或中度, 中/高度, 高度,
或极高度严谨条件下可杂交。合适的阴性对照是编码共有植酸酶-1 的 DNA。
25 合适的阳性对照是编码 CPl-thermo[8], 或 CPl-thermo[8]-Q50T-K91A 中任
一个的 DNA。

一种编码植酸酶的 DNA 序列, 所述酶含有与共有植酸酶-1-thermo[8](示
于图 7; 已添加回复突变 T50Q 和 A91K)中 1-467 位的氨基酸序列有至少 98.30
% ; 或至少 98.5, 99, 99.5 % 同一性的氨基酸序列。

30 (D)与共有植酸酶-1-thermo[8]-Q50T-K91A 相关的植酸酶及相应 DNA
序列

一种植酸酶，其含有与图 7 所示共有植酸酶-1-thermo[8]-Q50T-K91A 中 1-467 位的氨基酸序列有至少 97.87%；或至少 98，98.5，99，99.5% 同一性的氨基酸序列。

5 一种植酸酶，其含有与图 7 所示共有植酸酶-1-thermo[8]-Q50T-K91A 中 1-467 位的氨基酸序列有至少 98.08%；或至少 98.5，99，99.5% 相似性的氨基酸序列。

一种植酸酶，其由与图 7 所示共有植酸酶-1-thermo[8]-Q50T-K91A 之 DNA 序列的核苷酸 1-1407 有至少 98.37%；或至少 98.5，99，99.5% 同一性的 DNA 序列编码。

10 一种编码植酸酶的 DNA 序列，其与图 7 所示共有植酸酶-1-thermo[8]-Q50T-K91A 之 DNA 序列的核苷酸 1-1407：(i)至少 98.37%；或至少 98.5，99，99.5% 相同；或(ii)在低度，或中度，中/高度，高度，或极高度严谨条件下可杂交。合适的阴性对照是编码共有植酸酶-1 的 DNA。合适的阳性对照是编码 CPl-thermo[8]，或 CPl-thermo[8]-Q50T-K91A 中任一种的 DNA。

15 一种编码植酸酶的 DNA 序列，所述酶含有与图 7 所示的共有植酸酶-1-thermo[8]-Q50T-K91A 中 1-467 位的氨基酸序列有至少 97.87%；或至少 98，98.5，99，99.5% 同一性的氨基酸序列。

(E)与共有植酸酶-11 相关的植酸酶及相应 DNA 序列

20 一种植酸酶，其含有与图 6 所示共有植酸酶-11 中 1-482 位的氨基酸序列有至少 90.71%；或至少 91，91.5，92，92.5，93，93.5，94，94.5，95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 同一性的氨基酸序列。

一种植酸酶，其含有与图 6 所示共有植酸酶-11 中 1-482 位的氨基酸序列有至少 92.07%；或至少 92.5，93，93.5，94，94.5，95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 相似性的氨基酸序列。

25 一种编码植酸酶的 DNA 序列，所述酶含有与图 6 所示共有植酸酶-11 中 1-482 位的氨基酸序列有至少 90.71%；或至少 91，91.5，92，92.5，93，93.5，94，94.5，95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 同一性的氨基酸序列。

(F)与烟曲霉 α -突变体相关的植酸酶及相应 DNA 序列

30 一种植酸酶，其含有与图 9 所示烟曲霉 α -突变体(植酸酶)中 1-467 位的氨基酸序列有至少 97.17%；或至少 97.5，98，98.5，99，99.5% 同一性

的氨基酸序列。

一种植酸酶，其含有与图 9 所示烟曲霉 α -突变体(植酸酶)中 1-467 位的氨基酸序列有至少 97.82%；或至少 98, 98.5, 99, 99.5%相似性的氨基酸序列。

- 5 一种植酸酶，其由与图 9 所示烟曲霉 ATCC 13073 α -突变体(植酸酶)之 DNA 序列的核苷酸 1-1401 有至少 96.13%；或至少 96.5, 97, 97.5, 98, 98.5, 99, 99.5%同一性的 DNA 序列编码。

一种编码植酸酶的 DNA 序列，所述酶含有与图 9 所示烟曲霉 ATCC 13073 α -突变体(植酸酶)中 1-467 位的氨基酸序列有至少 97.17%；或至少 10 97.5, 98, 98.5, 99, 99.5%同一性的氨基酸序列。

一种编码植酸酶的 DNA 序列，其与图 9 所示烟曲霉 ATCC 13073 α -突变体(植酸酶)之 DNA 序列的核苷酸 1-1401:(i)有至少 96.13%；或至少 96.5, 97, 97.5, 98, 98.5, 99, 99.5%相同；或(ii)在低度，或中度，中/高度，高度，或极高度严谨条件下可杂交。合适的阴性对照是编码烟曲霉 13073 的 15 植酸酶的 DNA。合适的阳性对照是编码烟曲霉 ATCC 13073 α -突变体的植酸酶，或优化型 α -突变体的 DNA。

(G)与优化型烟曲霉 α -突变体相关的植酸酶及相应 DNA 序列

一种植酸酶，其含有与优化型烟曲霉 α -突变体的植酸酶序列至少 96.08%；或至少 96.5, 97, 97.5, 98, 98.5, 99, 99.5%相同的氨基酸序列。

- 20 一种植酸酶，其含有与优化型烟曲霉 α -突变体的植酸酶序列有至少 96.74%；或至少 97, 97.5, 98, 98.5, 99, 99.5%相似性的氨基酸序列。

一种植酸酶，其由与编码优化型烟曲霉 α -突变体植酸酶的 DNA 序列中核苷酸 1-1401 有至少 95.63%；或至少 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5%同一性的 DNA 序列编码。

- 25 一种 DNA 序列，其编码的植酸酶含有与优化型烟曲霉 α -突变体植酸酶有至少 96.08%；或至少 96.5, 97, 97.5, 98, 98.5, 99, 99.5%同一性的氨基酸序列。

一种编码植酸酶的 DNA 序列，其与编码优化型烟曲霉 α -突变体植酸酶的 DNA 序列中核苷酸 1-1401:(i)有至少 95.63%；或至少 96, 96.5, 97, 30 97.5, 98, 98.5, 99, 99.5%相同；或(ii)在低度，或中度，中/高度，高度，或极高度严谨条件下可杂交。

合适的阴性对照是编码烟曲霉 13073 植酸酶的 DNA。合适的阳性对照是编码烟曲霉 ATCC 13073 α -突变体植酸酶或优化型 α -突变体的 DNA。

(H)与共有植酸酶-7 相关的植酸酶及相应 DNA 序列

一种植酸酶，其含有与图 10 所示共有植酸酶-7 中 1-467 位的氨基酸序列有至少 94.87%；或至少 95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 同一性的氨基酸序列。

一种植酸酶，其含有与图 10 所示共有植酸酶-7 中 1-467 位的氨基酸序列有至少 95.30%；或至少 95.5，96，96.5，97，97.5，98，98.5，99，99.5% 相似性的氨基酸序列。

一种植酸酶，其由与图 10 所示共有植酸酶-7 之 DNA 序列的核苷酸 12-1412 有至少 96.38%；或至少 96.5，97，97.5，98，98.5，99，99.5% 同一性的 DNA 序列编码。

一种编码植酸酶的 DNA 序列，其与图 10 所示共有植酸酶-7 之 DNA 序列的核苷酸 12-1412：(i)有至少 96.38%；或至少 96.5，97，97.5，98，98.5，99，99.5% 相同；或(ii)在低度，或中度，中/高度，高度，或极高度严谨条件下可杂交。

一种编码植酸酶的 DNA 序列，所述酶含有与图 10 所示共有植酸酶-7 中 1-467 位的氨基酸序列有至少 94.87%；或至少 95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 同一性的氨基酸序列。

(I)与担子菌共有植酸酶相关的植酸酶

一种植酸酶，其含有与(i)图 3 所示担子菌共有植酸酶的氨基酸 1-441，和 (ii)图 5 所示氨基酸 1-26 的组合序列(将序列(ii)加在序列(i)的 N-末端)有 76.23%；或至少 77，78，79，80，81，82，83，84，85，86，87，88，89，90，91，92，93，94，94.5，95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 同一性的氨基酸序列。

一种植酸酶，其含有与图 3 所示担子菌共有植酸酶中 1-441 位的氨基酸序列有至少 79.50%；或至少 80，81，82，83，84，85，86，87，88，89，90，91，92，93，94，95，95.5，96，96.5，97，97.5，98，98.5，99，99.5% 相似性的氨基酸序列。

(J)与共有植酸酶-12 相关的植酸酶

一种植酸酶，其含有与图 21 所示共有植酸酶-12 的 1-467 位的氨基酸序列有至少 70, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5 % 同一性的氨基酸序列。

- 5 一种植酸酶，其含有与图 21 所示共有植酸酶-12 的 1-467 位的氨基酸序列有至少 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 95.5, 96, 96.5, 97, 97.5, 98, 98.5, 99, 99.5 % 相似性的氨基酸序列。

表 8

植酸酶氨基酸序列的比较

植酸酶	CP10	CP10-thermo[3]-Q50T-K91A	CP1-thermo[8]	CP1-thermo[8]-Q50T-K91A	CP11	CP7	Basidio	烟曲霉 α -突变体	烟曲霉 α -突变体(优化)
共有植酸酶-I	95.08/93.79	94.65/93.36	98.50/98.29	98.07/97.86	92.06/90.70	95.29/94.86	69.42/62.16	85.59/82.58	84.73/81.72
黑曲霉 NRRL3135	79.48/76.46	79.05/76.03	80.35/77.75	79.91/77.32	79.27/76.31	84.02/81.64	67.19/59.32	74.07/70.11	74.95/70.99
土曲霉 9-A1	76.04/72.11	75.82/71.90	76.47/72.33	76.25/72.11	76.51/73.02	75.76/71.18	65.39/58.02	69.67/64.84	69.45/64.84
土曲霉 cbs	79.04/75.11	78.82/74.89	79.48/75.76	79.26/75.55	77.19/73.27	79.17/75.00	66.92/59.65	72.59/67.76	72.37/67.76
E. nidulans	78.70/74.35	78.26/73.91	79.78/75.87	79.35/75.44	80.56/76.62	76.96/73.04	67.20/58.13	72.39/67.83	72.11/67.54
烟曲霉 13073	82.93/80.31	82.50/79.87	82.31/79.04	81.88/78.60	81.36/78.64	80.13/76.20	63.54/57.91	97.82/97.16	96.73/96.07
烟曲霉 32239	81.30/77.39	80.87/76.96	81.09/77.61	80.65/77.17	79.95/76.08	79.13/75.22	63.61/54.97	90.22/86.52	89.57/85.87
嗜热蹼节菌	77.83/73.84	77.38/73.39	78.67/74.89	78.22/74.44	78.47/74.76	76.51/73.15	61.54/54.36	72.01/66.82	72.69/67.49
嗜热毁丝霉	69.16/62.81	69.48/63.33	69.27/62.84	69.59/63.36	69.65/63.06	68.82/62.13	65.56/57.91	66.21/58.45	66.44/58.68
T. lanuginosus	73.52/66.70	73.06/66.44	71.92/64.61	71.46/64.16	74.21/68.86	69.50/62.62	67.20/57.41	68.91/61.02	69.61/61.72
P. lycii	64.92/59.10	64.91/59.37	64.46/58.09	64.46/58.36	65.03/59.84	63.13/56.50	77.75/73.07	64.08/57.11	62.47/55.91
平田头菇	64.51/51.81	64.86/51.94	62.98/51.41	63.33/51.54	64.50/52.30	63.05/51.15	78.92/74.71	61.64/52.38	62.13/53.07
P. involutus 1	66.67/58.07	66.67/58.33	64.84/56.51	64.84/56.77	63.30/54.52	65.33/56.53	79.49/76.22	59.59/51.81	59.95/52.20
P. involutus 2	65.54/55.70	65.30/55.53	66.85/56.87	66.58/56.68	66.30/56.35	64.27/54.13	78.09/74.59	61.26/52.62	61.04/52.47

绒毛栓菌	65.46/57.22	65.72/57.47	62.89/55.67	63.14/55.93	65.03/57.65	63.28/56.51	78.34/75.12	64.08/57.11	62.30/55.24
CP10	-	99.57/99.57	96.57/95.50	96.15/95.08	95.02/94.56	91.01/89.29	70.22/62.28	85.13/82.76	85.99/83.62
CP10t[3]Q50TK91A	99.57/99.57	-	96.15/95.08	96.57/95.50	94.56/94.10	90.58/88.87	70.47/62.28	85.13/82.76	85.99/83.62
CP1thermo[8]	96.57/95.50	96.15/95.08	-	99.57/99.57	93.42/92.29	94.43/93.79	68.40/60.74	84.52/81.94	85.38/82.80
CP1t[8]Q50TK91A	96.15/95.08	96.57/95.50	99.57/99.57	-	92.97/91.84	94.00/93.36	68.64/60.74	84.52/81.94	85.38/82.80
CP11	95.02/94.56	94.56/94.10	93.42/92.29	92.97/91.84	-	88.44/86.62	68.27/59.73	82.23/79.73	83.37/80.87
CP7	91.01/89.29	90.58/88.87	94.43/93.79	94.00/93.36	88.44/86.62	-	69.80/62.69	81.94/78.71	81.72/78.50
Basidio	70.22/62.28	70.47/62.28	68.40/60.74	68.64/60.74	68.27/59.73	69.80/62.69	-	65.97/60.52	66.41/60.68
烟曲霉 α -突变	85.13/82.76	85.13/82.76	84.52/81.94	84.52/81.94	82.23/79.73	81.94/78.71	65.97/60.52	-	98.93/98.93
烟曲霉 α -突变-优化	85.99/83.62	85.99/83.62	85.38/82.80	85.38/82.80	83.37/80.87	81.72/78.50	66.41/60.68	98.93/98.93	-

表 9
编码植酸酶的DNA序列的比较

植酸酶	CP10	CP10-thermo[3]-Q50T-K91A	CP1-thermo[8]	CP1-thermo[8]-Q50T-K91A	CP7	Basidio	烟曲霉 α -突变体	烟曲霉 α -突变体(优化)
共有植酸酶-1	95.87	95.87	98.72	98.36	96.37	65.46	66.88	66.88
黑曲霉 NRRL3135	65.10	64.82	66.10	65.74	67.52	50.68	65.88	66.17
土曲霉 9-A1	61.74	61.53	62.17	62.03	60.53	49.40	66.24	66.31
土曲霉 obs	62.52	62.30	63.02	62.88	61.45	49.74	68.17	68.24
E. nidulans	65.08	64.94	65.30	65.01	64.22	49.92	64.90	65.44
烟曲霉 13073	65.66	65.38	64.19	64.08	63.65	48.27	96.12	95.62
嗜热螺节菌	62.52	62.50	62.53	62.66	62.00	52.19	61.77	61.92
嗜热毁丝霉	55.51	55.15	55.36	55.22	53.91	48.44	58.17	58.24
T. lanuginosus	57.56	57.20	56.76	56.47	62.00	44.66	59.71	60.07
P. lycii	45.76	46.51	45.14	55.21	55.46	58.50	48.91	49.44
平田头菇	49.89	49.89	49.89	50.11	45.54	61.66	47.49	47.56
P. involutus 1	48.32	49.03	47.81	47.96	49.59	59.80	49.96	50.19
P. involutus 2	48.24	49.00	48.08	48.63	47.94	60.16	47.56	47.63
绒毛栓菌	47.00	47.17	46.46	47.62	46.83	60.37	49.89	49.96
CP10	-	99.43	96.40	96.05	93.73	66.40	67.81	68.24

CP10t [3]Q50TK91A	99.43	-	96.37	96.58	93.45	66.29	67.81	68.24
Cplthermo [8]	96.40	96.37	-	99.65	95.30	65.40	66.74	67.17
CP1t [8]Q50TK91A	96.05	96.58	99.65	-	94.94	65.47	66.74	67.17
CP7	93.73	93.45	95.30	94.94	-	64.56	65.88	65.88
Basidio	66.40	66.29	65.40	65.47	64.56	-	50.41	50.49
烟曲霉 α -突变	67.81	67.81	66.74	66.74	65.88	50.41	-	99.50
烟曲霉 α -突变-优化	68.24	68.24	67.17	67.17	65.88	50.49	99.50	-

序列表

<110> Novo Nordisk A/S

<120> 改进的植酸酶

<130> seqlist171299

<140>

<141>

<150> DK 99/00092

<151> 1999-01-22

<150> DK 99/01340

<151> 1999-09-21

<160> 89

<170> PatentIn Ver. 2.1

<210> 1

<211> 440

<212> PRT

<213> 土曲霉 9A-1

<400> 1

Lys	His	Ser	Asp	Cys	Asn	Ser	Val	Asp	His	Gly	Tyr	Gln	Cys	Phe	Pro
1				5						10				15	

Glu	Leu	Ser	His	Lys	Trp	Gly	Leu	Tyr	Ala	Pro	Tyr	Phe	Ser	Leu	Gln
			20					25					30		

Asp	Glu	Ser	Pro	Phe	Pro	Leu	Asp	Val	Pro	Glu	Asp	Cys	His	Ile	Thr
	35						40					45			

Phe	Val	Gln	Val	Leu	Ala	Arg	His	Gly	Ala	Arg	Ser	Pro	Thr	His	Ser
	50					55					60				

Lys	Thr	Lys	Ala	Tyr	Ala	Ala	Thr	Ile	Ala	Ala	Ile	Gln	Lys	Ser	Ala
65					70					75					80

Thr	Ala	Phe	Pro	Gly	Lys	Tyr	Ala	Phe	Leu	Gln	Ser	Tyr	Asn	Tyr	Ser
				85					90					95	

Leu	Asp	Ser	Glu	Glu	Leu	Thr	Pro	Phe	Gly	Arg	Asn	Gln	Leu	Arg	Asp
			100					105					110		

Leu	Gly	Ala	Gln	Phe	Tyr	Glu	Arg	Tyr	Asn	Ala	Leu	Thr	Arg	His	Ile
	115						120					125			

Asn	Pro	Phe	Val	Arg	Ala	Thr	Asp	Ala	Ser	Arg	Val	His	Glu	Ser	Ala
	130						135					140			

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

Gly Gly Asn Trp Ala Asp Cys Phe
435 440

<210> 2
<211> 440
<212> PRT
<213> 土曲霉 cbs

<400> 2
Asn His Ser Asp Cys Thr Ser Val Asp Arg Gly Tyr Gln Cys Phe Pro
1 5 10 15
Glu Leu Ser His Lys Trp Gly Leu Tyr Ala Pro Tyr Phe Ser Leu Gln
20 25 30
Asp Glu Ser Pro Phe Pro Leu Asp Val Pro Asp Asp Cys His Ile Thr
35 40 45
Phe Val Gln Val Leu Ala Arg His Gly Ala Arg Ser Pro Thr Asp Ser
50 55 60
Lys Thr Lys Ala Tyr Ala Ala Thr Ile Ala Ala Ile Gln Lys Asn Ala
65 70 75 80
Thr Ala Leu Pro Gly Lys Tyr Ala Phe Leu Lys Ser Tyr Asn Tyr Ser
85 90 95
Met Gly Ser Glu Asn Leu Thr Pro Phe Gly Arg Asn Gln Leu Gln Asp
100 105 110
Leu Gly Ala Gln Phe Tyr Arg Arg Tyr Asp Thr Leu Thr Arg His Ile
115 120 125
Asn Pro Phe Val Arg Ala Ala Asp Ser Ser Arg Val His Glu Ser Ala
130 135 140
Glu Lys Phe Val Glu Gly Phe Gln Asn Ala Arg Gln Gly Asp Pro His
145 150 155 160
Ala Asn Pro His Gln Pro Ser Pro Arg Val Asp Val Val Ile Pro Glu
165 170 175
Gly Thr Ala Tyr Asn Asn Thr Leu Glu His Ser Ile Cys Thr Ala Phe
180 185 190
Glu Ala Ser Thr Val Gly Asp Ala Ala Ala Asp Asn Phe Thr Ala Val
195 200 205
Phe Ala Pro Ala Ile Ala Lys Arg Leu Glu Ala Asp Leu Pro Gly Val
210 215 220
Gln Leu Ser Ala Asp Asp Val Val Asn Leu Met Ala Met Cys Pro Phe
225 230 235 240

Glu Thr Val Ser Leu Thr Asp Asp Ala His Thr Leu Ser Pro Phe Cys
 245 250 255
 Asp Leu Phe Thr Ala Ala Glu Trp Thr Gln Tyr Asn Tyr Leu Leu Ser
 260 265 270
 Leu Asp Lys Tyr Tyr Gly Tyr Gly Gly Gly Asn Pro Leu Gly Pro Val
 275 280 285
 Gln Gly Val Gly Trp Ala Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser
 290 295 300
 Pro Val His Asp His Thr Cys Val Asn Asn Thr Leu Asp Ala Asn Pro
 305 310 315 320
 Ala Thr Phe Pro Leu Asn Ala Thr Leu Tyr Ala Asp Phe Ser His Asp
 325 330 335
 Ser Asn Leu Val Ser Ile Phe Trp Ala Leu Gly Leu Tyr Asn Gly Thr
 340 345 350
 Lys Pro Leu Ser Gln Thr Thr Val Glu Asp Ile Thr Arg Thr Asp Gly
 355 360 365
 Tyr Ala Ala Ala Trp Thr Val Pro Phe Ala Ala Arg Ala Tyr Ile Glu
 370 375 380
 Met Met Gln Cys Arg Ala Glu Lys Gln Pro Leu Val Arg Val Leu Val
 385 390 395 400
 Asn Asp Arg Val Met Pro Leu His Gly Cys Ala Val Asp Asn Leu Gly
 405 410 415
 Arg Cys Lys Arg Asp Asp Phe Val Glu Gly Leu Ser Phe Ala Arg Ala
 420 425 430
 Gly Gly Asn Trp Ala Glu Cys Phe
 435 440

<210> 3

<211> 441

<212> PRT

<213> 黑曲霉泡盛变种

<400> 3

Asn Gln Ser Thr Cys Asp Thr Val Asp Gln Gly Tyr Gln Cys Phe Ser
 1 5 10 15

Glu Thr Ser His Leu Trp Gly Gln Tyr Ala Pro Phe Phe Ser Leu Ala
 20 25 30

Asn Glu Ser Ala Ile Ser Pro Asp Val Pro Ala Gly Cys Arg Val Thr
 35 40 45
 Phe Ala Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Glu Ser
 50 55 60
 Lys Gly Lys Lys Tyr Ser Ala Leu Ile Glu Glu Ile Gln Gln Asn Val
 65 70 75 80
 Thr Thr Phe Asp Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Ser
 85 90 95
 Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Glu Leu Val Asn
 100 105 110
 Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Glu Ser Leu Thr Arg Asn Ile
 115 120 125
 Ile Pro Phe Ile Arg Ser Ser Gly Ser Ser Arg Val Ile Ala Ser Gly
 130 135 140
 Glu Lys Phe Ile Glu Gly Phe Gln Ser Thr Lys Leu Lys Asp Pro Arg
 145 150 155 160
 Ala Gln Pro Gly Gln Ser Ser Pro Lys Ile Asp Val Val Ile Ser Glu
 165 170 175
 Ala Ser Ser Ser Asn Asn Thr Leu Asp Pro Gly Thr Cys Thr Val Phe
 180 185 190
 Glu Asp Ser Glu Leu Ala Asp Thr Val Glu Ala Asn Phe Thr Ala Thr
 195 200 205
 Phe Ala Pro Ser Ile Arg Gln Arg Leu Glu Asn Asp Leu Ser Gly Val
 210 215 220
 Thr Leu Thr Asp Thr Glu Val Thr Tyr Leu Met Asp Met Cys Ser Phe
 225 230 235 240
 Asp Thr Ile Ser Thr Ser Thr Val Asp Thr Lys Leu Ser Pro Phe Cys
 245 250 255
 Asp Leu Phe Thr His Asp Glu Trp Ile His Tyr Asp Tyr Leu Gln Ser
 260 265 270
 Leu Lys Lys Tyr Tyr Gly His Gly Ala Gly Asn Pro Leu Gly Pro Thr
 275 280 285
 Gln Gly Val Gly Tyr Ala Asn Glu Leu Ile Ala Arg Leu Thr His Ser
 290 295 300
 Pro Val His Asp Asp Thr Ser Ser Asn His Thr Leu Asp Ser Asn Pro
 305 310 315 320

Ala Thr Phe Pro Leu Asn Ser Thr Leu Tyr Ala Asp Phe Ser His Asp
 325 330 335

Asn Gly Ile Ile Ser Ile Leu Phe Ala Leu Gly Leu Tyr Asn Gly Thr
 340 345 350

Lys Pro Leu Ser Thr Thr Thr Val Glu Asn Ile Thr Gln Thr Asp Gly
 355 360 365

Phe Ser Ser Ala Trp Thr Val Pro Phe Ala Ser Arg Leu Tyr Val Glu
 370 375 380

Met Met Gln Cys Gln Ala Glu Gln Glu Pro Leu Val Arg Val Leu Val
 385 390 395 400

Asn Asp Arg Val Val Pro Leu His Gly Cys Pro Ile Asp Ala Leu Gly
 405 410 415

Arg Cys Thr Arg Asp Ser Phe Val Arg Gly Leu Ser Phe Ala Arg Ser
 420 425 430

Gly Gly Asp Trp Ala Glu Cys Ser Ala
 435 440

<210> 4

<211> 441

<212> PRT

<213> 黑曲霉 T213

<400> 4

Asn Gln Ser Ser Cys Asp Thr Val Asp Gln Gly Tyr Gln Cys Phe Ser
 1 5 10 15

Glu Thr Ser His Leu Trp Gly Gln Tyr Ala Pro Phe Phe Ser Leu Ala
 20 25 30

Asn Glu Ser Val Ile Ser Pro Asp Val Pro Ala Gly Cys Arg Val Thr
 35 40 45

Phe Ala Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Glu Ser
 50 55 60

Lys Gly Lys Lys Tyr Ser Ala Leu Ile Glu Glu Ile Gln Gln Asn Val
 65 70 75 80

Thr Thr Phe Asp Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Ser
 85 90 95

Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Glu Leu Val Asn
 100 105 110

Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Glu Ser Leu Thr Arg Asn Ile
 115 120 125

```

Ile Pro Phe Ile Arg Ser Ser Gly Ser Ser Arg Val Ile Ala Ser Gly
 130                      135                      140

Glu Lys Phe Ile Glu Gly Phe Gln Ser Thr Lys Leu Lys Asp Pro Arg
145                      150                      155                      160

Ala Gln Pro Gly Gln Ser Ser Pro Lys Ile Asp Val Val Ile Ser Glu
                      165                      170                      175

Ala Ser Ser Ser Asn Asn Thr Leu Asp Pro Gly Thr Cys Thr Val Phe
                      180                      185                      190

Glu Asp Ser Glu Leu Ala Asp Thr Val Glu Ala Asn Phe Thr Ala Thr
 195                      200                      205

Phe Ala Pro Ser Ile Arg Gln Arg Leu Glu Asn Asp Leu Ser Gly Val
 210                      215                      220

Thr Leu Thr Asp Thr Glu Val Thr Tyr Leu Met Asp Met Cys Ser Phe
225                      230                      235                      240

Asp Thr Ile Ser Thr Ser Thr Val Asp Thr Lys Leu Ser Pro Phe Cys
                      245                      250                      255

Asp Leu Phe Thr His Asp Glu Trp Ile His Tyr Asp Tyr Leu Arg Ser
 260                      265                      270

Leu Lys Lys Tyr Tyr Gly His Gly Ala Gly Asn Pro Leu Gly Pro Thr
 275                      280                      285

Gln Gly Val Gly Tyr Ala Asn Glu Leu Ile Ala Arg Leu Thr His Ser
 290                      295                      300

Pro Val His Asp Asp Thr Ser Ser Asn His Thr Leu Asp Ser Asn Pro
305                      310                      315                      320

Ala Thr Phe Pro Leu Asn Ser Thr Leu Tyr Ala Asp Phe Ser His Asp
                      325                      330                      335

Asn Gly Ile Ile Ser Ile Leu Phe Ala Leu Gly Leu Tyr Asn Gly Thr
 340                      345                      350

Lys Pro Leu Ser Thr Thr Thr Val Glu Asn Ile Thr Gln Thr Asp Gly
 355                      360                      365

Phe Ser Ser Ala Trp Thr Val Pro Phe Ala Ser Arg Leu Tyr Val Glu
 370                      375                      380

Met Met Gln Cys Gln Ala Glu Gln Glu Pro Leu Val Arg Val Leu Val
385                      390                      395                      400

Asn Asp Arg Val Val Pro Leu His Gly Cys Pro Ile Asp Ala Leu Gly
 405                      410                      415

```

Arg Cys Thr Arg Asp Ser Phe Val Arg Gly Leu Ser Phe Ala Arg Ser
420 425 430

Gly Gly Asp Trp Ala Glu Cys Phe Ala
435 440

<210> 5

<211> 441

<212> PRT

<213> 黑曲霉 NRRL3135

<400> 5

Asn Gln Ser Ser Cys Asp Thr Val Asp Gln Gly Tyr Gln Cys Phe Ser
1 5 10 15

Glu Thr Ser His Leu Trp Gly Gln Tyr Ala Pro Phe Phe Ser Leu Ala
20 25 30

Asn Glu Ser Val Ile Ser Pro Glu Val Pro Ala Gly Cys Arg Val Thr
35 40 45

Phe Ala Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Asp Ser
50 55 60

Lys Gly Lys Lys Tyr Ser Ala Leu Ile Glu Glu Ile Gln Gln Asn Ala
65 70 75 80

Thr Thr Phe Asp Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Ser
85 90 95

Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Glu Leu Val Asn
100 105 110

Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Glu Ser Leu Thr Arg Asn Ile
115 120 125

Val Pro Phe Ile Arg Ser Ser Gly Ser Ser Arg Val Ile Ala Ser Gly
130 135 140

Lys Lys Phe Ile Glu Gly Phe Gln Ser Thr Lys Leu Lys Asp Pro Arg
145 150 155 160

Ala Gln Pro Gly Gln Ser Ser Pro Lys Ile Asp Val Val Ile Ser Glu
165 170 175

Ala Ser Ser Ser Asn Asn Thr Leu Asp Pro Gly Thr Cys Thr Val Phe
180 185 190

Glu Asp Ser Glu Leu Ala Asp Thr Val Glu Ala Asn Phe Thr Ala Thr
195 200 205

```

Phe Val Pro Ser Ile Arg Gln Arg Leu Glu Asn Asp Leu Ser Gly Val
 210                215                220

Thr Leu Thr Asp Thr Glu Val Thr Tyr Leu Met Asp Met Cys Ser Phe
225                230                235                240

Asp Thr Ile Ser Thr Ser Thr Val Asp Thr Lys Leu Ser Pro Phe Cys
                245                250                255

Asp Leu Phe Thr His Asp Glu Trp Ile Asn Tyr Asp Tyr Leu Gln Ser
                260                265                270

Leu Lys Lys Tyr Tyr Gly His Gly Ala Gly Asn Pro Leu Gly Pro Thr
 275                280                285

Gln Gly Val Gly Tyr Ala Asn Glu Leu Ile Ala Arg Leu Thr His Ser
 290                295                300

Pro Val His Asp Asp Thr Ser Ser Asn His Thr Leu Asp Ser Ser Pro
305                310                315                320

Ala Thr Phe Pro Leu Asn Ser Thr Leu Tyr Ala Asp Phe Ser His Asp
                325                330                335

Asn Gly Ile Ile Ser Ile Leu Phe Ala Leu Gly Leu Tyr Asn Gly Thr
                340                345                350

Lys Pro Leu Ser Thr Thr Thr Val Glu Asn Ile Thr Gln Thr Asp Gly
 355                360                365

Phe Ser Ser Ala Trp Thr Val Pro Phe Ala Ser Arg Leu Tyr Val Glu
 370                375                380

Met Met Gln Cys Gln Ala Glu Gln Glu Pro Leu Val Arg Val Leu Val
385                390                395                400

Asn Asp Arg Val Val Pro Leu His Gly Cys Pro Val Asp Ala Leu Gly
                405                410                415

Arg Cys Thr Arg Asp Ser Phe Val Arg Gly Leu Ser Phe Ala Arg Ser
                420                425                430

Gly Gly Asp Trp Ala Glu Cys Phe Ala
 435                440

```

<210> 6

<211> 440

<212> PRT

<213> 烟曲霉 13073

<400> 6

```

Gly Ser Lys Ser Cys Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro
 1                5                10                15

```

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

Val Gln Asp His Thr Ser Thr Asn Ser Thr Leu Val Ser Asn Pro Ala
 305 310 315 320
 Thr Phe Pro Leu Asn Ala Thr Met Tyr Val Asp Phe Ser His Asp Asn
 325 330 335
 Ser Met Val Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Glu
 340 345 350
 Pro Leu Ser Arg Thr Ser Val Glu Ser Ala Lys Glu Leu Asp Gly Tyr
 355 360 365
 Ser Ala Ser Trp Val Val Pro Phe Gly Ala Arg Ala Tyr Phe Glu Thr
 370 375 380
 Met Gln Cys Lys Ser Glu Lys Glu Pro Leu Val Arg Ala Leu Ile Asn
 385 390 395 400
 Asp Arg Val Val Pro Leu His Gly Cys Asp Val Asp Lys Leu Gly Arg
 405 410 415
 Cys Lys Leu Asn Asp Phe Val Lys Gly Leu Ser Trp Ala Arg Ser Gly
 420 425 430
 Gly Asn Trp Gly Glu Cys Phe Ser
 435 440

<210> 7
 <211> 440
 <212> PRT
 <213> 烟曲霉 32722

<400> 7
 Gly Ser Lys Ser Cys Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro
 1 5 10 15
 Ala Thr Ser His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Glu
 20 25 30
 Asp Glu Leu Ser Val Ser Ser Lys Leu Pro Lys Asp Cys Arg Ile Thr
 35 40 45
 Leu Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser
 50 55 60
 Lys Ser Lys Lys Tyr Lys Lys Leu Val Thr Ala Ile Gln Ala Asn Ala
 65 70 75 80
 Thr Asp Phe Lys Gly Lys Phe Ala Phe Leu Lys Thr Tyr Asn Tyr Thr
 85 90 95

Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Gln Leu Val Asn
 100 105 110
 Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Lys Ala Leu Ala Arg Ser Val
 115 120 125
 Val Pro Phe Ile Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Gly
 130 135 140
 Glu Lys Phe Ile Glu Gly Phe Gln Gln Ala Lys Leu Ala Asp Pro Gly
 145 150 155 160
 Ala Thr Asn Arg Ala Ala Pro Ala Ile Ser Val Ile Ile Pro Glu Ser
 165 170 175
 Glu Thr Phe Asn Asn Thr Leu Asp His Gly Val Cys Thr Lys Phe Glu
 180 185 190
 Ala Ser Gln Leu Gly Asp Glu Val Ala Ala Asn Phe Thr Ala Leu Phe
 195 200 205
 Ala Pro Asp Ile Arg Ala Arg Ala Glu Lys His Leu Pro Gly Val Thr
 210 215 220
 Leu Thr Asp Glu Asp Val Val Ser Leu Met Asp Met Cys Ser Phe Asp
 225 230 235 240
 Thr Val Ala Arg Thr Ser Asp Ala Ser Gln Leu Ser Pro Phe Cys Gln
 245 250 255
 Leu Phe Thr His Asn Glu Trp Lys Lys Tyr Asn Tyr Leu Gln Ser Leu
 260 265 270
 Gly Lys Tyr Tyr Gly Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln
 275 280 285
 Gly Ile Gly Phe Thr Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro
 290 295 300
 Val Gln Asp His Thr Ser Thr Asn Ser Thr Leu Val Ser Asn Pro Ala
 305 310 315 320
 Thr Phe Pro Leu Asn Ala Thr Met Tyr Val Asp Phe Ser His Asp Asn
 325 330 335
 Ser Met Val Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Gly
 340 345 350
 Pro Leu Ser Arg Thr Ser Val Glu Ser Ala Lys Glu Leu Asp Gly Tyr
 355 360 365
 Ser Ala Ser Trp Val Val Pro Phe Gly Ala Arg Ala Tyr Phe Glu Thr
 370 375 380

Met Gln Cys Lys Ser Glu Lys Glu Pro Leu Val Arg Ala Leu Ile Asn
385 390 395 400

Asp Arg Val Val Pro Leu His Gly Cys Asp Val Asp Lys Leu Gly Arg
405 410 415

Cys Lys Leu Asn Asp Phe Val Lys Gly Leu Ser Trp Ala Arg Ser Gly
420 425 430

Gly Asn Trp Gly Glu Cys Phe Ser
435 440

<210> 8

<211> 440

<212> PRT

<213> 烟曲霉 58128

<400> 8

Gly Ser Lys Ser Cys Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro
1 5 10 15

Ala Thr Ser His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Glu
20 25 30

Asp Glu Leu Ser Val Ser Ser Lys Leu Pro Lys Asp Cys Arg Ile Thr
35 40 45

Leu Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser
50 55 60

Lys Ser Lys Lys Tyr Lys Lys Leu Val Thr Ala Ile Gln Ala Asn Ala
65 70 75 80

Thr Asp Phe Lys Gly Lys Phe Ala Phe Leu Lys Thr Tyr Asn Tyr Thr
85 90 95

Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Gln Leu Val Asn
100 105 110

Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Lys Ala Leu Ala Arg Ser Val
115 120 125

Val Pro Phe Ile Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Gly
130 135 140

Glu Lys Phe Ile Glu Gly Phe Gln Gln Ala Lys Leu Ala Asp Pro Gly
145 150 155 160

Ala Thr Asn Arg Ala Ala Pro Ala Ile Ser Val Ile Ile Pro Glu Ser
165 170 175

Glu Thr Phe Asn Asn Thr Leu Asp His Gly Val Cys Thr Lys Phe Glu
180 185 190

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

<210> 9

<211> 440

<212> PRT

<213> 烟曲霉 26906

<400> 9

```

Gly Ser Lys Ser Cys Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro
 1              5              10              15

Ala Thr Ser His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Glu
      20              25              30

Asp Glu Leu Ser Val Ser Ser Lys Leu Pro Lys Asp Cys Arg Ile Thr
      35              40              45

Leu Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser
      50              55              60

Lys Ser Lys Lys Tyr Lys Lys Leu Val Thr Ala Ile Gln Ala Asn Ala
      65              70              75              80

Thr Asp Phe Lys Gly Lys Phe Ala Phe Leu Lys Thr Tyr Asn Tyr Thr
      85              90              95

Leu Gly Ala Asp Asp Leu Thr Ala Phe Gly Glu Gln Gln Leu Val Asn
      100              105              110

Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Lys Ala Leu Ala Arg Ser Val
      115              120              125

Val Pro Phe Ile Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Gly
      130              135              140

Glu Lys Phe Ile Glu Gly Phe Gln Gln Ala Lys Leu Ala Asp Pro Gly
      145              150              155              160

Ala Thr Asn Arg Ala Ala Pro Ala Ile Ser Val Ile Ile Pro Glu Ser
      165              170              175

Glu Thr Phe Asn Asn Thr Leu Asp His Gly Val Cys Thr Lys Phe Glu
      180              185              190

Ala Ser Gln Leu Gly Asp Glu Val Ala Ala Asn Phe Thr Ala Leu Phe
      195              200              205

Ala Pro Asp Ile Arg Ala Arg Ala Lys Lys His Leu Pro Gly Val Thr
      210              215              220

Leu Thr Asp Glu Asp Val Val Ser Leu Met Asp Met Cys Ser Phe Asp
      225              230              235              240

Thr Val Ala Arg Thr Ser Asp Ala Ser Gln Leu Ser Pro Phe Cys Gln
      245              250              255

Leu Phe Thr His Asn Glu Trp Lys Lys Tyr Asn Tyr Leu Gln Ser Leu
      260              265              270

```

Gly Lys Tyr Tyr Gly Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln
 275 280 285
 Gly Ile Gly Phe Thr Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro
 290 295 300
 Val Gln Asp His Thr Ser Thr Asn Ser Thr Leu Val Ser Asn Pro Ala
 305 310 315 320
 Thr Phe Pro Leu Asn Ala Thr Met Tyr Val Asp Phe Ser His Asp Asn
 325 330 335
 Ser Met Val Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Glu
 340 345 350
 Pro Leu Ser Arg Thr Ser Val Glu Ser Ala Lys Glu Leu Asp Gly Tyr
 355 360 365
 Ser Ala Ser Trp Val Val Pro Phe Gly Ala Arg Ala Tyr Phe Glu Thr
 370 375 380
 Met Gln Cys Lys Ser Glu Lys Glu Pro Leu Val Arg Ala Leu Ile Asn
 385 390 395 400
 Asp Arg Val Val Pro Leu His Gly Cys Asp Val Asp Lys Leu Gly Arg
 405 410 415
 Cys Lys Leu Asn Asp Phe Val Lys Gly Leu Ser Trp Ala Arg Ser Gly
 420 425 430
 Gly Asn Trp Gly Glu Cys Phe Ser
 435 440

 <210> 10
 <211> 440
 <212> PRT
 <213> 烟曲霉 32239

 <400> 10
 Gly Ser Lys Ala Cys Asp Thr Val Glu Leu Gly Tyr Gln Cys Ser Pro
 1 5 10 15
 Gly Thr Ser His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Glu
 20 25 30
 Asp Glu Leu Ser Val Ser Ser Asp Leu Pro Lys Asp Cys Arg Val Thr
 35 40 45
 Phe Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ala Ser
 50 55 60
 Lys Ser Lys Lys Tyr Lys Lys Leu Val Thr Ala Ile Gln Lys Asn Ala
 65 70 75 80

Thr Glu Phe Lys Gly Lys Phe Ala Phe Leu Glu Thr Tyr Asn Tyr Thr
 85 90 95
 Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Gln Met Val Asn
 100 105 110
 Ser Gly Ile Lys Phe Tyr Gln Lys Tyr Lys Ala Leu Ala Gly Ser Val
 115 120 125
 Val Pro Phe Ile Arg Ser Ser Gly Ser Asp Arg Val Ile Ala Ser Gly
 130 135 140
 Glu Lys Phe Ile Glu Gly Phe Gln Gln Ala Asn Val Ala Asp Pro Gly
 145 150 155 160
 Ala Thr Asn Arg Ala Ala Pro Val Ile Ser Val Ile Ile Pro Glu Ser
 165 170 175
 Glu Thr Tyr Asn Asn Thr Leu Asp His Ser Val Cys Thr Asn Phe Glu
 180 185 190
 Ala Ser Glu Leu Gly Asp Glu Val Glu Ala Asn Phe Thr Ala Leu Phe
 195 200 205
 Ala Pro Ala Ile Arg Ala Arg Ile Glu Lys His Leu Pro Gly Val Gln
 210 215 220
 Leu Thr Asp Asp Asp Val Val Ser Leu Met Asp Met Cys Ser Phe Asp
 225 230 235 240
 Thr Val Ala Arg Thr Ala Asp Ala Ser Glu Leu Ser Pro Phe Cys Ala
 245 250 255
 Ile Phe Thr His Asn Glu Trp Lys Lys Tyr Asp Tyr Leu Gln Ser Leu
 260 265 270
 Gly Lys Tyr Tyr Gly Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln
 275 280 285
 Gly Ile Gly Phe Thr Asn Glu Leu Ile Ala Arg Leu Thr Asn Ser Pro
 290 295 300
 Val Gln Asp His Thr Ser Thr Asn Ser Thr Leu Asp Ser Asp Pro Ala
 305 310 315 320
 Thr Phe Pro Leu Asn Ala Thr Ile Tyr Val Asp Phe Ser His Asp Asn
 325 330 335
 Gly Met Ile Pro Ile Phe Phe Ala Met Gly Leu Tyr Asn Gly Thr Glu
 340 345 350
 Pro Leu Ser Gln Thr Ser Glu Glu Ser Thr Lys Glu Ser Asn Gly Tyr
 355 360 365

Ser Ala Ser Trp Ala Val Pro Phe Gly Ala Arg Ala Tyr Phe Glu Thr
370 375 380

Met Gln Cys Lys Ser Glu Lys Glu Pro Leu Val Arg Ala Leu Ile Asn
385 390 395 400

Asp Arg Val Val Pro Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg
405 410 415

Cys Lys Leu Lys Asp Phe Val Lys Gly Leu Ser Trp Ala Arg Ser Gly
420 425 430

Gly Asn Ser Glu Gln Ser Phe Ser
435 440

<210> 11

<211> 439

<212> PRT

<213> Emericella nidulans

<400> 11

Gln Asn His Ser Cys Asn Thr Ala Asp Gly Gly Tyr Gln Cys Phe Pro
1 5 10 15

Asn Val Ser His Val Trp Gly Gln Tyr Ser Pro Tyr Phe Ser Ile Glu
20 25 30

Gln Glu Ser Ala Ile Ser Glu Asp Val Pro His Gly Cys Glu Val Thr
35 40 45

Phe Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Glu Ser
50 55 60

Lys Ser Lys Ala Tyr Ser Gly Leu Ile Glu Ala Ile Gln Lys Asn Ala
65 70 75 80

Thr Ser Phe Trp Gly Gln Tyr Ala Phe Leu Glu Ser Tyr Asn Tyr Thr
85 90 95

Leu Gly Ala Asp Asp Leu Thr Ile Phe Gly Glu Asn Gln Met Val Asp
100 105 110

Ser Gly Ala Lys Phe Tyr Arg Arg Tyr Lys Asn Leu Ala Arg Lys Asn
115 120 125

Thr Pro Phe Ile Arg Ala Ser Gly Ser Asp Arg Val Val Ala Ser Ala
130 135 140

Glu Lys Phe Ile Asn Gly Phe Arg Lys Ala Gln Leu His Asp His Gly
145 150 155 160

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

<210> 12

<211> 443

<212> PRT

<213> 嗜热踝节菌

<400> 12

Asp Ser His Ser Cys Asn Thr Val Glu Gly Gly Tyr Gln Cys Arg Pro
 1 5 10 15

Glu Ile Ser His Ser Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Ala
 20 25 30

Asp Gln Ser Glu Ile Ser Pro Asp Val Pro Gln Asn Cys Lys Ile Thr
 35 40 45

Phe Val Gln Leu Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser
 50 55 60

Lys Thr Glu Leu Tyr Ser Gln Leu Ile Ser Arg Ile Gln Lys Thr Ala
 65 70 75 80

Thr Ala Tyr Lys Gly Tyr Tyr Ala Phe Leu Lys Asp Tyr Arg Tyr Gln
 85 90 95

Leu Gly Ala Asn Asp Leu Thr Pro Phe Gly Glu Asn Gln Met Ile Gln
 100 105 110

Leu Gly Ile Lys Phe Tyr Asn His Tyr Lys Ser Leu Ala Arg Asn Ala
 115 120 125

Val Pro Phe Val Arg Cys Ser Gly Ser Asp Arg Val Ile Ala Ser Gly
 130 135 140

Arg Leu Phe Ile Glu Gly Phe Gln Ser Ala Lys Val Leu Asp Pro His
 145 150 155 160

Ser Asp Lys His Asp Ala Pro Pro Thr Ile Asn Val Ile Ile Glu Glu
 165 170 175

Gly Pro Ser Tyr Asn Asn Thr Leu Asp Thr Gly Ser Cys Pro Val Phe
 180 185 190

Glu Asp Ser Ser Gly Gly His Asp Ala Gln Glu Lys Phe Ala Lys Gln
 195 200 205

Phe Ala Pro Ala Ile Leu Glu Lys Ile Lys Asp His Leu Pro Gly Val
 210 215 220

Asp Leu Ala Val Ser Asp Val Pro Tyr Leu Met Asp Leu Cys Pro Phe
 225 230 235 240

Glu Thr Leu Ala Arg Asn His Thr Asp Thr Leu Ser Pro Phe Cys Ala
 245 250 255

Leu Ser Thr Gln Glu Glu Trp Gln Ala Tyr Asp Tyr Tyr Gln Ser Leu
260 265 270

Gly Lys Tyr Tyr Gly Asn Gly Gly Gly Asn Pro Leu Gly Pro Ala Gln
275 280 285

Gly Val Gly Phe Val Asn Glu Leu Ile Ala Arg Met Thr His Ser Pro
290 295 300

Val Gln Asp Tyr Thr Thr Val Asn His Thr Leu Asp Ser Asn Pro Ala
305 310 315 320

Thr Phe Pro Leu Asn Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn
325 330 335

Thr Met Thr Ser Ile Phe Ala Ala Leu Gly Leu Tyr Asn Gly Thr Ala
340 345 350

Lys Leu Ser Thr Thr Glu Ile Lys Ser Ile Glu Glu Thr Asp Gly Tyr
355 360 365

Ser Ala Ala Trp Thr Val Pro Phe Gly Gly Arg Ala Tyr Ile Glu Met
370 375 380

Met Gln Cys Asp Asp Ser Asp Glu Pro Val Val Arg Val Leu Val Asn
385 390 395 400

Asp Arg Val Val Pro Leu His Gly Cys Glu Val Asp Ser Leu Gly Arg
405 410 415

Cys Lys Arg Asp Asp Phe Val Arg Gly Leu Ser Phe Ala Arg Gln Gly
420 425 430

Gly Asn Trp Glu Gly Cys Tyr Ala Ala Ser Glu
435 440

<210> 13

<211> 466

<212> PRT

<213> 嗜热毁丝霉

<400> 13

Glu Ser Arg Pro Cys Asp Thr Pro Asp Leu Gly Phe Gln Cys Gly Thr
1 5 10 15

Ala Ile Ser His Phe Trp Gly Gln Tyr Ser Pro Tyr Phe Ser Val Pro
20 25 30

Ser Glu Leu Asp Ala Ser Ile Pro Asp Asp Cys Glu Val Thr Phe Ala
35 40 45

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

Leu Gly Arg Pro Leu Tyr Ala Asp Phe Ser His Asp Asn Asp Met Met
 340 345 350
 Gly Val Leu Gly Ala Leu Gly Ala Tyr Asp Gly Val Pro Pro Leu Asp
 355 360 365
 Lys Thr Ala Arg Arg Asp Pro Glu Glu Leu Gly Gly Tyr Ala Ala Ser
 370 375 380
 Trp Ala Val Pro Phe Ala Ala Arg Ile Tyr Val Glu Lys Met Arg Cys
 385 390 395 400
 Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Glu Gly Arg Gln Glu Lys
 405 410 415
 Asp Glu Glu Met Val Arg Val Leu Val Asn Asp Arg Val Met Thr Leu
 420 425 430
 Lys Gly Cys Gly Ala Asp Glu Arg Gly Met Cys Thr Leu Glu Arg Phe
 435 440 445
 Ile Glu Ser Met Ala Phe Ala Arg Gly Asn Gly Lys Trp Asp Leu Cys
 450 455 460
 Phe Ala
 465

<210> 14
 <211> 441
 <212> PRT
 <213> 人工序列

<220>
 <223> 人工序列的描述：共有植酸酶

<400> 14
 Asn Ser His Ser Cys Asp Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro
 1 5 10 15
 Glu Ile Ser His Leu Trp Gly Gln Tyr Ser Pro Tyr Phe Ser Leu Glu
 20 25 30
 Asp Glu Ser Ala Ile Ser Pro Asp Val Pro Asp Asp Cys Arg Val Thr
 35 40 45
 Phe Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser
 50 55 60
 Lys Ser Lys Ala Tyr Ser Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala
 65 70 75 80

Thr Ala Phe Lys Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr
 85 90 95
 Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Asn Gln Met Val Asn
 100 105 110
 Ser Gly Ile Lys Phe Tyr Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile
 115 120 125
 Val Pro Phe Ile Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Ala
 130 135 140
 Glu Lys Phe Ile Glu Gly Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly
 145 150 155 160
 Ser Gln Pro His Gln Ala Ser Pro Val Ile Asp Val Ile Ile Pro Glu
 165 170 175
 Gly Ser Gly Tyr Asn Asn Thr Leu Asp His Gly Thr Cys Thr Ala Phe
 180 185 190
 Glu Asp Ser Glu Leu Gly Asp Asp Val Glu Ala Asn Phe Thr Ala Leu
 195 200 205
 Phe Ala Pro Ala Ile Arg Ala Arg Leu Glu Ala Asp Leu Pro Gly Val
 210 215 220
 Thr Leu Thr Asp Glu Asp Val Val Tyr Leu Met Asp Met Cys Pro Phe
 225 230 235 240
 Glu Thr Val Ala Arg Thr Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys
 245 250 255
 Ala Leu Phe Thr His Asp Glu Trp Arg Gln Tyr Asp Tyr Leu Gln Ser
 260 265 270
 Leu Gly Lys Tyr Tyr Gly Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala
 275 280 285
 Gln Gly Val Gly Phe Ala Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser
 290 295 300
 Pro Val Gln Asp His Thr Ser Thr Asn His Thr Leu Asp Ser Asn Pro
 305 310 315 320
 Ala Thr Phe Pro Leu Asn Ala Thr Leu Tyr Ala Asp Phe Ser His Asp
 325 330 335
 Asn Ser Met Ile Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr
 340 345 350
 Ala Pro Leu Ser Thr Thr Ser Val Glu Ser Ile Glu Glu Thr Asp Gly
 355 360 365

Tyr Ser Ala Ser Trp Thr Val Pro Phe Gly Ala Arg Ala Tyr Val Glu
370 375 380

Met Met Gln Cys Gln Ala Glu Lys Glu Pro Leu Val Arg Val Leu Val
385 390 395 400

Asn Asp Arg Val Val Pro Leu His Gly Cys Ala Val Asp Lys Leu Gly
405 410 415

Arg Cys Lys Arg Asp Asp Phe Val Glu Gly Leu Ser Phe Ala Arg Ser
420 425 430

Gly Gly Asn Trp Ala Glu Cys Phe Ala
435 440

<210> 15

<211> 1426

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 共有植酸酶

<220>

<221> CDS

<222> (12)..(1412)

<220>

<221> 单肽

<222> (12)..(89)

<220>

<221> 成熟肽

<222> (90)..(1412)

<400> 15

tatatgaatt c atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg 50
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu
-25 -20 -15

ttc ggt tcc aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac 98
Phe Gly Ser Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His
-10 -5 -1 1

tct tgt gac act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct 146
Ser Cys Asp Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser
5 10 15

cac ttg tgg ggt caa tac tct cca tac ttc tct ttg gaa gac gaa tct 194
His Leu Trp Gly Gln Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Ser
20 25 30 35

gct att tct cca gac gtt cca gac gac tgt aga gtt act ttc gtt caa	242
Ala Ile Ser Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln	
40 45 50	
ggt ttg tct aga cac ggt gct aga tac cca act tct tct aag tct aag	290
Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys	
55 60 65	
gct tac tct gct ttg att gaa gct att caa aag aac gct act gct ttc	338
Ala Tyr Ser Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe	
70 75 80	
aag ggt aag tac gct ttc ttg aag act tac aac tac act ttg ggt gct	386
Lys Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala	
85 90 95	
gac gac ttg act cca ttc ggt gaa aac caa atg gtt aac tct ggt att	434
Asp Asp Leu Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile	
100 105 110 115	
aag ttc tac aga aga tac aag gct ttg gct aga aag att gtt cca ttc	482
Lys Phe Tyr Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe	
120 125 130	
att aga gct tct ggt tct gac aga gtt att gct tct gct gaa aag ttc	530
Ile Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe	
135 140 145	
att gaa ggt ttc caa tct gct aag ttg gct gac cca ggt tct caa cca	578
Ile Glu Gly Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro	
150 155 160	
cac caa gct tct cca gtt att gac gtt att att cca gaa gga tcc ggt	626
His Gln Ala Ser Pro Val Ile Asp Val Ile Ile Pro Glu Gly Ser Gly	
165 170 175	
tac aac aac act ttg gac cac ggt act tgt act gct ttc gaa gac tct	674
Tyr Asn Asn Thr Leu Asp His Gly Thr Cys Thr Ala Phe Glu Asp Ser	
180 185 190 195	
gaa ttg ggt gac gac gtt gaa gct aac ttc act gct ttg ttc gct cca	722
Glu Leu Gly Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro	
200 205 210	
gct att aga gct aga ttg gaa gct gac ttg cca ggt gtt act ttg act	770
Ala Ile Arg Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr	
215 220 225	
gac gaa gac gtt gtt tac ttg atg gac atg tgt cca ttc gaa act gtt	818
Asp Glu Asp Val Val Tyr Leu Met Asp Met Cys Pro Phe Glu Thr Val	
230 235 240	
gct aga act tct gac gct act gaa ttg tct cca ttc tgt gct ttg ttc	866
Ala Arg Thr Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe	

245	250	255	
act cac gac gaa tgg aga caa tac gac tac ttg caa tct ttg ggt aag			914
Thr His Asp Glu Trp Arg Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys			
260	265	270	275
tac tac ggt tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt			962
Tyr Tyr Gly Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val			
	280	285	290
ggt ttc gct aac gaa ttg att gct aga ttg act aga tct cca gtt caa			1010
Gly Phe Ala Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln			
	295	300	305
gac cac act tct act aac cac act ttg gac tct aac cca gct act ttc			1058
Asp His Thr Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe			
	310	315	320
cca ttg aac gct act ttg tac gct gac ttc tct cac gac aac tct atg			1106
Pro Leu Asn Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Ser Met			
	325	330	335
att tct att ttc ttc gct ttg ggt ttg tac aac ggt act gct cca ttg			1154
Ile Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Ala Pro Leu			
	340	345	350
tct act act tct gtt gaa tct att gaa gaa act gac ggt tac tct gct			1202
Ser Thr Thr Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala			
	360	365	370
tct tgg act gtt cca ttc ggt gct aga gct tac gtt gaa atg atg caa			1250
Ser Trp Thr Val Pro Phe Gly Ala Arg Ala Tyr Val Glu Met Met Gln			
	375	380	385
tgt caa gct gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga			1298
Cys Gln Ala Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg			
	390	395	400
gtt gtt cca ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag			1346
Val Val Pro Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys			
	405	410	415
aga gac gac ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac			1394
Arg Asp Asp Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn			
	420	425	430
435			
tgg gct gaa tgt ttc gct taagaattca tata			1426
Trp Ala Glu Cys Phe Ala			
	440		

<210> 16

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶

<400> 16

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
  -25                -20                -15

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
-10                -5                -1    1                5

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
      10                15                20

Gly Gln Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Ser Ala Ile Ser
      25                30                35

Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
      40                45                50

Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Ala Tyr Ser
      55                60                65                70

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
      75                80                85

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
      90                95                100

Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
      105                110                115

Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
      120                125                130

Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
      135                140                145                150

Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
      155                160                165

Ser Pro Val Ile Asp Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn
      170                175                180

Thr Leu Asp His Gly Thr Cys Thr Ala Phe Glu Asp Ser Glu Leu Gly
      185                190                195

Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
      200                205                210

Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
      215                220                225                230

```

```
<210> 17
<211> 422
<212> PRT
<213> Paxillus involutus phyA1
```

<400> 17
Ser Val Pro Lys Asn Thr Ala Pro Thr Phe Pro Ile Pro Glu Ser Glu
1 5 10 15

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

Pro Leu Asn Lys Thr Phe Tyr Ala Asp Phe Ser His Asp Asn Leu Met
 305 310 315 320
 Val Ala Val Phe Ser Ala Met Gly Leu Phe Arg Gln Pro Ala Pro Leu
 325 330 335
 Ser Thr Ser Val Pro Asn Pro Trp Arg Thr Trp Arg Thr Ser Ser Leu
 340 345 350
 Val Pro Phe Ser Gly Arg Met Val Val Glu Arg Leu Ser Cys Phe Gly
 355 360 365
 Thr Thr Lys Val Arg Val Leu Val Gln Asp Gln Val Gln Pro Leu Glu
 370 375 380
 Phe Cys Gly Gly Asp Arg Asn Gly Leu Cys Thr Leu Ala Lys Phe Val
 385 390 395 400
 Glu Ser Gln Thr Phe Ala Arg Ser Asp Gly Ala Gly Asp Phe Glu Lys
 405 410 415
 Cys Phe Ala Thr Ser Ala
 420

<210> 18
 <211> 422
 <212> PRT
 <213> Paxillus involutus phyA2

<400> 18
 Ser Val Pro Arg Asn Ile Ala Pro Lys Phe Ser Ile Pro Glu Ser Glu
 1 5 10 15
 Gln Arg Asn Trp Ser Pro Tyr Ser Pro Tyr Phe Pro Leu Ala Glu Tyr
 20 25 30
 Lys Ala Pro Pro Ala Gly Cys Glu Ile Asn Gln Val Asn Ile Ile Gln
 35 40 45
 Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Ala Thr Arg Ile Lys
 50 55 60
 Ala Gly Leu Ser Lys Leu Gln Ser Val Gln Asn Phe Thr Asp Pro Lys
 65 70 75 80
 Phe Asp Phe Ile Lys Ser Phe Thr Tyr Asp Leu Gly Thr Ser Asp Leu
 85 90 95
 Val Pro Phe Gly Ala Ala Gln Ser Phe Asp Ala Gly Leu Glu Val Phe
 100 105 110
 Ala Arg Tyr Ser Lys Leu Val Ser Ser Asp Asn Leu Pro Phe Ile Arg
 115 120 125

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

Cys Leu Ala Thr Thr Val
420

<210> 19

<211> 420

<212> PRT

<213> 绒毛栓菌

<400> 19

His Ile Pro Leu Arg Asp Thr Ser Ala Cys Leu Asp Val Thr Arg Asp
1 5 10 15

Val Gln Gln Ser Trp Ser Met Tyr Ser Pro Tyr Phe Pro Ala Ala Thr
20 25 30

Tyr Val Ala Pro Pro Ala Ser Cys Gln Ile Asn Gln Val His Ile Ile
35 40 45

Gln Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Ala Lys Arg Ile
50 55 60

Gln Thr Ala Val Ala Lys Leu Lys Ala Ala Ser Asn Tyr Thr Asp Pro
65 70 75 80

Leu Leu Ala Phe Val Thr Asn Tyr Thr Tyr Ser Leu Gly Gln Asp Ser
85 90 95

Leu Val Glu Leu Gly Ala Thr Gln Ser Ser Glu Ala Gly Gln Glu Ala
100 105 110

Phe Thr Arg Tyr Ser Ser Leu Val Ser Ala Asp Glu Leu Pro Phe Val
115 120 125

Arg Ala Ser Gly Ser Asp Arg Val Val Ala Thr Ala Asn Asn Trp Thr
130 135 140

Ala Gly Phe Ala Leu Ala Ser Ser Asn Ser Ile Thr Pro Val Leu Ser
145 150 155 160

Val Ile Ile Ser Glu Ala Gly Asn Asp Thr Leu Asp Asp Asn Met Cys
165 170 175

Pro Ala Ala Gly Asp Ser Asp Pro Gln Val Asn Gln Trp Leu Ala Gln
180 185 190

Phe Ala Pro Pro Met Thr Ala Arg Leu Asn Ala Gly Ala Pro Gly Ala
195 200 205

Asn Leu Thr Asp Thr Asp Thr Tyr Asn Leu Leu Thr Leu Cys Pro Phe
210 215 220

Glu Thr Val Ala Thr Glu Arg Arg Ser Glu Phe Cys Asp Ile Tyr Glu
 225 230 235 240
 Glu Leu Gln Ala Glu Asp Ala Phe Ala Tyr Asn Ala Asp Leu Asp Lys
 245 250 255
 Phe Tyr Gly Thr Gly Tyr Gly Gln Pro Leu Gly Pro Val Gln Gly Val
 260 265 270
 Gly Tyr Ile Asn Glu Leu Ile Ala Arg Leu Thr Ala Gln Asn Val Ser
 275 280 285
 Asp His Thr Gln Thr Asn Ser Thr Leu Asp Ser Ser Pro Glu Thr Phe
 290 295 300
 Pro Leu Asn Arg Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Gln Met
 305 310 315 320
 Val Ala Ile Phe Ser Ala Met Gly Leu Phe Asn Gln Ser Ala Pro Leu
 325 330 335
 Asp Pro Thr Thr Pro Asp Pro Ala Arg Thr Phe Leu Val Lys Lys Ile
 340 345 350
 Val Pro Phe Ser Ala Arg Met Val Val Glu Arg Leu Asp Cys Gly Gly
 355 360 365
 Ala Gln Ser Val Arg Leu Leu Val Asn Asp Ala Val Gln Pro Leu Ala
 370 375 380
 Phe Cys Gly Ala Asp Thr Ser Gly Val Cys Thr Leu Asp Ala Phe Val
 385 390 395 400
 Glu Ser Gln Ala Tyr Ala Arg Asn Asp Gly Glu Gly Asp Phe Glu Lys
 405 410 415
 Cys Phe Ala Thr
 420

<210> 20
 <211> 435
 <212> PRT
 <213> 平田头菇

<400> 20
 Gly Gly Val Val Gln Ala Thr Phe Val Gln Pro Phe Phe Pro Pro Gln
 1 5 10 15
 Ile Gln Asp Ser Trp Ala Ala Tyr Thr Pro Tyr Tyr Pro Val Gln Ala
 20 25 30
 Tyr Thr Pro Pro Pro Lys Asp Cys Lys Ile Thr Gln Val Asn Ile Ile
 35 40 45

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

Pro Ser Phe Pro Asn Pro Lys Arg Thr Trp Val Thr Ser Arg Leu Thr
 340 345 350
 Pro Phe Ser Ala Arg Met Val Thr Glu Arg Leu Leu Cys Gln Arg Asp
 355 360 365
 Gly Thr Gly Ser Gly Gly Pro Ser Arg Ile Met Arg Asn Gly Asn Val
 370 375 380
 Gln Thr Phe Val Arg Ile Leu Val Asn Asp Ala Leu Gln Pro Leu Lys
 385 390 395 400
 Phe Cys Gly Gly Asp Met Asp Ser Leu Cys Thr Leu Glu Ala Phe Val
 405 410 415
 Glu Ser Gln Lys Tyr Ala Arg Glu Asp Gly Gln Gly Asp Phe Glu Lys
 420 425 430
 Cys Phe Asp
 435

<210> 21
 <211> 419
 <212> PRT
 <213> Peniophora lycii

<400> 21
 Ser Thr Gln Phe Ser Phe Val Ala Ala Gln Leu Pro Ile Pro Ala Gln
 1 5 10 15
 Asn Thr Ser Asn Trp Gly Pro Tyr Asp Pro Phe Phe Pro Val Glu Pro
 20 25 30
 Tyr Ala Ala Pro Pro Glu Gly Cys Thr Val Thr Gln Val Asn Leu Ile
 35 40 45
 Gln Arg His Gly Ala Arg Trp Pro Thr Ser Gly Ala Arg Ser Arg Gln
 50 55 60
 Val Ala Ala Val Ala Lys Ile Gln Met Ala Arg Pro Phe Thr Asp Pro
 65 70 75 80
 Lys Tyr Glu Phe Leu Asn Asp Phe Val Tyr Lys Phe Gly Val Ala Asp
 85 90 95
 Leu Leu Pro Phe Gly Ala Asn Gln Ser His Gln Thr Gly Thr Asp Met
 100 105 110
 Tyr Thr Arg Tyr Ser Thr Leu Phe Glu Gly Gly Asp Val Pro Phe Val
 115 120 125

Arg Ala Ala Gly Asp Gln Arg Val Val Asp Ser Ser Thr Asn Trp Thr
 130 135 140

Ala Gly Phe Gly Asp Ala Ser Gly Glu Thr Val Leu Pro Thr Leu Gln
 145 150 155 160

Val Val Leu Gln Glu Glu Gly Asn Cys Thr Leu Cys Asn Asn Met Cys
 165 170 175

Pro Asn Glu Val Asp Gly Asp Glu Ser Thr Thr Trp Leu Gly Val Phe
 180 185 190

Ala Pro Asn Ile Thr Ala Arg Leu Asn Ala Ala Ala Pro Ser Ala Asn
 195 200 205

Leu Ser Asp Ser Asp Ala Leu Thr Leu Met Asp Met Cys Pro Phe Asp
 210 215 220

Thr Leu Ser Ser Gly Asn Ala Ser Pro Phe Cys Asp Leu Phe Thr Ala
 225 230 235 240

Glu Glu Tyr Val Ser Tyr Glu Tyr Tyr Tyr Asp Leu Asp Lys Tyr Tyr
 245 250 255

Gly Thr Gly Pro Gly Asn Ala Leu Gly Pro Val Gln Gly Val Gly Tyr
 260 265 270

Val Asn Glu Leu Leu Ala Arg Leu Thr Gly Gln Ala Val Arg Asp Glu
 275 280 285

Thr Gln Thr Asn Arg Thr Leu Asp Ser Asp Pro Ala Thr Phe Pro Leu
 290 295 300

Asn Arg Thr Phe Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Pro
 305 310 315 320

Ile Phe Ala Ala Leu Gly Leu Phe Asn Ala Thr Ala Leu Asp Pro Leu
 325 330 335

Lys Pro Asp Glu Asn Arg Leu Trp Val Asp Ser Lys Leu Val Pro Phe
 340 345 350

Ser Gly His Met Thr Val Glu Lys Leu Ala Cys Ser Gly Lys Glu Ala
 355 360 365

Val Arg Val Leu Val Asn Asp Ala Val Gln Pro Leu Glu Phe Cys Gly
 370 375 380

Gly Val Asp Gly Val Cys Glu Leu Ser Ala Phe Val Glu Ser Gln Thr
 385 390 395 400

Tyr Ala Arg Glu Asn Gly Gln Gly Asp Phe Ala Lys Cys Gly Phe Val
 405 410 415

Pro Ser Glu

<210> 22

<211> 369

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：担子菌共有序列

<400> 22

Ser Pro Arg Thr Ala Ala Gln Leu Pro Ile Pro Gln Gln Trp Ser Pro
1 5 10 15

Tyr Ser Pro Tyr Phe Pro Val Ala Tyr Ala Pro Pro Ala Gly Cys Gln
20 25 30

Ile Gln Val Asn Ile Ile Gln Arg His Gly Ala Arg Phe Pro Thr Ser
35 40 45

Gly Ala Ala Thr Arg Ile Gln Ala Ala Val Ala Lys Leu Gln Ser Ala
50 55 60

Thr Asp Pro Lys Leu Asp Phe Leu Asn Thr Tyr Leu Gly Asp Asp Leu
65 70 75 80

Val Pro Phe Gly Ala Gln Ser Ser Gln Ala Gly Gln Glu Ala Phe Thr
85 90 95

Arg Tyr Ser Leu Val Ser Asp Asn Leu Pro Phe Val Arg Ala Ser Gly
100 105 110

Ser Asp Arg Val Val Asp Ser Ala Thr Asn Trp Thr Ala Gly Phe Ala
115 120 125

Ala Ser Asn Thr Pro Leu Val Ile Leu Ser Glu Gly Asn Asp Thr Leu
130 135 140

Asp Asp Asn Met Cys Pro Ala Gly Asp Ser Asp Pro Gln Asn Trp Leu
145 150 155 160

Ala Val Phe Ala Pro Pro Ile Thr Ala Arg Leu Asn Ala Ala Ala Pro
165 170 175

Gly Ala Asn Leu Thr Asp Asp Ala Asn Leu Leu Cys Pro Phe Glu Thr
180 185 190

Val Ser Glu Ser Phe Cys Asp Leu Phe Glu Pro Glu Glu Phe Ala Phe
195 200 205

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

Pro

<210> 23
 <211> 440
 <212> PRT
 <213> Thermomyces lanuginosus

<400> 23
 Asn Val Asp Ile Ala Arg His Trp Gly Gln Tyr Ser Pro Phe Phe Ser
 1 5 10 15
 Leu Ala Glu Val Ser Glu Ile Ser Pro Ala Val Pro Lys Gly Cys Arg
 20 25 30
 Val Glu Phe Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr
 35 40 45
 Ala His Lys Ser Glu Val Tyr Ala Glu Leu Leu Gln Arg Ile Gln Asp
 50 55 60
 Thr Ala Thr Glu Phe Lys Gly Asp Phe Ala Phe Leu Arg Asp Tyr Ala
 65 70 75 80

Tyr His Leu Gly Ala Asp Asn Leu Thr Arg Phe Gly Glu Glu Gln Met
 85 90 95
 Met Glu Ser Gly Arg Gln Phe Tyr His Arg Tyr Arg Glu Gln Ala Arg
 100 105 110
 Glu Ile Val Pro Phe Val Arg Ala Ala Gly Ser Ala Arg Val Ile Ala
 115 120 125
 Ser Ala Glu Phe Phe Asn Arg Gly Phe Gln Asp Ala Lys Asp Arg Asp
 130 135 140
 Pro Arg Ser Asn Lys Asp Gln Ala Glu Pro Val Ile Asn Val Ile Ile
 145 150 155 160
 Ser Glu Glu Thr Gly Ser Asn Asn Thr Leu Asp Gly Leu Thr Cys Pro
 165 170 175
 Ala Ala Glu Glu Ala Pro Asp Pro Thr Gln Pro Ala Glu Phe Leu Gln
 180 185 190
 Val Phe Gly Pro Arg Val Leu Lys Lys Ile Thr Lys His Met Pro Gly
 195 200 205
 Val Asn Leu Thr Leu Glu Asp Val Pro Leu Phe Met Asp Leu Cys Pro
 210 215 220
 Phe Asp Thr Val Gly Ser Asp Pro Val Leu Phe Pro Arg Gln Leu Ser
 225 230 235 240
 Pro Phe Cys His Leu Phe Thr Ala Asp Asp Trp Met Ala Tyr Asp Tyr
 245 250 255
 Tyr Tyr Thr Leu Asp Lys Tyr Tyr Ser His Gly Gly Gly Ser Ala Phe
 260 265 270
 Gly Pro Ser Arg Gly Val Gly Phe Val Asn Glu Leu Ile Ala Arg Met
 275 280 285
 Thr Gly Asn Leu Pro Val Lys Asp His Thr Thr Val Asn His Thr Leu
 290 295 300
 Asp Asp Asn Pro Glu Thr Phe Pro Leu Asp Ala Val Leu Tyr Ala Asp
 305 310 315 320
 Phe Ser His Asp Asn Thr Met Thr Gly Ile Phe Ser Ala Met Gly Leu
 325 330 335
 Tyr Asn Gly Thr Lys Pro Leu Ser Thr Ser Lys Ile Gln Pro Pro Thr
 340 345 350
 Gly Ala Ala Ala Asp Gly Tyr Ala Ala Ser Trp Thr Val Pro Phe Ala
 355 360 365

Ala Arg Ala Tyr Val Glu Leu Leu Arg Cys Glu Thr Glu Thr Ser Ser
370 375 380

Glu Glu Glu Glu Glu Gly Glu Asp Glu Pro Phe Val Arg Val Leu Val
385 390 395 400

Asn Asp Arg Val Val Pro Leu His Gly Cys Arg Val Asp Arg Trp Gly
405 410 415

Arg Cys Arg Arg Asp Glu Trp Ile Lys Gly Leu Thr Phe Ala Arg Gln
420 425 430

Gly Gly His Trp Asp Arg Cys Phe
435 440

<210> 24

<211> 441

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：共有序列 10

<400> 24

Asn Ser His Ser Cys Asp Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro
1 5 10 15

Glu Ile Ser His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Ala
20 25 30

Asp Glu Ser Ala Ile Ser Pro Asp Val Pro Lys Gly Cys Arg Val Thr
35 40 45

Phe Val Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser
50 55 60

Lys Ser Lys Lys Tyr Ser Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala
65 70 75 80

Thr Ala Phe Lys Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr
85 90 95

Leu Gly Ala Asp Asp Leu Thr Pro Phe Gly Glu Gln Gln Met Val Asn
100 105 110

Ser Gly Ile Lys Phe Tyr Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile
115 120 125

Val Pro Phe Val Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Ala
130 135 140

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

Gly Gly Asn Trp Glu Glu Cys Phe Ala
435 440

<210> 25

<211> 1426

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶 10

<220>

<221> CDS

<222> (12)..(1412)

<220>

<221> 成熟肽

<222> (90)..(1412)

<220>

<221> 单肽

<222> (12)..(89)

<400> 25

tatatgaatt c atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg 50
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu
-25 -20 -15

ttc ggt tcc aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac 98
Phe Gly Ser Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His
-10 -5 -1 1

tct tgt gac act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct 146
Ser Cys Asp Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser
5 10 15

cac ttg tgg ggt caa tac tct cca ttc ttc tct ttg gct gac gaa tct 194
His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser
20 25 30 35

gct att tct cca gac gtt cca aag ggt tgt aga gtt act ttc gtt caa 242
Ala Ile Ser Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln
40 45 50

gtt ttg tct aga cac ggt gct aga tac cca act tct tct aag tct aag 290
Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys
55 60 65

aag tac tct gct ttg att gaa gct att caa aag aac gct act gct ttc 338
Lys Tyr Ser Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe
70 75 80

aag ggt aag tac gct ttc ttg aag act tac aac tac act ttg ggt gct	386
Lys Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala	
85 90 95	
gac gac ttg act cca ttc ggt gaa caa caa atg gtt aac tct ggt att	434
Asp Asp Leu Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile	
100 105 110 115	
aag ttc tac aga aga tac aag gct ttg gct aga aag att gtt cca ttc	482
Lys Phe Tyr Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe	
120 125 130	
gtt aga gct tct ggt tct gac aga gtt att gct tct gct gaa aag ttc	530
Val Arg Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe	
135 140 145	
att gaa ggt ttc caa tct gct aag ttg gct gac cca ggt gct aac cca	578
Ile Glu Gly Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro	
150 155 160	
cac caa gct tct cca gtt att aac gtt att att cca gaa ggt gct ggt	626
His Gln Ala Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly	
165 170 175	
tac aac aac act ttg gac cac ggt ttg tgt act gct ttc gaa gaa tct	674
Tyr Asn Asn Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser	
180 185 190 195	
gaa ttg ggt gac gac gtt gaa gct aac ttc act gct gtt ttc gct cca	722
Glu Leu Gly Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro	
200 205 210	
cct att aga gct aga ttg gaa gct cac ttg cca ggt gtt aac ttg act	770
Pro Ile Arg Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr	
215 220 225	
gac gaa gac gtt gtt aac ttg atg gac atg tgt cca ttc gac act gtt	818
Asp Glu Asp Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val	
230 235 240	
gct aga act tct gac gct act caa ttg tct cca ttc tgt gac ttg ttc	866
Ala Arg Thr Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe	
245 250 255	
act cac gac gaa tgg att caa tac gac tac ttg caa tct ttg ggt aag	914
Thr His Asp Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys	
260 265 270 275	
tac tac ggt tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt	962
Tyr Tyr Gly Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val	
280 285 290	

```

ggt ttc gtt aac gaa ttg att gct aga ttg act cac tct cca gtt caa 1010
Gly Phe Val Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln
      295              300              305

gac cac act tct act aac cac act ttg gac tct aac cca gct act ttc 1058
Asp His Thr Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe
      310              315              320

cca ttg aac gct act ttg tac gct gac ttc tct cac gac aac act atg 1106
Pro Leu Asn Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met
      325              330              335

gtt tct att ttc ttc gct ttg ggt ttg tac aac ggt act aag cca ttg 1154
Val Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu
      340              345              350              355

tct act act tct gtt gaa tct att gaa gaa act gac ggt tac gct gct 1202
Ser Thr Thr Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ala Ala
      360              365              370

tct tgg act gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa 1250
Ser Trp Thr Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln
      375              380              385

tgt gaa gct gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga 1298
Cys Glu Ala Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg
      390              395              400

gtt gtt cca ttg cac ggt tgt ggt gtt gac aag ttg ggt aga tgt aag 1346
Val Val Pro Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys
      405              410              415

aga gac gac ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac 1394
Arg Asp Asp Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn
      420              425              430              435

tgg gaa gaa tgt ttc gct taagaattca tata 1426
Trp Glu Glu Cys Phe Ala
      440

```

<210> 26

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶 10

<400> 26

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
  -25              -20              -15

```

```

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
  -10              -5              -1      1              5

```

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

```

Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
295                      300                      305                      310

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
                      315                      320                      325

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
                      330                      335                      340

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
                      345                      350                      355

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ala Ala Ser Trp Thr
360                      365                      370

Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
375                      380                      385                      390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
                      395                      400                      405

Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
                      410                      415                      420

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
425                      430                      435

Cys Phe Ala
440

```

<210> 27

<211> 437

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶 11

<400> 27

```

Asn Ser His Ser Cys Asp Thr Val Asp Gly Tyr Gln Cys Pro Glu Ile
 1                      5                      10                      15

Ser His Leu Trp Gly Gln Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu
                20                      25                      30

Ser Ala Ile Ser Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val
35                      40                      45

Gln Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser
50                      55                      60

```

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

```

Ile Glu Thr Asp Gly Tyr Ala Ala Ser Trp Thr Val Pro Phe Ala Ala
    355                      360                      365

Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala Gly Gly Gly Gly Gly
    370                      375                      380

Glu Gly Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val
    385                      390                      395                      400

Val Pro Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Leu
    405                      410                      415

Asp Asp Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp
    420                      425                      430

Ala Glu Cys Phe Ala
    435

```

<210> 28

<211> 1404

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶 1 thermo 8 q50t, k91a

<220>

<221> CDS

<222> (1)..(1401)

<220>

<221> 成熟肽

<222> (79)..(1401)

<220>

<221> 单肽

<222> (1)..(78)

<400> 28

```

atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg ttc ggt tcc   48
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
    -25                      -20                      -15

```

```

aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac tct tgt gac   96
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
    -10                      -5                      -1    1                      5

```

```

act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct cac ttg tgg   144
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
    10                      15                      20

```

ggt acc tac tct cca tac ttc tct ttg gca gac gaa tct gct att tct	192
Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser	
25 30 35	
cca gac gtt cca gac gac tgt aga gtt act ttc gtt caa gtt ttg tct	240
Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser	
40 45 50	
aga cac ggt gct aga tac cca act tct tct gcg tct aag gct tac tct	288
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser	
55 60 65 70	
gct ttg att gaa gct att caa aag aac gct act gct ttc aag ggt aag	336
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys	
75 80 85	
tac gct ttc ttg aag act tac aac tac act ttg ggt gct gac gac ttg	384
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu	
90 95 100	
act cca ttc ggt gaa aac caa atg gtt aac tct ggt att aag ttc tac	432
Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr	
105 110 115	
aga aga tac aag gct ttg gct aga aag att gtt cca ttc att aga gct	480
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala	
120 125 130	
tct ggt tct gac aga gtt att gct tct gct gaa aag ttc att gaa ggt	528
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly	
135 140 145 150	
ttc caa tct gct aag ttg gct gac cca ggt tct caa cca cac caa gct	576
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala	
155 160 165	
tct cca gtt att aac gtg atc att cca gaa gga tcc ggt tac aac aac	624
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn	
170 175 180	
act ttg gac cac ggt act tgt act gct ttc gaa gac tct gaa tta ggt	672
Thr Leu Asp His Gly Thr Cys Thr Ala Phe Glu Asp Ser Glu Leu Gly	
185 190 195	
gac gac gtt gaa gct aac ttc act gct ttg ttc gct cca gct att aga	720
Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg	
200 205 210	
gct aga ttg gaa gct gac ttg cca ggt gtt act ttg act gac gaa gac	768
Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp	
215 220 225 230	

gtt gtt tac ttg atg gac atg tgt cca ttc gac act gtc gct aga act	816
Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	
tct gac gct act gaa ttg tct cca ttc tgt gct ttg ttc act cac gac	864
Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp	
250 255 260	
gaa tgg atc caa tac gac tac ttg caa agc ttg ggt aag tac tac ggt	912
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly	
265 270 275	
tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt ggt ttc gct	960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala	
280 285 290	
aac gaa ttg att gct aga ttg act cac tct cca gtt caa gac cac act	1008
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr	
295 300 305 310	
tct act aac cac act ttg gac tct aac cca gct act ttc cca ttg aac	1056
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn	
315 320 325	
gct act ttg tac gct gac ttc tct cac gac aac act atg ata tct att	1104
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile	
330 335 340	
ttc ttc gct ttg ggt ttg tac aac ggt acc aag cca ttg tct act act	1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr	
345 350 355	
tct gtt gaa tct att gaa gaa act gac ggt tac tct gct tct tgg act	1200
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr	
360 365 370	
gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa tgt caa gct	1248
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala	
375 380 385 390	
gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga gtt gtt cca	1296
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro	
395 400 405	
ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag aga gac gac	1344
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp	
410 415 420	
ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac tgg gct gaa	1392
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu	
425 430 435	

tgt ttc gct taa
Cys Phe Ala
440

1404

<210> 29
<211> 467
<212> PRT
<213> 人工序列
<223> 人工序列的描述：共有植酸酶 1 thermo 8 q50t, k91a

<400> 29
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
-25 -20 -15
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
-10 -5 -1 1 5
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
10 15 20
Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
25 30 35
Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
40 45 50
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
55 60 65 70
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
75 80 85
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
90 95 100
Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
105 110 115
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
120 125 130
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
135 140 145 150
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
155 160 165
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn
170 175 180
Thr Leu Asp His Gly Thr Cys Thr Ala Phe Glu Asp Ser Glu Leu Gly
185 190 195

```

Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
 200                      205                      210

Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
 215                      220                      225                      230

Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
                235                      240                      245

Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp
                250                      255                      260

Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
 265                      270                      275

Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala
 280                      285                      290

Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
 295                      300                      305                      310

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
                315                      320                      325

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile
                330                      335                      340

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
 345                      350                      355

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
 360                      365                      370

Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
 375                      380                      385                      390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
                395                      400                      405

Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
                410                      415                      420

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
 425                      430                      435

Cys Phe Ala
 440

```

```

<210> 30
<211> 1404
<212> DNA

```

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶 10 thermo 3 q50t, k91a

<220>

<221> CDS

<222> (1) .. (1401)

<220>

<221> 成熟肽

<222> (79) .. (1401)

<220>

<221> 单肽

<222> (1) .. (78)

<400> 30

atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg ttc ggt tcc	48
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser	
-25 -20 -15	

aca tcc ggt acc gcc ttg ggt cct cgt ggt aac tct cac tct tgt gac	96
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp	
-10 -5 -1 1 5	

act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct cac ttg tgg	144
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp	
10 15 20	

ggg aca tac tct cca ttc ttc tct ttg gct gac gaa tct gct att tct	192
Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser	
25 30 35	

cca gac gtt cca aag ggt tgt aga gtt act ttc gtt caa gtt ttg tct	240
Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser	
40 45 50	

aga cac ggt gct aga tac cca act tct tct gcg tct aag gcg tac tct	288
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser	
55 60 65 70	

gct ttg att gaa gct att caa aag aac gct act gct ttc aag ggt aag	336
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys	
75 80 85	

tac gct ttc ttg aag act tac aac tac act ttg ggt gct gac gac ttg	384
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu	
90 95 100	

act cca ttc ggt gaa caa caa atg gtt aac tct ggt att aag ttc tac	432
Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr	
105 110 115	

aga aga tac aag gct ttg gct aga aag att gtt cca ttc att aga gct	480
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala	
120 125 130	
tct ggt tct gac aga gtt att gct tct gct gaa aag ttc att gaa ggt	528
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly	
135 140 145 150	
ttc caa tct gct aag ttg gct gac cca ggt gct aac cca cac caa gct	576
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala	
155 160 165	
tct cca gtt att aac gtt att att cca gaa ggt gct ggt tac aac aac	624
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn	
170 175 180	
act ttg gac cac ggt ttg tgt act gct ttc gaa gaa tct gaa ttg ggt	672
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Glu Leu Gly	
185 190 195	
gac gac gtt gaa gct aac ttc act gct gtt ttc gct cca cca att aga	720
Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg	
200 205 210	
gct aga ttg gaa gct cac ttg cca ggt gtt aac ttg act gac gaa gac	768
Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp	
215 220 225 230	
gtt gtt aac ttg atg gac atg tgt cca ttc gac act gtt gct aga act	816
Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	
tct gac gct act caa ttg tct cca ttc tgt gac ttg ttc act cac gac	864
Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp	
250 255 260	
gaa tgg att caa tac gac tac ttg caa tct ttg ggt aag tac tac ggt	912
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly	
265 270 275	
tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt ggt ttc gtt	960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val	
280 285 290	
aac gaa ttg att gct aga ttg act cac tct cca gtt caa gac cac act	1008
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr	
295 300 305 310	
tct act aac cac act ttg gac tct aac cca gct act ttc cca ttg aac	1056
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn	
315 320 325	

```

gct act ttg tac gct gac ttc tct cac gac aac act atg gtt tct att 1104
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
          330                      335                      340

ttc ttc gct ttg ggt ttg tac aac ggt act aag cca ttg tct act act 1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
          345                      350                      355

tct gtt gaa tct att gaa gaa act gac ggt tac tct gct tct tgg act 1200
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
          360                      365                      370

gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa tgt gaa gct 1248
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
          375                      380                      385                      390

gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga gtt gtt cca 1296
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
          395                      400                      405

ttg cac ggt tgt ggt gtt gac aag ttg ggt aga tgt aag aga gac gac 1344
Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
          410                      415                      420

ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac tgg gaa gaa 1392
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
          425                      430                      435

tgt ttc gct taa 1404
Cys Phe Ala
          440

```

<210> 31

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶 10 thermo 3 q50t, k91a

<400> 31

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
-25                      -20                      -15

```

```

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
-10                      -5                      -1 1                      5

```

```

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
10                      15                      20

```

```

Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
25                      30                      35

```

```

Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser
  40                      45                      50

Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
  55                      60                      65                      70

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
                      75                      80                      85

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
                      90                      95                      100

Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
  105                      110                      115

Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
  120                      125                      130

Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
  135                      140                      145                      150

Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
                      155                      160                      165

Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
                      170                      175                      180

Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Glu Leu Gly
                      185                      190                      195

Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
  200                      205                      210

Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
  215                      220                      225                      230

Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
                      235                      240                      245

Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
                      250                      255                      260

Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
                      265                      270                      275

Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
  280                      285                      290

Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
  295                      300                      305                      310

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
                      315                      320                      325

```

```

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
      330                      335                      340

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
      345                      350                      355

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
      360                      365                      370

Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
      375                      380                      385                      390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
      395                      400                      405

Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
      410                      415                      420

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
      425                      430                      435

Cys Phe Ala
      440

```

<210> 32
 <211> 1404
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述：烟曲霉 α - 突变体

<220>
 <221> CDS
 <222> (1)..(1401)

<220>
 <221> 成熟肽
 <222> (79)..(1401)

<220>
 <221> 单肽
 <222> (1)..(78)

```

<400> 32
atg ggg gtt ttc gtc gtt cta tta tct atc gcg act ctg ttc ggc agc   48
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
      -25                      -20                      -15

```

```

aca tcg ggc act gcg ctg ggc ccc cgt gga aat cac tcc aag tcc tgc 96
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn His Ser Lys Ser Cys
-10 -5 -1 1 5

gat acg gta gac cta ggg tac cag tgc tcc cct gcg act tct cat cta 144
Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro Ala Thr Ser His Leu
10 15 20

tgg ggc acg tac tcg cca tac ttt tcg ctc gag gac gag ctg tcc gtg 192
Trp Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Leu Ser Val
25 30 35

tcg agt aag ctt ccc aag gat tgc cgg atc acc ttg gta cag gtg cta 240
Ser Ser Lys Leu Pro Lys Asp Cys Arg Ile Thr Leu Val Gln Val Leu
40 45 50

tcg cgc cat gga gcg cgg tac cca acc agc tcc aag agc aaa aag tat 288
Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Lys Tyr
55 60 65 70

aag aag ctt att acg gcg atc cag gcc aat gcc acc gac ttc aag ggc 336
Lys Lys Leu Ile Thr Ala Ile Gln Ala Asn Ala Thr Asp Phe Lys Gly
75 80 85

aag tac gcc ttt ttg aag acg tac aac tat act ctg ggt gcg gat gac 384
Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp
90 95 100

ctc act ccc ttt ggg gag cag cag ctg gtg aac tcg ggc atc aag ttc 432
Leu Thr Pro Phe Gly Glu Gln Gln Leu Val Asn Ser Gly Ile Lys Phe
105 110 115

tac cag agg tac aag gct ctg gcg cgc agt gtg gtg ccg ttt att cgc 480
Tyr Gln Arg Tyr Lys Ala Leu Ala Arg Ser Val Val Pro Phe Ile Arg
120 125 130

gcc tca ggc tcg gac cgg gtt att gct tcg gga gag aag ttc atc gag 528
Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Gly Glu Lys Phe Ile Glu
135 140 145 150

ggg ttc cag cag gcg aag ctg gct gat cct ggc gcg acg aac cgc gcc 576
Gly Phe Gln Gln Ala Lys Leu Ala Asp Pro Gly Ala Thr Asn Arg Ala
155 160 165

gct ccg gcg att agt gtg att att ccg gag agc gag acg ttc aac aat 624
Ala Pro Ala Ile Ser Val Ile Ile Pro Glu Ser Glu Thr Phe Asn Asn
170 175 180

acg ctg gac cac ggt gtg tgc acg aag ttt gag gcg agt cag ctg gga 672
Thr Leu Asp His Gly Val Cys Thr Lys Phe Glu Ala Ser Gln Leu Gly
185 190 195

```

gat gag gtt gcg gcc aat ttc act gcg ctc ttt gca ccc gac atc cga	720
Asp Glu Val Ala Ala Asn Phe Thr Ala Leu Phe Ala Pro Asp Ile Arg	
200 205 210	
gct cgc ctc gag aag cat ctt cct ggc gtg acg ctg aca gac gag gac	768
Ala Arg Leu Glu Lys His Leu Pro Gly Val Thr Leu Thr Asp Glu Asp	
215 220 225 230	
gtt gtc agt cta atg gac atg tgt ccg ttt gat acg gta gcg cgc acc	816
Val Val Ser Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	
agc gac gca agt cag ctg tca ccg ttc tgt caa ctc ttc act cac aat	864
Ser Asp Ala Ser Gln Leu Ser Pro Phe Cys Gln Leu Phe Thr His Asn	
250 255 260	
gag tgg aag aag tac gac tac ctt cag tcc ttg ggc aag tac tac ggc	912
Glu Trp Lys Lys Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly	
265 270 275	
tac ggc gca ggc aac cct ctg gga ccg gct cag ggg ata ggg ttc acc	960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Ile Gly Phe Thr	
280 285 290	
aac gag ctg att gcc cgg ttg acg cgt tcg cca gtg cag gac cac acc	1008
Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln Asp His Thr	
295 300 305 310	
agc act aac tcg act cta gtc tcc aac ccg gcc acc ttc ccg ttg aac	1056
Ser Thr Asn Ser Thr Leu Val Ser Asn Pro Ala Thr Phe Pro Leu Asn	
315 320 325	
gct acc atg tac gtc gac ttt tca cac gac aac agc atg gtt tcc atc	1104
Ala Thr Met Tyr Val Asp Phe Ser His Asp Asn Ser Met Val Ser Ile	
330 335 340	
ttc ttt gca ttg ggc ctg tac aac ggc act gaa ccc ttg tcc cgg acc	1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Glu Pro Leu Ser Arg Thr	
345 350 355	
tcg gtg gaa agc gcc aag gaa ttg gat ggg tat tct gca tcc tgg gtg	1200
Ser Val Glu Ser Ala Lys Glu Leu Asp Gly Tyr Ser Ala Ser Trp Val	
360 365 370	
gtg cct ttc ggc gcg cga gcc tac ttc gag acg atg caa tgc aag tcg	1248
Val Pro Phe Gly Ala Arg Ala Tyr Phe Glu Thr Met Gln Cys Lys Ser	
375 380 385 390	
gaa aag gag cct ctt gtt cgc gct ttg att aat gac cgg gtt gtg cca	1296
Glu Lys Glu Pro Leu Val Arg Ala Leu Ile Asn Asp Arg Val Val Pro	
395 400 405	

```

ctg cat ggc tgc gat gtg gac aag ctg ggg cga tgc aag ctg aat gac 1344
Leu His Gly Cys Asp Val Asp Lys Leu Gly Arg Cys Lys Leu Asn Asp
      410              415              420

ttt gtc aag gga ttg agt tgg gcc aga tct ggg ggc aac tgg gga gag 1392
Phe Val Lys Gly Leu Ser Trp Ala Arg Ser Gly Gly Asn Trp Gly Glu
      425              430              435

tgc ttt agt tga 1404
Cys Phe Ser
      440

```

<210> 33
 <211> 467
 <212> PRT
 <213> 人工序列
 <223> 人工序列的描述：烟曲霉 α -突变体

```

<400> 33
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
  -25              -20              -15

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn His Ser Lys Ser Cys
-10              -5              -1 1              5

Asp Thr Val Asp Leu Gly Tyr Gln Cys Ser Pro Ala Thr Ser His Leu
      10              15              20

Trp Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Leu Ser Val
      25              30              35

Ser Ser Lys Leu Pro Lys Asp Cys Arg Ile Thr Leu Val Gln Val Leu
      40              45              50

Ser Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Lys Tyr
      55              60              65              70

Lys Lys Leu Ile Thr Ala Ile Gln Ala Asn Ala Thr Asp Phe Lys Gly
      75              80              85

Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp
      90              95              100

Leu Thr Pro Phe Gly Glu Gln Gln Leu Val Asn Ser Gly Ile Lys Phe
      105              110              115

Tyr Gln Arg Tyr Lys Ala Leu Ala Arg Ser Val Val Pro Phe Ile Arg
      120              125              130

Ala Ser Gly Ser Asp Arg Val Ile Ala Ser Gly Glu Lys Phe Ile Glu
      135              140              145              150

```

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

Cys Phe Ser
440

<210> 34
<211> 1426
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：共有植酸酶 7

<220>
<221> CDS
<222> (12)..(1412)

<220>
<221> 成熟肽
<222> (90)..(1412)

<220>
<221> 单肽
<222> (12)..(89)

<400> 34
tatatgaatt c atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg 50
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu
-25 -20 -15

ttc ggt tcc aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac 98
Phe Gly Ser Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His
-10 -5 -1 1

tct tgt gac act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct 146
Ser Cys Asp Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser
5 10 15

cac ttg tgg ggt caa tac tct cca tac ttc tct ttg gaa gac gaa tct 194
His Leu Trp Gly Gln Tyr Ser Pro Tyr Phe Ser Leu Glu Asp Glu Ser
20 25 30 35

gct att tct cca gac gtt cca gac gac tgt aga gtt act ttc gtt caa 242
Ala Ile Ser Pro Asp Val Pro Asp Asp Cys Arg Val Thr Phe Val Gln
40 45 50

gtt ttg tct aga cac ggt gct aga tac cca act gac tct aag ggt aag 290
Val Leu Ser Arg His Gly Ala Arg Tyr Pro Thr Asp Ser Lys Gly Lys
55 60 65

aag tac tct gct ttg att gaa gct att caa aag aac gct act gct ttc 338
Lys Tyr Ser Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe
70 75 80

aag ggt aag tac gct ttc ttg aag act tac aac tac act ttg ggt gct	386
Lys Gly Lys Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala	
85 90 95	
gac gac ttg act cca ttc ggt gaa aac caa atg gtt aac tct ggt att	434
Asp Asp Leu Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile	
100 105 110 115	
aag ttc tac aga aga tac aag gct ttg gct aga aag att gtt cca ttc	482
Lys Phe Tyr Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe	
120 125 130	
att aga gct tct ggt tct tct aga gtt att gct tct gct gaa aag ttc	530
Ile Arg Ala Ser Gly Ser Ser Arg Val Ile Ala Ser Ala Glu Lys Phe	
135 140 145	
att gaa ggt ttc caa tct gct aag ttg gct gac cca ggt tct caa cca	578
Ile Glu Gly Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro	
150 155 160	
cac caa gct tct cca gtt att gac gtt att att tct gac gct tct tct	626
His Gln Ala Ser Pro Val Ile Asp Val Ile Ile Ser Asp Ala Ser Ser	
165 170 175	
tac aac aac act ttg gac cca ggt act tgt act gct ttc gaa gac tct	674
Tyr Asn Asn Thr Leu Asp Pro Gly Thr Cys Thr Ala Phe Glu Asp Ser	
180 185 190 195	
gaa ttg gct gac act gtt gaa gct aac ttc act gct ttg ttc gct cca	722
Glu Leu Ala Asp Thr Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro	
200 205 210	
gct att aga gct aga ttg gaa gct gac ttg cca ggt gtt act ttg act	770
Ala Ile Arg Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr	
215 220 225	
gac act gaa gtt act tac ttg atg gac atg tgt tct ttc gaa act gtt	818
Asp Thr Glu Val Thr Tyr Leu Met Asp Met Cys Ser Phe Glu Thr Val	
230 235 240	
gct aga act tct gac gct act gaa ttg tct cca ttc tgt gct ttg ttc	866
Ala Arg Thr Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe	
245 250 255	
act cac gac gaa tgg aga cac tac gac tac ttg caa tct ttg aag aag	914
Thr His Asp Glu Trp Arg His Tyr Asp Tyr Leu Gln Ser Leu Lys Lys	
260 265 270 275	
tac tac ggt cac ggt gct ggt aac cca ttg ggt cca act caa ggt gtt	962
Tyr Tyr Gly His Gly Ala Gly Asn Pro Leu Gly Pro Thr Gln Gly Val	
280 285 290	

```

ggt ttc gct aac gaa ttg att gct aga ttg act aga tct cca gtt caa 1010
Gly Phe Ala Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln
                295                      300                      305

gac cac act tct act aac cac act ttg gac tct aac cca gct act ttc 1058
Asp His Thr Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe
                310                      315                      320

cca ttg aac gct act ttg tac gct gac ttc tct cac gac aac ggt att 1106
Pro Leu Asn Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Gly Ile
                325                      330                      335

att tct att ttc ttc gct ttg ggt ttg tac aac ggt act gct cca ttg 1154
Ile Ser Ile Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Ala Pro Leu
                340                      345                      350                      355

tct act act tct gtt gaa tct att gaa gaa act gac ggt tac tct tct 1202
Ser Thr Thr Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ser
                360                      365                      370

gct tgg act gtt cca ttc gct tct aga gct tac gtt gaa atg atg caa 1250
Ala Trp Thr Val Pro Phe Ala Ser Arg Ala Tyr Val Glu Met Met Gln
                375                      380                      385

tgt caa gct gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga 1298
Cys Gln Ala Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg
                390                      395                      400

gtt gtt cca ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag 1346
Val Val Pro Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys
                405                      410                      415

aga gac gac ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac 1394
Arg Asp Asp Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn
                420                      425                      430                      435

tgg gct gaa tgt ttc gct taagaattca tata 1426
Trp Ala Glu Cys Phe Ala
                440

```

<210> 35

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶 7

<400> 35

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
-25                      -20                      -15

```

```

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
-10                      -5                      -1 1                      5

```

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

```

Asn Glu Leu Ile Ala Arg Leu Thr Arg Ser Pro Val Gln Asp His Thr
295              300              305              310

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
              315              320              325

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Gly Ile Ile Ser Ile
              330              335              340

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Ala Pro Leu Ser Thr Thr
              345              350              355

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ser Ala Trp Thr
360              365              370

Val Pro Phe Ala Ser Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
375              380              385              390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
              395              400              405

Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
              410              415              420

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
              425              430              435

Cys Phe Ala
440

```

<210> 36

<211> 467

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶 12

<400> 36

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
 1              5              10              15

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
              20              25              30

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser Ser Asn Trp
              35              40              45

Ser Pro Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
50              55              60

```

```

Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Gln
 65              70              75              80

Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Ala Thr Arg Ile Ser
              85              90              95

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
      100              105              110

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
      115              120              125

Val Pro Phe Gly Ala Asn Gln Ser Ser Gln Ala Gly Ile Lys Phe Tyr
      130              135              140

Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
      145              150              155              160

Ser Gly Ser Asp Arg Val Ile Asp Ser Ala Thr Asn Trp Ile Glu Gly
              165              170              175

Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
      180              185              190

Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
      195              200              205

Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Glu Leu Gly
      210              215              220

Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
      225              230              235              240

Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
              245              250              255

Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
      260              265              270

Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
      275              280              285

Glu Trp Ile Gln Tyr Asp Tyr Leu Gly Asp Leu Asp Lys Tyr Tyr Gly
      290              295              300

Thr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
      305              310              315              320

Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
              325              330              335

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
              340              345              350

```

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ala Ile
 355 360 365

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
 370 375 380

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Leu
 385 390 395 400

Val Pro Phe Ser Ala Arg Met Tyr Val Glu Met Met Gln Cys Glu Ala
 405 410 415

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
 420 425 430

Leu His Gly Cys Gly Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
 435 440 445

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
 450 455 460

Cys Phe Ala
 465

<210> 37
 <211> 26
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述：引物

<400> 37
 tatatgaatt catgggcgtg ttcgtc 26

<210> 38
 <211> 22
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述：引物

<400> 38
 tgaaaagttc attgaagggtt tc 22

<210> 39
 <211> 22
 <212> DNA
 <213> 人工序列

<220>	
<223> 人工序列的描述：引物	
<400> 39	
tcttcgaaag cagtacacaa ac	22
<210> 40	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：引物	
<400> 40	
tatatgaatt cttaagcgaa ac	22
<210> 41	
<211> 32	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：引物	
<400> 41	
cacttgtggg gtacctactc tccatacttc tc	32
<210> 42	
<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：引物	
<400> 42	
ggtcaatact ctccattctt ctctttggaa g	31
<210> 43	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：引物	
<400> 43	
catacttctc tttggcagac gaatctgc	28

<210> 44
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 44
ctccagacgt cccaaaggac ttagagtta c 31

<210> 45
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 45
ctccagacgt ccagacggc ttagagtta c 31

<210> 46
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 46
gatacccaac ttcttctgcg tctaaggctt actctg 36

<210> 47
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 47
cttctaagtc taagaagtac tctgctttg 29

<210> 48
<211> 41
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 48
gcttactctg ctttgattga acggattcaa aagaacgcta c 41

<210> 49
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 49
ccattcgggtg aacagcaaatt ggttaactc 29

<210> 50
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 50
gatacaaggc tctcgcgaga aacattgttc 30

<210> 51
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 51
gattgttcca ttcgtgcgcg cttctggttc 30

<210> 52
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 52
ctccagttat taacgtgatc attccagaag g 31

<210> 53
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 53
ggctgaccca ggggcccaac cacaccaagc 30

<210> 54
<211> 32
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 54
cactttggac catggtcttt gtactgcttt cg 32

<210> 55
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 55
gctttcgaag actctaccct aggtgacgac gttg 34

<210> 56
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 56
ggtgacgacg ctgaagctaa cttcac 26

<210> 57
<211> 27
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 57

ctaacttcac cgcggtgttc gctccag

27

<210> 58

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 58

gctttgttcg ctccacctat tagagctaga ttgg

34

<210> 59

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 59

gccaggtgtt aacttgactg acgaag

26

<210> 60

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 60

gacgaagacg tcgttaactt gatggac

27

<210> 61

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 61

gtccattcga cactgtcgct agaacttc

28

<210> 62
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 62
ctgacgctac tcagctgtct ccattc 26

<210> 63
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 63
gtctccattc tgtgatttgt tcactcac 28

<210> 64
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 64
gctttgttca ccgcggacga atggag 26

<210> 65
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 65
cacgacgaat ggatccaata cgactac 27

<210> 66
<211> 27
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 66

gacgaatgga gagcgtacga ctacttg

27

<210> 67

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 67

ggtgttggtt tcgttaacga attgattgc

29

<210> 68

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 68

gctagattga ctactctcc agttcaag

28

<210> 69

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 69

ctcacgacaa cactatgata tctattttct tc

32

<210> 70

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 70

cgacaactcc atggtttcta ttttcttcgc

30

<210> 71
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 71
gtacaacggt accaagccat tgtctac 27

<210> 72
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 72
ctgacggtta cgctgcttct tggac 25

<210> 73
<211> 26
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 73
ctgttccatt cgctgctaga gcttac 26

<210> 74
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 74
gatgcaatgt gaagctgaaa aggaacc 27

<210> 75
<211> 26
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 75

cacggttggtg gtgtcgacaa gttggg

26

<210> 76

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 76

gatctggtgg caattgggag gaatgtttcg

30

<210> 77

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 77

cacgtactcg ccatactttt cgctcgag

28

<210> 78

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 78

ccatactttt cgctcgcgga cgagctgtcc gtg

33

<210> 79

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 79

gtataagaag cttattacgg cgatccaggc c

31

<210> 80
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 80
cttcaagggc aagtacgcct ttttgaagac g 31

<210> 81
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 81
catccgagct cgcctcgaga agcatcttc 29

<210> 82
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 82
ctaattggatg tgtccgtttg atacggtag 29

<210> 83
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：引物

<400> 83
gtggaagaag tacgactacc ttcagtc 27

<210> 84
<211> 28
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 84

gcccggttga cgcattcgcc agtgcagg

28

<210> 85

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 85

cacacgacaa caccatgggtt tccatcttc

29

<210> 86

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 86

gtggtgcctt tcgccgcgcg agcctacttc

30

<210> 87

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 87

tatatcatga gcgtgttcgt cgtgctactg ttc

33

<210> 88

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：引物

<400> 88

accgactta caaagcgaat tctatagata tat

33

<210> 89
 <211> 33
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述：引物

<400> 89
 acccttctta caaagcgaat tctatagata tat 33

<210> 90
 <211> 1404
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述：共有植酸酶-3-thermo-11-Q50T

<220>
 <221> CDS
 <222> (1)..(1401)

<220>
 <221> 单肽
 <222> (1)..(69)

<220>
 <221> 成熟肽
 <222> (70)..(1401)

<400> 90
 atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg ttc ggt tcc 48
 Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
 -20 -15 -10

aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac tct tgt gac 96
 Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
 -5 -1 1 5

act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct cac ttg tgg 144
 Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
 10 15 20 25

ggg acc tac tct cca tac ttc tct ttg gca gac gaa tct gct att tct 192
 Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
 30 35 40

cca gac gtc cca aag gac tgt aga gtt act ttc gtt caa gtt ttg tct 240
 Pro Asp Val Pro Lys Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
 45 50 55

aga cac ggt gct aga tac cca act tct tct aag tct aag gct tac tct	288
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Ala Tyr Ser	
60 65 70	
gct ttg att gaa gct att caa aag aac gct act gct ttc aag ggt aag	336
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys	
75 80 85	
tac gct ttc ttg aag act tac aac tac act ttg ggt gct gac gac ttg	384
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu	
90 95 100 105	
act cca ttc ggt gaa aac caa atg gtt aac tct ggt att aag ttc tac	432
Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr	
110 115 120	
aga aga tac aag gct ttg gct aga aag att gtt cca ttc att aga gct	480
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala	
125 130 135	
tct ggt tct gac aga gtt att gct tct gct gaa aag ttc att gaa ggt	528
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly	
140 145 150	
ttc caa tct gct aag ttg gct gac cca ggt tct caa cca cac caa gct	576
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala	
155 160 165	
tct cca gtt att aac gtg atc att cca gaa gga tcc ggt tac aac aac	624
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn	
170 175 180 185	
act ttg gac cat ggt ctt tgt act gct ttc gaa gac tct acc cta ggt	672
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Asp Ser Thr Leu Gly	
190 195 200	
gac gac gtt gaa gct aac ttc act gct ttg ttc gct cca gct att aga	720
Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg	
205 210 215	
gct aga ttg gaa gct gac ttg cca ggt gtt act ttg act gac gaa gac	768
Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp	
220 225 230	
gtt gtt tac ttg atg gac atg tgt cca ttc gac act gtc gct aga act	816
Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	
tct gac gct act gaa ttg tct cca ttc tgt gct ttg ttc act cac gac	864
Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp	
250 255 260 265	

```

gaa tgg atc caa tac gac tac ttg caa agc ttg ggt aag tac tac ggt 912
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
      270      275      280

tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt ggt ttc gct 960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala
      285      290      295

aac gaa ttg att gct aga ttg act cac tct cca gtt caa gac cac act 1008
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
      300      305      310

tct act aac cac act ttg gac tct aac cca gct act ttc cca ttg aac 1056
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
      315      320      325

gct act ttg tac gct gac ttc tct cac gac aac act atg ata tct att 1104
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile
      330      335      340      345

ttc ttc gct ttg ggt ttg tac aac ggt acc aag cca ttg tct act act 1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
      350      355      360

tct gtt gaa tct att gaa gaa act gac ggt tac tct gct tct tgg act 1200
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
      365      370      375

gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa tgt caa gct 1248
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
      380      385      390

gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga gtt gtt cca 1296
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
      395      400      405

ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag aga gac gac 1344
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
      410      415      420      425

ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac tgg gct gaa 1392
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
      430      435      440

tgt ttc gct taa
Cys Phe Ala 1404

```

<210> 91

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述: 共有植酸酶-3-thermo-11-Q50T

<400> 91

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
      -20                      -15                      -10

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
      -5                      -1  1                      5

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
  10                      15                      20                      25

Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
      30                      35                      40

Pro Asp Val Pro Lys Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
      45                      50                      55

Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Ala Tyr Ser
      60                      65                      70

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
  75                      80                      85

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
  90                      95                      100                      105

Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
      110                      115                      120

Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
      125                      130                      135

Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
      140                      145                      150

Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
      155                      160                      165

Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn
      170                      175                      180                      185

Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Asp Ser Thr Leu Gly
      190                      195                      200

Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
      205                      210                      215

Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
      220                      225                      230

Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
      235                      240                      245

```

Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp
 250 255 260 265
 Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
 270 275 280
 Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala
 285 290 295
 Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
 300 305 310
 Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
 315 320 325
 Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile
 330 335 340 345
 Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
 350 355 360
 Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
 365 370 375
 Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
 380 385 390
 Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
 395 400 405
 Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
 410 415 420 425
 Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
 430 435 440
 Cys Phe Ala

<210> 92

<211> 1404

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶-3-thermo-11-Q50T-K91A

<220>

<221> 单肽

<222> (1)..(69)

<220>

<221> CDS

<222> (1)..(1401)

<220>

<221> 成熟肽

<222> (70)..(1401)

<400> 92

```

atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg ttc ggt tcc 48
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
      -20                -15                -10

aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac tct tgt gac 96
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
      -5                -1  1                5

act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct cac ttg tgg 144
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
  10                15                20                25

ggt acc tac tct cca tac ttc tct ttg gca gac gaa tct gct att tct 192
Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
      30                35                40

cca gac gtc cca aag gac tgt aga gtt act ttc gtt caa gtt ttg tct 240
Pro Asp Val Pro Lys Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
      45                50                55

aga cac ggt gct aga tac cca act tct tct gcg tct aag gct tac tct 288
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
      60                65                70

gct ttg att gaa gct att caa aag aac gct act gct ttc aag ggt aag 336
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
      75                80                85

tac gct ttc ttg aag act tac aac tac act ttg ggt gct gac gac ttg 384
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
      90                95                100                105

act cca ttc ggt gaa aac caa atg gtt aac tct ggt att aag ttc tac 432
Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
      110                115                120

aga aga tac aag gct ttg gct aga aag att gtt cca ttc att aga gct 480
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
      125                130                135

tct ggt tct gac aga gtt att gct tct gct gaa aag ttc att gaa ggt 528
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
      140                145                150

```

ttc caa tct gct aag ttg gct gac cca ggt tct caa cca cac caa gct	576
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala	
155 160 165	
tct cca gtt att aac gtg atc att cca gaa gga tcc ggt tac aac aac	624
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn	
170 175 180 185	
act ttg gac cat ggt ctt tgt act gct ttc gaa gac tct acc cta ggt	672
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Asp Ser Thr Leu Gly	
190 195 200	
gac gac gtt gaa gct aac ttc act gct ttg ttc gct cca gct att aga	720
Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg	
205 210 215	
gct aga ttg gaa gct gac ttg cca ggt gtt act ttg act gac gaa gac	768
Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp	
220 225 230	
gtt gtt tac ttg atg gac atg tgt cca ttc gac act gtc gct aga act	816
Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	
tct gac gct act gaa ttg tct cca ttc tgt gct ttg ttc act cac gac	864
Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp	
250 255 260 265	
gaa tgg atc caa tac gac tac ttg caa agc ttg ggt aag tac tac ggt	912
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly	
270 275 280	
tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt ggt ttc gct	960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala	
285 290 295	
aac gaa ttg att gct aga ttg act cac tct cca gtt caa gac cac act	1008
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr	
300 305 310	
tct act aac cac act ttg gac tct aac cca gct act ttc cca ttg aac	1056
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn	
315 320 325	
gct act ttg tac gct gac ttc tct cac gac aac act atg ata tct att	1104
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile	
330 335 340 345	
ttc ttc gct ttg ggt ttg tac aac ggt acc aag cca ttg tct act act	1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr	
350 355 360	

```

tct gtt gaa tct att gaa gaa act gac ggt tac tct gct tct tgg act 1200
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
      365                      370                      375

gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa tgt caa gct 1248
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
      380                      385                      390

gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga gtt gtt cca 1296
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
      395                      400                      405

ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag aga gac gac 1344
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
      410                      415                      420                      425

ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac tgg gct gaa 1392
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
      430                      435                      440

tgt ttc gct taa 1404
Cys Phe Ala

```

<210> 93

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶-3-thermo-11-Q50T-K91A

<400> 93

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
      -20                      -15                      -10

```

```

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
      -5                      -1 1                      5

```

```

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
      10                      15                      20                      25

```

```

Gly Thr Tyr Ser Pro Tyr Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
      30                      35                      40

```

```

Pro Asp Val Pro Lys Asp Cys Arg Val Thr Phe Val Gln Val Leu Ser
      45                      50                      55

```

```

Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
      60                      65                      70

```

```

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
      75                      80                      85

```

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
 90 95 100 105
 Thr Pro Phe Gly Glu Asn Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
 110 115 120
 Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
 125 130 135
 Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
 140 145 150
 Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ser Gln Pro His Gln Ala
 155 160 165
 Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ser Gly Tyr Asn Asn
 170 175 180 185
 Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Asp Ser Thr Leu Gly
 190 195 200
 Asp Asp Val Glu Ala Asn Phe Thr Ala Leu Phe Ala Pro Ala Ile Arg
 205 210 215
 Ala Arg Leu Glu Ala Asp Leu Pro Gly Val Thr Leu Thr Asp Glu Asp
 220 225 230
 Val Val Tyr Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
 235 240 245
 Ser Asp Ala Thr Glu Leu Ser Pro Phe Cys Ala Leu Phe Thr His Asp
 250 255 260 265
 Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
 270 275 280
 Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Ala
 285 290 295
 Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
 300 305 310
 Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
 315 320 325
 Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Ile Ser Ile
 330 335 340 345
 Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
 350 355 360
 Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
 365 370 375

Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Gln Ala
380 385 390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
395 400 405

Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
410 415 420 425

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Ala Glu
430 435 440

Cys Phe Ala

<210> 94

<211> 1404

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶-10-thermo-5-Q50T

<220>

<221> 单肽

<222> (1)..(69)

<220>

<221> CDS

<222> (1)..(1401)

<220>

<221> 成熟肽

<222> (70)..(1401)

<400> 94

atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg ttc ggt tcc 48
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
-20 -15 -10

aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac tct tgt gac 96
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
-5 -1 1 5

act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct cac ttg tgg 144
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
10 15 20 25

ggg aca tac tct cca ttc ttc tct ttg gct gac gaa tct gct att tct 192
Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
30 35 40

cca gac gtt cca aag ggt tgt aga gtt act ttc gtt caa gtt ttg tct	240
Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser	
45 50 55	
aga cac ggt gct aga tac cca act tct tct aag tct aag gct tac tct	288
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Ala Tyr Ser	
60 65 70	
gct ttg att gaa gct att caa aag aac gct act gct ttc aag ggt aag	336
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys	
75 80 85	
tac gct ttc ttg aag act tac aat tac act ttg ggt gct gac gac ttg	384
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu	
90 95 100 105	
act cca ttc ggt gaa caa caa atg gtt aac tct ggt att aag ttc tac	432
Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr	
110 115 120	
aga aga tac aag gct ttg gct aga aag att gtt cca ttc att aga gct	480
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala	
125 130 135	
tct ggt tct gac aga gtt att gct tct gcc gaa aag ttc att gaa ggt	528
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly	
140 145 150	
ttc caa tct gct aag ttg gct gac cca ggt gct aac cca cac caa gct	576
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala	
155 160 165	
tct cca gtt att aac gtt att att cca gaa ggt gct ggt tac aac aac	624
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn	
170 175 180 185	
act ttg gac cac ggt ttg tgt act gct ttc gaa gaa tct acc cta ggt	672
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Thr Leu Gly	
190 195 200	
gac gac gtt gaa gct aac ttc act gct gtt ttc gct cca cca att aga	720
Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg	
205 210 215	
gct aga ttg gaa gct cac ttg cca ggt gtt aac ttg act gac gaa gac	768
Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp	
220 225 230	
gtt gtt aac ttg atg gac atg tgt cca ttc gac act gtt gct aga act	816
Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	

tct gac gct act caa ttg tct cca ttc tgt gac ttg ttc act cac gac	864
Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp	
250 255 260 265	
 gaa tgg att caa tac gac tac ttg caa tct ttg ggt aag tac tac ggt	912
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly	
270 275 280	
 tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt ggt ttc gtt	960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val	
285 290 295	
 aac gaa ttg att gct aga ttg act cac tct cca gtt caa gac cac act	1008
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr	
300 305 310	
 tct act aac cac act ttg gac tct aac cca gct act ttc cca ttg aac	1056
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn	
315 320 325	
 gct act ttg tac gct gac ttc tct cac gac aac act atg gtt tct att	1104
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile	
330 335 340 345	
 ttc ttc gct ttg ggt ttg tac aac ggt act aag cca ttg tct act act	1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr	
350 355 360	
 tct gtt gaa tct att gaa gaa act gac ggt tac tct gct tct tgg act	1200
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr	
365 370 375	
 gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa tgt gaa gct	1248
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala	
380 385 390	
 gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga gtt gtt cca	1296
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro	
395 400 405	
 ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag aga gac gac	1344
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp	
410 415 420 425	
 ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac tgg gaa gaa	1392
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu	
430 435 440	
 tgt ttc gct taa	1404
Cys Phe Ala	

<210> 95

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶-10-thermo-5-Q50T

<400> 95

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
 -20 -15 -10

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
 -5 -1 1 5

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
 10 15 20 25

Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
 30 35 40

Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser
 45 50 55

Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Lys Ser Lys Ala Tyr Ser
 60 65 70

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
 75 80 85

Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
 90 95 100 105

Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
 110 115 120

Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
 125 130 135

Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
 140 145 150

Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
 155 160 165

Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
 170 175 180 185

Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Thr Leu Gly
 190 195 200

Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
 205 210 215

Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
 220 225 230

```

Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
 235                240                245

Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
250                255                260                265

Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
                270                275                280

Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
                285                290                295

Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
 300                305                310

Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
 315                320                325

Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
330                335                340                345

Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
                350                355                360

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
                365                370                375

Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
 380                385                390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
 395                400                405

Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
410                415                420                425

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
                430                435                440

Cys Phe Ala

```

<210> 96

<211> 1404

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：共有植酸酶-10-thermo-5-Q50T-K91A

<220>

<221> 单肽

<222> (1)..(69)

<220>

<221> CDS

<222> (1)..(1401)

<220>

<221> 成熟肽

<222> (70)..(1401)

<400> 96

```
atg ggc gtg ttc gtc gtg cta ctg tcc att gcc acc ttg ttc ggt tcc 48
Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
      -20                -15                -10
```

```
aca tcc ggt acc gcc ttg ggt cct cgt ggt aat tct cac tct tgt gac 96
Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
      -5                -1 1                5
```

```
act gtt gac ggt ggt tac caa tgt ttc cca gaa att tct cac ttg tgg 144
Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
      10                15                20                25
```

```
ggg aca tac tct cca ttc ttc tct ttg gct gac gaa tct gct att tct 192
Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
      30                35                40
```

```
cca gac gtt cca aag ggt tgt aga gtt act ttc gtt caa gtt ttg tct 240
Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser
      45                50                55
```

```
aga cac ggt gct aga tac cca act tct tct gcg tct aag gct tac tct 288
Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
      60                65                70
```

```
gct ttg att gaa gct att caa aag aac gct act gct ttc aag ggt aag 336
Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
      75                80                85
```

```
tac gct ttc ttg aag act tac aat tac act ttg ggt gct gac gac ttg 384
Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
      90                95                100                105
```

```
act cca ttc ggt gaa caa caa atg gtt aac tct ggt att aag ttc tac 432
Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
      110                115                120
```

```
aga aga tac aag gct ttg gct aga aag att gtt cca ttc att aga gct 480
Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
      125                130                135
```

tct ggt tct gac aga gtt att gct tct gcc gaa aag ttc att gaa ggt	528
Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly	
140 145 150	
ttc caa tct gct aag ttg gct gac cca ggt gct aac cca cac caa gct	576
Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala	
155 160 165	
tct cca gtt att aac gtt att att cca gaa ggt gct ggt tac aac aac	624
Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn	
170 175 180 185	
act ttg gac cac ggt ttg tgt act gct ttc gaa gaa tct acc cta ggt	672
Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Thr Leu Gly	
190 195 200	
gac gac gtt gaa gct aac ttc act gct gtt ttc gct cca cca att aga	720
Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg	
205 210 215	
gct aga ttg gaa gct cac ttg cca ggt gtt aac ttg act gac gaa gac	768
Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp	
220 225 230	
gtt gtt aac ttg atg gac atg tgt cca ttc gac act gtt gct aga act	816
Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr	
235 240 245	
tct gac gct act caa ttg tct cca ttc tgt gac ttg ttc act cac gac	864
Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp	
250 255 260 265	
gaa tgg att caa tac gac tac ttg caa tct ttg ggt aag tac tac ggt	912
Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly	
270 275 280	
tac ggt gct ggt aac cca ttg ggt cca gct caa ggt gtt ggt ttc gtt	960
Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val	
285 290 295	
aac gaa ttg att gct aga ttg act cac tct cca gtt caa gac cac act	1008
Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr	
300 305 310	
tct act aac cac act ttg gac tct aac cca gct act ttc cca ttg aac	1056
Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn	
315 320 325	
gct act ttg tac gct gac ttc tct cac gac aac act atg gtt tct att	1104
Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile	
330 335 340 345	

```

ttc ttc gct ttg ggt ttg tac aac ggt act aag cca ttg tct act act 1152
Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
      350                      355                      360

tct gtt gaa tct att gaa gaa act gac ggt tac tct gct tct tgg act 1200
Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
      365                      370                      375

gtt cca ttc gct gct aga gct tac gtt gaa atg atg caa tgt gaa gct 1248
Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
      380                      385                      390

gaa aag gaa cca ttg gtt aga gtt ttg gtt aac gac aga gtt gtt cca 1296
Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
      395                      400                      405

ttg cac ggt tgt gct gtt gac aag ttg ggt aga tgt aag aga gac gac 1344
Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
410                      415                      420                      425

ttc gtt gaa ggt ttg tct ttc gct aga tct ggt ggt aac tgg gaa gaa 1392
Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
      430                      435                      440

tgt ttc gct taa 1404
Cys Phe Ala

```

<210> 97

<211> 467

<212> PRT

<213> 人工序列

<223> 人工序列的描述：共有植酸酶-10-thermo-5-Q50T-K91A

<400> 97

```

Met Gly Val Phe Val Val Leu Leu Ser Ile Ala Thr Leu Phe Gly Ser
      -20                      -15                      -10

```

```

Thr Ser Gly Thr Ala Leu Gly Pro Arg Gly Asn Ser His Ser Cys Asp
      -5                      -1 1                      5

```

```

Thr Val Asp Gly Gly Tyr Gln Cys Phe Pro Glu Ile Ser His Leu Trp
10                      15                      20                      25

```

```

Gly Thr Tyr Ser Pro Phe Phe Ser Leu Ala Asp Glu Ser Ala Ile Ser
      30                      35                      40

```

```

Pro Asp Val Pro Lys Gly Cys Arg Val Thr Phe Val Gln Val Leu Ser
      45                      50                      55

```

```

Arg His Gly Ala Arg Tyr Pro Thr Ser Ser Ala Ser Lys Ala Tyr Ser
      60                      65                      70

```

Ala Leu Ile Glu Ala Ile Gln Lys Asn Ala Thr Ala Phe Lys Gly Lys
 75 80 85
 Tyr Ala Phe Leu Lys Thr Tyr Asn Tyr Thr Leu Gly Ala Asp Asp Leu
 90 95 100 105
 Thr Pro Phe Gly Glu Gln Gln Met Val Asn Ser Gly Ile Lys Phe Tyr
 110 115 120
 Arg Arg Tyr Lys Ala Leu Ala Arg Lys Ile Val Pro Phe Ile Arg Ala
 125 130 135
 Ser Gly Ser Asp Arg Val Ile Ala Ser Ala Glu Lys Phe Ile Glu Gly
 140 145 150
 Phe Gln Ser Ala Lys Leu Ala Asp Pro Gly Ala Asn Pro His Gln Ala
 155 160 165
 Ser Pro Val Ile Asn Val Ile Ile Pro Glu Gly Ala Gly Tyr Asn Asn
 170 175 180 185
 Thr Leu Asp His Gly Leu Cys Thr Ala Phe Glu Glu Ser Thr Leu Gly
 190 195 200
 Asp Asp Val Glu Ala Asn Phe Thr Ala Val Phe Ala Pro Pro Ile Arg
 205 210 215
 Ala Arg Leu Glu Ala His Leu Pro Gly Val Asn Leu Thr Asp Glu Asp
 220 225 230
 Val Val Asn Leu Met Asp Met Cys Pro Phe Asp Thr Val Ala Arg Thr
 235 240 245
 Ser Asp Ala Thr Gln Leu Ser Pro Phe Cys Asp Leu Phe Thr His Asp
 250 255 260 265
 Glu Trp Ile Gln Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly
 270 275 280
 Tyr Gly Ala Gly Asn Pro Leu Gly Pro Ala Gln Gly Val Gly Phe Val
 285 290 295
 Asn Glu Leu Ile Ala Arg Leu Thr His Ser Pro Val Gln Asp His Thr
 300 305 310
 Ser Thr Asn His Thr Leu Asp Ser Asn Pro Ala Thr Phe Pro Leu Asn
 315 320 325
 Ala Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Thr Met Val Ser Ile
 330 335 340 345
 Phe Phe Ala Leu Gly Leu Tyr Asn Gly Thr Lys Pro Leu Ser Thr Thr
 350 355 360

Ser Val Glu Ser Ile Glu Glu Thr Asp Gly Tyr Ser Ala Ser Trp Thr
365 370 375

Val Pro Phe Ala Ala Arg Ala Tyr Val Glu Met Met Gln Cys Glu Ala
380 385 390

Glu Lys Glu Pro Leu Val Arg Val Leu Val Asn Asp Arg Val Val Pro
395 400 405

Leu His Gly Cys Ala Val Asp Lys Leu Gly Arg Cys Lys Arg Asp Asp
410 415 420 425

Phe Val Glu Gly Leu Ser Phe Ala Arg Ser Gly Gly Asn Trp Glu Glu
430 435 440

Cys Phe Ala

	1		50		
土曲霉 9A-1	KhsDCNSVDh	GYQCFPELSH	kWGLYAPYFS	LQDESPPFLD	VPEDChITFV
土曲霉 cbs	NhsDCTSVDr	GYQCFPELSH	kWGLYAPYFS	LQDESPPFLD	VPDDChITFV
黑曲霉泡盛变种	NqsTCDTVDQ	GYQCFSETSH	LWGQYAPFFS	LANESAISPD	VPAGCrVTFA
黑曲霉 T213	NqsSCDTVDQ	GYQCFSETSH	LWGQYAPFFS	LANESVISPD	VPAGCrVTFA
黑曲霉 NRRL3135	NqsSCDTVDQ	GYQCFSETSH	LWGQYAPFFS	LANESVISPE	VPAGCrVTFA
烟曲霉 13073	GSKSCDTVDl	GYQCsPATSH	LWGQYSPFFS	LEDElSVSSK	LPKDCrITLV
烟曲霉 32722	GSKSCDTVDl	GYQCsPATSH	LWGQYSPFFS	LEDElSVSSK	LPKDCrITLV
烟曲霉 58128	GSKSCDTVDl	GYQCsPATSH	LWGQYSPFFS	LEDElSVSSK	LPKDCrITLV
烟曲霉 26906	GSKSCDTVDl	GYQCsPATSH	LWGQYSPFFS	LEDElSVSSK	LPKDCrITLV
烟曲霉 32239	GSKACDTVEL	GYQCsPGTSH	LWGQYSPFFS	LEDElSVSSD	LPKDCrVTFV
E. nidulans	QNHSCNTADG	GYQCFPNVSH	VWGQYSPYFS	IEQESAISeD	VPHGCEvTFV
嗜热踝节菌	DSHSCNTVEG	GYQCrPEISH	sWGQYSPFFS	LADQSEISPD	VPQNCkITFV
嗜热毁丝霉	ESRPCDTpDl	GFQCgTAISH	FWGQYSPYFS	VpSElDaS..	IPDDCeVTFa
共有	NSHSCDTVDG	GYQCFPEISH	LWGQYSPYFS	LEDESAISPD	VPDDC-VTFV
共有植酸酶	NSHSCDTVDG	GYQCFPEISH	LWGQYSPYFS	LEDESAISPD	VPDDCrVTFV
	51		100		
土曲霉 9A-1	QVLARHGARS	PThSKtKAYA	AtIAAIQKSA	TaFpGKYAFL	QSYNYSLDSE
土曲霉 cbs	QVLARHGARS	PTDSKtKAYA	AtIAAIQKNA	TaLpGKYAFL	KSYNYSMGSE
黑曲霉泡盛变种	QVLSRHGARY	PTESKqKkYS	ALIEEIQQNV	TtFDGKYAFL	KTYNYSLGAD
黑曲霉 T213	QVLSRHGARY	PTESKqKkYS	ALIEEIQQNV	TtFDGKYAFL	KTYNYSLGAD
黑曲霉 NRRL3135	QVLSRHGARY	PTDSKqKkYS	ALIEEIQQNA	TtFDGKYAFL	KTYNYSLGAD
烟曲霉 13073	QVLSRHGARY	PTSSKsKkYK	kLVTAIQaNA	TdFKGKFAFL	KTYNYTLGAD
烟曲霉 32722	QVLSRHGARY	PTSSKsKkYK	kLVTAIQaNA	TdFKGKFAFL	KTYNYTLGAD
烟曲霉 58128	QVLSRHGARY	PTSSKsKkYK	kLVTAIQaNA	TdFKGKFAFL	KTYNYTLGAD
烟曲霉 26906	QVLSRHGARY	PTSSKsKkYK	kLVTAIQaNA	TdFKGKFAFL	KTYNYTLGAD
烟曲霉 32239	QVLSRHGARY	PTASKsKkYK	kLVTAIQKNA	TeFKGKFAFL	ETNYNYTLGAD
E. nidulans	QVLSRHGARY	PTESKsKAYS	GLIEAIQKNA	TsFwGQYAFI	ESYNYTLGAD
嗜热踝节菌	QVLSRHGARY	PTSSKtELYS	QLISrIQKTA	TaYKGyYAFL	KDYrYqLGAN
嗜热毁丝霉	QVLSRHGARA	PTlKRAaSvY	DLIDrIHhGA	IsYgPgYEFL	RTYDYTLGAD
共有	QVLSRHGARY	PTSSK-KAYS	ALIEAIQKNA	T-FKGKYAFL	KTYNYTLGAD
共有植酸酶	QVLSRHGARY	PTSSKSKAYS	ALIEAIQKNA	TAFKGKYAFL.	KTYNYTLGAD
	101		150		
土曲霉 9A-1	ELTPFGrNQL	rDlGaQFYeR	YNALTRhInP	FVRATDASRV	hESAekFVEG
土曲霉 cbs	NLTPFGrNQL	qDlGaQFYRR	YDTLTRhInP	FVRAADSSRV	hESAekFVEG
黑曲霉泡盛变种	DLTPFGEQEL	VNSGIKFYQR	YESLTRNIIP	FIRSSGSSRV	IASGEKFIEG
黑曲霉 T213	DLTPFGEQEL	VNSGIKFYQR	YESLTRNIIP	FIRSSGSSRV	IASGEKFIEG
黑曲霉 NRRL3135	DLTPFGEQEL	VNSGIKFYQR	YESLTRNIIP	FIRSSGSSRV	IASGEKFIEG
烟曲霉 13073	DLTPFGEQQL	VNSGIKFYQR	YKALARSVVP	FIRASGSDRV	IASGEKFIEG
烟曲霉 32722	DLTPFGEQQL	VNSGIKFYQR	YKALARSVVP	FIRASGSDRV	IASGEKFIEG
烟曲霉 58128	DLTPFGEQQL	VNSGIKFYQR	YKALARSVVP	FIRASGSDRV	IASGEKFIEG
烟曲霉 26906	DLTPFGEQQL	VNSGIKFYQR	YKALARSVVP	FIRASGSDRV	IASGEKFIEG
烟曲霉 32239	DLTPFGEQQM	VNSGIKFYQK	YKALAgSVVP	FIRSSGSDRV	IASGEKFIEG
E. nidulans	DLTiFGENQM	VDSGaKFYRR	YKNLARKnTP	FIRASGSDRV	VASAekFING
嗜热踝节菌	DLTPFGENQM	IQlGIKFYnH	YKSLARNaVP	FVRCsGSDRV	IASGrIFIEG
嗜热毁丝霉	ELTRtGQQQM	VNSGIKFYRR	YRALARKsIP	FVRTAGqDRV	VhSAENFTQG
共有	DLTPFGENQM	VNSGIKFYRR	YKALARK-VP	FVRASGSDRV	IASAEKFIEG
共有植酸酶	DLTPFGENQM	VNSGIKFYRR	YKALARKIVP	FIRASGSDRV	IASAEKFIEG

图 1a

土曲霉 9A-1	151	FQTARqDDHh	ANpHQPSPrV	DVaIPEGSAY	NNTLEHS1CT	AFES...STV
土曲霉 cbs		FQNARqGDPh	ANpHQPSPrV	DVVIPEGTAY	NNTLEHS1CT	AFEA...STV
黑曲霉 泡盛变种		FQSTKLkDPr	AqpgQSSPkI	DVVISEASSs	NNTLDPGTCT	VFED...SEL
黑曲霉 T213		FQSTKLkDPr	AqpgQSSPkI	DVVISEASSs	NNTLDPGTCT	VFED...SEL
黑曲霉 NRRL3135		FQSTKLkDPr	AqpgQSSPkI	DVVISEASSs	NNTLDPGTCT	VFED...SEL
烟曲霉 13073		FQqAKLADPG	A.TNRAAPAI	SVIIPESETF	NNTLDHGVCT	kFEA...SQL
烟曲霉 32722		FQqAKLADPG	A.TNRAAPAI	SVIIPESETF	NNTLDHGVCT	kFEA...SQL
烟曲霉 58128		FQqAKLADPG	A.TNRAAPAI	SVIIPESETF	NNTLDHGVCT	kFEA...SQL
烟曲霉 26906		FQqAKLADPG	A.TNRAAPAI	SVIIPESETF	NNTLDHGVCT	kFEA...SQL
烟曲霉 32239		FQqANVADPG	A.TNRAAPVI	SVIIPESETY	NNTLDHSVCT	NFEA...SEL
<i>E. nidulans</i>		FRKAQLHDHG	S.gQATPVV	NVIIPEIdGF	NNTLDHSTCV	SPEN...DER
嗜热蹼节菌		FQSAKVIDPh	SDkHDAPPTI	NVIIeEGPSY	NNTLDtGSCP	VFED...SSg
嗜热毁丝霉		FHSALLADRG	StvRPTlPyd	mVVIPETAGa	NNTLHND1CT	AFEEgpySTI
共有		FQSAKLADPG	S-PHQASPMI	NVIIPEGSY	NNTLDHGTCT	AFED---SEL
共有植酸酶		FQSAKLADPG	SQPHQASPMI	DVIIPEGSY	NNTLDHGTCT	AFED...SEL
土曲霉 9A-1	201	GDDAVANFTA	VFAPAIaQRL	EADLPGVqLS	TDDVVnLMAM	CPFETVS1TD
土曲霉 cbs		GDAAADNFTA	VFAPAIakRL	EADLPGVqLS	ADDVVnLMAM	CPFETVS1TD
黑曲霉 泡盛变种		ADTVEANFTA	TFAPSIRQRL	ENDLSGVTLT	DTEVTyLMMD	CSFDTIS1ST
黑曲霉 T213		ADTVEANFTA	TFAPSIRQRL	ENDLSGVTLT	DTEVTyLMMD	CSFDTIS1ST
黑曲霉 NRRL3135		ADTVEANFTA	TFVPSIRQRL	ENDLSGVTLT	DTEVTyLMMD	CSFDTIS1ST
烟曲霉 13073		GDEVAANFTA	lFAPDIRara	EkHLPGVTLT	DEDVVslMDM	CSFDTVARTS
烟曲霉 32722		GDEVAANFTA	lFAPDIRara	EkHLPGVTLT	DEDVVslMDM	CSFDTVARTS
烟曲霉 58128		GDEVAANFTA	lFAPDIRara	EkHLPGVTLT	DEDVVslMDM	CSFDTVARTS
烟曲霉 26906		GDEVAANFTA	lFAPDIRara	EkHLPGVTLT	DEDVVslMDM	CSFDTVARTS
烟曲霉 32239		GDEVEANFTA	lFAPAIRARI	EkHLPGVqLT	DDDVVslMDM	CSFDTVARTA
<i>E. nidulans</i>		ADEiEANFTA	IMGPPIRkRL	ENDLPGIKLT	NENViyLMMD	CSFDTMARTA
嗜热蹼节菌		GHDAOEKFAK	qFAPAIleKI	KDHLPGVDLA	vSDVpyLMDL	CPFETLARNh
嗜热毁丝霉		GDDAQDTY1S	TFAGPitARV	NANLPGANLT	DADTVaLMDL	CPFETVAsSS
共有		GDDAEANFTA	TFAPAIRARL	EADLPGVTLT	DEDVV-LMDM	CPFETVARTS
共有植酸酶		GDDVEANFTA	LFAPAIRARL	EADLPGVTLT	DEDVVYLMDM	CPFETVARTS
土曲霉 9A-1	251		..DAhTLSPFC	DLFTateWtq	YNYL1SLDKY	YGYGGGNPLG
土曲霉 cbs			..DAhTLSPFC	DLFTAaEWtq	YNYL1SLDKY	YGYGGGNPLG
黑曲霉 泡盛变种			..vDTKLSPFC	DLFTHdEWih	YDYLQSLkKY	YGHGAGNPLG
黑曲霉 T213			..vDTKLSPFC	DLFTHdEWih	YDYLRLSLkKY	YGHGAGNPLG
黑曲霉 NRRL3135			..vDTKLSPFC	DLFTHdEWin	YDYLQSLkKY	YGHGAGNPLG
烟曲霉 13073			..DASQLSPFC	QLFTHnEWkk	YNYLQSLGKY	YGYGAGNPLG
烟曲霉 32722			..DASQLSPFC	QLFTHnEWkk	YNYLQSLGKY	YGYGAGNPLG
烟曲霉 58128			..DASQLSPFC	QLFTHnEWkk	YNYLQSLGKY	YGYGAGNPLG
烟曲霉 26906			..DASQLSPFC	QLFTHnEWkk	YNYLQSLGKY	YGYGAGNPLG
烟曲霉 32239			..DASELSPFC	AlFTHnEWkk	YDYLQSLGKY	YGYGAGNPLG
<i>E. nidulans</i>			..HGTELSPFC	AlFTEkEW1q	YDYLQSLSKY	YGYGAGSP1G
嗜热蹼节菌			..TDT.LSPFC	ALsTQeEWqa	YDYYQSLGKY	YGNGGGNPLG
嗜热毁丝霉		sdpatadagg	gNGrPLSPFC	rLFSEsEWra	YDYLQSVGKW	YGYGPGNPLG
共有		-----	-DATELSPFC	ALFTE-EW--	YDYLQSLGKY	YGYGAGNPLG
共有植酸酶			..DATELSPFC	ALFTHdEWRQ	YDYLQSLGKY	YGYGAGNPLG

图 1b

土曲霉 9A-1	301	PVQGVGWaNE	LMARLTRAPV	HDHTCVNNTL	DASPATFPLN	ATLYADFSHD	350
土曲霉 cbs		PVQGVGWaNE	LIARLTRSPV	HDHTCVNNTL	DANPATFPLN	ATLYADFSHD	
黑曲霉 泡盛变种		PTQGVGYaNE	LIARLTHSPV	HDDTSSNHTL	DSNPATFPLN	STLYADFSHD	
黑曲霉 T213		PTQGVGYaNE	LIARLTHSPV	HDDTSSNHTL	DSNPATFPLN	STLYADFSHD	
黑曲霉 NRRL3135		PTQGVGYaNE	LIARLTHSPV	HDDTSSNHTL	DSSPATFPLN	STLYADFSHD	
烟曲霉 13073		PAQGIGFtNE	LIARLTRSPV	QDHTSTNsTL	vSNPATFPLN	ATMYVDFSHD	
烟曲霉 32722		PAQGIGFtNE	LIARLTRSPV	QDHTSTNsTL	vSNPATFPLN	ATMYVDFSHD	
烟曲霉 58128		PAQGIGFtNE	LIARLTRSPV	QDHTSTNsTL	vSNPATFPLN	ATMYVDFSHD	
烟曲霉 26906		PAQGIGFtNE	LIARLTRSPV	QDHTSTNsTL	vSNPATFPLN	ATMYVDFSHD	
烟曲霉 32239		PAQGIGFtNE	LIARLTNSPV	QDHTSTNsTL	DSDPATFPLN	ATIYVDFSHD	
E. nidulans		PAQGIGFtNE	LIARLTQSPV	QDNTSTNHTL	DSNPATFPLD	rKLYADFSHD	
嗜热踝节菌		PAQGVGFvNE	LIARMTTHSPV	QDYTTVNHTL	DSNPATFPLN	ATLYADFSHD	
嗜热毁丝霉		PTQGVGFvNE	LLARLAgvPV	RDgTSTNRTL	DGDPrTFPLG	rPLYADFSHD	
共有		PAQGVGF-NE	LIARLTHSPV	QDHTSTNHTL	DSNPATFPLN	ATLYADFSHD	
共有植酸酶		PAQGVGFANE	LIARLTRSPV	QDHTSTNHTL	DSNPATFPLN	ATLYADFSHD	
土曲霉 9A-1	351	SNLVSIFWAL	GLYNGTAPLS	qTSVESVSQT	DGYAAAWTVP	FAARAYVEMM	400
土曲霉 cbs		SNLVSIFWAL	GLYNGTkPLS	qTTVEDITYT	DGYAAAWTVP	FAARAYVEMM	
黑曲霉 泡盛变种		NGIISILFAL	GLYNGTkPLS	TTTVENITQT	DGFSSAWTVP	FASRLYVEMM	
黑曲霉 T213		NGIISILFAL	GLYNGTkPLS	TTTVENITQT	DGFSSAWTVP	FASRLYVEMM	
黑曲霉 NRRL3135		NGIISILFAL	GLYNGTkPLS	TTTVENITQT	DGFSSAWTVP	FASRLYVEMM	
烟曲霉 13073		NSMVSIFFAL	GLYNGTEPLS	rTSVESaKEl	DGYSASWVVP	FGARAYFetM	
烟曲霉 32722		NSMVSIFFAL	GLYNGTEPLS	rTSVESaKEl	DGYSASWVVP	FGARAYFetM	
烟曲霉 58128		NSMVSIFFAL	GLYNGTEPLS	rTSVESaKEl	DGYSASWVVP	FGARAYFetM	
烟曲霉 26906		NSMVSIFFAL	GLYNGTEPLS	rTSVESaKEl	DGYSASWVVP	FGARAYFetM	
烟曲霉 32239		NGMIPiFFAM	GLYNGTEPLS	qTSeESTKES	NGYSASWAVP	FGARAYFetM	
E. nidulans		NSMISiFFAM	GLYNGTQPLS	mDSVESIQEm	DGYAASWTVP	FGARAYFELM	
嗜热踝节菌		NTMTSiFaAL	GLYNGTAKLS	TTEIKSiEET	DGYAASWTVP	FGGRAYIEMM	
嗜热毁丝霉		NDMMGVLGAL	GayDGVPPLD	KTArTdPeEl	GGYAASWAVP	FAARiYVEKM	
共有		NSMISiFFAL	GLYNGTAPLS	TTSVESiEET	DGYAASWTVP	FGARAYVEMM	
共有植酸酶		NSMISiFFAL	GLYNGTAPLS	TTSVESiEET	DGYAASWTVP	FGARAYVEMM	
土曲霉 9A-1	401	QC.....	RAEKE	PLVRVLVNDR	VMPLHGCPID	KLGRCKrDAF	450
土曲霉 cbs		QC.....	RAEKQ	PLVRVLVNDR	VMPLHGCAVD	NLGRCKrDDF	
黑曲霉 泡盛变种		QC.....	QAEQE	PLVRVLVNDR	VVPLHGCPID	aLGRCTrDSF	
黑曲霉 T213		QC.....	QAEQE	PLVRVLVNDR	VVPLHGCPID	aLGRCTrDSF	
黑曲霉 NRRL3135		QC.....	QAEQE	PLVRVLVNDR	VVPLHGCPVD	aLGRCTrDSF	
烟曲霉 13073		QC.....	KSEKE	PLVRALINDR	VVPLHGCDVD	KLGRCKLNDf	
烟曲霉 32722		QC.....	KSEKE	PLVRALINDR	VVPLHGCDVD	KLGRCKLNDf	
烟曲霉 58128		QC.....	KSEKE	PLVRALINDR	VVPLHGCDVD	KLGRCKLNDf	
烟曲霉 26906		QC.....	KSEKE	PLVRALINDR	VVPLHGCAVD	KLGRCKLKDF	
烟曲霉 32239		QC.....	E.KE	PLVRVLVNDR	VVPLHGCAVD	KFGRCTLDfW	
E. nidulans		QC.....	DDSDE	FVVRVLVNDR	VVPLHGCEVD	SLGRCKrDDF	
嗜热踝节菌		RCsgggggggg	gggrQEKE	eMVRVLVNDR	VMTLkGCGAD	ErGMCTLErF	
嗜热毁丝霉		QC-----	QAEKE	PLVRVLVNDR	VVPLHGCAVD	KLGRCKLDDF	
共有		QC.....	QAEKE	PLVRVLVNDR	VVPLHGCAVD	KLGRCKRDDF	
共有植酸酶							

图 1c

	451	471
土曲霉 9A-1	VAGLSFAQAG	GNWADCF--- -
土曲霉 cbs	VEGLSFARAG	NWAECF---
黑曲霉 泡盛变种	VrGLSFARSG	GDWAECsA-- -
黑曲霉 T213	VrGLSFARSG	GDWAECFA-- -
黑曲霉 NRRL3135	VrGLSFARSG	DWAECFA--
烟曲霉 13073	VKGLSWARSG	GNWGECFS-- -
烟曲霉 32722	VKGLSWARSG	GNWGECFS-- -
烟曲霉 58128	VKGLSWARSG	GNWGECFS-- -
烟曲霉 26906	VKGLSWARSG	GNWGECFS-- -
烟曲霉 32239	VKGLSWARSG	NSEQSFS--
E. nidulans	VEGLNFARSG	GNWkTCFT1- -
嗜热蹠节菌	VrGLSFARqG	GNWEGCYAas e
嗜热毁丝霉	IESMAFARGN	GKWD1CFA-- -
共有	VEGLSFARSG	GNWAECFA-- -
共有植酸酶	VEGLSFARSG	GNWAECFA.. .

图 1d

CP-1
 Eco RI M G V F V V L L S I A T L F G S T
 TATATGAATTCATGGGCGTGTTCGTCTGCTACTGTCCATTGCCACCTTGTTCCGGTTCCA
 1 -----+-----+-----+-----+-----+-----+ 60
 ATATACTTAAGTACCCGCACAAGCAGCAGCATGACAGGTAACGGTGGAACAAGCCAAGGT

 S G T A L G P R G N S H S C D T V D G G
 CATCCGGTACCGCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG
 61 -----+-----+-----+-----+-----+-----+ 120
 GTAGGCCATGGCGGAACCCAGGAGCACCATTAAAGAGTGAGAACACTGTGACAACTGCCAC
 CP-2
 CP-3
 Y Q C F P E I S H L W G Q Y S P Y F S L
 GTTACCAATGTTTCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT
 121 -----+-----+-----+-----+-----+-----+ 180
 CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCAGTTATGAGAGGTATGAAGAGAA

 E D E S A I S P D V P D D C R V T F V Q
 TGGAAGACGAATCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCTGTT
 181 -----+-----+-----+-----+-----+-----+ 240
 ACCTTCTGCTTAGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG
 CP-4
 CP-5
 V L S R H G A R Y P T S S K S K A Y S A
 AAGTTTGTCTAGACACGGTGCTAGATACCCAACTTCTTCTAAGTCTAAGGCTTACTCTG
 241 -----+-----+-----+-----+-----+-----+ 300
 TTCAAACAGATCTGTGCCACGATCTATGGGTTGAAGAAGATTGAGATTCCGAATGAGAC

 L I E A I Q K N A T A F K G K Y A F L K
 CTTTGATTGAAGCTATTCAAAGAACGCTACTGCTTCAAGGGTAAGTACGCTTTCTTGA
 301 -----+-----+-----+-----+-----+-----+ 360
 GAACTAAGTTCGATAAGTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAAGT
 CP-6
 CP-7
 T Y N Y T L G A D D L T P F G E N Q M V
 AGACTTACAACTACACTTTGGGTGCTGACGACTTGACTCCATTGCGTGAAAACCAAATGG
 361 -----+-----+-----+-----+-----+-----+ 420
 TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACC

 N S G I K F Y R R Y K A L A R K I V P F
 TTAAGTCTGGTATTAAAGTTCTACAGAAGATAAAGGCTTTGGCTAGAAAGATTGTTCCAT
 421 -----+-----+-----+-----+-----+-----+ 480
 AATTGAGACCATAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTA
 CP-8
 CP-9
 I R A S G S D R V I A S A E K F I E G F
 TCATTAGAGCTTCTGGTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTT
 481 -----+-----+-----+-----+-----+-----+ 540
 AGTAATCTCGAAGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAAGTTCCAA

 Q S A K L A D P G S Q P H Q A S P V I D
 TCCAATCTGCTAAGTTGGCTGACCCAGGTTCTCAACCACCAAGCTTCTCCAGTTATTG
 541 -----+-----+-----+-----+-----+-----+ 600
 AGGTTAGACGATTCAACCGACTGGGTCCAAGAGTTGGTGTGTTTGAAGAGGTCAATAAC

图 2a

CP-10

CP-11

V I I P E G S G Y N N T L D H G T C T A
 ACGTTATTATTCCAGAAGGaTCCGGTTACAACAACACTTTGGACCACGGTACTTGTACTG
 601 -----+-----+-----+-----+-----+ 660
 TGCAATAATAAGGTCTTCCTAGgCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGAC

F E D S E L G D D V E A N F T A L F A P
 CTTTCGAAGACTCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTGCTC
 661 -----+-----+-----+-----+-----+ 720
 GAAAGCTTCTGAGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAG

CP-12

A I R A R L E A D L P G V T L T D E D V
 CAGCTATTAGAGCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACG
 721 -----+-----+-----+-----+-----+ 780
 GTCGATAATCTCGATCTAACCTTCGACTGAACGGTCCACAATGAACTGACTGCTTCTGC

CP-13

V Y L M D M C P F E T V A R T S D A T E
 TTGTTTACTTGATGGACATGTGTCCATTGCAAACCTGTTGCTAGAACTTCTGACGCTACTG
 781 -----+-----+-----+-----+-----+ 840
 AACAAATGAACTACCTGTACACAGGTAAGCTTTGACAACGATCTTGAAGACTGCGATGAC

L S P F C A L F T H D E W R Q Y D Y L Q
 AATTGTCTCCATTCTGTGCTTTGTTCACTCACGACGAATGGAGACAATACGACTACTTGC
 841 -----+-----+-----+-----+-----+ 900
 TTAACAGAGGTAAAGACACGAAACAAGTGAGTGCTGCTTACCTCTGTTATGCTGATGAACG

CP-14

CP-15

S L G K Y Y G Y G A G N P L G P A Q G V
 AATCTTTGGGTAAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTG
 901 -----+-----+-----+-----+-----+ 960
 TTAGAAACCCATTTCATGATGCCAATGCCACGACCATTTGGGTAACCCAGGTGAGTTCCAC

G F A N E L I A R L T R S P V Q D H T S
 TTGGTTTCGCTAACGAATTGATTGCTAGATTGACTAGATCTCCAGTTCAAGACCACACTT
 961 -----+-----+-----+-----+-----+ 1020
 AACCAAAGCGATTGCTTAACCTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA

CP-16

CP-17

T N H T L D S N P A T F P L N A T L Y A
 CTACTAACACACTTTGGACTCTAACCCAGCTACTTTCCCATTTGAACGCTACTTTGTACG
 1021 -----+-----+-----+-----+-----+ 1080
 GATGATTGGTGTGAAACCTGAGATTGGGTGATGAAAGGGTAACCTTGCATGAAACATGC

D F S H D N S M I S I F F A L G L Y N G
 CTGACTTCTCTCACGACAACCTCTATGATTTCTATTTCTTCGCTTTGGGTTTGTACAACG
 1081 -----+-----+-----+-----+-----+ 1140
 GACTGAAGAGAGTGCTGTTGAGATACTAAAGATAAAAGAAGCGAAACCCAAACATGTTGC

CP-18

CP-19

T A P L S T T S V E S I E E T D G Y S A
 GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTG
 1141 -----+-----+-----+-----+-----+ 1200
 CATGACGAGGTAACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGAC

图 2b

```

      S W T V P F G A R A Y V E M M Q C Q A E
CTTCTTGGACTGTTCCATTGGGTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG
1201 -----+-----+-----+-----+-----+ 1260
      GAAGAACCTGACAAGGTAAGCCACGATCTCGAATGCAACTTTACTACGTTACAGTTCGAC
                                CP-20
                                CP-21
      K E P L V R V L V N D R V V P L H G C A
AAAAGGAACCATTTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
1261 -----+-----+-----+-----+-----+ 1320
      TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC

      V D K L G R C K R D D F V E G L S F A R
CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA
1321 -----+-----+-----+-----+-----+ 1380
      GACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGAT
                                CP-22
      S G G N W A E C F A * Eco RI
GATCTGGTGGTAACTGGGCTGAATGTTTCGCTTAAGAATTCATATA
1381 -----+-----+-----+-----+----- 1426
      CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT

```

图 2c

	1				50
<i>P. involutus</i> (phyA1)	SvP.KnTAPt	FPIPeseQrn	WSPYSPYFPL	AeYkAPPAGC	QInQVNIQR
<i>P. involutus</i> (phyA2)	SvP.RniAPK	FSIPeseQrn	WSPYSPYFPL	AeYkAPPAGC	EInQVNIQR
绒毛栓菌	hiPlRdTSAc	LdVTrDvQqs	WSmYSPYFPa	AtYvAPPASC	QInQVHIIQR
平田头菇	GgvvQaTfvQ	pfFPpQiQds	WAAYTPYYPV	qaYtPPPkDC	KitQVNIQR
<i>P. lycii</i>	StQfsfvAAQ	LPIPaQntsn	WGPYdPFFPV	EpYaAPPEGC	tVtQVNIQR
Basidio	S-P-R-TAAQ	LPIP-Q-Q--	WSPYSPYFPV	A-Y-APPAGC	QI-QVNIQR
	51				100
<i>P. involutus</i> (phyA1)	HGARFPTSGA	TTRIKAGLTK	LQVqnfTDA	KFNFIksfky	dLGnsDLVFP
<i>P. involutus</i> (phyA2)	HGARFPTSGA	ATRIKAGLSK	LQSVqnfTDP	KFDFIksfTY	dLGtsDLVFP
绒毛栓菌	HGARFPTSGA	AkRIQTAVAK	LKAAsnyTDP	lLAFVtNyTY	sLGqDsLVeL
平田头菇	HGARFPTSGA	GTRIQAAsvK	LQSAktyTDP	RLDFltNyTY	tLGhDDLVPF
<i>P. lycii</i>	HGARWPTSGA	rSRqvaAAVAK	IQmArpfTDP	KYEFLnDfvY	kFGvADLLPF
Basidio	HGARFPTSGA	ATRIQAAsvK	LQSA---TDP	KLDfL-N-TY	-LG-DDLVFP
	101				150
<i>P. involutus</i> (phyA1)	GAaQSfDAGQ	EAFARYSkLV	SkNNLPFIRA	dGSDRVVDSA	TNWTAGFAsA
<i>P. involutus</i> (phyA2)	GAaQSfDAGl	EvFARYSkLV	SsDNLPFIRS	dGSDRVVDTA	TNWTAGFAsA
绒毛栓菌	GATQSSEAGQ	EAFTRYsLV	saDELpFVRA	SGSDRVVATA	nNWTAGFAlA
平田头菇	GALQSSQAGE	ETfQrYSfLV	SkENLPFVRA	SSSNRVVDSA	TNWTGEGFSaA
<i>P. lycii</i>	GAnQShQTGt	DmYTRYStLf	egGDVPFVRA	AGdQRVVDSS	TNWTAGFGdA
Basidio	GA-QSSQAGQ	EAFTRYs-LV	S-DNLpFVRA	SGSDRVVDSA	TNWTAGFA-A
	151				200
<i>P. involutus</i> (phyA1)	ShNTvqPkLn	LILPQtGNDT	LEDNMCPaAG	DSDPQvNaWL	AVafPSITAR
<i>P. involutus</i> (phyA2)	SrNAiqPkLd	LILPQtGNDT	LEDNMCPaAG	ESDPQvDaWL	AsafPSVTAQ
绒毛栓菌	SsNSitPvLs	VIIEAgNDT	LDDNMCPaAG	DSDPQvNqWL	AqFAPPMTAR
平田头菇	ShHvlnPiLf	VILSEsLNDT	LDDaMCPnAG	sSDPQtGiWt	SIYGTPIAnR
<i>P. lycii</i>	SgETvlPtLq	VVLqEeGNcT	LcNNMCPnEv	DGDest.tWL	GVFAPnITAR
Basidio	S-NT--P-L-	VILSE-GNDT	LDDNMCP-AG	DSDPQ-N-WL	AVFAPPITAR
	201				250
<i>P. involutus</i> (phyA1)	LNAAAPsvNL	TDtDAfNLvs	LCAFlTVSke	kkSdFctLFE	giPGsFeAFa
<i>P. involutus</i> (phyA2)	LNAAAPGANL	TDaDAfNLvs	LCPFmTVSke	qkSdFctLFE	giPGsFeAFa
绒毛栓菌	LNAGAPGANL	TDtDTyNLlt	LCPFETVate	rrSeFCDIYE	elQAE.dAfa
平田头菇	LNqgAPGANI	TAAdvsNLip	LCAFETivke	tpSpFCNLf.	.tPEEFaqFe
<i>P. lycii</i>	LNAAAPSANL	SDsDAltLmd	MCPFDTLsSG	naSpFCDLf.	.tAEeYvSYe
Basidio	LNAAAPGANL	TD-DA-NL--	LCPFETVS-E	--S-FCDLFE	--PEEF-AF-

图 3a

	251		300
<i>P. involutus</i> (phyA1)	YgGDLDFYg	TGYGQeLGPV	QGVGYVNELI ARLTnsAVRD NTQTNRtLDA
<i>P. involutus</i> (phyA2)	YaGDLDFYg	TGYGQALGPV	QGVGYINELL ARLTnsAVnD NTQTNRtLDA
绒毛栓菌	YnADLDFYg	TGYGQPLGPV	QGVGYINELI ARLTaQnVsD HTQTNsTLDS
平田头菇	YfGDLDFYg	TGYGQPLGPV	QGVGYINELL ARLTemPVrD NTQTNRtLDS
<i>P. lycii</i>	YyyDLDFYg	TGpGNALGPV	QGVGYVNELL ARLTgQAVRD ETQTNRtLDS
Basidio	Y-GDLDFYg	TGYGQPLGPV	QGVGYINELL ARLT-QAVRD NTQTNRtLDS
	301		350
<i>P. involutus</i> (phyA1)	SPvTFPLNKT	FYADFSDNl	MVAVFSAMGL FrQPAPLsTS vPNPwRTWrT
<i>P. involutus</i> (phyA2)	APdTFPLNKT	MYADFSDNl	MVAVFSAMGL FrQSAPLsTS tPDPNRtLWT
绒毛栓菌	SPeTFPLNRT	LYADFSDNQ	MVAIFSAMGL FNQSAPLDPT tPDPaRTFLv
平田头菇	SPlTFPLDRS	IYADLSHDNQ	MIAIFSAMGL FNQSSPLDPS fPNPKRTWVT
<i>P. lycii</i>	dPaTFPLNRT	FYADFSDNt	MVPIFAALGL FNaTA.LDPl kPDENRlWVd
Basidio	SP-TFPLNRT	FYADFSDNQ	MVAIFSAMGL FNQSAPLDPS -PDPNRtWVT
	351		400
<i>P. involutus</i> (phyA1)	SsLVPFSGRM	VVERLsC..f	GT..... tkV RVLVQDqVQP
<i>P. involutus</i> (phyA2)	SsVVPFSARM	aVERLsC..a	GT..... tkV RVLVQDqVQP
绒毛栓菌	kKIVPFsARM	VVERLdC..g	GA..... qSv RLLVNDAVQP
平田头菇	SRLtPFsARM	VtERLlCqrd	GTsgggsri mrngnvqtFv RILVNDAVQP
<i>P. lycii</i>	SKLVPFSGHM	tVEKLsC...	sgkeaV RVLVNDAVQP
Basidio	SKLVPFsARM	VVERL-C---	GT-----V RVLVNDAVQP
	401		441
<i>P. involutus</i> (phyA1)	LEFCGGDrNG	lCTLakFVES	QtFARsDGaG DFEKCFATSa -
<i>P. involutus</i> (phyA2)	LEFCGGDqDG	lCALDkFVES	QaYARsGGaG DFEKCLATTv -
绒毛栓菌	LAFCGADtsG	vCTLDaFVES	QaYARNDGEG DFEKCFAT-- -
平田头菇	LKFCGGDmDS	lCTLEaFVES	QkYAREdGQG DFEKCFD--- -
<i>P. lycii</i>	LEFCGG.vDG	vCeLsAFVES	QtYARENGQG DfAKCgfvPs e
Basidio	LEFCGGD-DG	-CTLDaFVES	Q-YAREdGQG DFEKCFATP- -

图 3b

	1		50
土曲霉 9a1	KhSDCNSVDh	GYQCFPELSH	kWGLYAPYFS LqDESFPFLD VPdDCHITFV
土曲霉 cbs	NhSDCtSVDr	GYQCFPELSH	kWGLYAPYFS LqDESFPFLD VPdDCHITFV
黑曲霉 泡盛变种	NqSTCDTVDq	GYQCFSEtSH	LWGQYAPFFS LANESAISPD VPAGCRVTFa
黑曲霉 NRRL3135	NqSsCDTVDq	GYQCFSEtSH	LWGQYAPFFS LANESvISPE VPAGCRVTFa
烟曲霉 13073	GskSCDTVDl	GYQCsPatSH	LWGQYSPFFS LEDELsvSSK LPkDCRITLV
烟曲霉 32722	GskSCDTVDl	GYQCsPatSH	LWGQYSPFFS LEDELsvSSK LPkDCRITLV
烟曲霉 58128	GskSCDTVDl	GYQCsPatSH	LWGQYSPFFS LEDELsvSSK LPkDCRITLV
烟曲霉 26906	GskSCDTVDl	GYQCsPatSH	LWGQYSPFFS LEDELsvSSK LPkDCRITLV
烟曲霉 32239	GskACDTVEl	GYQCsPGtSH	LWGQYSPFFS LEDELsvSSD LPkDCRVTFV
E. nidulans	QNHSCNTaDG	GYQCFPNVSH	VWGQYSPYFS IEQESAISED VPhGCeVTFV
嗜热蹼节菌	DSHSCNTVEG	GYQCrPEISH	sWGQYSPFFS LADQSEISPD VPqNCKITFV
T. lanuginosus	-----	-----nvDIAR	hWGQYSPFFS LAEvSEISPA VPkGCRVeFv
嗜热毁丝霉	ESRPCDTpDl	GFQCgTAISH	FWGQYSPYFS VPseIdaS... IPdDCeVTFa
Basidio	xSxPxrxxtAA	qLPipxQxqx	xWSPYSPYFP VAXyxA.... pPaGQCIXqV
共有	NSHSCDTVDG	GYQC-PEISH	LWGQYSPFFS LADESAISPD VP-GCRVTfV
Fcp10	NSHSCDTVDG	GYQCfPEISH	LWGQYSPFFS LADESAISPD VPkGCRVTfV
	51		100
土曲霉 9a1	QVLARHGArS	PThSKTKaYA	AtIaAIQKSA TaFpGKYAFL QSYNYSLDSE
土曲霉 cbs	QVLARHGArS	PTdSKTKaYA	AtIaAIQKNA TaLpGKYAFL KSYNYSMGSE
黑曲霉 泡盛变种	QVLSRHGARY	PTeSKGKKYS	ALIEEIQQNv TtFDGKYAFL KTYNYSLGAD
黑曲霉 NRRL3135	QVLSRHGARY	PTdSKGKKYS	ALIEEIQQNA TtFDGKYAFL KTYNYSLGAD
烟曲霉 13073	QVLSRHGARY	PTSSKSkkYk	kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
烟曲霉 32722	QVLSRHGARY	PTSSKSkkYk	kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
烟曲霉 58128	QVLSRHGARY	PTSSKSkkYk	kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
烟曲霉 26906	QVLSRHGARY	PTSSKSkkYk	kLVtAIQaNA TdFKGKFAFL KTYNYTLGAD
烟曲霉 32239	QVLSRHGARY	PTASKSKkYk	kLVtAIQKNA TeFKGKFAFL ETYNYTLGAD
E. nidulans	QVLSRHGARY	PTeSKSKaYS	GLIEAIQKNA TsFwGQYAFI ESYNYTLGAD
嗜热蹼节菌	QLLSRHGARY	PTSSKTELYS	qLIserIQKtA TaYKGyYAFI KdYrYqLGAN
T. lanuginosus	QVLSRHGARY	PTAHKSEvYA	ELlqrIQDtA TeFKGDFAFI RdYaYhLGAD
嗜热毁丝霉	QVLSRHGARA	PTlkRAAsYv	DLIdrIHhGA isYgPgYEFL RTYDYTLGAD
Basidio	NIIqRHGARF	PTSGaAtrIq	AaVakLQsax xxtDPKLDfL xnxTYxLGxD
共有	QVLSRHGARY	PTSSKSkkYS	ALI-AIQKNA T-FKGKYAFL KTYNYTLGAD
Fcp10	QVLSRHGARY	PTSSKSkkYS	ALIEAIQKNA TaFKGKYAFL KTYNYTLGAD
	101		150
土曲霉 9a1	ELTPFGGrNQL	rDlGaQFYeR	YNAL.TRhIn PFVRATDAeR VhESAeKFVE
土曲霉 cbs	NLTPFGGrNQL	qDlGaQFYRR	YDTL.TRhIn PFVRAADSsR VhESAeKFVE
黑曲霉 泡盛变种	DLTPFGEQEL	VNSGIKFYQR	YESL.TRnII PFIRSSGSsR VIASGEKFIE
黑曲霉 NRRL3135	DLTPFGEQEL	VNSGIKFYQR	YESL.TRnIV PFIRSSGSsR VIASGKKFIE
烟曲霉 13073	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARsVV PFIRASGSDR VIASGEKFIE
烟曲霉 32722	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARsVV PFIRASGSDR VIASGEKFIE
烟曲霉 58128	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARsVV PFIRASGSDR VIASGEKFIE
烟曲霉 26906	DLTAFGEQQL	VNSGIKFYQR	YKAL.ARsVV PFIRASGSDR VIASGEKFIE
烟曲霉 32239	DLTPFGEQQM	VNSGIKFYQK	YKAL.AgsVV PFIRSSGSDR VIASGEKFIE
E. nidulans	DLTiFGENQM	VDSGaKFYRR	YKnL.ARKnt PFIRASGSDR VVASAEKFIn
嗜热蹼节菌	DLTPFGENQM	IQlGIKFYnH	YKSL.ARnAv PFVRCSGSDR VIASGrLPiE
T. lanuginosus	NLTRFGEEQM	MESGrQFYHR	YREq.AReIV PFVRAAGSAR VIASAEffnR
嗜热毁丝霉	ELTRtGQQQM	VNSGIKFYRR	YRAL.ARksI PFVRTAGQDR VVhSAENftQ
Basidio	DLvPFGAxQs	sQAGqEaFtR	YsxLvSxdnL PFVRASGSDR VVDSatNWtA
共有	DLTPFGEQQM	VNSGIKFYRR	YKAL-AR-IV PFVRASGSDR VIASAEKFIE
Fcp10	DLTPFGEQQM	VNSGIKFYRR	YKAL.ARKIV PFVRASGSDR VIASAEKFIE

图 4a

图 4b

	301		350
土曲霉 9a1	PLGpVQGVGW aNELMARLTR	A. PVHDHTCv	NNTLDASPAT FPLNATLYAD
土曲霉 cbs	PLGpVQGVGW aNELIARLTH	S. PVHDHTCv	NNTLDANPAT FPLNATLYAD
黑曲霉 泡盛变种	PLGPTQGVGY aNELIARLTH	S. PVHDDTSS	NHTLDSNPAT FPLNSTLYAD
黑曲霉 NRRL3135	PLGPTQGVGY aNELIARLTH	S. PVHDDTSS	NHTLDSSPAT FPLNSTLYAD
烟曲霉 13073	PLGPAQGIGF tNELIARLTH	S. PVQDHTST	NeTLvSNPAT FPLNATMYvD
烟曲霉 32722	PLGPAQGIGF tNELIARLTH	S. PVQDHTST	NeTLvSNPAT FPLNATMYvD
烟曲霉 58128	PLGPAQGIGF tNELIARLTH	S. PVQDHTST	NeTLvSNPAT FPLNATMYvD
烟曲霉 26906	PLGPAQGIGF tNELIARLTH	S. PVQDHTST	NeTLvSNPAT FPLNATMYvD
烟曲霉 32239	PLGPAQGIGF tNELIARLTH	S. PVQDHTST	NeTLvSNPAT FPLNATMYvD
E. nidulans	PLGPAQGIGF tNELIARLTH	S. PVQDHTST	NeTLvSNPAT FPLNATMYvD
嗜热蹠节菌	PLGPAQGVGF vNELIARLTH	S. PVQDHTST	NHTLDSNPAT FPLNATLYAD
T. lanuginosus	AFGpSRGVGF vNELIARLTH	S. PVQDHTST	NHTLDSDPAT FPLNATLYAD
嗜热毁丝霉	PLGPTQGVGF vNELIARLTH	S. PVQDHTST	NHTLDSNPAT FPLNATLYAD
Basidio	PLGpVQGVGW iNELIARLTH	qa. VRDNTQT	NRTLDSSPAT FPLNATLYAD
共有	PLGPAQGVGF -NELIARLTH	S. PVQDHTST	NHTLDSNPAT FPLNATLYAD
Fcp10	PLGPAQGVGF vNELIARLTH	S. PVQDHTST	NHTLDSNPAT FPLNATLYAD
	351		400
土曲霉 9a1	FSHDSnLVSI FWALGLYNGT	aPLSQTSE.	.SvsQTDGYA AAWTVPFAAR
土曲霉 cbs	FSHDSnLVSI FWALGLYNGT	kPLSQTTE.	.ditrTDGYA AAWTVPFAAR
黑曲霉 泡盛变种	FSHDNGIISI LFALGLYNGT	kPLSTTTVE.	.NitQTDGFS SAWTVPFASR
黑曲霉 NRRL3135	FSHDNGIISI LFALGLYNGT	kPLSTTTVE.	.NitQTDGFS SAWTVPFASR
烟曲霉 13073	FSHDNSMVIS FFALGLYNGT	ePLSrTSVE.	.SaKELDGYS ASWvVPFGAR
烟曲霉 32722	FSHDNSMVIS FFALGLYNGT	gPLSrTSVE.	.SaKELDGYS ASWvVPFGAR
烟曲霉 58128	FSHDNSMVIS FFALGLYNGT	ePLSrTSVE.	.SaKELDGYS ASWvVPFGAR
烟曲霉 26906	FSHDNSMVIS FFALGLYNGT	ePLSrTSVE.	.SaKELDGYS ASWvVPFGAR
烟曲霉 32239	FSHDNGMIPI FFAMGLYNGT	ePLSQTSE.	.StKESNGYS ASWAVPFGAR
E. nidulans	FSHDNSMISI FFAMGLYNGT	qPLSmdSVE.	.SiQEmDGYA ASWTVPFAAR
嗜热蹠节菌	FSHDNTMTSI FaALGLYNGT	akLSTTeIK.	.SiEETDGYS AAWTVPFGGR
T. lanuginosus	FSHDNTMTGI FaAMGLYNGT	kPLSTSKIQP	pTgAAADGYA ASWTVPFAAR
嗜热毁丝霉	FSHDNDMMGV LgALGaYDGV	pPLdkTA.	.R rdpEELGGYA ASWAVPFAAR
Basidio	FSHDNQMAI FaAMGLFMQS	aPLdPSxpDP	nrt....Wv TSKlVPFSAR
共有	FSHDNTMVIS FFALGLYNGT	-PLSTTSVEP	-S-EETDGYS ASWTVPFAAR
Fcp10	FSHDNTMVIS FFALGLYNGT	KPLSTTSVE.	.SiEETDGYS ASWTVPFAAR
	401		450
土曲霉 9a1	AYVEMMQC.. ra.....	EKEPL	VRVLVNDVRV PLHGCPDVKL
土曲霉 cbs	AYIEMMQC.. ra.....	EKQPL	VRVLVNDVRV PLHGCAVDNL
黑曲霉 泡盛变种	LYVEMMQC.. Qa.....	EQEPL	VRVLVNDVRV PLHGCPIDaL
黑曲霉 NRRL3135	LYVEMMQC.. Qa.....	EQEPL	VRVLVNDVRV PLHGCPIDaL
烟曲霉 13073	AYfEtMQC.. Ks.....	EKEPL	VRaLINDRVV PLHGCDVDKL
烟曲霉 32722	AYfEtMQC.. Ks.....	EKEPL	VRaLINDRVV PLHGCDVDKL
烟曲霉 58128	AYfEtMQC.. Ks.....	EKESL	VRaLINDRVV PLHGCDVDKL
烟曲霉 26906	AYfEtMQC.. Ks.....	EKEPL	VRaLINDRVV PLHGCDVDKL
烟曲霉 32239	AYfEtMQC.. Ks.....	EKEPL	VRaLINDRVV PLHGCAVDKL
E. nidulans	AYfEtMQC.. E.....	KKEPL	VRVLVNDRVV PLHGCAVDKL
嗜热蹠节菌	AYIEMMQC.. Dd.....	sDEPV	VRVLVNDRVV PLHGCEVDsL
T. lanuginosus	AYVELLRC.. Etetsseeee	EG...EDEPF	VRVLVNDRVV PLHGCEVDsL
嗜热毁丝霉	iYVEkMRC.. sgggggggggg	EGrqeKDEeM	VRVLVNDRVV TLkGCGaDer
Basidio	mvVErLxCxx	xgtxxxxxxx	xxxxxxx VRVLVNDaVq PLefCGDxd
共有	AYVEMMQC-- E-----	EG---EKEPL	VRVLVNDRVV PLHGCGVDKL
Fcp10	AYVEMMQC.. EA.....	EKEPL	VRVLVNDRVV PLHGCGVDKL

图 4c

	451		482
土曲霉 9a1	GRCKrDAFVA GLSFAQAG..	GNWADCF---	--
土曲霉 cbs	GRCKrDDFVE GLSFARAG..	GNWAECF---	--
黑曲霉 泡盛变种	GRCtrDsFVr GLSFARSG..	GDWAECsA--	--
黑曲霉 NRRL3135	GRCtrDsFVr GLSFARSG..	GDWAECFA--	--
烟曲霉 13073	GRCKlNDFVK GLSWARSG..	GNWGECFS--	--
烟曲霉 32722	GRCKlNDFVK GLSWARSG..	GNWGECFS--	--
烟曲霉 58128	GRCKlNDFVK GLSWARSG..	GNWGECFS--	--
烟曲霉 26906	GRCKlNDFVK GLSWARSG..	GNWGECFS--	--
烟曲霉 32239	GRCKlKDFVK GLSWARSG..	GNSEQSFS--	--
E. nidulans	GRCtlDDWVE GLNFARSG..	GNWKtCFTl-	--
嗜热踝节菌	GRCKrDDFVr GLSFARqG..	GNWEGCYAas e-	
T. lanuginosus	GRCRrDEWIK GLTFARqG..	GHWDrCF---	--
嗜热毁丝霉	GmCtlErFIE SMAFARGN..	GKWDlCFA--	--
Basidio	GxCtlDAFVE SqxYAReDgq	GDFEKCFAtp	xx
共有	GRCK-DDFVE GLSFARSG--	GNWEECFa--	--
Fcp10	GRCKRDDFVE GLSFARSG..	GNWEECFa..	..

图 4d

CP-1
 Eco RI M G V F V V L L S I A T L F G S T 17
 TATATGAATTCATGGGCGTGTTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA
 1 -----+-----+-----+-----+-----+ 60
 ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGAACAAGCCAAGGT

 S G T A L G P R G N S H S C D T V D G G 37
 CATCCGGTACCGCCTTGGGTCCTCGTGGAATTCTCACTCTTGTGACACTGTTGACGGTG
 61 -----+-----+-----+-----+-----+ 120
 GTAGGCCATGGCGGAACCCAGGAGCACCATTAAAGAGTGAGAACACTGTGACAACTGCCAC
 CP-2
CP-3.10
 Y Q C F P E I S H L W G Q Y S P F F S L 57
 GTTACCAATGTTTCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATTCTTCTCTT
 121 -----+-----+-----+-----+-----+ 180
 CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCAGTTATGAGAGGTAAGAAGAGAA

 A D E S A I S P D V P K G C R V T F V Q 77
 TGGCTGACGAATCTGCTATTTCTCCAGACGTTCCAAAGGGTGTAGAGTTACTTTTCGTTT
 181 -----+-----+-----+-----+-----+ 240
 ACCGACTGCTTAGACGATAAAGAGGTCTGCAAGGTTTCCCGACATCTCAATGAAAGCAAG
CP-4.10
CP-5.10
 V L S R H G A R Y P T S S K S K K Y S A 97
 AAGTTTTGTCTAGACACGGTGCTAGATACCCAACTTCTTCTAAGTCTAAGAAGTACTCTG
 241 -----+-----+-----+-----+-----+ 300
 TTCAAACAGATCTGTGCCACGATCTATGGGTTGAAGAAGATTCAAGTTCTTCATGAGAC

 L I E A I Q K N A T A F K G K Y A F L K 117
 CTTTGATTGAAGCTATTCAAAGAACGCTACTGCTTCAAGGGTAAGTACGCTTTCTTGA
 301 -----+-----+-----+-----+-----+ 360
 GAAACTAACTTCGATAAGTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAAGT
 CP-6
CP-7.10
 T Y N Y T L G A D D L T P F G E Q Q M V 137
 AGACTTACAACTACACTTTGGGTGCTGACGACTTGACTCCATTTCGGTGAACAACAAATGG
 361 -----+-----+-----+-----+-----+ 420
 TCTGAATGTTGATGTGAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTGTTTACC

 N S G I K F Y R R Y K A L A R K I V P F 157
 TTAAGTCTGGTATTAAGTTCTACAGAAGATAACAAGGCTTTGGCTAGAAAGATTGTTCCAT
 421 -----+-----+-----+-----+-----+ 480
 AATTGAGACCATAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTCTAACAAGGTA
CP-8.10
CP-9.10
 V R A S G S D R V I A S A E K F I E G F 177
 TCGTTAGAGCTTCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCAATTGAAGGTT
 481 -----+-----+-----+-----+-----+ 540
 AGCAATCTCGAAGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

 Q S A K L A D P G A N P H Q A S P V I N 197
 TCCAATCTGCTAAGTTGGCTGACCCAGGTGCTAACCACACCAAGCTTCTCCAGTTATTA
 541 -----+-----+-----+-----+-----+ 600
 AGGTTAGACGATTCAACCGACTGGGTCCACGATTGGGTGTGGTTCGAAGAGGTCAATAAT

图 5a

CP-10.10
CP-11.10

V I I P E G A G Y N N T L D H G L C T A 217
ACGTTATTATTCCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGTTTGTGTACTG
601 -----+-----+-----+-----+-----+ 660
TGCAATAATAAGGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGAC

F E E S E L G D D V E A N F T A V F A P 237
CTTTCGAAGAATCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTGTTTTCGCTC
661 -----+-----+-----+-----+-----+ 720
GAAAGCTTCTTAGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGACAAAAGCGAG

CP-12.10

P I R A R L E A H L P G V N L T D E D V 257
CACCTATTAGAGCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACG
721 -----+-----+-----+-----+-----+ 780
GTGGATAATCTCGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGC

CP-13.10

V N L M D M C P F D T V A R T S D A T Q 277
TTGTTAACTTGATGGACATGTGTCCATTGACACTGTTGCTAGAACTTCTGACGCTACTC
781 -----+-----+-----+-----+-----+ 840
AACAATTGAACTACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAG

L S P F C D L F T H D E W I Q Y D Y L Q 297
AATTGTCTCCATTCTGTGACTTGTTCCTCAACGACGAATGGATTCAATACGACTACTTGC
841 -----+-----+-----+-----+-----+ 900
TTAACAGAGGTAAGACACTGAACAAGTGAGTGCTGCTTACCTAAGTTATGCTGATGAACG

CP-14.10
CP-15.10

S L G K Y Y G Y G A G N P L G P A Q G V 317
AATCTTTGGGTAAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTG
901 -----+-----+-----+-----+-----+ 960
TTAGAAACCCATTCAATGATGCCAATGCCACGACCATTTGGGTAACCCAGGTGAGTTCCAC

G F V N E L I A R L T H S P V Q D H T S 337
TTGGTTTTCGTTAACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTT
961 -----+-----+-----+-----+-----+ 1020
AACCAGCAATTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAA

CP-16.10
CP-17.10

T N H T L D S N P A T F P L N A T L Y A 357
CTACTAACCCACACTTTGGACTCTAACCCAGCTACTTTCCATTGAACGCTACTTTGTACG
1021 -----+-----+-----+-----+-----+ 1080
GATGATTGGTGTGAAACCTGAGATTGGGTGATGAAAGGTAACCTTGCATGAAACATGC

D F S H D N T M V S I F F A L G L Y N G 377
CTGACTTCTCTCACGACAACACTATGGTTTCTATTTTCTTCGCTTTGGGTTTGTACAACG
1081 -----+-----+-----+-----+-----+ 1140
GACTGAAGAGAGTGCTGTTGTGATACCAAGATAAAGAAAGCGAAACCCAAACATGTTGC

CP-18.10
CP-19.10

T K P L S T T S V E S I E E T D G Y A A 397
GTACTAAGCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACGCTG
1141 -----+-----+-----+-----+-----+ 1200
CATGATTCGGTAACAGATGATGAAGACAACTTAGATAAAGTTCTTTGACTGCCAATGCGAC

图 5b

```

      S W T V P F A A R A Y V E M M Q C E A E 417
      CTTCTGGACTGTTCCATTGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTG
1201 -----+-----+-----+-----+-----+-----+ 1260
      GAAGAACCTGACAAGGTAAGCCACGATCTCGAATGCAACTTTACTACGTTACACTTCGAC
                               CP-20.10
                               CP-21.10
      K E P L V R V L V N D R V V P L H G C G 437
      AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
1261 -----+-----+-----+-----+-----+-----+ 1320
      TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC

      V D K L G R C K R D D F V E G L S F A R 457
      GTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA
1321 -----+-----+-----+-----+-----+-----+ 1380
      CACAACCTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGAT
                               CP-22.10
      S G G N W E E C F A * Eco RI 467
      GATCTGGTGGTAACTGGGAAGAATGTTTCGCTTAAGAATTCATATA
1381 -----+-----+-----+-----+-----+-----+ 1426
      CTAGACCACCATTGACCCTTCTTACAAAGCGAATTCTTAAGTATAT

```

图 5c

	1		50
P. involutus (phyA1)	-----	-FPipeseqR nWSPYSPYFP LAEyKA....	pPaGCQInqV
P. involutus (phyA2)	-----	-FsipeseqR nWSPYSPYFP LAEyKA....	pPaGCEInqV
绒毛栓菌	-----	-LDvtRDVqQ sWsmYSPYFP aAtyvA....	pPaSCQInqV
平田头菇	-----	-pffpPQIQD sWAAyTPYYP VqAyTP....	pPKDCKITqV
P. lycii	-----	-LPipAQnTs nWGPYdPFFP VEpyAA....	pPEGCTVTqV
土曲霉 9a1	KhsdCNSVDh	GYQCfPELSH kWGLYAPYFS LqDESPFP1D	VPEDCHITFV
土曲霉 cbs	NhsdCtSVDr	GYQCfPELSH kWGLYAPYFS LqDESPFP1D	VPDDCHITFV
黑曲霉泡盛变种	NqsTCDTVDq	GYQCfSEtSH LWGQYAPFFS LANESAISPD	VPAGCRVTFa
黑曲霉 T213	NqsSCDTVDq	GYQCfSEtSH LWGQYAPFFS LANESvISPD	VPAGCRVTFa
黑曲霉 NRRL3135	NqsSCDTVDq	GYQCfSEtSH LWGQYAPFFS LANESvISPE	VPAGCRVTFa
烟曲霉 ATCC13073	GskSCDTVD1	GYQCsPATSH LWGQYSPFFS LEDELSVSSK	LPKDCRITLV
烟曲霉 ATCC32722	GskSCDTVD1	GYQCsPATSH LWGQYSPFFS LEDELSVSSK	LPKDCRITLV
烟曲霉 ATCC58128	GskSCDTVD1	GYQCsPATSH LWGQYSPFFS LEDELSVSSK	LPKDCRITLV
烟曲霉 ATCC26906	GskSCDTVD1	GYQCsPATSH LWGQYSPFFS LEDELSVSSK	LPKDCRITLV
烟曲霉 ATCC32239	GskACDTVEL	GYQCsPGtSH LWGQYSPFFS LEDELSVSSD	LPKDCRVTfV
E. nidulans	QNHSCNTaDg	GYQCfPNVSH VWGQYSPYFS IEQESAISd	VPhGCeVTFV
嗜热蹠节菌	DSHSCNTVEg	GYQCrPEISH sWGQYSPFFS LADQSEISPD	VPQNCKITfV
T. lanuginosus	-----	----nVDIAR hWGQYSPFFS LAEvSEISPA	VPKGCRVeFV
嗜热毁丝霉	ESRPCDTpD1	GFQCgTAISH FWGQYSPYFS VPSElDaS..	IPDDCeVTFa
共有序列 11	NSHSCDTVD-	GYQC-PEISH LWGQYSPFFS LADESAISPD	VPKGCRVTfV
	51		100
P. involutus (phyA1)	NIIqRHGARF	PTSGaTtRik AgLtKLQgvq nftDAKFNF	KSFKYdLGns
P. involutus (phyA2)	NIIqRHGARF	PTSGaAtRik AgLsKLQsvq nftDPKFDFI	KSFTYdLGts
绒毛栓菌	HIIqRHGARF	PTSGaAKRiq TaVAKLKaaS nytdP1LAFV	tnYtYSLGqD
平田头菇	NIIqRHGARF	PTSGaGtRiq AaVKKLQsAk TytdPRLDfL	tnYtYTLGhd
P. lycii	NLIqRHGARW	PTSGarsRqv AaVAKIQmar PftDPKYEFL	NdFvYkFGvA
土曲霉 9a1	QVLARHGARS	PTHsKTKaYA AtIAaIQKSA TaFpGKYAFL	QSYNYSLDSE
土曲霉 cbs	QVLARHGARS	PTdSKTKaYA AtIAaIQKNA TaLpGKYAFL	KSYNYSMGSE
黑曲霉泡盛变种	QVLSRHGARY	PTeSKGKKYS ALIEeIQQNv TtFDGKYAFL	KTYNYSLGAD
黑曲霉 T213	QVLSRHGARY	PTeSKGKKYS ALIEeIQQNv TtFDGKYAFL	KTYNYSLGAD
黑曲霉 NRRL3135	QVLSRHGARY	PTdSKGKKYS ALIEeIQQNA TtFDGKYAFL	KTYNYSLGAD
烟曲霉 ATCC13073	QVLSRHGARY	PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL	KTYNYTLGAD
烟曲霉 ATCC32722	QVLSRHGARY	PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL	KTYNYTLGAD
烟曲霉 ATCC58128	QVLSRHGARY	PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL	KTYNYTLGAD
烟曲霉 ATCC26906	QVLSRHGARY	PTSSKSKKYk kLVtaIQaNA TdFKGKFAFL	KTYNYTLGAD
烟曲霉 ATCC32239	QVLSRHGARY	PTASKSKKYk kLVtaIQKNA TeFKGKFAFL	ETNYNYTLGAD
E. nidulans	QVLSRHGARY	PTeSKSKaYS GLIEaIQKNA TsFwGQYAFL	ESYNYTLGAD
嗜热蹠节菌	QLLSRHGARY	PTSSKTElyS qLIeRIQKtA TaYKGyYAFL	KdYrYqLGAN
T. lanuginosus	QVLSRHGARY	PTAhKSEvYA ELLQRIQDtA TeFKGDFaFL	RdYaYnLGAD
嗜热毁丝霉	QVLSRHGARA	PTlkRAasYv DLIDRIHhGA isYgPgYEFL	RTYDYTLGAD
共有序列 11	QVLSRHGARY	PTSSKSKKYs ALIERIQKNA T-FKGKYAFL	KTYNYTLGAD

图 6a

	101		150		
P. involutus (phyA1)	DLvPFGAaQs	fDAGqEaFaR	YskLvSKNnL PFIRAdGSDR	VVDSAtNwTA	
P. involutus (phyA2)	DLvPFGAaQs	fDAGLEvFaR	YskLvSsDnL PFIRsdGSDR	VVDTAtNwTA	
绒毛栓菌	sLveLGAtQs	sEAGqEaFtR	YsSLvSaDeL PFVRASGSDR	VVATANNwTA	
平田头菇	DLvPFGAlQs	sQAGeEtFQR	YsfLvSKEnL PFVRASSNR	VVDSAtNwTE	
P. lycii	DLlPFGANQs	hQTGTDMYtR	YsTLfEgGdV PFVRAAGdQR	VVDSStNwTA	
土曲霉 9a1	ELTPFGGrNQL	rDlGaQFYeR	YNAL.TRHIn PFVRATDaSR	VhESAeKFVE	
土曲霉 cbs	NLTPFGGrNQL	qDlGaQFYRR	YDTL.TRHIn PFVRAADsSR	VhESAeKFVE	
黑曲霉泡盛变种	DLTPFGEQEL	VNSGIKFYQR	YESL.TRNII PFIRSSGSsR	VIASGEKFIE	
黑曲霉 T213	DLTPFGEQEL	VNSGIKFYQR	YESL.TRNII PFIRSSGSsR	VIASGEKFIE	
黑曲霉 NRRL3135	DLTPFGEQEL	VNSGIKFYQR	YESL.TRNIV PFIRSSGSsR	VIASGKKFIE	
烟曲霉 ATCC13073	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARSVV PFIRASGSDR	VIASGEKFIE	
烟曲霉 ATCC32722	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARSVV PFIRASGSDR	VIASGEKFIE	
烟曲霉 ATCC58128	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARSVV PFIRASGSDR	VIASGEKFIE	
烟曲霉 ATCC26906	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARSVV PFIRASGSDR	VIASGEKFIE	
烟曲霉 ATCC32239	DLTPFGEQQL	VNSGIKFYQR	YKAL.ARSVV PFIRASGSDR	VIASGEKFIE	
E. nidulans	DLTiFGENQM	VDSGaKFYRR	YKnL.ARKnt PFIRASGSDR	VVASAEKFIn	
嗜热蹼节菌	DLTPFGENQM	IQlGKFYnH	YKSL.ARNv PFVRCSGSDR	VIASGrLFIE	
T. lanuginosus	NLTRPGEEQM	MESGrQFYHR	YREq.AREIV PFVRAAGSAR	VIASAEfFnr	
嗜热毁丝霉	ELTRtGQQQM	VNSGIKFYRR	YRAL.ARKSr PFVRTAGqDR	VVhSAENftQ	
共有序列 11	DLTPFGENQM	VNSGIKFYRR	YKAL-ARNIV PFVRASGSDR	VIASAEKFIE	
	151		200		
P. involutus (phyA1)	GFaSA.....	..shNtvqPk	LNlILPQ..T	gNDTLEDNMC	PAaGD.....
P. involutus (phyA2)	GFaSA.....	..srNaiqPk	LDLILPQ..T	gNDTLEDNMC	PAaGE.....
绒毛栓菌	GFaLA.....	..ssNsITPV	LSVIIE..A	gNDTLDDNMC	PAaGD.....
平田头菇	GFsAA.....	..shHvlnPI	LfVILSE..S	LNDTLDDAMC	PnaGs.....
P. lycii	GFgda.....	..sgEtvlpT	LQVVLQE..E	gNcTLcNNMC	PnevD.....
土曲霉 9a1	GFQTARqDDh	hAnpHQPSPr	VDVaIPEGSA	YNNTLEHSLC	TAFes...ST
土曲霉 cbs	GFQNaRqGDP	hAnpHQPSPr	VDVVIPEGTA	YNNTLEHSIC	TAFEA...ST
黑曲霉泡盛变种	GFQSTKLkDP	rAqpGQSSPk	IDVVISeASS	sNNTLDpGtC	TvFED...Se
黑曲霉 T213	GFQSTKLkDP	rAqpGQSSPk	IDVVISeASS	sNNTLDpGtC	TvFED...Se
黑曲霉 NRRL3135	GFQSTKLkDP	rAqpGQSSPk	IDVVISeASS	sNNTLDpGtC	TvFED...Se
烟曲霉 ATCC13073	GFQqAKLADP	gAt.NRAAPa	ISVIIPESeT	FNNTLDHGVC	TkFEA...Sq
烟曲霉 ATCC32722	GFQqAKLADP	gAt.NRAAPa	ISVIIPESeT	FNNTLDHGVC	TkFEA...Sq
烟曲霉 ATCC58128	GFQqAKLADP	gAt.NRAAPa	ISVIIPESeT	FNNTLDHGVC	TkFEA...Sq
烟曲霉 ATCC26906	GFQqAKLADP	gAt.NRAAPa	ISVIIPESeT	FNNTLDHGVC	TkFEA...Sq
烟曲霉 ATCC32239	GFQqANVADP	gAt.NRAAPV	ISVIIPESeT	YNNTLDHSVC	TnFEA...Se
E. nidulans	GFRkaQLhDh	g.s.gQATPV	VNVIIEIdG	FNNTLDHStC	vSFEN...de
嗜热蹼节菌	GFQSAKVlDP	hSdKHDAPpt	INVIIeEGPS	YNNTLDtGsC	PvFED...SS
T. lanuginosus	GFQdAKdrDP	rSnkDQAePV	INVIISeETG	sNNTLDgltC	PAaEE...AP
嗜热毁丝霉	GFHSALLADR	gStvRPTlPy	dmVVIPEtAG	aNNTLDHNDLC	TAFEEgpyST
共有序列 11	GFQSAKLADP	-A--HQASPV	INVIIPEGSG	YNNTLDHGVC	TAFED---ST

图 6b

	201		250		
P. involutus (phyA1)	.SDpqvnaWl	AVafPSItAR	LNAaaPSVNL	TDtDafNLVs	LCAFlTVSK.
P. involutus (phyA2)	.SDpqvDaWl	AsafPSVtAQ	LNAaaPGaNL	TDADafNLVs	LCPFmTVSK.
绒毛栓菌	.SDpqvnQWl	AqFAPPMtAR	LNagaPGaNL	TDtDtyNLLt	LCPFETVAt.
平田头菇	.SDpqtGiWT	SIYGTPIanR	LNqqaPGaNI	TAADVaNLIp	LCAFETivK.
P. lycii	.GDESt.tWl	GVFAPnItAR	LNAaaPSaNL	SDsDaLtLMD	MCPFDTLSS.
土曲霉 9a1	VGDDAvANFT	AVFAPAiaqR	LEAdLPGVQL	StDDVVNLMA	MCPFETVSlT
土曲霉 cbs	VGDAADNFT	AVFAPAiaKR	LEAdLPGVQL	SADDVVNLMA	MCPFETVSlT
黑曲霉泡盛变种	LADtveANFT	AtFAPSIRqR	LEndLSGVtL	TDtEVtyLMD	MCSFDTISStS
黑曲霉 T213	LADtveANFT	AtFAPSIRqR	LEndLSGVtL	TDtEVtyLMD	MCSFDTISStS
黑曲霉 NRRL3135	LADtveANFT	AtFvPSIRqR	LEndLSGVtL	TDtEVtyLMD	MCSFDTISStS
烟曲霉 ATCC13073	LGDeVAANFT	ALFAPdIRAR	aEkhlPGVtL	TDEDVVSIMD	MCSFDTVART
烟曲霉 ATCC32722	LGDeVAANFT	ALFAPdIRAR	aEkhlPGVtL	TDEDVVSIMD	MCSFDTVART
烟曲霉 ATCC58128	LGDeVAANFT	ALFAPdIRAR	aEkhlPGVtL	TDEDVVSIMD	MCSFDTVART
烟曲霉 ATCC26906	LGDeVAANFT	ALFAPdIRAR	aKKhLPGVtL	TDEDVVSIMD	MCSFDTVART
烟曲霉 ATCC32239	LGDeVAANFT	ALFAPdIRAR	IEkhLPGVQL	TDDDVSIMD	MCSFDTVART
E. nidulans	rADEiEANFT	AIMGPPIRkR	LEndLPGIKL	TNENVIyLMD	MCSFDTMART
嗜热蹼节菌	gGHDAQEKFA	kqFAPAIEK	IKDhLPGVDL	AvsDVpyLMD	LCPFETLARN
T. lanuginosus	DptqpAEFl	qVFGPRVlkk	ItkhMPGVNL	TLEDVplFMD	LCPFDTVGsd
嗜热毁丝霉	IGDDAQDtYl	StFAGPItAR	VNanLPGaNL	TDADtVaLMD	LCPFETVASS
共有序列 11	LGDDAEANFT	AVFAPPiRAR	LEA-LPGVNL	TDEDVVNLMD	MCPFDTVART
	251		300		
P. involutus (phyA1)		...ekkSdF	CtLFegiPGs	FeaFAYggdL	dKFYGtGyGQ
P. involutus (phyA2)		...eqkSdF	CtLFegiPGs	FeaFAYagdL	dKFYGtGyGQ
绒毛栓菌		...errSeF	CDIYeelqAE	.daFAYnadL	dKFYGtGyGQ
平田头菇		...etpSPF	CNLf..TPEE	FaQFEYFgdL	dKFYGtGyGQ
P. lycii		...gnaSPF	CDLf..TAEE	YvsYEYYdL	dKYYGtGPGN
土曲霉 9a1	dD..Aht...	LSPF	CDLf..TAtE	WtQYNYLlSL	dKYYGYGGGN
土曲霉 cbs	dD..Aht...	LSPF	CDLf..TAAE	WtQYNYLlSL	dKYYGYGGGN
黑曲霉泡盛变种	Tv..DTK...	LSPF	CDLf..ThDE	WiHYDYlQSL	kKYYGHGAGN
黑曲霉 T213	Tv..DTK...	LSPF	CDLf..ThDE	WiHYDYlRSL	kKYYGHGAGN
黑曲霉 NRRL3135	Tv..DTK...	LSPF	CDLf..ThDE	WiNYDYlQSL	kKYYGHGAGN
烟曲霉 ATCC13073	SD..ASQ...	LSPF	CQLF..ThNE	WkKYNYlQSL	gKYYGYGAGN
烟曲霉 ATCC32722	SD..ASQ...	LSPF	CQLF..ThNE	WkKYNYlQSL	gKYYGYGAGN
烟曲霉 ATCC58128	SD..ASQ...	LSPF	CQLF..ThNE	WkKYNYlQSL	gKYYGYGAGN
烟曲霉 ATCC26906	SD..ASQ...	LSPF	CQLF..ThNE	WkKYNYlQSL	gKYYGYGAGN
烟曲霉 ATCC32239	AD..ASE...	LSPF	CAIF..ThNE	WkKYDYlQSL	gKYYGYGAGN
E. nidulans	AH..GTE...	LSPF	CAIF..TEKE	WlQYDYlQSL	sKYYGYGAGS
嗜热蹼节菌	ht..DT....	LSPF	CALs..TqEE	WqaYDYyQSL	gKYYGnGGGN
T. lanuginosus	PvlfPrQ...	LSPF	CHLf..TADD	WmaYDYyTL	dKYYSHGGGS
嗜热毁丝霉	SsdpaTadag	ggngRPLSPF	CrLf..SEsE	WraYDYlQSV	gKYYGYGPGN
共有序列 11	SD--ATQ---	-----LSPF	CDLf--TADe	W-QYDYlQSL	-KYYGYGAGN

图 6c

	301		350
P. involutus (phyA1)	eLGPvQGVGY vNELIARLTN S.AVRDNTqT NRTLDAASPvT FPLNkTFYAD		
P. involutus (phyA2)	ALGPvQGVGY iNELIARLTN S.AVNDNTqT NRTLDAApDT FPLNkTMYAD		
绒毛栓菌	PLGPvQGVGY iNELIARLTa q.nVsdHTqT NeTLdSSPET FPLNrTLYAD		
平田头菇	PLGPvQGVGY iNELIARLTm m.FVRDNTqT NRTLdSSPlT FPLDrSIYAD		
P. lycii	ALGPvQGVGY vNELIARLTg q.AVRDETqT NRTLdSDPAT FPLNrTFYAD		
土曲霉 9a1	PLGPvQGVGW aNELMARLT R A.PVHDHTCv NNTLDASPAT FPLNATLYAD		
土曲霉 cbs	PLGPvQGVGW aNELIARLT R S.PVHDHTCv NNTLDANPAT FPLNATLYAD		
黑曲霉泡盛变种	PLGPTQGVGY aNELIARLT H S.PVHDDTSS NHTLdSNPAT FPLNSTLYAD		
黑曲霉 T213	PLGPTQGVGY aNELIARLT H S.PVHDDTSS NHTLdSNPAT FPLNSTLYAD		
黑曲霉 NRRL3135	PLGPTQGVGY aNELIARLT H S.PVHDDTSS NHTLdSSPAT FPLNSTLYAD		
烟曲霉 ATCC13073	PLGPAQGIGF tNELIARLT R S.PVQDHTST NeTLvSNPAT FPLNATMYvD		
烟曲霉 ATCC32722	PLGPAQGIGF tNELIARLT R S.PVQDHTST NeTLvSNPAT FPLNATMYvD		
烟曲霉 ATCC58128	PLGPAQGIGF tNELIARLT R S.PVQDHTST NeTLvSNPAT FPLNATMYvD		
烟曲霉 ATCC26906	PLGPAQGIGF tNELIARLT R S.PVQDHTST NeTLvSNPAT FPLNATMYvD		
烟曲霉 ATCC32239	PLGPAQGIGF tNELIARLT N S.PVQDHTST NeTLdSDPAT FPLNATTYvD		
E. nidulans	PLGPAQGIGF tNELIARLT Q S.PVQDNTST NHTLdSNPAT FPLDrkLYAD		
嗜热蹼节菌	PLGPAQGVGF vNELIARMT H S.PVQDYTTv NHTLdSNPAT FPLNATLYAD		
T. lanuginosus	AFGpSRGVGF vNELIARMTg NlPVKDHTTv NHTLdDNPET FPLDAvLYAD		
嗜热毁丝霉	PLGPTQGVGF vNELIARLTa. GvPVRDgTST NRTLdGDP rT FPLGrPLYAD		
共有序列 11	PLGPAQGVGF -NELIARLT H S-PVQDHTST NHTLdSNPAT FPLNATLYAD		
	351		400
P. involutus (phyA1)	FSHDNlMVAV FsAMGLFrqP aPLSTSVpNP wrt.....Wr TSSlVPFSGR		
P. involutus (phyA2)	FSHDNlMVAV FsAMGLFrqS aPLSTSTpDP nrt.....Wl TSSvVPFSAR		
绒毛栓菌	FSHDNgMVAI FsAMGLFNqS aPLdPTTpDP art.....Fl vkkiVPFSAR		
平田头菇	LSHDNgMIAI FsAMGLFNqS sPLdPSfpNP krt.....Wv TSRLtPFASAR		
P. lycii	FSHDNTMVPI FaALGLFNAT a.LdPlkpDe nrl.....Wv DSKlVPFSGH		
土曲霉 9a1	FSHDSnLVSI FWALGLYNGT aPLSQtSVES Vs...QTDGYA AAWTVPFaAR		
土曲霉 cbs	FSHDSnLVSI FWALGLYNGT KPLSQtTVE d It...rTDGYA AAWTVPFaAR		
黑曲霉泡盛变种	FSHDNGIISI LFALGLYNGT KPLSTTTVEN It...QTDGFS SAWTVPFASR		
黑曲霉 T213	FSHDNGIISI LFALGLYNGT KPLSTTTVEN It...QTDGFS SAWTVPFASR		
黑曲霉 NRRL3135	FSHDNGIISI LFALGLYNGT KPLSTTTVEN It...QTDGFS SAWTVPFASR		
烟曲霉 ATCC13073	FSHDNSMVISI FFALGLYNGT EPLSrTSVES ak...ElDGYS ASWvVPFGAR		
烟曲霉 ATCC32722	FSHDNSMVISI FFALGLYNGT gPLSrTSVES ak...ElDGYS ASWvVPFGAR		
烟曲霉 ATCC58128	FSHDNSMVISI FFALGLYNGT EPLSrTSVES ak...ElDGYS ASWvVPFGAR		
烟曲霉 ATCC26906	FSHDNSMVISI FFALGLYNGT EPLSrTSVES ak...ElDGYS ASWvVPFGAR		
烟曲霉 ATCC32239	FSHDNGMIPI FFAMGLYNGT EPLSQtSeES tk...ESNGYS ASWAVPFGAR		
E. nidulans	FSHDNSMISI FFAMGLYNGT QPLSmdSVES Iq...EmDGYA ASWTVPFGAR		
嗜热蹼节菌	FSHDNTMtSI FaALGLYNGT akLSTTeIKS Ie...ETDGYS AAWTVPFGGR		
T. lanuginosus	FSHDNTMtGI FaAMGLYNGT KPLSTSkIQP ptgaAADGYA ASWTVPFAAR		
嗜热毁丝霉	FSHDNdMMGV LgALGaYDgv pPLdkTAr d ..peElGGYA ASWAVPFaAR		
共有序列 11	FSHDNTMVSI FFALGLYNGT KPLSTTSVES I---ETDGYS ASWTVPFAAR		

图 6d

	401		450
P. involutus (phyA1)	mvVERLsC.. fGt..... Tk	VRVLVQDQVq	PLEfCGgDRn
P. involutus (phyA2)	maVERLsC.. AGt..... Tk	VRVLVQDQVq	PLEfCGgDQd
绒毛栓菌	mvVERLDC.. GGa..... Qs	VRLLVNDaVq	PLafCGaDts
平田头菇	mvTErLLCQr	DGTGsGGpsr imrNgnvQTF	VRILVNDaLq
P. lycii	mtVEkLaC.. ..sgKea	VRVLVNDaVq	PLEfCGg.vd
土曲霉 9a1	AYVEMMQCrA	..EK...EPL	VRVLVNDRVm
土曲霉 cbs	AYIEMMQCrA	..EK...QPL	VRVLVNDRVm
黑曲霉泡盛变种	LYVEMMQCQA	..EQ...EPL	VRVLVNDRVV
黑曲霉 T213	LYVEMMQCQA	..EQ...EPL	VRVLVNDRVV
黑曲霉 NRRL3135	LYVEMMQCQA	..EQ...EPL	VRVLVNDRVV
烟曲霉 ATCC13073	AYfEtMQCKS	..EK...EPL	VraLINDRVV
烟曲霉 ATCC32722	AYfEtMQCKS	..EK...EPL	VraLINDRVV
烟曲霉 ATCC58128	AYfEtMQCKS	..EK...ESL	VraLINDRVV
烟曲霉 ATCC26906	AYfEtMQCKS	..EK...EPL	VraLINDRVV
烟曲霉 ATCC32239	AYfEtMQCKS	..EK...EPL	VraLINDRVV
E. nidulans	AYfELMQCE.	..KK...EPL	VRVLVNDRVV
嗜热蹼节菌	AYIEMMQCDD	..sD...EPV	VRVLVNDRVV
T. lanuginosus	AYVELLRcET ETsSeEeEG ..	..ED...EPF	VRVLVNDRVV
嗜热毁丝霉	iyVEkMRCsG GGgGgGGgEG ..	..rQekdEem	VRVLVNDRVm
共有序列 11	AYVEMMQCEA GG-G-GG-EG	--EK---EPL	VRVLVNDRVV
			PLHGCGVDKL
	451		482
P. involutus (phyA1)	GlCtLAKFVE SqTFARSDga	GDFEKCFats	a-
P. involutus (phyA2)	GlCaLDKFVE SqAYARSGga	GDFEKCLAtt	v-
绒毛栓菌	GvCtLDAFVE SqAYARNDge	GDFEKCFat~	--
平田头菇	SlCtLEAFVE SqkYAReDgq	GDFEKCFD--	--
P. lycii	GvCELSAFVE SqTYAReNgq	GDFAKCgfv	se
土曲霉 9a1	GRCKrDAFVA GLSFAQAG..	GNWADCF---	--
土曲霉 cbs	GRCKrDDFVE GLSFARAG..	GNWAECF---	--
黑曲霉泡盛变种	GRCtrDsFVr GLSFARSG..	GDWAECsa--	--
黑曲霉 T213	GRCtrDsFVr GLSFARSG..	GDWAECFA--	--
黑曲霉 NRRL3135	GRCtrDsFVr GLSFARSG..	GDWAECFA--	--
烟曲霉 ATCC13073	GRCKLNDfVK GLSWARSG..	GNWGEcfs--	--
烟曲霉 ATCC32722	GRCKLNDfVK GLSWARSG..	GNWGEcfs--	--
烟曲霉 ATCC58128	GRCKLNDfVK GLSWARSG..	GNWGEcfs--	--
烟曲霉 ATCC26906	GRCKLNDfVK GLSWARSG..	GNWGEcfs--	--
烟曲霉 ATCC32239	GRCKLNDfVK GLSWARSG..	GNSEQsfs--	--
E. nidulans	GRCTLDdWVE GLNFARSG..	GNWktCFT1-	--
嗜热蹼节菌	GRCKrDDFVr GLSFARqG..	GNWEGCYAas	e-
T. lanuginosus	GRCRrDEWIK GLTFARqG..	GHWDrcf---	--
嗜热毁丝霉	GmCtLErFIE SMAFARGN..	GKWDlCFA--	--
共有序列 11	GRCKLDDFVE GLSFARSG--	GNWAECFA--	--

图 6e

```

M G V F V V L L S I A T L F G S T S G T      20
ATGGGCGTGTTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCACATCCGGTACC
1 -----+-----+-----+-----+-----+-----+-----+ 60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGAACAAGCCAAGGTGTAGGCCATGG

A L G P R G N S H S C D T V D G G Y Q C      40
GCCTTGGGTCCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61 -----+-----+-----+-----+-----+-----+-----+ 120
CGGAACCCAGGAGCACCATTAAAGAGTGAGAACACTGTGACAACCTGCCACCAATGGTTACA

F P E I S H L W G T Y S P Y F S L A D E      60
TTCCCAGAAATTTCTCACTTGTGGGTACCTACTCTCCATACTTCTCTTTGGCAGACGAA
121 -----+-----+-----+-----+-----+-----+-----+ 180
AAGGGTCTTTAAAGAGTGAACACCCCATGGATGAGAGGTATGAAGAGAAACCGTCTGCTT

S A I S P D V P D D C R V T F V Q V L S      80
TCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTTCGTTCAAGTTTTGTCT
187 -----+-----+-----+-----+-----+-----+-----+ 240
AGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAGTTCAAAACAGA

R H G A R Y P T S S A S K A Y S A L I E      100
AGACACGGTGCTAGATACCCAACCTTCTTCTGCGTCTAAGGCTTACTCTGCTTTGATTGAA
241 -----+-----+-----+-----+-----+-----+-----+ 300
TCTGTGCCACGATCTATGGGTGAAGAAGACGCAGATTCCGAATGAGACGAAACTAAGTT

A I Q K N A T A F K G K Y A F L K T Y N      120
GCTATTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
301 -----+-----+-----+-----+-----+-----+-----+ 360
CGATAAGTTTTCTTGCATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

Y T L G A D D L T P F G E N Q M V N S G      140
TACACTTTGGGTGCTGACGACTTGACTCCATTCCGGTGAACCAATGGTTAACTCTGGT
361 -----+-----+-----+-----+-----+-----+-----+ 420
ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACCAATTGAGACCA

I K F Y R R Y K A L A R K I V P F I R A      160
ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTTCATTAGAGCT
421 -----+-----+-----+-----+-----+-----+-----+ 480
TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

S G S D R V I A S A E K F I E G F Q S A      180
TCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT
481 -----+-----+-----+-----+-----+-----+-----+ 540
AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACCTTCCAAGGTTAGACGA

K L A D P G S Q P H Q A S P V I N V I I      200
AAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTAAACGTGATCATT
541 -----+-----+-----+-----+-----+-----+-----+ 600
TTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCGAAGAGGTCAATAATTGCACTAGTAA

P E G S G Y N N T L D H G T C T A F E D      220
CCAGAAGGATCCGGTTACAACAACACTTTGGACCACGGTACTTGTACTGCTTTTGAAGAC
601 -----+-----+-----+-----+-----+-----+-----+ 660
GGTCTTCTAGGCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGACGAAAGCTTCTG

```

图 7a

```

S E L G D D V E A N F T A L F A P A I R      240
TCTGAATTAGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTTCGCTCCAGCTATTAGA
661 -----+-----+-----+-----+-----+-----+-----+ 720
AGACTTAATCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAGGTCGATAATCT

A R L E A D L P G V T L T D E D V V Y L      260
GCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACGTTGTTTACTTG
721 -----+-----+-----+-----+-----+-----+-----+ 780
CGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGCAACAAATGAAC

M D M C P F D T V A R T S D A T E L S P      280
ATGGACATGTGTCCATTGACACTGTGCTAGAACTTCTGACGCTACTGAATTGTCTCCA
781 -----+-----+-----+-----+-----+-----+-----+ 840
TACCTGTACACAGGTAAGCTGTGACAGCGATCTTGAAGACTGCGATGACTTAACAGAGGT

F C A L F T H D E W I Q Y D Y L Q S L G      300
TTCTGTGCTTTGTTCACTCACGACGAATGGATCCAATACGACTACTTGCAAAGCTTGGGT
841 -----+-----+-----+-----+-----+-----+-----+ 900
AAGACACGAAACAAGTGAGTGCTGCTTACCTAGGTTATGCTGATGAACGTTTCGAACCCA

K Y Y G Y G A G N P L G P A Q G V G F A      320
AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGCT
901 -----+-----+-----+-----+-----+-----+-----+ 960
TTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTCGAGTTCCACAACCAAAGCGA

N E L I A R L T H S P V Q D H T S T N H      340
AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAGACCACTTCTACTAACCAC
961 -----+-----+-----+-----+-----+-----+-----+ 1020
TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

T L D S N P A T F P L N A T L Y A D F S      360
ACTTTGGACTCTAACCAGCTACTTTCCATTGAACGCTACTTTGTACGCTGACTTCTCT
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
TGAAACCTGAGATTGGGTCGATGAAAGGGTAACCTTGCGATGAAACATGCGACTGAAGAGA

H D N T M I S I F F A L G L Y N G T K P      380
CACGACAACACTATGATATCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACCAAGCCA
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
GTGCTGTTGTGATACTATAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGTTTCGGT

L S T T S V E S I E E T D G Y S A S W T      400
TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGTTTACTCTGCTTCTTGGACT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
AACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

V P F A A R A Y V E M M Q C Q A E K E P      420
GTTCCATTGCTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTGAAAAGGAACCA
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACAGTTTCGACTTTTCCTTGGT

L V R V L V N D R V V P L H G C A V D K      440
TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGCTGTTGACAAG
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACGACAACCTGTTC

```

图 7b

```
      L G R C K R D D F V E G L S F A R S G G      460
      TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 -----+-----+-----+-----+-----+-----+ 1380
      AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

      N W A E C F A *      467
      AACTGGGCTGAATGTTTCGCTTAA
1381 -----+-----+----- 1410
      TTGACCCGACTTACAAAGCGAATT
```

图 7c

```

M G V F V V L L S I A T L F G S T S G T      20
ATGGGCGTGTTCGTGCTACTGTCCATTGCCACCTTGTTTCGGTTCACATCCGGTACC
1 -----+-----+-----+-----+-----+-----+-----+ 60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG

A L G P R G N S H S C D T V D G G Y Q C      40
GCCTTGGGTCTCTGTTGTAACCTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61 -----+-----+-----+-----+-----+-----+-----+ 120
CGGAACCCAGGAGCACCATTGAGAGTGAGAACACTGTGACAACTGCCACCAATGGTTACA

F P E I S H L W G T Y S P F F S L A D E      60
TTCCAGAAATTTCTCACTTGTGGGTACATACTCTCCATTCTTCTCTTTGGCTGACGAA
121 -----+-----+-----+-----+-----+-----+-----+ 180
AAGGGTCTTTAAAGAGTGAACACCCCATGTATGAGAGGTAAGAAGAGAAACCGACTGCTT

S A I S P D V P K G C R V T F V Q V L S      80
TCTGCTATTTCTCCAGACGTTCCAAAGGGTGTAGAGTTACTTTTCGTTCAAGTTTGTCT
181 -----+-----+-----+-----+-----+-----+-----+ 240
AGACGATAAAGAGGTCTGCAAGGTTTCCCAACATCTCAATGAAAGCAAGTTCAAAACAGA

R H G A R Y P T S S A S K A Y S A L I E      100
AGACACGGTGCTAGATACCCAACCTTCTTCTGCGTCTAAGGCGTACTCTGCTTTGATTGAA
241 -----+-----+-----+-----+-----+-----+-----+ 300
TCTGTGCCACGATCTATGGGTGAAGAAGACGCAGATTCCGCATGAGACGAACTAATCTT

A I Q K N A T A F K G K Y A F L K T Y N      120
GCTATTCAAAGAAGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
301 -----+-----+-----+-----+-----+-----+-----+ 360
CGATAAGTTTTCTTGCATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

Y T L G A D D L T P F G E Q Q M V N S G      140
TACACTTTGGGTGCTGACGACTTGACTCCATTTCGGTGAACAACAAATGGTTAACTCTGGT
361 -----+-----+-----+-----+-----+-----+-----+ 420
ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTTGTTTACCAATTGAGACCA

I K F Y R R Y K A L A R K I V P F I R A      160
ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTCATTAGAGCT
421 -----+-----+-----+-----+-----+-----+-----+ 480
TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

S G S D R V I A S A E K F I E G F Q S A      180
TCTGGTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT
481 -----+-----+-----+-----+-----+-----+-----+ 540
AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAAGGTTAGACGA

K L A D P G A N P H Q A S P V I N V I I      200
AAGTTGGCTGACCCAGGTGCTAACCACACCAAGCTTCTCCAGTTATTAACGTTATTATT
541 -----+-----+-----+-----+-----+-----+-----+ 600
TTCAACCGACTGGGTCCACGATTGGGTGTGGTTTGAAGAGGTCAATAATTGCAATAATAA

P E G A G Y N N T L D H G L C T A F E E      220
CCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGTTTGTGTACTGCTTTCGAAGAA
601 -----+-----+-----+-----+-----+-----+-----+ 660
GGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGACGAAAGCTTCTT

```

图 8a

```

S E L G D D V E A N F T A V F A P P I R      240
TCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTGTTTTTCGCTCCACCAATTAGA
661 -----+-----+-----+-----+-----+-----+-----+ 720
AGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGACAAAAGCGAGGTGGTTAATCT

A R L E A H L P G V N L T D E D V V N L      260
GCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACGTTGTTAACTTG
721 -----+-----+-----+-----+-----+-----+-----+ 780
CGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGCAACAATTGAAC

M D M C P F D T V A R T S D A T Q L S P      280
ATGGACATGTGTCCATTGACACTGTTGCTAGAACTTCTGACGCTACTCAATTGTCTCCA
781 -----+-----+-----+-----+-----+-----+-----+ 840
TACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAGTTAACAGAGGT

F C D L F T H D E W I Q Y D Y L Q S L G      300
TTCTGTGACTTGTTCACTCACGACGAATGGATTCAATACGACTACTTGCAATCTTTGGGT
841 -----+-----+-----+-----+-----+-----+-----+ 900
AAGACACTGAACAAGTGAGTGCTTGCTTACCTAAGTTATGCTGATGAACGTTAGAAACCCA

K Y Y G Y G A G N P L G P A Q G V G F V      320
AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGTT
901 -----+-----+-----+-----+-----+-----+-----+ 960
TTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTGAGTTCCACAACCAAAGCAA

N E L I A R L T H S P V Q D H T S T N H      340
AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC
961 -----+-----+-----+-----+-----+-----+-----+ 1020
TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

T L D S N P A T F P L N A T L Y A D F S      360
ACTTTGGACTCTAACCAGCTACTTTCCCATTTGAACGCTACTTTGTACGCTGACTTCTCT
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
TGAAACCTGAGATTGGGTGCGATGAAAGGGTAACCTTGCGATGAAACATGCGACTGAAGAGA

H D N T M V S I F F A L G L Y N G T K P      380
CAGGACAACACTATGGTTTCTATTTCTTCGCTTTGGGTTTGTACAACGGTACTAAGCCA
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
GTGCTGTTGTGATACCAAAGATAAAAGAAGCGAAACCAAACATGTTGCCATGATTGCGGT

L S T T S V E S I E E T D G Y S A S W T      400
TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTTCTTGGACT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
AACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

V P F A A R A Y V E M M Q C E A E K E P      420
GTTCCATTGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTGAAAAGGAACCA
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACACTTCGACTTTTCCTGGT

L V R V L V N D R V V P L H G C G V D K      440
TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGGTGTGACAAG
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACCACAACCTGTTTC

```

图 8b

```
      L G R C K R D D F V E G L S F A R S G G      460
      TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 -----+-----+-----+-----+-----+-----+ 1380
      AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

      N W E E C F A *      467
      AACTGGGAAGAATGTTTCGCTTAA
1381 -----+-----+-----+-----+-----+ 1404
      TTGACCCTTCTTACAAAGCGAATT
```

图 8c

```

M G V F V V L L S I A T L F G S T S G T 20
ATGGGGGTTTTTCGTCGTTCTATTATCTATCGCGACTCTGTTCCGGCAGCACATCGGGCACT
1 -----+-----+-----+-----+-----+-----+-----+ 60
TACCCCCAAAAGCAGCAAGATAATAGATAGCGCTGAGACAAGCCGTCGTGTAGCCCGTGA

A L G P R G N H S K S C D T V D L G Y Q 40
GCGCTGGGCCCCCGTGGAAATCACTCCAAGTCTTGCATACGGTAGACCTAGGGTACCAG
61 -----+-----+-----+-----+-----+-----+-----+ 120
CGCGACCCGGGGGCACCTTTAGTGAGGTTCAAGACGCTATGCCATCTGGATCCCATGGTC

C S P A T S H L W G T Y S P Y F S L E D 60
TGCTCCCTGCGACTTCTCATCTATGGGGCACGTACTCGCCATaCTTTTCGCTCGAGGAC
121 -----+-----+-----+-----+-----+-----+-----+ 180
ACGAGGGGACGCTGAAGAGTAGATACCCCGtgCATGAGCGGTatGAAAAGCGAGTCTCTG

E L S V S S K L P K D C R I T L V Q V L 80
GAGCTGTCCGTGTCGAGTAAGCTTCCCAAGGATTGCCGGATCACCTTGGTACAGGTGCTA
181 -----+-----+-----+-----+-----+-----+-----+ 240
CTCGACAGGCACAGCTCATTGGAAGGGTTCCTAACGGCCTAGTGAACCATGTCCACGAT

S R H G A R Y P T S S K S K K Y K K L I 100
TCGCGCCATGGAGCGCGGTACCCAACAGCTCCAAGAGCAAAAAGTATAAGAAGCTTaTt
241 -----+-----+-----+-----+-----+-----+-----+ 300
AGCGCGGTACCTCGCGCCATGGGTTGGTCGAGGTTCTCGTTTTTCATATTCTTCGAAtAa

T A I Q A N A T D F K G K Y A F L K T Y 120
ACGGCGATCCAGGCCAATGCCACCGACTTCAAGGGCAAGTAcGCCTTTTTTGAAGACGTAC
301 -----+-----+-----+-----+-----+-----+-----+ 360
TGCCGCTAGGTCCGGTTACGGTGGCTGAAGTTCCCGTTCatgCGGAAAACTTCTGCATG

N Y T L G A D D L T P F G E Q Q L V N S 140
AACTATACTCTGGGTGCGGATGACCTCACTCCCTTTGGGGAGCAGCAGCTGGTGAACCTG
361 -----+-----+-----+-----+-----+-----+-----+ 420
TTGATATGAGACCCACGCCTACTGGAGTGAGGGAAACCCCTCGTCGTCGACCACTTGAGC

G I K F Y Q R Y K A L A R S V V P F I R 160
GGCATCAAGTTCTACCAGAGGTACAAGGCTCTGGCGCGCAGTGTGGTGCCGTTTATTTCGC
421 -----+-----+-----+-----+-----+-----+-----+ 480
CCGTAGTTCAAGATGGTCTCCATGTTCCGAGACCGCGGTCACACCACGGCAAATAAGCG

A S G S D R V I A S G E K F I E G F Q Q 180
GCCTCAGGCTCGGACCGGGTTATTGCTTCGGGAGAGAAGTTCATCGAGGGGTTCCAGCAG
481 -----+-----+-----+-----+-----+-----+-----+ 540
CGGAGTCCGAGCCTGGCCCAATAACGAAGCCCTCTCTTCAAGTAGCTCCCCAAGTCTGTC

A K L A D P G A T N R A A P A I S V I I 200
GCGAAGCTGGCTGATCCTGGCGCGACGAACCGCGCGCTCCGGCGATTAGTGTGATTATT
541 -----+-----+-----+-----+-----+-----+-----+ 600
CGCTTCGACCGACTAGGACCGCGCTGCTTGGCGCGGCGAGGCCGCTAATCACACTAATAA

P E S E T F N N T L D H G V C T K F E A 220
CCGGAGAGCGAGACGTTCAACAATACGCTGGACCACGGTGTGTGCACGAAGTTTGAGGCG
601 -----+-----+-----+-----+-----+-----+-----+ 660
GGCCTCTCGCTCTGCAAGTTGTTATGCGACCTGGTGCCACACACGTGCTTCAAACCTCCGC

```

图 9a

```

S Q L G D E V A A N F T A L F A P D I R 240
AGTCAGCTGGGAGATGAGGTTGCGGCCAATTTCACTGCGCTCTTTGCACCCGACATCCGA
661 -----+-----+-----+-----+-----+-----+ 720
TCAGTCGACCCTCTACTCCAACGCCGGTTAAAGTGACGCGAGAAACGTGGGCTGTAGGCT

A R L E K H L P G V T L T D E D V V S L 260
GCTCGCctCGAGAAGCATCTTCCTGGCGTGACGCTGACAGACGAGGACGTTGTTCAGTCTA
721 -----+-----+-----+-----+-----+-----+ 780
CGAGCGgaGCTCTTCGTAGAAGGACCGCACTGCGACTGTCTGCTCCTGCAACAGTCAGAT

M D M C P F D T V A R T S D A S Q L S P 280
ATGGACATGTGTcCGTTTGATACGGTAGCGCGCACCAGCGACGCAAGTCAGCTGTCACCG
781 -----+-----+-----+-----+-----+-----+ 840
TACCTGTACACagGCAAACCTATGCCATCGCGCGTGGTTCGCTGCGTTTCAGTCGACAGTGGC

F C Q L F T H N E W K K Y D Y L Q S L G 300
TTCTGTCAACTCTTCACTCACAATGAGTGAAGAAGTACgACTACCTTCAGTCCTTGGGC
841 -----+-----+-----+-----+-----+-----+ 900
AAGACAGTTGAGAAGTGAGTGTACTCACCTTCTTCATGcTGATGGAAGTCAGGAACCCG

K Y Y G Y G A G N P L G P A Q G I G F T 320
AAGTACTACGGCTACGGCGCAGGCAACCCTCTGGGACCGGCTCAGGGGATAGGGTTCCACC
901 -----+-----+-----+-----+-----+-----+ 960
TTCATGATGCCGATGCCGCGTCCGTTGGGAGACCCTGGCCGAGTCCCCTATCCCAAGTGG

N E L I A R L T R S P V Q D H T S T N S 340
AACGAGCTGATTGCCCCGTTGACgCGTTCGCCAGTGACGAGCACACCAGCACTAACTCG
961 -----+-----+-----+-----+-----+-----+ 1020
TTGCTCGACTAACGGGCCAACTGcGCAAGCGGTCACGTCCTGGTGTGGTTCGTGATTGAGC

T L V S N P A T F P L N A T M Y V D F S 360
ACTCTAGTCTCCAACCCGGCCACCTTCCCGTTGAACGCTACCATGTACGTCGACTTTTCA
1021 -----+-----+-----+-----+-----+-----+ 1080
TGAGATCAGAGGTTGGGCGCGTGGAAGGGCAACTTGCGATGGTACATGCAGCTGAAAAGT

H D N S M V S I F F A L G L Y N G T E P 380
CACGACAACAGCATGGTTTCCATCTTCTTTGCATTGGGCCTGTACAACGGCACTGAACCC
1081 -----+-----+-----+-----+-----+-----+ 1140
GTGCTGTTGTCTACCAAAGGTAGAAGAAACGTAACCCGACATGTTGCCGTGACTTGGG

L S R T S V E S A K E L D G Y S A S W V 400
TTGTCCCGGACCTCGGTGGAAGCGCCAAGGAATTGGATGGGTATTCTGCATCCTGGGTG
1141 -----+-----+-----+-----+-----+-----+ 1200
AACAGGGCCTGGAGCCACCTTTCGCGGTTCCCTTAACCTACCCATAAGACGTAGGACCCAC

V P F G A R A Y F E T M Q C K S E K E P 420
GTGCCTTTCGGCGCGCGAGCCTACTTCGAGACGATGCAATGCAAGTCGGAAAAGGAGCCT
1201 -----+-----+-----+-----+-----+-----+ 1260
CACGGAAAGCCGCGCGCTCGGATGAAGCTCTGCTACGTTACGTTTCAGCCTTTTCCTCGGA

L V R A L I N D R V V P L H G C D V D K 440
CTTGTTTCGCGCTTTGATTAATGACCGGGTTGTGCCACTGCATGGCTGCGATGTGGACAAG
1261 -----+-----+-----+-----+-----+-----+ 1320
GAACAAGCGGAAACTAATTACTGGCCCAACACGGTGACGTACCGACGCTACACCTGTTTC

```

图 9b

```
      L G R C K L N D F V K G L S W A R S G G      460
      CTGGGGCGATGCAAGCTGAATGACTTTGTCAAGGGATTGAGTTGGGCCAGATCTGGGGGC
1321 -----+-----+-----+-----+-----+-----+ 1380
      GACCCCGCTACGTTGACTTACTGAAACAGTTCCCTAACTCAACCCGGTCTAGACCCCG

      N W G E C F S *      467
      AACTGGGGAGAGTGCTTTAGTTGA
1381 -----+-----+----- 1404
      TTGACCCCTCTCACGAAATCAACT
```

图 9c

CP-1
 Eco RI M G V F V V L L S I A T L F G S T
 TATATGAATTCATGGGCGTGTTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA
 1 -----+-----+-----+-----+-----+ 60
 ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGAACAAGCCAAGGT

 S G T A L G P R G N S H S C D T V D G G
 CATCCGGTACCGCCTTGGGTCTCGTGGAATTCTCACTCTGTGACACTGTTGACGGTG
 61 -----+-----+-----+-----+-----+ 120
 GTAGGCCATGGCGGAACCCAGGAGCACCATTAAAGAGTGAGAACACTGTGACAACCTGCCAC
 CP-2
 CP-3
 Y Q C F P E I S H L W G Q Y S P Y F S L
 GTTACCAATGTTTCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT
 121 -----+-----+-----+-----+-----+ 180
 CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCAGTTATGAGAGGTATGAAGAGAA

 E D E S A I S P D V P D D C R V T F V Q
 TGAAGACGAATCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCGTTT
 181 -----+-----+-----+-----+-----+ 240
 ACCTTCTGCTTAGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG
 CP-4.7
 CP-5.7
 V L S R H G A R Y P T D S K G K K Y S A
 AAGTTTGTCTAGACACGGTGCTAGATACCCAACTgacTCTAAGggtAAGaagTACTCTG
 241 -----+-----+-----+-----+-----+ 300
 TTCAAAACAGATCTGTGCCACGATCTATGGGTTGActgAGATTCCAATTCcttCATGAGAC

 L I E A I Q K N A T A F K G K Y A F L K
 CTTTGATTGAAGCTATTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTCTTGA
 301 -----+-----+-----+-----+-----+ 360
 GAAACTAACTTCGATAAGTTTTCTTGCAGTACGAAAGTTCCCATTCATGCGAAAGAACT
 CP-6
 CP-7
 T Y N Y T L G A D D L T P F G E N Q M V
 AGACTTACAACCTACACTTTGGGTGCTGACGACTTGACTCCATTCCGGTGAACCAATGG
 361 -----+-----+-----+-----+-----+ 420
 TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACC

 N S G I K F Y R R Y K A L A R K I V P F
 TTAACCTCTGGTATTAAGTTCTACAGAAGATAACAAGGCTTTGGCTAGAAAGATTGTTCCAT
 421 -----+-----+-----+-----+-----+ 480
 AATTGAGACCATAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTCTAACAAGGTA
 CP-8.7
 CP-9
 I R A S G S S R V I A S A E K F I E G F
 TCATTAGAGCTTCTGGTTCTTctAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTT
 481 -----+-----+-----+-----+-----+ 540
 AGTAATCTCGAAGACCAAGAgaTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

 Q S A K L A D P G S Q P H Q A S P V I D
 TCCAATCTGCTAAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTG
 541 -----+-----+-----+-----+-----+ 600
 AGGTTAGACGATTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTTCAAGAGGTCAATAAC

图 10a

CP-10.7
CP-11.7

V I I S E A S S Y N N T L D P G T C T A
ACGTTATTATTtctGAcgctTCTtctTACAACAACACTTTGGACccaGGTACTTGTACTG
601 -----+-----+-----+-----+-----+ 660
TGCAATAATAAagaCTgcgaAGGagaATGTTGTTGTGAAACCTGggtCCATGAACATGAC

F E D S E L A D T V E A N F T A L F A P
CTTTCGAAGACTCTGAATTGgctGACactGTTGAAGCTAACTTCACTGCTTTGTTCGCTC
661 -----+-----+-----+-----+-----+ 720
GAAAGCTTCTGAGACTTAACcgaCTGtgaCAACTTCGATTGAAGTGACGAAACAAGCGAG

CP-12.7

A I R A R L E A D L P G V T L T D T E V
CAGCTATTAGAGCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACactgaaG
721 -----+-----+-----+-----+-----+ 780
GTGCATAATCTCGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGtgacttc

CP-13.7

T Y L M D M C S F E T V A R T S D A T E
TTactTACTTGATGGACATGTGTtctTTTCGAAACTGTTGCTAGAACTTCTGACGCTACTG
781 -----+-----+-----+-----+-----+ 840
AatgaATGAACTACCTGTACACAagaAAGCTTTGACAACGATCTTGAAGACTGCGATGAC

L S P F C A L F T H D E W R H Y D Y L Q
AATTGTCTCCATTCTGTGCTTTGTTCACTCACGACGAATGGAGacacTACGACTACTTGC
841 -----+-----+-----+-----+-----+ 900
TTAACAGAGGTAAGACACGAAACAAGTGAGTGCTGCTTACCTCTgtgATGCTGATGAACG

CP-14.7
CP-15.7

S L K K Y Y G H G A G N P L G P T Q G V
AATCTTTGaagAAGTACTACGGTcacGGTGCTGGTAACCCATTGGGTCCAactCAAGGTG
901 -----+-----+-----+-----+-----+ 960
TTAGAAACTtctTTCATGATGCCAgtagCCACGACCATTGGGTAACCCAGGTtgaGTTCCAC

G F A N E L I A R L T R S P V Q D H T S
TTGGTTTCGCTAACGAATTGATTGCTAGATTGACTAGATCTCCAGTTCAAGACCACACTT
961 -----+-----+-----+-----+-----+ 1020
AACCAAAGCGATTGCTTAACTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA

CP-16
CP-17.7

T N H T L D S N P A T F P L N A T L Y A
CTACTAACACACTTTGGACTCTAACCCAGCTACTTTCCCATGAACGCTACTTTGTACG
1021 -----+-----+-----+-----+-----+ 1080
GATGATTGGTGTGAAACCTGAGATTGGGTGATGAAAGGGTAACCTTGCGATGAAACATGC

D F S H D N G I I S I F F A L G L Y N G
CTGACTTCTCTCACGACAACggtattATTTCTATTTCTTCGCTTTGGGTTTGTACAACG
1081 -----+-----+-----+-----+-----+ 1140
GACTGAAGAGAGTGCTGTTGccataaTAAAGATAAAAGAAGCGAAACCCAAACATGTTGC

CP-18.7
CP-19.7

T A P L S T T S V E S I E E T D G Y S S
GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTt
1141 -----+-----+-----+-----+-----+ 1200
CATGACGAGGTAACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGAA

图 10b

```

      A W T V P F A S R A Y V E M M Q C Q A E
ctgctTGGACTGTTCCATTCgcttctAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG
1201 -----+-----+-----+-----+-----+-----+ 1260
gacgaACCTGACAAGGTAAGcgaagaTCTCGAATGCAACTTTACTACGTTACAGTTCGAC
                                CP-20
                                CP-21
      K E P L V R V L V N D R V V P L H G C A
AAAAGGAACCATTTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
1261 -----+-----+-----+-----+-----+-----+ 1320
TTTTCTTGGTAACCAATCTCAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC

      V D K L G R C K R D D F V E G L S F A R
CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA
1321 -----+-----+-----+-----+-----+-----+ 1380
GACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGAT
                                CP-22
      S G G N W A E C F A * Eco RI
GATCTGGTGGTAACTGGGCTGAATGTTTCGCTTAAGAATTCATATA
1381 -----+-----+-----+-----+-----+-----+ 1426
CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT

```

图 10c

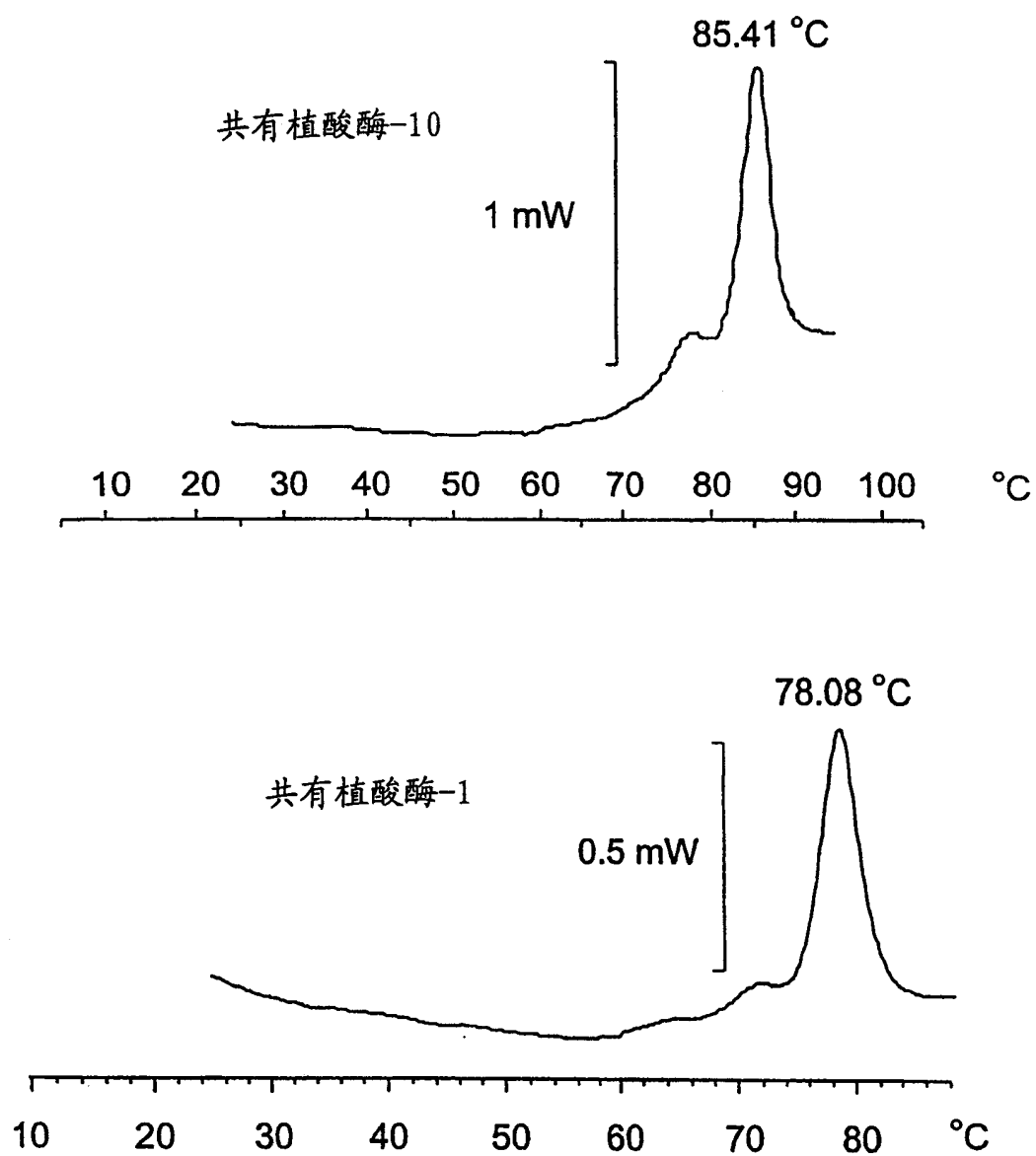


图 11

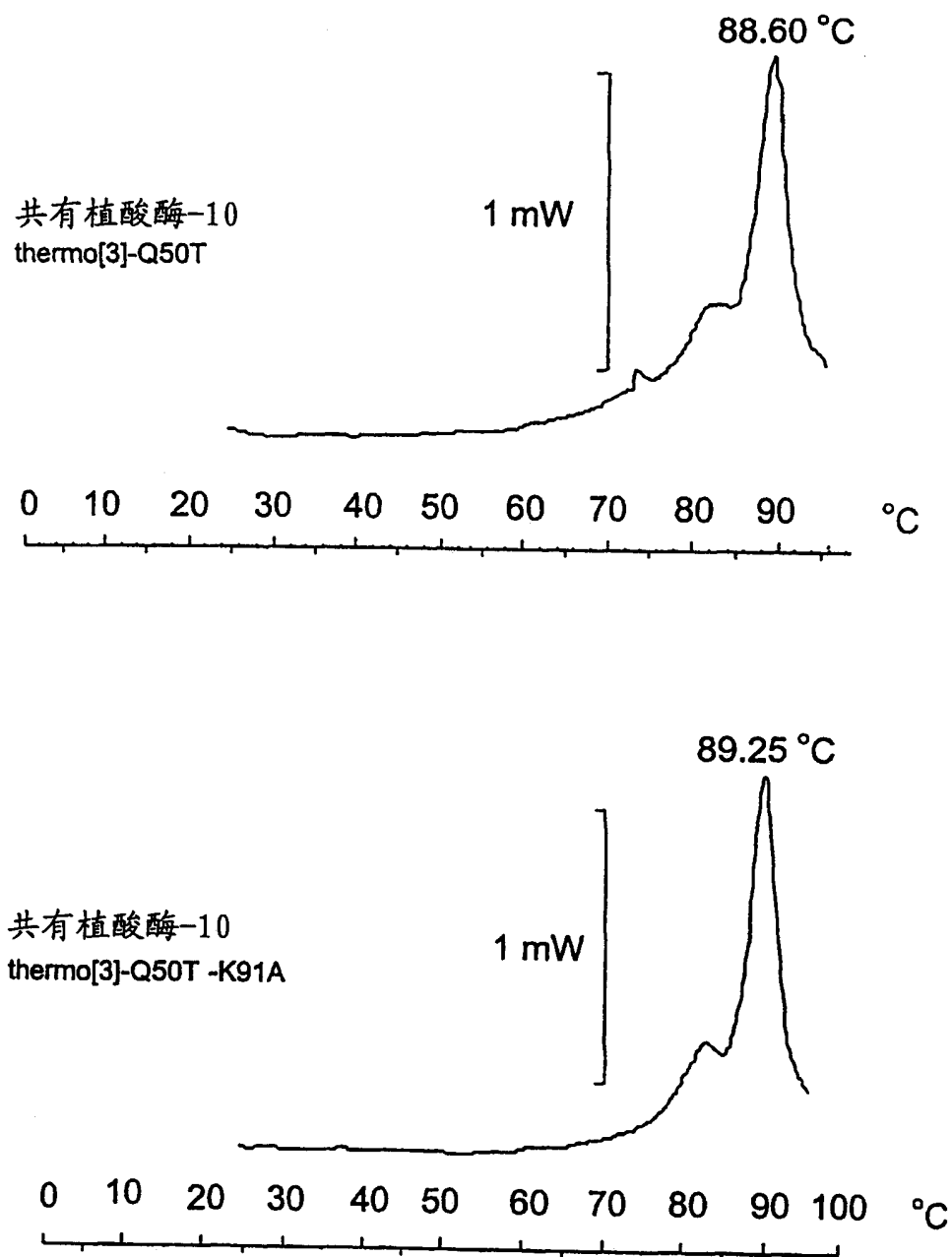


图 12

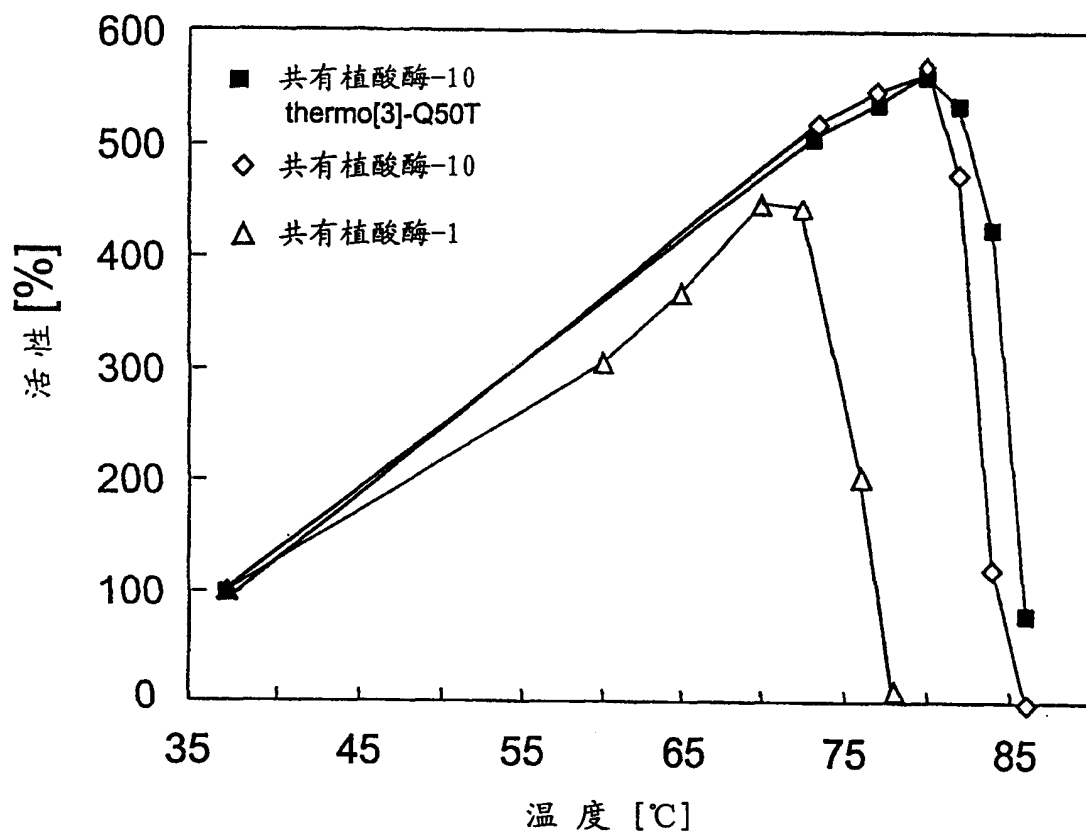


图 13

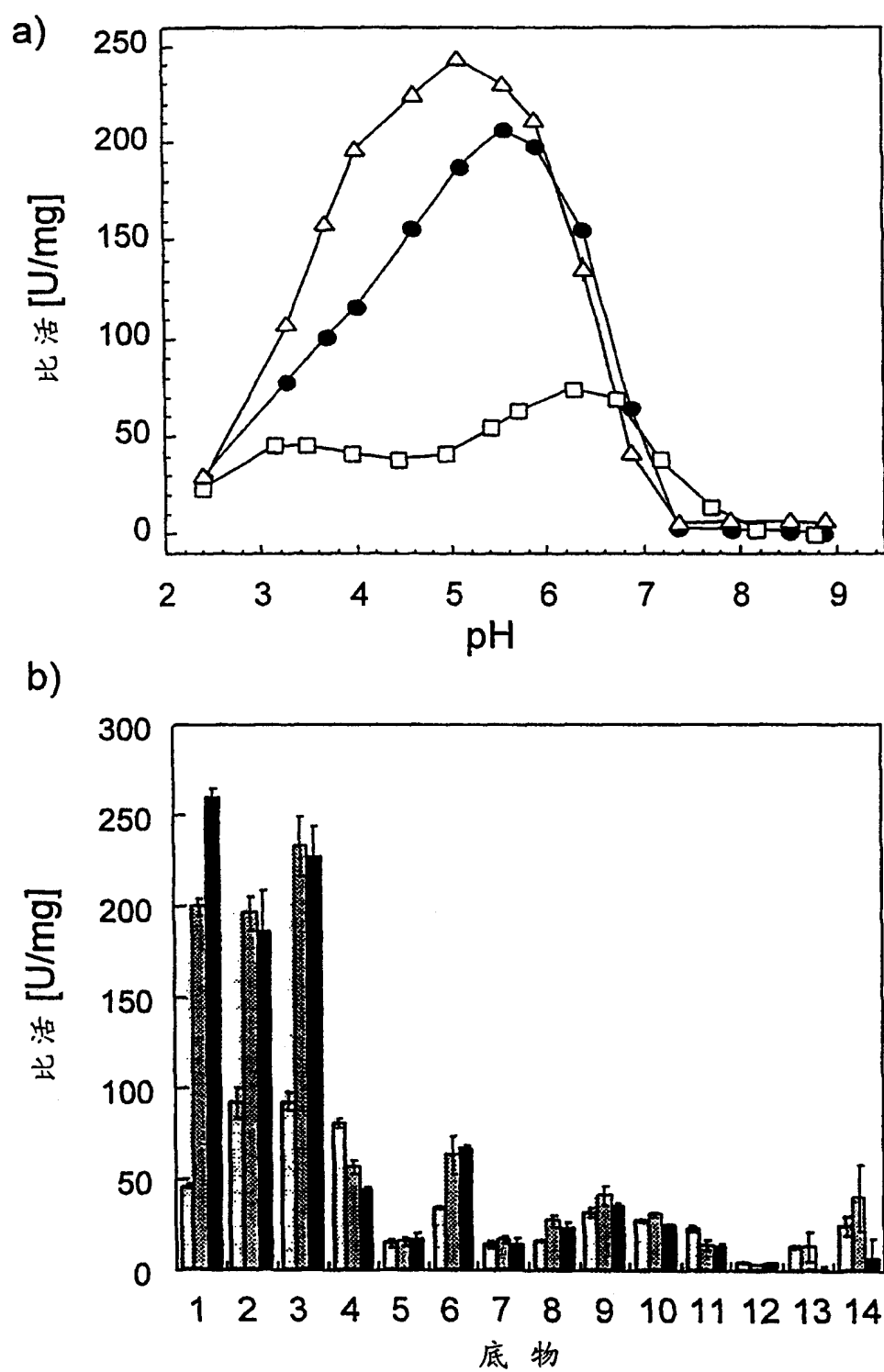


图 14

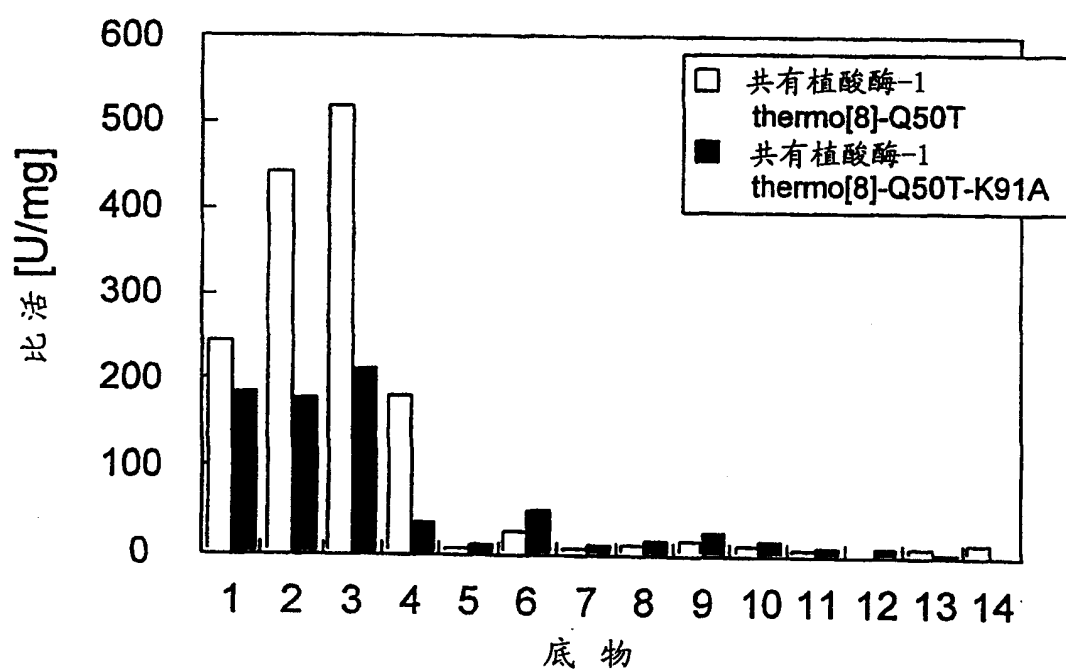
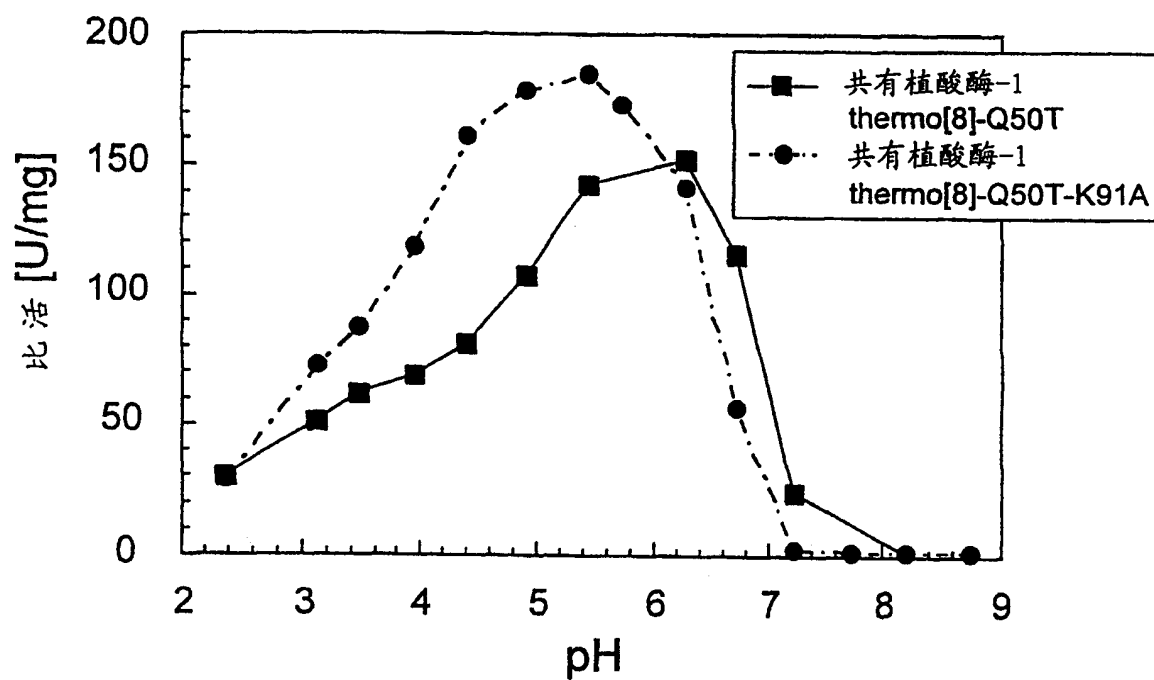


图 15

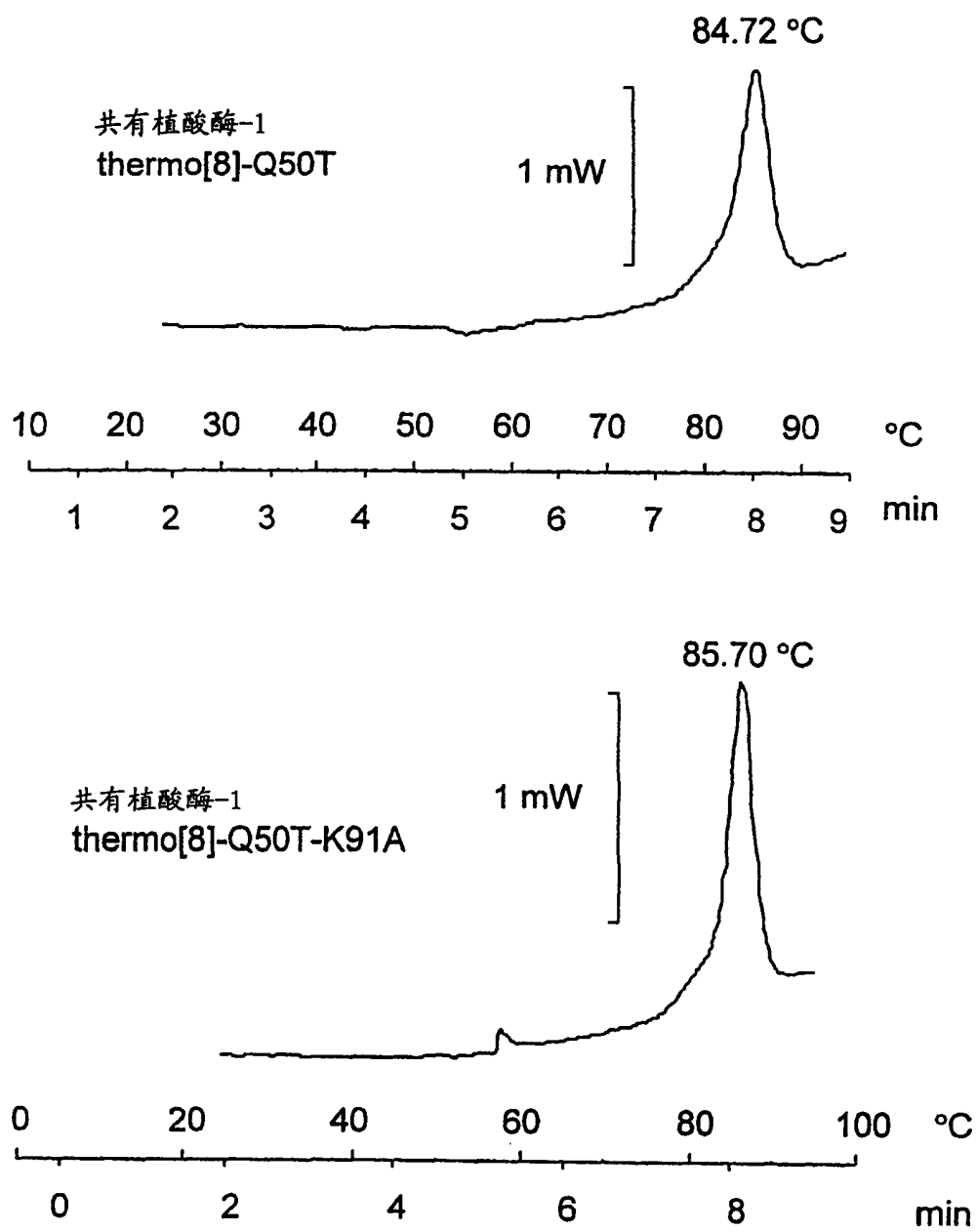


图 16

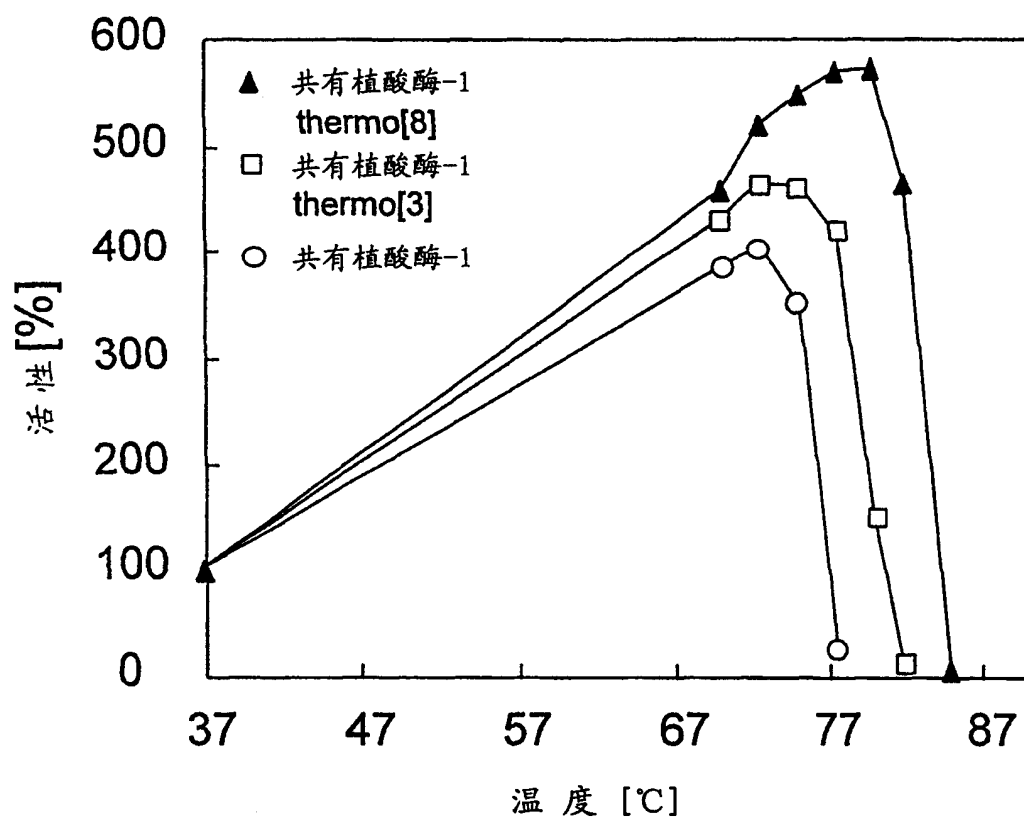


图 17

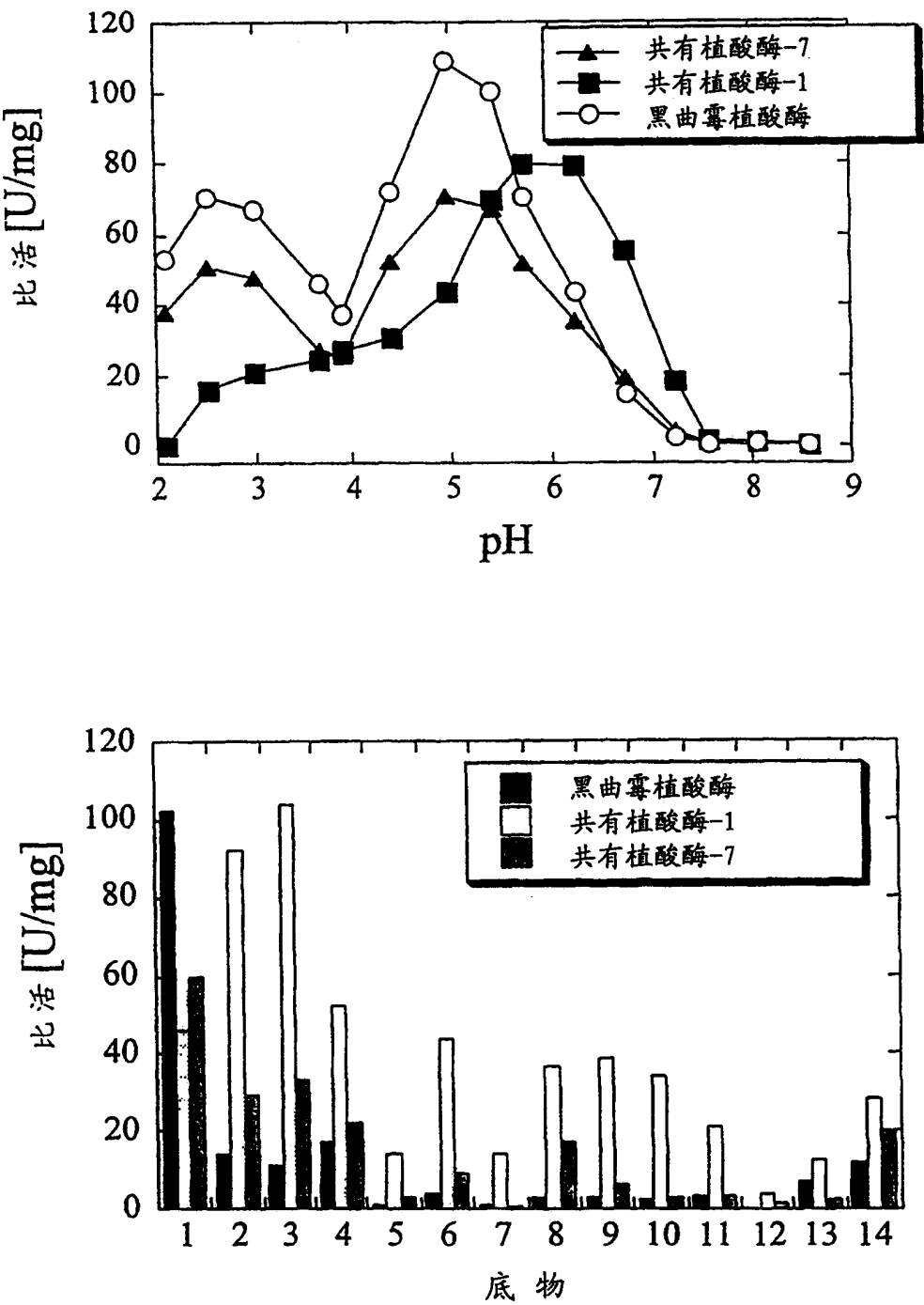


图 18

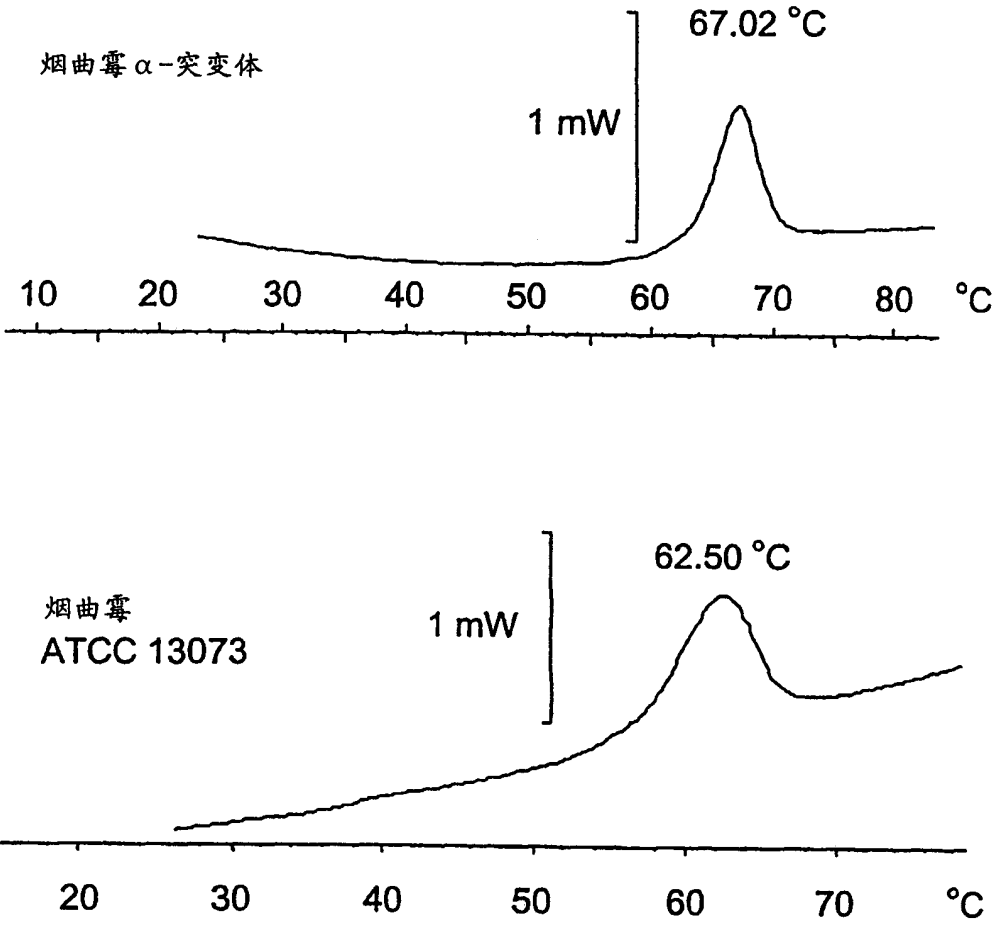


图 19

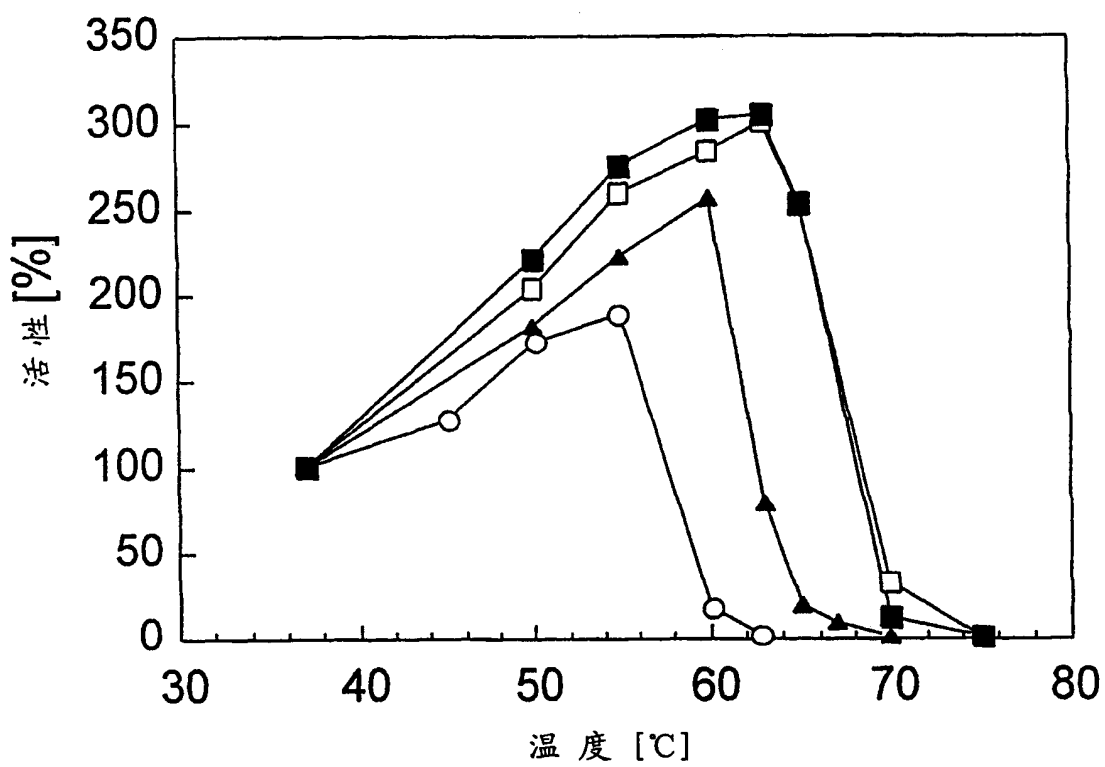


图 20

1 MGVFVLLSI ATLFGSTSGT ALGPRGNSHS CDTVDGGYQC FPEISSNWSP
51 YSPYFLADE SAISPDVPKG CRVTFVQVLQ RHGARFPTSG AATRISALIE
101 AIQKNATAFK GKYAFLKTYN YTLGADDLVP FGANQSSQAG IKFYRRYKAL
151 ARKIVPFIRA SGSDRVIDSA TNWIEGFQSA KLADPGANPH QASPVINVII
201 PEGAGYNNTL DHGLCTAFEE SELGDDVEAN FTAVFAPPIR ARLEAHLPGV
251 NLTDEDVVNL MDMCPFDIVA RTSDATELSP FCDLFTHDEW IQDYLGDL
301 KYYGTGAGNP LGPAQGVGFV NELIARLTHS PVQDHTSTNH TLDSNPATFP
351 LNATLYADFS HDNTMVAIFF ALGLYNGTKP LSTTSVESIE ETDGYSASWL
401 VPFSARMYVE MMQCEAEKEP LVRVLVNDRV VPLHGCGVDK LGRCKRDDFV
451 EGLSFARSGG NWEECFA

图 21

```

ATGGGCGTGTTCGTGCTGCTACTGTCCATTGCCACCTTGTTCGGTTCACATCCGGTACC
1  -----+-----+-----+-----+-----+-----+ 60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG

M G V F V V L L S I A T L F G S T S G T -

GCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61  -----+-----+-----+-----+-----+-----+ 120
CGGAACCCAGGAGCACCATTAAAGAGTGAGAACACTGTGACAACTGCCACCAATGGTTACA

A L G P R G N S H S C D T V D G G Y Q C -

TTCCCAGAAATTTCTCACTTGTGGGGTACCTACTCTCCATACTTCTCTTTGGCAGACGAA
121 -----+-----+-----+-----+-----+-----+ 180
AAGGGTCTTTAAAGAGTGAACACCCCATGGATGAGAGGTATGAAGAGAAACCGTCTGCTT

F P E I S H L W G T Y S P Y F S L A D E -

TCTGCTATTTCTCCAGACGTCCCAAAGGACTGTAGAGTTACTTTTCGTTCAAGTTTGTCT
181 -----+-----+-----+-----+-----+-----+ 240
AGACGATAAAGAGGTCTGCAGGGTTTCTGACATCTCAATGAAAGCAAGTTCAAAACAGA

S A I S P D V P K D C R V T F V Q V L S -

AGACACGGTGCTAGATACCCAACTTCTTCTAAGTCTAAGGCTTACTCTGCTTTGATTGAA
241 -----+-----+-----+-----+-----+-----+ 300
TCTGTGCCACGATCTATGGGTTGAAGAAGATTCCAGATTCCGAATGAGACGAACTAACTT

R H G A R Y P T S S K S K A Y S A L I E -

GCTATTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
301 -----+-----+-----+-----+-----+-----+ 360
CGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTG

A I Q K N A T A F K G K Y A F L K T Y N -

TACACTTTGGGTGCTGACGACTTGACTCCATTCCGGTGAAAACCAAATGGTTAACTCTGGT
361 -----+-----+-----+-----+-----+-----+ 420
ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACCAATTGAGACCA

Y T L G A D D L T P F G E N Q M V N S G -

ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTCAATAGAGCT
421 -----+-----+-----+-----+-----+-----+ 480
TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

I K F Y R R Y K A L A R K I V P F I R A -

TCTGGTTCCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTTTCCAATCTGCT
481 -----+-----+-----+-----+-----+-----+ 540
AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAAGGTTAGACGA

S G S D R V I A S A E K F I E G F Q S A -

```

图 22a

```

AAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTAACGTGATCATT
541 -----+-----+-----+-----+-----+-----+ 600
TTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTTCAAGAGGTCAATAATTGCACTAGTAA

K L A D P G S Q P H Q A S P V I N V I I -

CCAGAAGGATCCGGTTACAACAACACTTTGGACCATGGTCTTTGTACTGCTTTTGAAGAC
601 -----+-----+-----+-----+-----+-----+ 660
GGTCTTCCTAGGCCAATGTTGTTGTGAAACCTGGTACCAGAAACATGACGAAAGCTTCTG

P E G S G Y N N T L D H G L C T A F E D -

TCTACCCTAGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTCGCTCCAGCTATTAGA
661 -----+-----+-----+-----+-----+-----+ 720
AGATGGGATCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAGGTGCGATAATCT

S T L G D D V E A N F T A L F A P A I R -

GCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACGTTGTTTACTTG
721 -----+-----+-----+-----+-----+-----+ 780
CGATCTAACCTTCGACTGAACGGTCCACAATGAACTGACTGCTTCTGCAACAAATGAAC

A R L E A D L P G V T L T D E D V V Y L -

ATGGACATGTGTCCATTTCGACACTGTGCTAGAACTTCTGACGCTACTGAATTGTCTCCA
781 -----+-----+-----+-----+-----+-----+ 840
TACCTGTACACAGGTAAGCTGTGACAGCGATCTTGAAGACTGCGATGACTTAACAGAGGT

M D M C P F D T V A R T S D A T E L S P -

TTCTGTGCTTTGTTCACTCACGACGAATGGATCCAATACGACTACTTGCAAAGCTTGGGT
841 -----+-----+-----+-----+-----+-----+ 900
AAGACACGAAACAAGTGAGTGCTGCTTACCTAGGTTATGCTGATGAACGTTTCGAACCCA

F C A L F T H D E W I Q Y D Y L Q S L G -

AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGCT
901 -----+-----+-----+-----+-----+-----+ 960
TTCATGATGCCAATGCCACGACCATTGGGTAAACCCAGGTGCGAGTTCCACAACCAAAGCGA

K Y Y G Y G A G N P L G P A Q G V G F A -

AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC
961 -----+-----+-----+-----+-----+-----+ 1020
TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

N E L I A R L T H S P V Q D H T S T N H -

ACTTTGGACTCTAACCCAGCTACTTTCCCATTTGAACGCTACTTTGTACGCTGACTTCTCT
1021 -----+-----+-----+-----+-----+-----+ 1080
TGAAACCTGAGATTGGGTGCGATGAAAGGGTAACTTGCATGAAACATGCGACTGAAGAGA

T L D S N P A T F P L N A T L Y A D F S -

```

图 22b

```

      CACGACAACACTATGATATCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACCAAGCCA
1081 -----+-----+-----+-----+-----+-----+ 1140
      GTGCTGTTGTGATACTATAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGGTTCGGT

      H D N T M I S I F F A L G L Y N G T K P -

      TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTTCTTGGA
1141 -----+-----+-----+-----+-----+-----+ 1200
      AACAGATGATGAAGACAACCTTAGATAAATTCTTTGACTGCCAATGAGACGAAGAACCTGA

      L S T T S V E S I E E T D G Y S A S W T -

      GTTCCATTTCGCTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTGAAAAGGAACCA
1201 -----+-----+-----+-----+-----+-----+ 1260
      CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACAGTTCGACTTTTCCTTGGT

      V P F A A R A Y V E M M Q C Q A E K E P -

      TTGGTTAGAGTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGCTGTTGACAAG
1261 -----+-----+-----+-----+-----+-----+ 1320
      AACCAATCTCAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACGACAACCTGTT

      L V R V L V N D R V V P L H G C A V D K -

      TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 -----+-----+-----+-----+-----+-----+ 1380
      AACCCATCTACATTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

      L G R C K R D D F V E G L S F A R S G G -

      AACTGGGCTGAATGTTTCGCTTAA
1381 -----+-----+----- 1404
      TTGACCCGACTTACAAAGCGAATT

      N W A E C F A *

```

图 22c

```

ATGGGCGTGTTCGTGCTACTGTCCATTGCCACCTTGTTTCGGTTCACATCCGGTACC
1 -----+-----+-----+-----+-----+-----+ 60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGAACAAGCCAAGGTGTAGGCCATGG

M G V F V V L L S I A T L F G S T S G T -

GCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61 -----+-----+-----+-----+-----+-----+ 120
CGGAACCCAGGAGCACCATTAAAGAGTGAGAACACTGTGACAACCTGCCACCAATGGTTACA

A L G P R G N S H S C D T V D G G Y Q C -

TTCCAGAAATTTCTCACTTGTGGGGTACCTACTCTCCATACTTCTCTTTGGCAGACGAA
121 -----+-----+-----+-----+-----+-----+ 180
AAGGGTCTTTAAAGAGTGAACACCCCATGGATGAGAGGTATGAAGAGAAACCGTCTGCCT

F P E I S H L W G T Y S P Y F S L A D E -

TCTGCTATTTCTCCAGACGTCCCAAAGGACTGTAGAGTTACTTTTCGTTCAAGTTTTGTCT
181 -----+-----+-----+-----+-----+-----+ 240
AGACGATAAAGAGGTCTGCAGGGTTTCTGACATCTCAATGAAAGCAAGTTCAAAACAGA

S A I S P D V P K D C R V T F V Q V L S -

AGACACGGTGCTAGATACCCAACCTTCTTCTGCGTCTAAGGCTTACTCTGCTTTGATTGAA
241 -----+-----+-----+-----+-----+-----+ 300
TCTGTGCCACGATCTATGGGTTGAAGAAGACGCAGATTCCGAATGAGACGAAACTAACTT

R H G A R Y P T S S A S K A Y S A L I E -

GCTATTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAC
301 -----+-----+-----+-----+-----+-----+ 360
CGATAAGTTTTCTTGCATGACGAAAGTTCCCATTCATGCGAAAGAAGTTCTGAATGTTG

A I Q K N A T A F K G K Y A F L K T Y N -

TACACTTTGGGTGCTGACGACTTGACTCCATTCCGGTGAAAACCAAATGGTTAACTCTGGT
361 -----+-----+-----+-----+-----+-----+ 420
ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACCAATTGAGACCA

Y T L G A D D L T P F G E N Q M V N S G -

ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTTCATTAGAGCT
421 -----+-----+-----+-----+-----+-----+ 480
TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

I K F Y R R Y K A L A R K I V P F I R A -

TCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTTCATTGAAGGTTTCCAATCTGCT
481 -----+-----+-----+-----+-----+-----+ 540
AGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAAGTTCCAAAGGTTAGACGA

S G S D R V I A S A E K F I E G F Q S A -

```

图 23a

```

AAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTAACGTGATCATT
541 -----+-----+-----+-----+-----+ 600
TTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCGAAGAGGTCAATAATTGCACTAGTAA

K L A D P G S Q P H Q A S P V I N V I I -

CCAGAAGGATCCGGTTACAACAACACTTTGGACCATGGTCTTTGTACTGCTTTCGAAGAC
601 -----+-----+-----+-----+-----+ 660
GGTCTTCCTAGGCCAATGTTGTTGTGAAACCTGGTACCAGAAACATGACGAAAGCTTCTG

P E G S G Y N N T L D H G L C T A F E D -

TCTACCCTAGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTGCTCCAGCTATTAGA
661 -----+-----+-----+-----+-----+ 720
AGATGGGATCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAGGTCGATAATCT

S T L G D D V E A N F T A L F A P A I R -

GCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACGTTGTTTACTTG
721 -----+-----+-----+-----+-----+ 780
CGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGCAACAAATGAAC

A R L E A D L P G V T L T D E D V V Y L -

ATGGACATGTGTCCATTTCGACACTGTCGCTAGAACTTCTGACGCTACTGAATTGTCTCCA
781 -----+-----+-----+-----+-----+ 840
TACCTGTACACAGGTAAGCTGTGACAGCGATCTTGAAGACTGCGATGACTTAACAGAGGT

M D M C P F D T V A R T S D A T E L S P -

TTCTGTGCTTTGTTCACTCACGACGAATGGATCCAATACGACTACTTGCAAAGCTTGGGT
841 -----+-----+-----+-----+-----+ 900
AAGACACGAAACAAGTGAGTGCTGCTTACCTAGGTTATGCTGATGAACGTTTCGAACCCA

F C A L F T H D E W I Q Y D Y L Q S L G -

AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGCT
901 -----+-----+-----+-----+-----+ 960
TTCATGATGCCAATGCCACGACCATTTGGGTAACCCAGGTTCGAGTTCCACAACCAAAGCGA

K Y Y G Y G A G N P L G P A Q G V G F A -

AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAGACCACTTCTACTAACCAC
961 -----+-----+-----+-----+-----+ 1020
TTGCTTAACCTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

N E L I A R L T H S P V Q D H T S T N H -

ACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACGCTGACTTCTCT
1021 -----+-----+-----+-----+-----+ 1080
TGAAACCTGAGATTGGGTTCGATGAAAGGGTAACCTGCGATGAAACATGCGACTGAAGAGA

T L D S N P A T F P L N A T L Y A D F S -

```

图 23b

```

      CACGACAACACTATGATATCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACCAAGCCA
1081 -----+-----+-----+-----+-----+-----+ 1140
      GTGCTGTTGTGATACTATAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGGTTCCGGT

      H D N T M I S I F F A L G L Y N G T K P -

      TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTTCTTGGACT
1141 -----+-----+-----+-----+-----+-----+ 1200
      AACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

      L S T T S V E S I E E T D G Y S A S W T -

      GTTCCATTTCGCTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTGAAAAGGAACCA
1201 -----+-----+-----+-----+-----+-----+ 1260
      CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACAGTTCGACTTTTCCTTGGT

      V P F A A R A Y V E M M Q C Q A E K E P -

      TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGCTGTTGACAAG
1261 -----+-----+-----+-----+-----+-----+ 1320
      AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACGACAACCTGTTT

      L V R V L V N D R V V P L H G C A V D K -

      TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 -----+-----+-----+-----+-----+-----+ 1380
      AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

      L G R C K R D D F V E G L S F A R S G G -

      AACTGGGCTGAATGTTTCGCTTAA
1381 -----+-----+-----+-----+-----+ 1404
      TTGACCCGACTTACAAAGCGAATT

      N W A E C F A *

```

图 23c

```

ATGGGCGTGTTCGTGCTGCTACTGTCCATTGCCACCTTGTTTCGGTTCCACATCCGGTACC
1  -----+-----+-----+-----+-----+-----+ 60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG

M G V F V V L L S I A T L F G S T S G T -

GCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61  -----+-----+-----+-----+-----+-----+ 120
CGGAACCCAGGAGCACCATTAAAGAGTGAGAACTGTGACAACTGCCACCAATGGTTACA

A L G P R G N S H S C D T V D G G Y Q C -

TTCACAGAAATTTCTCACTTGTGGGTACATACTCTCCATTCTTCTTTGGCTGACGAA
121 -----+-----+-----+-----+-----+-----+ 180
AAGGTCCTTTAAAGAGTGAACACCCCATGTATGAGAGGTAAGAAGAGAAACCGACTGCTT

F P E I S H L W G T Y S P F F S L A D E -

TCTGCTATTTCTCCAGACGTTCCAAAGGGTTGTAGAGTTACTTTCGTTCAAGTTTTGTCT
181 -----+-----+-----+-----+-----+-----+ 240
AGACGATAAAGAGGTCTGCAAGGTTTCCCAACATCTCAATGAAAGCAAGTTCAAAACAGA

S A I S P D V P K G C R V T F V Q V L S -

AGACACGGTGCTAGATACCCAACTTCTTCTAAGTCTAAGGCTTACTCTGCTTTGATTGAA
241 -----+-----+-----+-----+-----+-----+ 300
TCTGTGCCACGATCTATGGGTTGAAGAAGATTCAAGATTCCGAATGAGACGAACTAACTT

R H G A R Y P T S S K S K A Y S A L I E -

GCTATTTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAT
301 -----+-----+-----+-----+-----+-----+ 360
CGATAAGTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTA

A I Q K N A T A F K G K Y A F L K T Y N -

TACACTTTGGGTGCTGACGACTTGACTCCATTCCGGTGAACAACAAATGGTTAACTCTGGT
361 -----+-----+-----+-----+-----+-----+ 420
ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTTGTTTACCAATTGAGACCA

Y T L G A D D L T P F G E Q Q M V N S G -

ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTCTTAGAGCT
421 -----+-----+-----+-----+-----+-----+ 480
TAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

I K F Y R R Y K A L A R K I V P F I R A -

TCTGGTTCTGACAGAGTTATTGCTTCTGCCGAAAAGTTCATTGAAGGTTTCCAATCTGCT
481 -----+-----+-----+-----+-----+-----+ 540
AGACCAAGACTGTCTCAATAACGAAGACGGCTTTTCAAGTAACTTCCAAAGGTTAGACGA

S G S D R V I A S A E K F I E G F Q S A -

```

图 24a

```

AAGTTGGCTGACCCAGGTGCTAACCCACACCAAGCTTCTCCAGTTATTAACGTTATTATT
541 -----+-----+-----+-----+-----+-----+ 600
TTCAACCGACTGGGTCCACGATTGGGTGTGGTTCGAAGAGGTCAATAATTGCAATAATAA

K L A D P G A N P H Q A S P V I N V I I -

CCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGTTTGTGTACTGCTTTCGAAGAA
601 -----+-----+-----+-----+-----+-----+ 660
GGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGACGAAAGCTTCTT

P E G A G Y N N T L D H G L C T A F E E -

TCTACCCCTAGGTGACGACGTTGAAGCTAACTTCACTGCTGTTTTCGCTCCACCAATTAGA
661 -----+-----+-----+-----+-----+-----+ 720
AGATGGGATCCACTGCTGCAACTTCGATTGAAGTGACGACAAAAGCGAGGTGGTTAATCT

S T L G D D V E A N F T A V F A P P I R -

GCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACGTTGTTAACTTG
721 -----+-----+-----+-----+-----+-----+ 780
CGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGCAACAATTGAAC

A R L E A H L P G V N L T D E D V V N L -

ATGGACATGTGTCCATTCGACACTGTTGCTAGAACTTCTGACGCTACTCAATTGTCTCCA
781 -----+-----+-----+-----+-----+-----+ 840
TACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAGTTAACAGAGGT

M D M C P F D T V A R T S D A T Q L S P -

TTCTGTGACTTGTTCACTCACGACGAATGGATTCAATACGACTACTTGCAATCTTTGGGT
841 -----+-----+-----+-----+-----+-----+ 900
AAGACACTGAACAAGTGAGTGCTGCTTACCTAAGTTATGCTGATGAACGTTAGAAACCCA

F C D L F T H D E W I Q Y D Y L Q S L G -

AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGTT
901 -----+-----+-----+-----+-----+-----+ 960
TTCATGATGCCAATGCCACGACCATTTGGGTAACCCAGGTCGAGTTCCACAACCAAAGCAA

K Y Y G Y G A G N P L G P A Q G V G F V -

AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC
961 -----+-----+-----+-----+-----+-----+ 1020
TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

N E L I A R L T H S P V Q D H T S T N H -

ACTTTGGACTCTAACCCAGCTACTTTCCCATGAAACGCTACTTTGTACGCTGACTTCTCT
1021 -----+-----+-----+-----+-----+-----+ 1080
TGAAACCTGAGATTGGGTGCGATGAAAGGGTAACTTGCGATGAAACATGCGACTGAAGAGA

T L D S N P A T F P L N A T L Y A D F S -

```

图 24b

```

      CACGACAACACTATGGTTTCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACTAAGCCA
1081 -----+-----+-----+-----+-----+-----+ 1140
      GTGCTGTTGTGATACCAAAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGATTGCGT

      H D N T M V S I F F A L G L Y N G T K P -

      TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTTCTTGGA
1141 -----+-----+-----+-----+-----+-----+ 1200
      AACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

      L S T T S V E S I E E T D G Y S A S W T -

      GTTCCATTGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTGAAAAGGAACCA
1201 -----+-----+-----+-----+-----+-----+ 1260
      CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACACTTCGACTTTTCCTTGGT

      V P F A A R A Y V E M M Q C E A E K E P -

      TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGCTGTTGACAAG
1261 -----+-----+-----+-----+-----+-----+ 1320
      AACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACGACAACTGTTC

      L V R V L V N D R V V P L H G C A V D K -

      TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 -----+-----+-----+-----+-----+-----+ 1380
      AACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGATCTAGACCACCA

      L G R C K R D D F V E G L S F A R S G G -

      AACTGGGAAGAATGTTTCGCTTAA
1381 -----+-----+----- 1404
      TTGACCCTTCTTACAAAGCGAATT

      N W E E C F A *

```

图 24c

```

ATGGGCGTGTTCGTCTGCTACTGTCCATTGCCACCTTGTTTCGGTTCACATCCGGTACC
1  -----+-----+-----+-----+-----+ 60
TACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGTGTAGGCCATGG

M G V F V V L L S I A T L F G S T S G T -

GCCTTGGGTCTCTGTTGAATTCTCACTCTTGTGACACTGTTGACGGTGGTTACCAATGT
61  -----+-----+-----+-----+-----+ 120
CGGAACCCAGGAGCACCATTAAAGAGTGAGAACTGTGACAACTGCCACCAATGGTTACA

A L G P R G N S H S C D T V D G G Y Q C -

TTCCAGAAATTTCTCACTTGTGGGGTACATACTCTCCATTCTTCTCTTTGGCTGACGAA
121 -----+-----+-----+-----+-----+ 180
AAGGGTCTTTAAAGAGTGAACACCCCATGTATGAGAGGTAAGAAGAGAAACCGACTGCTT

F P E I S H L W G T Y S P F F S L A D E -

TCTGCTATTTCTCCAGACGTTCCAAAGGGTGTAGAGTTACTTTTCGTTCAAGTTTTGTCT
181 -----+-----+-----+-----+-----+ 240
AGACGATAAAGAGGTCTGCAAGGTTTCCCAACATCTCAATGAAAGCAAGTTCAAAACAGA

S A I S P D V P K G C R V T F V Q V L S -

AGACACGGTGCTAGATACCCAACTTCTTCTGCGTCTAAGGCTTACTCTGCTTTGATTGAA
241 -----+-----+-----+-----+-----+ 300
TCTGTGCCACGATCTATGGGTTGAAGAAGACGCAGATTCCGAATGAGACGAACTAATT

R H G A R Y P T S S A S K A Y S A L I E -

GCTATTCAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGAAGACTTACAAT
301 -----+-----+-----+-----+-----+ 360
CGATAAGTTTTCTTGCATGACGAAAGTTCCCATTCATGCGAAAGAACTTCTGAATGTTA

A I Q K N A T A F K G K Y A F L K T Y N -

TACACTTTGGGTGCTGACGACTTGACTCCATTTCGGTGAACAACAAATGGTTAACTCTGGT
361 -----+-----+-----+-----+-----+ 420
ATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTGTGTGTTACCAATTGAGACCA

Y T L G A D D L T P F G E Q Q M V N S G -

ATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCATTTCATTAGAGCT
421 -----+-----+-----+-----+-----+ 480
TAATTCAAGATGTCTTCTATGTTCGAAACCGATCTTTCTAACAAGGTAAGTAATCTCGA

I K F Y R R Y K A L A R K I V P F I R A -

TCTGGTCTGACAGAGTTATTGCTTCTGCCGAAAAGTTCATTGAAGGTTTCCAATCTGCT
481 -----+-----+-----+-----+-----+ 540
AGACCAAGACTGTCTCAATAACGAAGACGGCTTTTCAAGTAACTTCCAAAGGTTAGACGA

S G S D R V I A S A E K F I E G F Q S A -

```

图 25a

```

AAGTTGGCTGACCCAGGTGCTAACCACACCAAGCTTCTCCAGTTATTAACGTTATTATT
541 -----+-----+-----+-----+-----+-----+ 600
TTCAACCGACTGGGTCCACGATTGGGTGTGGTTCGAAGAGGTCAATAATTGCAATAATAA

K L A D P G A N P H Q A S P V I N V I I -

CCAGAAGGTGCTGGTTACAACAACACTTTGGACCACGGTTTGTGTACTGCTTTCGAAGAA
601 -----+-----+-----+-----+-----+-----+ 660
GGTCTTCCACGACCAATGTTGTTGTGAAACCTGGTGCCAAACACATGACGAAAGCTTCTT

P E G A G Y N N T L D H G L C T A F E E -

TCTACCCTAGGTGACGACGTTGAAGCTAAGTTCCTGCTGTTTTCGCTCCACCAATTAGA
661 -----+-----+-----+-----+-----+-----+ 720
AGATGGGATCCACTGCTGCAACTTCGATTGAAGTGACGACAAAAGCGAGGTGGTTAATCT

S T L G D D V E A N F T A V F A P P I R -

GCTAGATTGGAAGCTCACTTGCCAGGTGTTAACTTGACTGACGAAGACGTTGTTAACTTG
721 -----+-----+-----+-----+-----+-----+ 780
CGATCTAACCTTCGAGTGAACGGTCCACAATTGAACTGACTGCTTCTGCAACAATTGAAC

A R L E A H L P G V N L T D E D V V N L -

ATGGACATGTGTCCATTCGACACTGTTGCTAGAACTTCTGACGCTACTCAATTGTCTCCA
781 -----+-----+-----+-----+-----+-----+ 840
TACCTGTACACAGGTAAGCTGTGACAACGATCTTGAAGACTGCGATGAGTTAACAGAGGT

M D M C P F D T V A R T S D A T Q L S P -

TTCTGTGACTTGTTCACTCACGACGAATGGATTCAATACGACTACTTGCAATCTTTGGGT
841 -----+-----+-----+-----+-----+-----+ 900
AAGACACTGAACAAGTGAGTGCTGCTTACCTAAGTTATGCTGATGAACGTTAGAAACCCA

F C D L F T H D E W I Q Y D Y L Q S L G -

AAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTGTTGGTTTCGTT
901 -----+-----+-----+-----+-----+-----+ 960
TTCATGATGCCAATGCCACGACCATTTGGGTAAACCCAGGTGCGAGTTCCACAACCAAAGCAA

K Y Y G Y G A G N P L G P A Q G V G F V -

AACGAATTGATTGCTAGATTGACTCACTCTCCAGTTCAAGACCACACTTCTACTAACCAC
961 -----+-----+-----+-----+-----+-----+ 1020
TTGCTTAACTAACGATCTAACTGAGTGAGAGGTCAAGTTCTGGTGTGAAGATGATTGGTG

N E L I A R L T H S P V Q D H T S T N H -

ACTTTGGACTCTAACCCAGCTACTTTCCCATTTGAACGCTACTTTGTACGCTGACTTCTCT
1021 -----+-----+-----+-----+-----+-----+ 1080
TGAAACCTGAGATTGGGTGCGATGAAAGGGTAACTTGGCGATGAAACATGCGACTGAAGAGA

T L D S N P A T F P L N A T L Y A D F S -

```

图 25b

```

CACGACAACACTATGGTTTCTATTTTCTTCGCTTTGGGTTTGTACAACGGTACTAAGCCA
1081 -----+-----+-----+-----+-----+-----+ 1140
GTGCTGTTGTGATACCAAAGATAAAAGAAGCGAAACCCAAACATGTTGCCATGATTGGT

H D N T M V S I F F A L G L Y N G T K P -

TTGTCTACTACTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTTCTTGGA
1141 -----+-----+-----+-----+-----+-----+ 1200
AACAGATGATGAAGACAACCTTAGATAACTTCTTTGACTGCCAATGAGACGAAGAACCTGA

L S T T S V E S I E E T D G Y S A S W T -

GTTCCATTGCTGCTAGAGCTTACGTTGAAATGATGCAATGTGAAGCTGAAAAGGAACCA
1201 -----+-----+-----+-----+-----+-----+ 1260
CAAGGTAAGCGACGATCTCGAATGCAACTTTACTACGTTACACTTCGACTTTTCCTTGGT

V P F A A R A Y V E M M Q C E A E K E P -

TTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTGCTGTTGACAAG
1261 -----+-----+-----+-----+-----+-----+ 1320
AACCAATCTCAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACACGACAACCTGTC

L V R V L V N D R V V P L H G C A V D K -

TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTAGATCTGGTGGT
1321 -----+-----+-----+-----+-----+-----+ 1380
AACCCATCTACATTCTCTCTGCTGAAGCAACTTCAAACAGAAAGCGATCTAGACCACCA

L G R C K R D D F V E G L S F A R S G G -

AACTGGGAAGAATGTTTCGCTTAA
1381 -----+-----+-----+-----+ 1404
TTGACCCCTTCTTACAAAGCGAATT

N W E E C F A *

```

图 25c