



(12) 发明专利

(10) 授权公告号 CN 102666570 B

(45) 授权公告日 2015. 12. 02

(21) 申请号 201080054717. 4

(22) 申请日 2010. 12. 01

(30) 优先权数据

09177593. 2 2009. 12. 01 EP

61/266711 2009. 12. 04 US

(85) PCT国际申请进入国家阶段日
2012. 06. 01

(86) PCT国际申请的申请数据
PCT/EP2010/068630 2010. 12. 01

(87) PCT国际申请的公布数据
W02011/067283 EN 2011. 06. 09

(73) 专利权人 诺沃—诺迪斯克有限公司
地址 丹麦鲍斯韦

(72) 发明人 A. C. 肖

(74) 专利代理机构 中国专利代理(香港)有限公
司 72001

代理人 孔青 郭文洁

(51) Int. Cl.

C07K 14/195(2006. 01)

C12N 9/88(2006. 01)

(56) 对比文件

EP 0465404 A3, 1992. 04. 08, 全文.

CN 101287840 A, 2008. 10. 15, 全文.

Lucas S.. EBI accession
no. UNIPROT:C4L4B8. 《Uniprot 数据库》. 2009, 全
文.

刘深基等. 多肽酰胺化酶的结构与功能. 《生
物化学与生物物理进展》. 1999, 第 26 卷 (第 03
期), 第 195-197 页.

审查员 段珊

权利要求书1页 说明书44页

序列表25页 附图21页

(54) 发明名称

新的肽基 α -羟基甘氨酸 α -酰胺化裂合酶

(57) 摘要

本专利申请涉及能够催化 α -羟基甘氨酸转
化为 α -酰胺的酶以及这样的酶用于产生 C 端
 α -酰胺化肽的用途。

1. 一种用于产生 α -酰胺化肽的方法,其包括以下步骤:(i) 在适合于酶促活性的条件下,使得具有 C 端甘氨酸残基的靶肽与肽基甘氨酸 α -羟化单加氧酶 (PHM) 和来自原核生物体的具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶 (PAL) 活性的重组多肽两者能够反应,其中用所述 PHM 和所述重组多肽对所述靶肽的反应在两个单独的步骤中或同时进行;(ii) 回收 C 端 α -酰胺化肽。

2. 权利要求 1 的方法,其中所述来自原核生物体的多肽具有不含半胱氨酸残基或含至多 1 个半胱氨酸残基或至多 2 个半胱氨酸残基的氨基酸序列。

3. 权利要求 1-2 中任一项的方法,其中所述多肽是野生型酶。

4. 权利要求 1-2 中任一项的方法,其中相比于野生型多肽所述多肽的至少一个 Cys 残基缺失或被 Ala、Ser 或 Val 取代。

5. 权利要求 1-2 中任一项的方法,其中所述来自原核生物体的具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶 (PAL) 活性的多肽已经在重组大肠杆菌中表达。

6. 权利要求 1-2 中任一项的方法,其中所述来自原核生物体的多肽在没有重折叠以获得催化活性的步骤的情况下而制备。

7. 权利要求 1-2 中任一项的方法,其中当使用所述多肽时,存在 Cu^{2+} 离子和 / 或抗坏血酸。

8. 权利要求 1-2 中任一项的方法,其中 PHM 酶来自哺乳动物生物体。

9. 权利要求 1-2 中任一项的方法,其中回收的 α -酰胺化肽被用作药物。

10. 权利要求 1-2 中任一项的方法,其中所述靶肽选自:胰淀粉样肽、神经肽 Y (NPY)、肽 YY (PYY)、胰多肽 (PP)、胰高血糖素样肽、胃泌素、降钙素、降钙素相关肽、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、胰抑释素、阿黑皮质素原 (POMC)、 α -促黑素细胞激素 (α -MSH)、 γ -促黑素细胞激素 (γ -MSH) 和酰胺化铰链肽 (HP-N)。

11. 权利要求 1-2 中任一项的方法,其中所述靶肽选自 PYY-3-36 和 GLP-1。

新的肽基 α -羟基甘氨酸 α -酰胺化裂合酶

技术领域

[0001] 本发明涉及具有肽基 α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽、用于制备这样的多肽的方法和这样的多肽在用于产生 C 端 α -酰胺化肽的方法中的用途。

[0002] 发明背景

[0003] 在多细胞生物体中,某些肽(“前体”),像神经肽,是在一连串切割并进一步修饰肽底物来产生完整功能的生物反应性肽的酶促步骤中翻译后修饰的。该过程在反式-高尔基体中开始并作为未成熟的分泌颗粒继续。对于多数这类肽,非常重要的晚期翻译后修饰是羧基端的 α 酰胺化。

[0004] C 端残基的 α 酰胺化对于参与人或动物代谢的数种肽激素的活性是关键。数种肽激素现今被用作药物治疗人类(例如用于控制肥胖症和/或糖尿病)或作为潜在药物在开发中。这种肽激素的一个实例是胰淀粉样肽(例如 Symlin®, 醋酸普兰林肽,其为人胰淀粉样肽的类似物)。人胰淀粉样肽是 37 个氨基酸残基的肽,其可用于治疗或预防肥胖症和/或糖尿病。因此,为了获得完整生物学活性,胰淀粉样肽的 C 端需要被酰胺化。同样地肽 YY (PYY) 应该被 α 酰胺化以获得完整生物学活性。

[0005] 大肠杆菌 (*E. coli*) 和酵母被广泛用于真核生物来源的肽的重组表达。然而,由于 C 端 α -酰胺基的性质,使用基于大肠杆菌和酵母的现有技术微生物表达系统不能将肽激素表达为活性形式,因为 α 酰胺化酶在这些生物体中不能被天然表达。因此,C 端 α 酰胺必须被引入到重组表达肽中,这使用例如具有 α 酰胺化酶的离体修饰。

[0006] 在数种真核生物体中发现,凭借双功能肽基 α -酰胺化单加氧酶 (PAM) 将具有 C 端 Gly 的肽前体酶修饰为 α -酰胺。关于该酶,已描述了多种编码不同同工型的可变剪接转录物变体。关于多细胞生物体,已专门描述该酶 (Metazoa)。将肽中 C 端 Gly 残基转化为 α -酰胺是两步过程,其中 PAM 的 N 端结构域(命名为 PHM)催化甘氨酸转化为 α -羟基甘氨酸,然后 PAM 的 C 端结构域(命名为 PAL)催化 α -羟基甘氨酸转化为 α -酰胺。在真核生物体中,两个催化结构域顺序工作来催化神经内分泌肽为活性 α -酰胺化产物。在来自真核生物体的 PAL 结构域中,两条二硫桥是高度保守的。

[0007] 虽然通过化学方法合成包含 C 末端上酰胺基(α 酰胺)的小肽是可能的,但是产生较大的 α -酰胺化肽是困难和昂贵的。因此, α 酰胺化酶可用于转化重组前体肽为成熟肽。

[0008] US4708934 描述从甲状腺髓样癌细胞系和组织样品中提取的肽基-甘氨酸 α -酰胺化单加氧酶。

[0009] US5789234 描述通过重组 DNA 技术产生 α -酰胺化酶。

[0010] W090/08194 涉及一种用于通过使用真核生物 C 端 α 酰胺化酶从前体肽产生 C 端 α 酰胺化肽的方法。还描述了一种用于真核表达这些 C 端 α 酰胺化酶的方法。

[0011] W089/02460 描述一种牛来源的 PAM 酶、其克隆、cDNA 和通过重组 DNA 技术的表达。

[0012] EP0448513 描述一种用于重组表达来源于非洲爪蟾 (*Xenopus laevis*) 的肽基甘氨酸 α -羟化单加氧酶的方法,其包括培养用重组杆状病毒转染的昆虫细胞来产生该酶,

其中编码肽基甘氨酸 α -羟化单加氧酶的 DNA 已掺入到重组杆状病毒中。

[0013] EP0465404 描述一种来源于非洲爪蟾的酶 (PHL; PAL)、酶的克隆及其在昆虫细胞中的重组表达,该酶催化切割在 C 端 α -羟化肽的 α -羟基甘氨酸部分中的 N-C 键。

[0014] US 20060292672 描述一种用于表达 PAM 或其两个催化结构域的一个的细胞系。

[0015] EP2172550 描述一种重组 C 端 α -酰胺化酶衍生物,其不形成来源于非洲爪蟾的 C 端 α -酰胺化酶中通常存在的 5 个二硫键中的至少 1 个,和描述一种在大肠杆菌中重组产生所述衍生物的方法,其中获得的包含体是增溶的并经历重折叠步骤。

[0016] 发明概述

[0017] 本发明涉及新的酶,其能够催化肽中 α -羟基甘氨酸转化为 α -酰胺(肽基 α -羟基甘氨酸 α -酰胺化裂合酶活性)。

[0018] 新的肽来源于原核生物体并具有与对真核生物 PAL 酶所述的不同的物理化学和结构特性。

[0019] 因此,本发明提供具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的酶,其特征在于它们来源于原核生物体。

[0020] 本发明还提供具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的酶,其特征在于它们在大肠杆菌中可被表达为可溶性酶促活性蛋白。

[0021] 本发明还提供具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的酶,其特征在于它们具有不包含半胱氨酸残基或者包含至多 1 或至多 2 个半胱氨酸残基的氨基酸序列。

[0022] 本发明还提供一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的酶,其特征在于其可通过包括以下步骤的方法产生:(i) 在适合于酶表达的条件下培养重组大肠杆菌菌株宿主细胞,该细胞包含含有编码酶的核苷酸序列的核酸构建体;和(ii) 在宿主细胞破碎和离心后从上清液中回收酶。

[0023] 本发明还提供能够催化肽中 α -羟基甘氨酸转化为 α -酰胺的酶,其中所述酶具有包含下述基序(命名为基序 1)的氨基酸序列:

[0024] Xaa₁ Val Xaa₂ Asp Arg Xaa₃ Xaa₄ Xaa₅ Arg Xaa₆ Gln Xaa₇ Xaa₈ Xaa₉ Xaa₁₀ Xaa₁₁ Gly Xaa₁₂ Xaa₁₃ Xaa₁₄ Xaa₁₅ Xaa₁₆ Trp;其中 Xaa₁-Xaa₁₆独立选自天然存在氨基酸,条件是 Xaa₁和 Xaa₇不是 Cys。

[0025] 本发明还提供能够催化肽中 α -羟基甘氨酸转化为 α -酰胺的酶,其中所述酶具有包含下述基序(命名为基序 2)的氨基酸序列:

[0026] Asp Gly Tyr Xaa₁₇ Asn Xaa₁₈ Arg Xaa₁₉ Xaa₂₀ Xaa₂₁ Phe Xaa₂₂ Xaa₂₃ Xaa₂₄ Gly Xaa₂₅ Xaa₂₆ Xaa₂₇ Xaa₂₈ Xaa₂₉ Xaa₃₀ Xaa₃₁ Xaa₃₂ Xaa₃₃ Gly Xaa₃₄ Xaa₃₅ Xaa₃₆ Gly Xaa₃₇ Phe;其中 Xaa₁₇-Xaa₃₇独立选自天然存在氨基酸,条件是 Xaa₁₇不是 Cys。

[0027] 本发明还提供具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的酶,其包含与选自以下序列的氨基酸序列有至少 80% 同一性的氨基酸序列:(a) SEQ ID NO: 1 的氨基酸 2-306;(b) SEQ ID NO: 2 的氨基酸 3-336;(c) SEQ ID NO: 3 的氨基酸 3-305;(d) SEQ ID NO: 4 的氨基酸 3-279;(e) SEQ ID NO: 19;(f) SEQ ID NO: 20;(g) SEQ ID NO: 21;(h) SEQ ID NO: 22 和 (i) SEQ ID NO: 23。

[0028] 本发明另外提供本发明的酶在用于通过催化肽中 C 端 α -羟基甘氨酸残基转化为 α -酰胺残基来制备 α -酰胺化肽的方法中的用途。本发明还提供使用本发明的酶产生

α -酰胺化肽的方法。此外,本发明涉及一种用于通过提供编码本发明酶的分离核酸、包含核酸的载体和宿主细胞由重组技术产生本发明酶的方法。

附图说明

[0029] 图 1:质粒 C 的典型 pET11a 载体图谱,其编码来自赤细菌属 (*Erythrobacter*) 的细菌 PAL 样结构域 (SEQ ID NO: 1) (不具有预测的信号肽) 和 N 端融合伴侣 (SEQ ID NO: 7) 以及介于中间的接头区 (SEQ ID NO: 8) 中的 HRV14 3C 蛋白酶切割位点 (用特征蛋白 3 标记整个融合蛋白)。描绘了 NdeI、XhoI 和 BamHI 限制酶位点。T7 启动子区、氨苄青霉素抗性基因、lacI 阻抑区和复制起点位点也在载体图谱中示出。

[0030] 图 2:SDS-PAGE,显示加上 RL9_THEMA TAP 标签的赤细菌属 PAL 样结构域 (A) 和大鼠 PAL (B) 的表达概况。Uind:未诱导的细胞培养物对照 (阴性对照),Ind:在 30℃用 0.5mM IPTG 诱导 3 小时后,Sup:可溶性级分,PeI:不溶性级分。箭头指示凝胶上 PAL 条带的分子量位置。

[0031] 虽然表达水平是可比的,但是赤细菌属 PAL 样结构域被表达为高度可溶性蛋白,相比之下大鼠 PAL 被表达为不溶的并将需要重折叠来获得功能性酶。

[0032] 图 3:A:FPLC 色谱图,显示加上 RL9_THEMA 标签的赤细菌属 PAL 样结构域的 SP Sepharose FF 纯化概况。用箭头所示的约 0.25 M 浓度的 NaCl 洗脱融合蛋白。虚线表示电导率曲线,实线表示在 280nm 的 UV 吸收信号。在 X 轴上的数字表示级分和 ml。

[0033] B:SDS-PAGE 凝胶,显示来自图 3A 色谱图上主峰的级分。Ap1:装载到柱子上的装载物 (application),Ft:流通级分。在一个阳离子交换色谱步骤后,TAP 标签提供有效的捕获和高纯度。

[0034] 图 4A:FPLC 色谱图,显示用 HRV14 3C 蛋白酶加工融合蛋白后,释放 RL9_THEMA 标签的成熟微小杆菌属 (*Exiguobacterium*) PAL 样结构域的 Q Sepharose HP 分离。以箭头所示的约 0.4M NaCl 洗脱主峰。虚线表示电导率曲线,实线表示在 280nm 的 UV 信号。在 X 轴上的数字表示级分和 ml。

[0035] 图 4B:来自图 4A 所示的 FPLC 分离的洗脱级分的 SDS-PAGE 分析。图 4A. Ap1:装载到柱子上的装载物,Ft:流通级分。在装载物泳道中,两条条带代表成熟微小杆菌属 PAL 样结构域 (约 30 kDa) 和释放的纯化标签 (约 18 kDa)。来自纯化的主峰几乎只包括成熟 PAL (实箭头),而释放的 TAP 标签不结合阴离子交换柱并在流通级分中存在 (虚箭头)。

[0036] 图 5:分别代表 α -羟基甘氨酸 (Gly(OH)) 和 α -酰胺 ($-\text{NH}_2$) (在相关峰下标记) 的合成的 α -羟基苯甲酰甘氨酸和苯甲酰胺的 UPLC 分析。用在 N 端具有或不具有 TAP 标签的细菌 PAL 样酶加上相关辅因子在 37℃孵育 α -羟基苯甲酰甘氨酸 3 小时。加粗箭头指示最显著的酶促转化为苯甲酰胺。

[0037] 图 5. A:使用酸性 (MES pH5.5) 或碱性 (Tris pH7.5) 缓冲条件测试的具有 (+TAP) (SEQ ID NO: 7) 或不具有 (-TAP) 纯化标签的赤细菌属 PAL 样结构域的分析。阴性 (neg.) 对照 (ctrl) 是不加入酶的样品的色谱图。 α -羟基苯甲酰甘氨酸转化为苯甲酰胺可在如之前文献所述的高 pH (图 5A 6/7) 下自发发生。然而,从 α -羟基苯甲酰甘氨酸和苯甲酰胺底物峰面积的相对面积得出的结论是,加入酶可催化转化,具有 (图 5A 2/7) 和不具有 TAP 标签 (图 5A 4/7) 两者都在 pH7.5 时最有效。

[0038] 图 5. B:使用酸性 (MES pH5.5) 或碱性 (Tris pH7.5) 缓冲条件测试的具有 (+TAP) (SEQ ID NO: 7) 的食土性细菌 (*Chthoniobacter*) PAL 样结构域的分析。与赤细菌属酶相比,食土性细菌酶在较低的 pH 范围内更有活性,因为与阴性对照 (图 5B 3/7) 相比,在 pH5.5 用酶孵育后苯甲酰胺峰面积显著增加 (图 5B1/7)。

[0039] 图 6:胰淀粉样肽类似物模型肽的提取和去卷积 MS 谱,通过与大鼠 PHM 共孵育证实细菌 PAL 样结构域的 PAL 活性。X 轴:质荷比,以道尔顿计。Y 轴:MS 峰的相对强度。显著的 MS 峰的质量在谱中示出;前体肽对照 (包含 C 端 Gly 残基):13934.66 Da, α -羟基甘氨酸中间体肽 (包含 C 端 α -羟基甘氨酸):13951 Da, α -酰胺化肽:13876.62Da

[0040] A:未处理的对照,B:大鼠 PHM 单独 (2 小时),C:浮霉状菌属 (*Planctomyces*) PAL 样结构域,蛋白 10 (2 小时),D:赤细菌属 PAL 样结构域,蛋白 3 (2 小时),E:浮霉状菌属 PAL 样结构域,蛋白 10 (5 小时),F:赤细菌属 PAL 样结构域,蛋白 3 (5 小时)。

[0041] 图 7:与四种不同的细菌 PAL 样结构域 (SEQ ID NO: 1-4) 孵育的纯化的 C 端 α -羟基甘氨酸延伸的普兰林肽的提取和去卷积 MS 谱。X 轴:质荷比,以道尔顿计。Y 轴:MS 峰的相对强度。两个显著 MS 峰的质量在谱中示出; α -羟基甘氨酸中间体肽 (包含 C 端 α -羟基甘氨酸:约 4023Da) 峰和 α -酰胺化肽 (约 3949Da) 峰都确定为非常接近于计算的平均同位素质量。对于各反应,示出酶/底物比 (w/w)。

[0042] A:赤细菌属 (SEQ ID NO: 1) i) PHM 处理的对照 (ctrl), ii) 1:1250, iii) 1:500, iv) 1:100

[0043] B:浮霉状菌属 (SEQ ID NO: 4) i) PHM 处理的对照 (ctrl), ii) 1:1250, iii) 1:500, iv) 1:100

[0044] C:食土性细菌 (SEQ ID NO: 3) i) PHM 处理的对照 (ctrl), ii) 1:100, iii) 1:50, iv) 1:25

[0045] D:微小杆菌属 (SEQ ID NO: 2) i) PHM 处理的对照 (ctrl), ii) 1:100, iii) 1:50, iv) 1:25

[0046] 定义

[0047] “PAM”或“肽基-甘氨酸 α -酰胺化单加氧酶”指的是催化 C 端甘氨酸残基转化为 α -羟基甘氨酸并催化 α -羟基甘氨酸转化为 α -酰胺的双功能性酶。该酶亦被称为肽基甘氨酸 2-羟化酶、肽基 α -酰胺化酶、肽基甘氨酸 α -羟化单加氧酶和肽基- α -羟基甘氨酸 α -酰胺化裂合酶。

[0048] “PHM”或“肽基甘氨酸 α -羟化单加氧酶”是能够催化 C 端甘氨酸残基转化为 α -羟基甘氨酸的酶。PHM 的其他术语有:肽基甘氨酸 2-羟化酶、肽基甘氨酸 α -酰胺化单加氧酶、肽基甘氨酸 α -羟化酶、肽基甘氨酸 α -羟化单加氧酶、肽基甘氨酸 α -羟化单加氧酶、肽基甘氨酸 α -单加氧酶、EC 1.14.17.3 和肽基甘氨酸单加氧酶。

[0049] 本文所用的“PAL 酶”、“PAL”或“肽基- α -羟基甘氨酸 α -酰胺化裂合酶”是能够催化 α -羟基甘氨酸转化为 α -酰胺的酶。PAL 的同义词有:肽基酰胺基羟乙酸裂合酶 α -羟基甘氨酸酰胺化脱烷基酶、HGAD、PGL、肽基酰胺基羟乙酸肽基酰胺裂合酶、EC 4.3.2.5 和肽基羟基甘氨酸 N-C 裂合酶 (PHL)。PAL 酶的活性可如测定 (I) 中描述来证明。

[0050] “能够催化 C 端 α -羟基甘氨酸转化为 α -酰胺的酶”意指,能够催化反应 $R-Gly(OH) \rightarrow R-NH_2$ 的酶,其中 R 是肽、蛋白或化合物。

[0051] 本文所用的表述“本发明的酶”意指具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的本发明多肽。

[0052] 非肽底物（例如 α -羟基苯甲酰甘氨酸）也充当 PAL 的底物。

[0053] 术语“PAL样”意指与已知真核生物 PAL 酶具有相同活性的酶。

[0054] “靶肽”意指在 α -酰胺化过程中被修饰而获得 C 端 α -酰胺基的肽。靶肽应该在 C 端中包含 Gly 残基。靶肽可被描述为具有式 $R'-X-Gly$ ，其中 X 代表任何氨基酸且 X 是将被转化为氨基酸酰胺的氨基酸，即对于该氨基酸，在酶促 α -酰胺化过程反应中 $-COOH$ 被转化为 $CO-NH_2$ ， R' 代表肽的剩余部分，和 Gly 代表 C 端甘氨酸残基。

[0055] 靶肽的一个实例是胰淀粉样肽前体，其除了胰淀粉样肽序列外还在 C 端包含 Gly 残基。与本发明相关的 Gly- 延伸肽前体的其他非限制性实例包括神经肽 Y (NPY)、肽 YY (PYY)、PYY-3-36、胰多肽 (PP)、胰高血糖素样肽 (GLP-1)、胃泌素、降钙素、降钙素相关肽 (CGRP)、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、胰抑释素、阿黑皮质素原 (POMC)、 α -促黑素细胞激素 (α MSH)、 γ -促黑素细胞激素 (γ 1MSH) 和酰胺化铰链肽 (HP-N)，或者它们的功能类似物。

[0056] 本文所用的术语“分离的多肽”或“分离的多核苷酸”指的是从其天然来源分离的多肽或多核苷酸。

[0057] 本文所用的术语“纯化”或“回收”指的是从样品中移除污染物。例如，本发明的 PAL 酶是通过移除溶液或制品中的污染蛋白和其他化合物纯化。在一些方面，使用细菌表达本发明的 PAL 酶，并通过移除其他宿主细胞成分来纯化这些重组 PAL 酶，从而增加样品中重组 PAL 的百分比。

[0058] 术语“实质上纯的多肽”在本文中指，多肽制品包含至多 20%、至多 10% 或至多 5% 重量的与其天然或重组缔合的其他多肽材料。因此，以制品中存在的总多肽材料的重量计，优选实质上纯的多肽是至少 80% 纯、至少 90% 纯或优选至少 95% 纯。本发明的多肽优选为实质上纯的形式，即多肽制品基本上不含与其天然或重组缔合的其他多肽材料。这可例如通过公知的重组方法或通过经典的纯化方法制备多肽来实现。

[0059] 在一个方面，“%纯度”定义为目标蛋白的量 / (目标蛋白的量 + 宿主细胞污染物的量) $\times 100$ 。其可通过 SDS-PAGE 分析或确定量的 HPLC 分离来确定。

[0060] 在一个方面，本发明多肽为至少 1% 纯，例如至少 5% 纯、至少 10% 纯、至少 20% 纯、至少 40% 纯、至少 60% 纯、至少 80% 纯、至少 90% 纯。

[0061] 本文使用的关于本发明多肽的术语“回收”意指，在一个实施方案中，多肽 / 酶不与显著水平（例如，至多 1%、至多 2%、至多 3%、至多 5%、至多 10% 或至多 25%）的产生多肽的系统（例如，细胞培养物）内所包含的任何外来和不需要生物分子缔合。回收的多肽指的是，由于人类干预（不管是自动、手动还是两者）而通过纯净阶段的本发明的多肽。应当理解，在组合物中存在的本发明的回收多肽和本发明的分离多肽也在本发明内。换言之，术语“回收”或“分离”并不意在排除有其他化合物或材料的人工或合成混合物。

[0062] 本文可交换使用术语“蛋白”、“肽”和“多肽”。其中肽是蛋白的一部分，本领域技术人员理解该术语在上下文中的使用。本文可交换使用表述“本发明的酶”和“本发明的多肽”。

[0063] 氨基酸序列的“修饰”意指，在序列中置换、缺失和 / 或添加（包括插入）一个或

多个氨基酸。在另外的方面中,它也包括一条或多条氨基酸侧链的置换。

[0064] 术语“表达载体”本文定义为这样的线性或环状 DNA 分子,其包含编码本发明多肽的多核苷酸,并与提供其表达的另外的核苷酸有效连接。术语“质粒”、“表达载体”和“载体”可交换使用,因为质粒是目前最常用的载体形式。然而,本发明旨在包括这样的起等价功能的其他形式的表达载体。本文使用的“表达载体”或“载体”指的是这样的 DNA 构建体,其包含与能够影响 DNA 在合适宿主中表达的合适控制序列有效连接的 DNA 序列。这样的控制序列例如可包括影响转录的启动子、控制这种转录的任选的操纵基因序列、编码合适 mRNA 核糖体结合位点的序列和控制转录和翻译终止的序列。载体例如可为质粒、噬菌体或仅潜在的基因组插入片段。一旦转化到合适宿主中,载体可例如复制并独立于宿主基因组起作用,或在一些实例中可将其自身整合到基因组中。

[0065] 两条氨基酸序列之间的关联性由参数“同一性”(“% 同一性”)描述。在本发明氨基酸序列的上下文中,同一性可通过任何合适的技术/程序确定,通常通过使用具有起始空位罚分为-12 和延伸罚分为-2 的 BLOSUM50 评分矩阵的 Needleman-Wunsch 比对分析(参见 Needleman 和 Wunsch, J. Mol. Biol. (1970) 48:443-453) 确定。使用 Needle 标记的“最长同一性”(使用 -nobrief 选项获得)的输出,百分比同一性可如下计算:(相同的残基 x 100)/(比对的长度 - 比对中空位的总数)。因为 Needleman-Wunsch 比对提供两条序列之间的总体或全局同一性测定,所以应该认识到,靶序列(其可为较大肽序列的部分或子序列)可以类似于完整序列的方式应用,或备选地,可将局部比对值用于评估子序列之间的关系,例如通过 Smith-Waterman 比对(J. Mol. Biol. (1981) 147:195-197) 确定,其可通过可得的程序获得。适合于分析同一性的其他局部比对方法可包括应用启发式局部比对算法的程序,例如 FastA 和 BLAST 程序。

[0066] 当本文使用的术语“编码序列”意指核苷酸序列时,其直接指定其蛋白产物的氨基酸序列。编码序列的边界通常由开放阅读框确定,开放阅读框通常以 ATG 起始密码子或备选起始密码子(例如 GTG 和 TTG) 开始并以终止密码子(例如 TAA、TAG 和 TGA) 结束。编码序列可为 DNA、cDNA、合成或重组核苷酸序列。术语“cDNA”本文定义为可通过逆转录从获自真核细胞的成熟、剪接 mRNA 分子中制备的 DNA 分子。cDNA 没有通常存在于相应基因组 DNA 中的内含子序列。最初的初级 RNA 转录物是 mRNA 的前体,其在成为成熟剪接 mRNA 之前通过一系列的步骤加工。这些步骤包括通过称为剪接的过程移除内含子序列。cDNA 来源于 mRNA,因此没有任何内含子序列。

[0067] 术语“编码...的核酸分子”、“编码...的核酸序列”、“编码...的 DNA 序列”和“编码...的 DNA”指的是,沿着一连串脱氧核糖核酸的脱氧核糖核苷酸的顺序或序列。这些脱氧核糖核苷酸的顺序决定沿着蛋白链的氨基酸的顺序。DNA 序列因此编码蛋白(例如酶)的氨基酸序列。本文使用的术语“核酸构建体”指的是单链或双链的核酸分子,其从天然存在基因中分离,或者其经修饰而以非天然方式包含核酸区段,或者其为合成的。当核酸构建体包含表达本发明的编码序列所需的控制序列时,术语核酸构建体与术语“表达盒”同义。

[0068] 术语“有效连接”在本文中指一种结构,其中控制序列位于相对于多核苷酸序列的编码序列的合适位置,使得控制序列指导多肽的编码序列的表达。术语“表达”包括涉及多肽产生的任何步骤,其包括例如转录、转录后修饰、翻译、翻译后修饰和分泌。

[0069] 本文与从给定来源(即生物体)中衍生的多肽或多核苷酸相关地使用的术语“来

源于”意指,与来源生物体相比,多核苷酸(编码多肽的多核苷酸)与该来源生物体中天然存在的多核苷酸序列相同或为其变体,而不管该多核苷酸序列是否已被插入到另一生物体中或者该多肽由另一生物体产生。“来源于”在本发明的上下文中还意指“从…中鉴定”。“来源于”在本发明的上下文中还意指用细菌核苷酸/蛋白序列从数据库中鉴定本发明的酶,即通过在蛋白数据库(例如 Uniprot、trEMBL 或 RefSeqP)中进行计算机辅助检索。

[0070] 本文在本发明的上下文中使用的术语野生型意指,其天然存在的形式(例如基因或蛋白序列)。也包括由通过在包括细菌核苷酸/蛋白序列数据的数据库中检索而推测的核苷酸序列所编码的蛋白。在一个实施方案中,术语野生型包括不具有信号肽和前导肽的肽序列。

[0071] 本文使用的术语“成熟”(本发明的酶、多肽或氨基酸序列)意指,推定的多肽最小功能序列,其中未添加天然或人工氨基酸延伸(例如信号肽或融合伴侣)。

[0072] 本文使用的术语“宿主细胞”包括,具有包含本发明多核苷酸的核酸构建体或表达载体的易受转化、转染、转导等影响的任何细胞类型。

[0073] 表述“纯化标签”意指,在酶的 N 或 C 端与酶融合并用于纯化该酶的肽序列。

[0074] 表述“TAP”标签指的是来源于嗜热细菌的热稳定碱性蛋白标签,当其与酶的肽序列在 N 或 C 端融合时可用于纯化该酶,如在以编号 WO 2006/108826 和 WO 2008/043847 公布的国际专利申请中所揭露。

[0075] 表述“融合酶”、“融合蛋白”或“加上标签的酶”意指具有与酶的 C 端或 N 端连接的“融合伴侣”的酶。融合伴侣的一个实例是蛋白标签,其可增加融合蛋白的表达水平、溶解度或纯化。

[0076] 表述“接头”意指将融合伴侣(例如纯化标签)和酶连接在一起的氨基酸序列。接头序列可例如包括促进靶蛋白更好折叠的序列和/或用于切除纯化标签的切割位点。

[0077] “螺旋结构”的特征在于具有导致由链间氢键稳定的卷曲结构的氨基酸序列。

[0078] “%溶解度”定义为,来自宿主细胞裂解物的可溶性蛋白的量/(来自宿主细胞裂解物的可溶性+不溶性蛋白的量) X 100。它可通过基于 SDS-PAGE 分析来比较细胞裂解物的不溶性和可溶性级分而确定。

[0079] 在本上下文中,术语“功能性酶”意在表示与天然酶具有相似功能的蛋白。蛋白可与天然酶结构上相似并通过如下方式衍生自天然酶:向天然酶的 C 端和 N 端之一或两者中添加一个或多个氨基酸,在天然氨基酸序列中的一个或多个不同位点上置换一个或多个氨基酸,或在天然酶的一端或两端或者在氨基酸序列中的一个或数个位点上缺失一个或多个氨基酸,或在天然氨基酸序列中的一个或多个位点上插入一个或多个氨基酸。

[0080] “装载物”意指装载在纯化柱上的包含酶的样品。“流通”意指不与纯化柱结合的包含宿主细胞蛋白和污染物的部分装载物。“主峰”指的是具有最强 UV 强度并包含蛋白的纯化色谱图中的峰。“mAU”是毫吸光度单位。“UV280强度”是以毫吸光度单位计量的在 280nm 波长下的吸光度,在该波长下蛋白将吸收。“IPTG”是异丙基-β-D-硫代吡喃半乳糖苷。SDS-PAGE 是十二烷基硫酸钠-聚丙烯酰胺凝胶电泳。FPLC(快速蛋白液相色谱)是类似于高效液相色谱(HPLC)的一种液相色谱形式,其用于从复杂混合物中分离或纯化蛋白。LC-MS(液相色谱-质谱)是结合液相色谱(或 HPLC)的物理分离能力和质谱的质量分析能力的分析技术。

[0081] 氨基酸:在本上下文中,3 字母或 1 字母表示的氨基酸用作如表 1 中所示的其常规意思。除非明确指出,本文提及的氨基酸为 L- 氨基酸。另外,除非另外指定,肽的氨基酸序列的左端和右端分别是 N 端和 C 端。

[0082] 表 1:氨基酸的缩写:

[0083]

氨基酸	3 字母代码	1 字母代码
甘氨酸	Gly	G
脯氨酸	Pro	P
丙氨酸	Ala	A
缬氨酸	Val	V
亮氨酸	Leu	L
异亮氨酸	Ile	I
甲硫氨酸	Met	M
半胱氨酸	Cys	C
苯丙氨酸	Phe	F
酪氨酸	Tyr	Y
色氨酸	Trp	W
组氨酸	His	H
赖氨酸	Lys	K
精氨酸	Arg	R
谷氨酰胺	Gln	Q
天冬酰胺	Asn	N
谷氨酸	Glu	E
天冬氨酸	Asp	D
丝氨酸	Ser	S
苏氨酸	Thr	T

[0084] 发明详述

[0085] 在数种真核生物体中发现,凭借双功能性肽基 α - 酰胺化单加氧酶 (PAM) 将具有 C 端 Gly 的肽前体酶促修饰为 α - 酰胺化肽。对于多细胞生物体,已专门描述参与肽 α 酰胺化的属于神经激素或神经毒素的 PAM 天然底物和酶 (PAM、PHM 和 PHL) (Metazoa)。属于细菌或古细菌的生物体迄今为止尚未被证明拥有可进行所需的酶促步骤以将肽中 C 端 Gly 转化为 α 酰胺部分的多肽。

[0086] 当在其他表达系统 (例如大肠杆菌) 中表达时,其中酶被表达为需要重折叠的不溶性蛋白,由于不足的活性酶产率,已在哺乳动物细胞表达系统中通过重组表达提供了用于工业体外转化 C 端甘氨酸延伸的肽前体为 α 酰胺化肽的真核生物 PAL、PHM 和 PAM 酶。

[0087] 相比于在细菌、真菌或酵母来源的系统中表达,在哺乳动物细胞表达系统中表达提供有限的产率。另外,在哺乳动物细胞中表达过程的可持续性是有限的,培养时间通常长,而每时间单位的表达产出低。相比于哺乳动物表达系统,蛋白在大肠杆菌中的表达通常导致较短的发酵循环和数倍的产出/细胞。

[0088] 对于真核生物 PAL、PHM 或 PAM 酶,由于产生正确折叠和有活性的酶需要哺乳动物细胞表达系统,设计提供高数量的 α -酰胺化酶的经济可持续方法用于在工业生产重组 α -酰胺化肽中使用是困难的。

[0089] 真核生物 PAM、PHM 和 PAL 酶具有许多 Cys 残基作为其结构完整性的基础,当这些酶在大肠杆菌细胞的还原性胞质溶胶环境中表达时,它们不能像天然酶一样形成正确的二硫键。Kolhekar A. S 等 *Biochemistry* 2002, 41, 12384-12394 描述了大鼠 PAL 的催化结构域在二硫桥方面的表征。PAL 中的两个二硫桥在来自真核生物的特征化 PAL 结构域之间是完全保守的(第 12388 页),表明二硫桥对酶的结构完整性重要。从 Kolhekar A. S 等 *Biochemistry* 2002, 41, 12384-12394 来看,用 B-巯基乙醇还原二硫桥导致酶活性的降低,因此支持二硫桥在真核生物 PAL 结构域中的结构重要性。

[0090] 在 W090/08194 中描述了一种增加酶活性的尝试,其通过用还原剂(例如二硫苏糖醇或 2-巯基乙醇)与变性剂的组合处理大肠杆菌中表达的 α -酰胺化酶,然后使还原的蛋白氧化,但该方法没能增加酶活性。

[0091] EP2172550 描述一种重组 C 端 α -酰胺化酶衍生物,其不形成在来源于非洲爪蟾的 C 端 α -酰胺化酶(PHM 活性)中通常存在的 5 个二硫键中的至少一个,还描述了在大肠杆菌中重组产生所述衍生物的方法。

[0092] 本文所述的本发明提供新的肽(本发明的多肽),其能够催化 α -羟基甘氨酸转化为 α -酰胺。本发明的新的肽来源于原核生物体并具有与对真核生物 PAL 酶所述的不同的物理化学性质和结构特性。

[0093] 本发明尤其提供产生对重组过程的方法经济关键的 α -酰胺化酶的可持续方法。相比于哺乳动物表达 PAM/PAL 酶,本发明新的酶在大肠杆菌中表达产生更高的终产率。

[0094] 本发明酶的特征是它们可在大肠杆菌中以高产率表达为可溶性蛋白并在下游加工中对比于其真核生物对应物非常容易操作,因为它们不需要重折叠来获得活性。

[0095] 本发明酶的另一重要特征是尽管与 N 端融合伴侣融合,其活性还是被维持。这些融合伴侣的一个实例是在本发明中使用的 TAP 标签,其使得酶容易纯化和酶在酶促步骤后容易移除。另一实例是 His6 标签。在大肠杆菌中克隆并表达许多加上 TAP 标签的酶。因此,本发明的酶可在大肠杆菌中以高产率表达为加上 TAP 标签的可溶性融合蛋白,并在下游加工中对比于其真核生物对应物非常容易操作。

[0096] 因此,本发明酶不需要潜在耗时并昂贵的重折叠。因此,在由大肠杆菌例示的非哺乳动物宿主细胞中提供酶的高产率产生方法。

[0097] 下面描述本发明的不同方面。

[0098] 在本发明一个方面,酶包含下述基序 1: Xaa₁ Val Xaa₂ Asp Arg Xaa₃ Xaa₄ Xaa₅ Arg Xaa₆ Gln Xaa₇ Xaa₈ Xaa₉ Xaa₁₀ Xaa₁₁ Gly Xaa₁₂ Xaa₁₃ Xaa₁₄ Xaa₁₅ Xaa₁₆ Trp,

[0099] 其中 Xaa₁和 Xaa₇可为除 Cys 外的任何天然存在氨基酸,而 Xaa₂、Xaa₃、Xaa₄、Xaa₅、Xaa₆、Xaa₈、Xaa₉、Xaa₁₀、Xaa₁₁、Xaa₁₂、Xaa₁₃、Xaa₁₄、Xaa₁₅和 Xaa₁₆可为任何天然存在氨基酸。

[0100] 在一个方面, Xaa₄是 Asn。在一个方面, Xaa₁₄是 Leu。在一个方面, Xaa₇是 Val 或 Ile。在一个方面, Xaa₈是 Phe 或 Leu。在一个方面, Xaa₉是 Asp 或 Ser。

[0101] 在本发明一个方面, 酶包含下述基序 2: Asp Gly Tyr Xaa₁₇ Asn Xaa₁₈ Arg Xaa₁₉ Xaa₂₀ Xaa₂₁ Phe Xaa₂₂ Xaa₂₃ Xaa₂₄ Gly Xaa₂₅ Xaa₂₆ Xaa₂₇ Xaa₂₈ Xaa₂₉ Xaa₃₀ Xaa₃₁ Xaa₃₂ Xaa₃₃ Gly Xaa₃₄ Xaa₃₅ Xaa₃₆ Gly Xaa₃₇ Phe, 其中 Xaa₁₇-Xaa₃₇独立选自天然存在氨基酸, 条件是 Xaa₁₇不是 Cys。

[0102] 基序 1 位于活性位点区中并在与大鼠 (褐家鼠 (*Rattus norvegicus*)) PAL 的活性位点区中的 Arg706 类似的位置中包含 Arg 残基。由基序 1 所涵盖的围绕 Arg706 残基的区中的氨基酸在大鼠 PAL 和本发明酶之间显示有限的序列同一性, 但 Arg706 本身在本发明的特征化酶 (SEQ ID NO: 1-4) 中保守。Chufan 等 2009 确定了大鼠 PALcc R706A 突变体的相对 V_{max} 仅为野生型大鼠 PALcc 的 3%, 并显示该残基位于活性位点口袋中, 因此证实 Arg706 (第 969 页) 对于 PAL 酶的酶促催化的重要作用 (Chufán EE, De M, Eipper BA, Mains RE, Amzel LM. Amidation of Bioactive Peptides: The Structure of the Lyase Domain of the Amidating Enzyme (生物活性肽的酰胺化: 酰胺化酶的裂合酶结构域的结构). Structure. 2009 年 7 月 15 日; 17(7): 965-73)。

[0103] 基序 2 位于另一活性位点区中, 其在大鼠 PAL 的 Tyr654 相似的位置中包含 Tyr 残基, 其是高度保守的并为酶促催化的必需残基, 如通过诱变和结构研究所示 (De, M., Bell, J., Blackburn, N.J., Mains, R.E., and Eipper, B.A. (2006). Role for an essential tyrosine in peptide amidation (在肽酰胺化中必需酪氨酸的作用). J. Biol. Chem. 281, 20873 - 20882.)。

[0104] 在一个方面, 本发明涉及一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽选自: 包含与选自以下的氨基酸序列有至少 70%、或至少 75%、例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96%、或至少 98%、或 100% 同一性的氨基酸序列或由之组成的多肽: (a) SEQ ID NO: 1 的氨基酸 2-306; (b) SEQ ID NO: 2 的氨基酸 3-336; (c) SEQ ID NO: 3 的氨基酸 3-305; (d) SEQ ID NO: 4 的氨基酸 3-279; (e) SEQ ID NO: 13; (f) SEQ ID NO: 15; (g) SEQ ID NO: 19; (h) SEQ ID NO: 20; (i) SEQ ID NO: 21; (j) SEQ ID NO: 22; 和 (j) SEQ ID NO: 23。

[0105] 在一个方面, 酶来源于原核生物体, 例如细菌。

[0106] 在本发明的一个方面, 酶来源于伯克氏菌属 (*Burkholderia*)、 α 变形细菌 (*alpha proteobacterium*)、甲烷八叠球菌属 (*Methanosarcina*)、堆囊菌属 (*Sorangium*)、中慢生根瘤菌属 (*Mesorhizobium*)、慢生根瘤菌属 (*Bradyrhizobium*) 或土细菌 (*Solibacter*) 种。

[0107] 在本发明一个方面, 酶来源于赤细菌属种。在一个方面, 酶来源于微小杆菌属种。在一个方面, 酶来源于食土性细菌种。在一个方面, 酶来源于浮霉状菌属种。

[0108] 在本发明一个方面, 酶来源于赤细菌属、微小杆菌属、食土性细菌或浮霉状菌属种。在一个方面, 本发明的酶来源于选自以下细菌的原核生物体: 赤细菌属、微小杆菌属、食土性细菌、浮霉状菌属、伯克氏菌属种、 α 变形细菌、甲烷八叠球菌属种、堆囊菌属、海洋放线菌 (*Salinispora*) 种、中慢生根瘤菌属、慢生根瘤菌属和土细菌种。

[0109] 在一个方面, 酶是野生型酶。在一个方面, 本发明的多肽是无信号肽的成熟序列。

[0110] 在一个方面, 本发明的酶相比于野生型酶包含多达 30 个修饰。在一个方面, 酶相

比于野生型酶包含 1-30 个修饰。在一个方面,酶相比于野生型酶包含 5-25 个修饰。在一个方面,酶相比于野生型酶包含 10-20 个修饰。在一个方面,酶相比于野生型酶包含 12-18 个修饰。在一个方面,本文所述的修饰的数量是相比于成熟野生型序列。

[0111] 在一个方面,本发明的酶包含多达 30 个修饰。在一个方面,酶包含 1-30 个修饰。在一个方面,酶包含 5-25 个修饰。在一个方面,酶包含 10-20 个修饰。在一个方面,酶包含 12-18 个修饰。应当理解,作为来源于原核来源的本文例示的序列可形成这样的修饰的基础(例如 SEQ ID NO: 1)。

[0112] 在本发明一个实施方案中,酶被修饰,使得至少一个 Cys 残基被取代或缺失,例如通过其他合适的氨基酸(例如 Ala、Ser 或 Val)。

[0113] 在一个实施方案中,本发明的多肽包括融合伴侣。

[0114] 这些融合伴侣的一个实例是 TAP 标签,例如本发明中使用的 TAP 标签,其使得酶容易纯化和在酶促步骤后酶容易移除。另一实例是 His 标签,例如 His6 标签。

[0115] 在一个实施方案中,本发明的酶可包含纯化标签。在一个实施方案中,本发明的酶包含纯化标签,该标签含有来源于如在以编号 WO 2006/108826 和 WO 2008/043847 公布的国际专利申请中所述的嗜热细菌的高度碱性核糖体蛋白。

[0116] 在实施方案中,纯化后标签保留在酶上。

[0117] 在本发明的一个实施方案中,纯化标签是固定化金属亲和标签(例如 His6 或 His8)、谷胱甘肽转移酶标签、用抗体回收的标签(例如 FLAG 标签、HA 标签、MYC 标签)、生物素或链霉亲和素。

[0118] 在一个实施方案中,标签将包含接头序列,该接头序列包含用于体外切割纯化标签以提供酶的切割位点。

[0119] 在本发明的一个实施方案中,本发明的多肽包含融合伴侣 SEQ ID NO: 7 或 SEQ ID NO: 11。

[0120] 接头可例如具有 1-30、1-25、1-20 或 1-15 个氨基酸残基,在一个方面,接头可包含增加 α 螺旋形成或导致结构刚性的其他特征的氨基酸残基,例如 Leu、Pro 和 Ala。在另一方面,接头可包含为接头提供柔性的 Gly 和 Ser 残基。

[0121] 在一个实施方案中,接头可选自:

[0122] Ser Ser Ser Gly Gly Ser Gly Ser Glu Val Leu Phe Gln (SEQ ID NO: 8)

[0123] Ser Ser Ser Gly Ser Gly Glu Val Leu Phe Gln (SEQ ID NO: 9)

[0124] SER SER SER GLY GLY SER GLY GLY SER GLY (SEQ ID NO: 10)

[0125] Ser Ser Gly Gly Ser Gly Ser Glu Val Leu Phe Gln (SEQ ID NO: 12)

[0126] Ser Ser Ser Gly Gly Ser Gly Ser Glu Thr Leu Phe Gln (SEQ ID NO: 14)

[0127] SER SER SER GLY GLY SER GLY GLY SER (SEQ ID NO: 16)

[0128] 接头可与酶的 C 端或 N 端连接。切割位点可为能够使纯化标签从酶上体外切割下来的任何切割位点。

[0129] 在本发明的一个实施方案中,酶包含 SEQ ID NO: 1 并能够催化 α -羟基甘氨酸转化为 α -酰胺。在本发明的一个方面,酶包含 SEQ ID NO: 2 并能够催化 α -羟基甘氨酸转化为 α -酰胺。在本发明的一个方面,酶包含 SEQ ID NO: 3 并能够催化 α -羟基甘氨酸转化为 α -酰胺。在本发明的一个方面,酶包含 SEQ ID NO: 4 并能够催化 α -羟基甘氨酸转

化为 α -酰胺。在本发明的一个方面,酶包含 SEQ ID NO: 13 并能够催化 α -羟基甘氨酸转化为 α -酰胺。在本发明的一个方面,酶包含 SEQ ID NO: 15 并能够催化 α -羟基甘氨酸转化为 α -酰胺。

[0130] 在本发明的一个方面,本发明的多肽包含至多 2 个半胱氨酸残基。在本发明的一个方面,本发明的多肽不包含半胱氨酸残基。在本发明的一个方面,本发明的多肽包含 1 个半胱氨酸残基。

[0131] 在本发明一个方面,进行 α -酰胺化过程,其中本发明的酶起作用以催化靶肽的 C 端 α -羟基甘氨酸转化为 α -酰胺。

[0132] 在本发明的一个方面,本发明的酶在一种用于制备 α -酰胺化肽的方法中使用。

[0133] 在本发明的一个方面,通过使蛋白经受具有 PHM 活性的酶并藉此使得 C 端 Gly 残基被转化为 α -羟基甘氨酸,进行 α -酰胺化方法的第一步,然后使得本发明的酶催化 α -羟基甘氨酸转化为 α -酰胺。

[0134] 真核生物 PHM 活性取决于抗坏血酸和 Cu^{2+} , 而真核生物 PAL 活性取决于作为辅因子的 Zn^{2+} 和其他二价离子。对于本发明的 PAL 样酶,观察到活性对于 Zn^{2+} 的类似依赖性。

[0135] 在本发明的一个方面,通过包括以下步骤的方法进行靶肽的 α -酰胺化:i) 在加入 Cu^{2+} 和 / 或抗坏血酸的条件下使具有 C 端 Gly 残基的靶肽经受具有 PHM 活性的酶,藉此使得肽的 C 端 Gly 残基被转化为 α -羟基甘氨酸;和 ii) 使得本发明的酶催化所述肽的 α -羟基甘氨酸转化为 α -酰胺,其中用所述 PHM 和本发明所述酶进行的对所述肽的反应在两个单独的步骤中进行或同时进行。

[0136] 在本发明的一个方面,通过包括以下步骤的方法进行靶肽的 α -酰胺化:i) 在使得 PHM 酶将 C 端 Gly 残基转化为 α -羟基甘氨酸的条件下,使具有 C 端 Gly 残基的靶肽经受具有 PHM 活性的酶;和 ii) 使用本发明的酶催化所述肽的 α -羟基甘氨酸转化为 α -酰胺。在一个方面,在 Cu^{2+} 和 / 或抗坏血酸存在下使靶肽经受具有 PHM 活性的酶。在一个方面,使靶肽同时经受具有 PHM 活性的酶和本发明的酶。在一个方面,在 Cu^{2+} 和 / 或抗坏血酸存在下使靶肽同时经受具有 PHM 活性的酶和本发明的酶,并。

[0137] 在一个方面,本发明的酶能够将 α -羟基苯甲酰甘氨酸转化为苯甲酰胺。通过如 Katopodis AG 等, Biochemistry. 1990, 29(26):6115-6120 所述或如测定 (I) 中所述,测量 α -羟基苯甲酰甘氨酸向苯甲酰胺的转化,可证明酶的活性。

[0138] 在本发明的一个方面,酶在一种用于制备 α -酰胺化肽的方法中使用。当肽在 C 端 α -酰胺化时,某些肽的生物学活性显著增加。获益于 C 端 α -酰胺化的靶肽的实例是胰淀粉样肽、神经肽 Y (NPY)、肽 YY (PYY)、PYY-3-36、胰多肽 (PP)、胃泌素、降钙素、降钙素相关肽 (CGRP)、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、胰抑释素、阿黑皮质素原 (POMC)、 α -促黑素细胞激素 (α MSH)、 γ -促黑素细胞激素 (γ MSH) 和酰胺化铰链肽 (HP-N)。

[0139] 在一个方面,肽靶标是胰淀粉样肽或其功能类似物的 C 端 Gly 延伸的前体。在一个方面,肽靶标是 GLP-1 或其功能类似物的 C 端 Gly 延伸的前体。在一个方面,肽靶标是 PYY 或其功能类似物的 C 端 Gly 延伸的前体。

[0140] 在本发明的一个方面,酶和具有 PHM 活性的酶一起使用。

[0141] 在一个方面,酶和具有 PHM 活性的酶一起在 Cu^{2+} 和 / 或抗坏血酸存在下使用。

[0142] 当用本发明的酶制备 α -酰胺化肽时,使酶与具有 C 端 Gly 残基的肽前体反应。在抗坏血酸和 / 或 Cu^{2+} 存在下,通过酶促方法可将 Gly 残基转化为 α -羟基甘氨酸。能够将 Gly 转化为 α -羟基甘氨酸的酶的一个实例是具有 PHM 活性的酶。本发明的酶催化 α -羟基甘氨酸转化为 α -酰胺。

[0143] 在一个方面, α -酰胺化肽用于制备药物,例如用于治疗或预防肥胖症、高血糖症(包括应激性高血糖症)、2 型糖尿病、葡萄糖耐量减低、1 型糖尿病、烧伤、手术创伤、在治疗中需要合成代谢作用的其他疾病或损伤、心肌梗死、中风、冠心病、其他心血管疾病,治疗病危糖尿病和非糖尿病患者和多神经病。

[0144] 在另一方面, α -酰胺化肽作为药物用于延缓或预防 2 型糖尿病中的疾病进展。

[0145] 在本发明的一个方面,提供 α -酰胺化肽用作药物,用于治疗或预防肥胖症、高血糖症(包括应激性高血糖症)、2 型糖尿病、葡萄糖耐量减低、1 型糖尿病、和烧伤、手术创伤和在治疗中需要合成代谢作用的其他疾病或损伤、心肌梗死、中风、冠心病和其他心血管疾病。

[0146] 在再一方面,本发明涉及一种治疗或预防以下疾病的方法:肥胖症、高血糖症(包括应激性高血糖症)、2 型糖尿病、葡萄糖耐量减低、1 型糖尿病、和烧伤、手术创伤和在治疗中需要合成代谢作用的其他疾病或损伤、心肌梗死、冠心病和其他心血管疾病、中风,该方法包括对需要这种治疗的患者给予用于这种治疗的有效量的 α -酰胺化肽。

[0147] 在一个方面,本发明包括一种用于产生 α -酰胺化肽的方法,其包括使得靶肽与本发明的酶反应并纯化 α -酰胺化肽。在一个方面,该方法包括使用具有 PHM 活性的酶。

[0148] 在一个方面,该方法包括从以下中选择靶肽:胰淀粉样肽、神经肽 Y (NPY)、肽 YY (PYY)、PYY-3-36、胰多肽 (PP)、胰高血糖素样肽 (GLP-1)、胃泌素、降钙素、降钙素相关肽 (CGRP)、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、胰抑释素、阿黑皮质素原 (POMC)、 α -促黑素细胞激素 (α MSH)、 γ -促黑素细胞激素 (γ MSH) 和酰胺化铰链肽 (HP-N),或者它们的功能类似物。

[0149] 在本发明的一个方面,本发明涉及一种编码本发明酶的分离核酸。在一个方面,提供一种包含启动子和所述分离核酸的重组核酸。

[0150] 在一个方面,提供所述重组核酸,其中该核酸还编码纯化标签。

[0151] 在一个方面,提供一种包含所述重组核酸的载体。在一个方面,提供一种包含所述重组核酸的宿主细胞。在本发明的一个方面,重组核酸存在于宿主细胞的基因组中或者在宿主细胞中自主复制的载体中。

[0152] 在一个实施方案中,宿主细胞是哺乳动物细胞,例如 CHO 细胞。一个实施方案,宿主细胞是非哺乳动物宿主细胞。在一个方面,细菌、真菌(例如酵母)被用作宿主细胞用于产生本发明的多肽。合适的宿主细胞有例如埃希氏菌属 (*Escherichia*) (例如大肠杆菌)、芽孢杆菌属 (*Bacillus*)、酵母属 (*Saccharomyces*) (例如啤酒酵母 (*Saccharomyces cerevisiae*)) 或毕赤酵母属 (*Pichia*) 的种。在一个方面,宿主细胞是大肠杆菌。

[0153] 在本发明的一个方面,提供一种用于产生能够催化 α -羟基甘氨酸转化为 α -酰胺的本发明酶的方法,其包括在适合于产生本发明酶的条件下维持如本文所述的本发明的宿主细胞。

[0154] 本发明的一个实施方案涉及一种用于产生具有肽基- α -羟基甘氨酸 α -酰胺化

裂合酶 (PAL) 活性的酶的方法,其包括以下步骤:(i) 在适合于酶表达的条件下培养非哺乳动物来源的重组表达宿主细胞,该宿主细胞包含含有编码如本文所述的本发明酶的核苷酸序列的核酸构建体,和(ii) 从(a) 细胞破碎和离心后的上清液和/或(b) 生长培养基中回收酶;其中宿主细胞是非哺乳动物来源,例如大肠杆菌菌株,且其中当在步骤(ii) 中回收时,该酶是可溶的。在一个实施方案中,当在步骤(ii) 中回收时,酶为催化活性的形式,使得酶不需要重折叠的步骤来获得催化活性(如本文所述的 PAL 样活性)。因此,一种用于产生本发明多肽的方法在本发明内,例如在大肠杆菌中产生,该方法不包括重折叠多肽的步骤。在一个方面,在本文所述的重组宿主细胞(例如大肠杆菌)中产生的本发明多肽是至少 50% 或至少 80% 可溶的,即不需要重折叠来获得催化活性。

[0155] 在一个方面,宿主细胞是大肠杆菌,其包含含有编码本发明酶、纯化标签和启动子的重组核酸的载体,并将宿主细胞在适合于产生所述酶的条件下维持。

[0156] 在一个方面,通过使用一个阳离子交换色谱步骤从发酵液中纯化酶,产生的纯度为约 80%,例如为至少 80%。

[0157] 在大肠杆菌中,大多数丰度蛋白在范围为 pI 4-7 和 8-10 的 pI 簇中发现,其中大多数蛋白和最高丰度蛋白在 pI 4-7 的范围内发现。可用于纯化本发明酶的 TAP 标签是高度碱性的,并当与酶融合时将显著增加酶的总正电荷和 pI,使得其明显区别于大块的宿主细胞污染物。这将使得酶能够在宿主细胞污染物不能与给定阳离子交换基质结合的盐浓度或 pI 下被洗脱。在一个实施方案中,纯化标签具有高于约 9 的 pI,例如高于 9。在一个实施方案中,纯化标签具有高于约 10 的 pI,即高于 10。在一个实施方案中,纯化标签的 pI 可在约 9 和约 12.5 之间,即在 9-12.5 的范围内,和在再一方面,pI 为约 10,即 pI 为 10。

[0158] 任何合适的阳离子交换基质可在本发明的方法中使用,合适阳离子交换柱材料的非限制性列表为:SP-Sepharose XL Amersham 产品目录号 17-5073-01;Streamline SP XL Amersham 产品目录号 17-5076-01;Streamline Direct CST Amersham 产品目录号 17-5266-03;Obelix SP Amersham 产品目录号 11-0010-86;S-Support Unosphere, BioRad 产品目录号 156-0113;SP-Sepharose High Performance Amersham 产品目录号 17-1087-03;Source30S Amersham 产品目录号 17-1273-02 和 Toyopearl SP650S TosoHaas 产品目录号 08437。

[0159] 取决于所选择的纯化标签的 pI 和电荷,TAP 标签将不同地有助于特定酶的总电荷。因此,通过选择能够使融合蛋白在仅最小量的宿主细胞污染物将共洗脱的盐浓度或 pH 下洗脱的纯化标签,可优化特定靶蛋白的纯化。

[0160] 接头中的氨基酸残基可选自这样的氨基酸残基:其将对加上标签的酶提供柔性较低的结构。由此可使酶和纯化标签之间的干扰最小。在一个实施方案中,接头可包含结构元件,例如 α 螺旋结构。

[0161] 由细胞产生的所表达的加上标签的酶可通过常规步骤从培养基中回收,该步骤包括:通过离心或过滤从培养基中分离宿主细胞,通过机械细胞破碎(例如超声处理或压力)释放融合蛋白,通过盐(例如硫酸铵)沉淀上清液或滤液的蛋白水性成分。在阳离子交换色谱捕获步骤后,可以盐梯度洗脱加上标签的酶并收集包含融合蛋白的洗脱级分。

[0162] 以未加标签的酶表示的本发明多肽可通过纯化回收,例如阴离子交换色谱、疏水作用色谱和凝胶过滤色谱(参见,例如,Scopes, R., Protein Purification,

Springer-Verlag, N.Y., 1982)。

[0163] 本发明多肽的纯度可例如通过使用凝胶图分析软件分析考马斯染色的 PAGE 凝胶或分析 HPLC UV214nm 色谱图来评价。

[0164] 在第一纯化步骤后,可用合适的加工酶(例如 EK)将纯化标签直接切除。如果盐浓度太高,那么在切割前可将酶脱盐。切割位点可为分离纯化的酶后能使体外切割有效的任何切割位点。最常用的肠激酶切割位点具有序列 DDDDK,其中切割在 K 之后发生。其他非限制性的加工酶切割位点包括因子 Xa 切割位点,其最常见为 IEGR,其中切割在 R 之后发生;凝血酶切割位点,其最常见为 LVPRG 或 LVPRGS,其中切割在 R 之后发生;烟草蚀纹病毒 (TEV) 蛋白酶切割位点,其最常见为 ENLYFQG/S,其中切割在 Q 之后发生;和 HRV14 3C 蛋白酶切割位点,其最常见为 LEVLFQ/GP,其中切割在 Q 之后发生。对专利申请 WO 2006/108826 进行了参考。

[0165] 切割之后的步骤可包括如第一步中的进一步的阳离子交换柱纯化。在这种方案下,通过加工酶所释放的纯化标签将具有极高的 pI,导致与阳离子交换基质非常有效的结合。切割的酶于是可在流通中从柱子上收集,而切下的纯化标签和来自生产细胞系的剩余高电荷污染物将保留在阳离子交换柱上。

[0166] 切割后的纯化步骤也可包括其他纯化方法,例如阴离子交换色谱、疏水作用色谱和凝胶过滤色谱(参见,例如,Scopes, R., Protein Purification, Springer-Verlag, N.Y., 1982)。

[0167] 在本发明的一个方面,酶被纯化至至少约 90-95% 同质性(即至 90-95% 同质性)、至至少约 98% 同质性(即至 98% 同质性)。通过例如凝胶电泳、氨基酸分析或其他基于 HPLC 的方法可评估纯化。

[0168] 编码具有或不具有标签的酶的重组核酸也可通过已建立的标准方法来合成制备,例如由 Beaucage 和 Caruthers, Tetrahedron Letters 22 (1981), 1859 - 1869 所述的氨基亚磷酸酯法或由 Matthes 等, EMBO Journal 3 (1984), 801 - 805 所述的方法。根据氨基亚磷酸酯法,例如在自动 DNA 合成仪中合成寡核苷酸、纯化、退火、在合适载体中连接并克隆。编码具有或不具有标签的酶的 DNA 序列也可通过聚合酶链式反应来制备,例如通过使用特异性引物的重叠延伸 PCR 剪接,例如描述于 US 4,683,202, Saiki 等, Science 239 (1988), 487-491 或 Sambrook 等, 同上。

[0169] 此外,根据标准技术,通过连接合成、基因组或 cDNA 来源(视情况而定)的对应于完整核酸构建体的不同部分的片段,可制备合成和基因组来源混合的、合成和 cDNA 来源混合的或基因组和 cDNA 来源混合的重组核酸。

[0170] 编码酶的 DNA 序列通常被插入到可为任何载体的重组载体中,其可合宜地经受重组 DNA 步骤,和载体的选择将通常取决于其被导入到其中的宿主细胞。因此,载体可为自主复制的载体,即载体作为染色体外实体存在,其复制不依赖染色体复制,例如质粒。备选地,载体可为,当导入到宿主细胞中时,其被整合到宿主细胞基因组中并与将其整合的染色体一起复制。

[0171] 载体优选为其中编码酶的 DNA 序列与 DNA 转录所需的另外区段有效连接的表达载体。通常,表达载体来源于质粒或病毒 DNA,或可包含两者的元件。术语“有效连接”表示,排列区段使得它们为其预期目的而协调地起作用,例如转录在启动子中起始并前进通过编

码本发明多肽（例如融合蛋白）的 DNA 序列。

[0172] 用于在表达酶中使用的表达载体将包含能够指导转录克隆的基因或 cDNA 的启动子。启动子可为任何 DNA 序列，其在选择的宿主细胞中显示转录活性并可来源于编码与宿主细胞同源或异源的蛋白的基因。

[0173] 用于在哺乳动物细胞中指导转录 DNA 的合适启动子的实例为 SV40 启动子 (Subramani 等, Mol. Cell Biol. 1 (1981), 854-864)、MT-1 (金属硫蛋白基因) 启动子 (Palmiter 等, Science 222 (1983), 809 - 814)、CMV 启动子 (Boshart 等, Cell 41:521-530, 1985) 或腺病毒 2 主要晚期启动子 (Kaufman 和 Sharp, Mol. Cell. Biol, 2:1304-1319, 1982)。

[0174] 用于在酵母宿主细胞中使用的合适启动子的实例包括来自酵母糖酵解基因 (Hitzeman 等, J. Biol. Chem. 255 (1980), 12073 - 12080 ;Alber 和 Kawasaki, J. Mol. Appl. Gen. 1 (1982), 419 - 434) 或乙醇脱氢酶基因 (Young 等, 载于 Genetic Engineering of Microorganisms for Chemicals (Hollaender 等, 主编), Plenum Press, New York, 1982) 的启动子,或 TPI1 (US 4,599,311) 或 ADH2-4c (Russell 等, Nature 304 (1983), 652 - 654) 启动子。

[0175] 用于在丝状真菌宿主细胞中使用的合适启动子的实例为,例如 ADH3 启动子 (McKnight 等, The EMBO J. 4 (1985), 2093 - 2099) 或 *tpiA* 启动子。其他有用启动子的实例是来源于编码米曲霉 (*A. oryzae*) TAKA 淀粉酶、米黑根毛霉 (*Rhizomucor miehei*) 天冬氨酸蛋白酶、黑曲霉 (*A. niger*) 中性 α 淀粉酶、黑曲霉酸稳定 α 淀粉酶、黑曲霉或泡盛曲霉 (*A. awamori*) 葡糖淀粉酶 (*gluA*)、米黑根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉丙糖磷酸异构酶或构巢曲霉 (*A. nidulans*) 乙酰胺酶的基因的那些启动子。优选为 TAKA 淀粉酶和 *gluA* 启动子。合适启动子在例如 EP 238 023 和 EP 383 779 中提及。

[0176] 用于在细菌宿主细胞中使用的合适启动子的实例包括,嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*) 产麦芽糖淀粉酶基因、地衣芽孢杆菌 (*Bacillus licheniformis*) α -淀粉酶基因、解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*) BAN 淀粉酶基因、枯草芽孢杆菌 (*Bacillus subtilis*) 碱性蛋白酶基因或短小芽孢杆菌木糖苷酶基因的启动子,或噬菌体 λ P_R 或 P_L 启动子,或用于在大肠杆菌中表达的启动子,例如 *lac*、*trp*、*phoA*、*araBAD*、*tac*, 噬菌体 T7 和 *cspA*。

[0177] 载体也可包含选择性标记,例如补充宿主细胞中缺陷的基因产物,例如编码二氢叶酸还原酶 (DHFR) 的基因,或粟酒裂殖酵母 (*Schizosaccharomyces pombe*) TPI 基因 (由 P. R. Russell, Gene 40, 1985, 第 125-130 页所述),或赋予药物抗性的标记基因,药物例如氨苄青霉素、卡那霉素、四环素、氯霉素、新霉素、潮霉素或甲氨蝶呤。对于丝状真菌,选择性标记包括 *amdS*、*pyrG*、*argB*、*niaD*和 *sC*。

[0178] 若需要,编码酶的 DNA 序列也可有效连接于合适终止子,例如人生长激素终止子 (Palmiter 等, Science 222, 1983, 第 809-814 页) 或 TPI1 (Alber 和 Kawasaki, J. Mol. Appl. Gen. 1, 1982, 第 419-434 页) 或 ADH3 (McKnight 等, The EMBO J. 4, 1985, 第 2093-2099 页) 终止子。表达载体也可包含位于启动子下游和多肽序列自身 (例如融合蛋白) 的插入位点上游的一系列 RNA 剪接位点。优选的 RNA 剪接位点可获自腺病毒和 / 或免疫球蛋白基因。在表达载体中也包含位于插入位点下游的聚腺苷酸化信号。

尤其优选的聚腺苷酸化信号包括,来自 SV40 的早期或晚期聚腺苷酸化信号 (Kaufman 和 Sharp, 同上),来自腺病毒 5' E1b 区的聚腺苷酸化信号,人生长激素基因终止子 (DeNoto 等 Nucl. Acids Res. 9:3719-3730, 1981)。表达载体也可包含位于启动子和 RNA 剪接位点之间的非编码病毒前导序列,例如腺病毒 2 三联前导序列;和增强子序列,例如 SV40 增强子。

[0179] 为了将酶导向宿主细胞的分泌途径中,在重组载体中可提供分泌信号序列(亦被称为前导序列、前原序列(prepro sequence)或前序列(pre sequence))。分泌信号序列在正确阅读框中与编码酶的 DNA 序列连接。分泌信号序列通常位于编码肽的 DNA 序列的 5'。分泌信号序列可为通常与酶相关的信号序列,或可来自编码另一分泌蛋白的基因。

[0180] 关于从酵母细胞中分泌,分泌信号序列可编码任何确保将表达的多肽有效导向细胞分泌途径中的信号肽。信号肽可为天然存在信号肽或其功能性部分,或者它可为合成肽。已发现的合适信号肽为, α -因子信号肽(参见 US 4,870,008)、小鼠唾液淀粉酶的信号肽(参见 O. Hagenbuchle 等, Nature 289, 1981, 第 643-646 页)、修饰的羧肽酶信号肽(参见 L. A. Valls 等, Cell 48, 1987, 第 887-897 页)、酵母 BAR1 信号肽(参见 WO 87/02670)或酵母天冬氨酸蛋白酶 3 (YAP3) 信号肽(参见 M. Egel-Mitani 等, Yeast 6, 1990, 第 127-137 页)。

[0181] 为了在酵母中有效分泌,也可将编码前导肽的序列插入到信号序列的下游和编码多肽的 DNA 序列的上游。前导肽的功能是使得表达的肽能够从内质网被导向高尔基体中并进一步将用于分泌的分泌囊泡导向培养基中(即多肽穿过细胞壁或至少通过细胞膜进入到酵母细胞的壁膜间隙中的输出)。前导肽可为酵母 α -因子前导序列(其用途例如在 US 4,546,082、US 4,870,008、EP 16201、EP 123 294、EP 1235 44 和 EP 163 529 中描述)。备选地,前导肽可为合成前导肽,即非天然存在的前导肽。合成前导肽可例如如 WO 89/02463 或 WO 92/11378 中所述构建。

[0182] 为了在丝状真菌中使用,信号肽可合宜地来源于编码曲霉属 (*Aspergillus* sp.) 淀粉酶或葡糖淀粉酶的基因、编码米黑根毛霉脂肪酶或蛋白酶或柔毛腐质霉 (*Humicola lanuginosa*) 脂肪酶的基因。信号肽优选来源于编码米曲霉 TAKA 淀粉酶、黑曲霉中性 α 淀粉酶、黑曲霉酸稳定淀粉酶或黑曲霉葡糖淀粉酶的基因。合适信号肽在例如 EP 238 023 和 EP 215 594 中公开。

[0183] 被编码酶的 DNA 导入到其中的宿主细胞可为能够产生本发明酶的任何细胞,并包括细菌、酵母、真菌和高等真核生物细胞。

[0184] 能够基于培养产生本发明多肽的细菌宿主细胞的实例为革兰氏阳性菌,例如芽孢杆菌属的菌株,例如枯草芽孢杆菌、地衣芽孢杆菌、迟缓芽孢杆菌 (*B. lentus*)、短芽孢杆菌 (*B. brevis*)、嗜热脂肪芽孢杆菌、嗜碱芽孢杆菌 (*B. alkalophilus*)、解淀粉芽孢杆菌、凝结芽孢杆菌 (*B. coagulans*)、环状芽孢杆菌 (*B. circulans*)、灿烂芽孢杆菌 (*B. lautus*)、巨大芽孢杆菌 (*B. megatherium*) 或苏云金芽孢杆菌 (*B. thuringiensis*) 的菌株,或链霉菌属 (*Streptomyces*) 的菌株,例如浅青紫链霉菌 (*S. lividans*) 或鼠灰链霉菌 (*S. murinus*);或革兰氏阴性细菌,例如大肠杆菌的菌株。通过原生质体转化或通过以本质上已知的方法使用感受态细胞实现细菌的转化(参见 Sambrook 等, 同上)。

[0185] 合适酵母细胞的实例包括酵母属或裂殖酵母属的细胞,尤其是酿酒酵母或克鲁

弗酵母 (*Saccharomyces kluyveri*) 菌株。用异源 DNA 转化酵母细胞并藉此产生异源多肽的方法已描述在例如 US 4,599,311、US 4,931,373、US 4,870,008、5,037,743 和 US 4,845,075 中,所有这些文献通过引用结合到本文中。通过由选择性标记所确定的表型(通常为药物抗性或在特定养分(例如亮氨酸)不存在下有能力生长)选择转化的细胞。用于在酵母中使用的优选的载体是在 US 4,931,373 中揭露的 POT1。合适酵母细胞的另外的实例是克鲁维酵母属 (*Kluyveromyces*) (例如乳酸克鲁维酵母 (*K. lactis*))、汉逊酵母属 (*Hansenula*) (例如多形汉逊酵母 (*H. polymorpha*)) 或毕赤酵母属 (例如巴斯德毕赤酵母 (*P. pastoris*)) 的菌株(参见 Gleeson 等, J. Gen. Microbiol. 132, 1986, 第 3459-3465 页; US 4,882,279)。

[0186] 其他真菌细胞的实例有丝状真菌细胞,例如曲霉菌属、脉孢菌属 (*Neurospora spp.*)、镰孢属 (*Fusarium spp.*) 或木霉属 (*Trichoderma spp.*),尤其是米曲霉、构巢曲霉或黑曲霉的菌株。使用曲霉菌属表达蛋白在例如 EP 272 277、EP 238 023、EP 184 438 中有述。尖镰孢 (*F. oxysporum*) 的转化可例如按照 Malardier 等, 1989, Gene 78: 147-156 所述进行。木霉属的转化可例如按照 EP 244 234 中所述进行。

[0187] 当丝状真菌用作宿主细胞时,它可用本发明重组核酸来转化,合宜地通过在宿主染色体中整合核酸来获得重组宿主细胞。因为重组核酸更可能在细胞中稳定维持,所以该整合通常被认为是优势。根据常规方法,例如通过同源或异源重组,可进行重组核酸到宿主染色体中的整合。

[0188] 然后在允许酶表达的条件下,将转化或转染的宿主细胞在合适营养培养基中培养,在此之后可从培养物中回收全部或部分所产生的酶,例如未加标签的酶。用于培养细胞的培养基可为适合于培养宿主细胞的任何常规培养基,例如含有恰当补充物的基本培养基或复合培养基。合适培养基可获自工业供应商或可根据公开的配方(例如美国典型培养物保藏中心 (American Type Culture Collection) 的目录)制备。

[0189] 本发明进一步在下述段落(实施方案)中概述:

[0190] 1a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽,其中所述多肽具有含有下述基序 1 的氨基酸序列:

[0191] Xaa₁ Val Xaa₂ Asp Arg Xaa₃ Xaa₄ Xaa₅ Arg Xaa₆ Gln Xaa₇ Xaa₈ Xaa₉ Xaa₁₀ Xaa₁₁ Gly Xaa₁₂ Xaa₁₃ Xaa₁₄ Xaa₁₅ Xaa₁₆ Trp;

[0192] 其中 Xaa₁、Xaa₂、Xaa₃、Xaa₄、Xaa₅、Xaa₆、Xaa₇、Xaa₈、Xaa₉、Xaa₁₀、Xaa₁₁、Xaa₁₂、Xaa₁₃、Xaa₁₄、Xaa₁₅ 和 Xaa₁₆ (Xaa₁-Xaa₁₆) 独立选自天然存在氨基酸,条件是 Xaa₁ 和 Xaa₇ 不是 Cys。

[0193] 2a. 段落 1a 的多肽,其中 Xaa₄ 是 Asn。

[0194] 3a. 段落 1a 或 2a 的多肽,其中 Xaa₁₄ 是 Leu。

[0195] 4a. 段落 1a-3a 中任一项的多肽,其中 Xaa₇ 是 Val 或 Ile。

[0196] 5a. 段落 1a-4a 中任一项的多肽,其中 Xaa₈ 是 Phe 或 Leu。

[0197] 6a. 段落 1a-5a 中任一项的多肽,其中 Xaa₉ 是 Asp 或 Ser。

[0198] 7a. 段落 1a 的多肽,其中 Xaa₄ 是 Asn, Xaa₁₄ 是 Leu, Xaa₇ 是 Val 或 Ile, Xaa₈ 是 Phe 或 Leu 且 Xaa₉ 是 Asp 或 Ser。

[0199] 8a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽,其中所述多肽具有含有下述基序 2 的氨基酸序列:

[0200] Asp Gly Tyr Xaa₁₇ Asn Xaa₁₈ Arg Xaa₁₉ Xaa₂₀ Xaa₂₁ Phe Xaa₂₂ Xaa₂₃ Xaa₂₄ Gly Xaa₂₅ Xaa₂₆ Xaa₂₇ Xaa₂₈ Xaa₂₉ Xaa₃₀ Xaa₃₁ Xaa₃₂ Xaa₃₃ Gly Xaa₃₄ Xaa₃₅ Xaa₃₆ Gly Xaa₃₇ Phe

[0201] 其中 Xaa₁₇ Xaa₁₈ Xaa₁₉ Xaa₂₀ Xaa₂₁ Xaa₂₂ Xaa₂₃ Xaa₂₄ Xaa₂₅ Xaa₂₆ Xaa₂₇ Xaa₂₈ Xaa₂₉ Xaa₃₀ Xaa₃₁ Xaa₃₂ Xaa₃₃ Xaa₃₄ Xaa₃₅ Xaa₃₆ Xaa₃₇ (Xaa₁₇-Xaa₃₇) 独立选自天然存在氨基酸,条件是 Xaa₁₇不是 Cys。

[0202] 9a. 段落 1a-8a 中任一项的多肽,其中多肽来源于原核生物体。

[0203] 10a. 段落 9a 的多肽,其中所述原核生物体选自:赤细菌属、微小杆菌属、食土性细菌和浮霉状菌属种,或者其中所述原核生物体选自赤细菌属、微小杆菌属、食土性细菌、浮霉状菌属、伯克氏菌属种、 α 变形细菌、甲烷八叠球菌属种、堆囊菌属、海洋放线菌种、中慢生根瘤菌属、慢生根瘤菌属和土细菌种。

[0204] 11a. 段落 9a 或 10a 的多肽,其中所述多肽是野生型序列。

[0205] 12a. 段落 11a 的多肽,其中所述多肽是无信号肽的成熟序列。

[0206] 13a. 段落 1a-12a 中任一项的多肽,其中所述多肽包含 0、1 或 2 个半胱氨酸残基。

[0207] 14a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽,该多肽选自:

[0208] (a) 包含与 SEQ ID NO: 1 的氨基酸 2-306 有至少 75% 同一性的氨基酸序列或由之组成的多肽;

[0209] (b) 包含与 SEQ ID NO: 2 的氨基酸 3-336 有至少 75% 同一性的氨基酸序列或由之组成的多肽;

[0210] (c) 包含与 SEQ ID NO: 3 的氨基酸 3-305 有至少 75% 同一性的氨基酸序列或由之组成的多肽;和

[0211] (d) 包含与 SEQ ID NO: 4 的氨基酸 3-279 有至少 75% 同一性的氨基酸序列或由之组成的多肽;

[0212] (e) 包含与 SEQ ID NO: 13 的氨基酸有至少 75% 同一性的氨基酸序列或由之组成的多肽;和

[0213] (f) 包含与 SEQ ID NO: 15 的氨基酸有至少 75% 同一性的氨基酸序列或由之组成的多肽。

[0214] 15a. 段落 14a 的多肽,其中多肽包含与 SEQ ID NO: 1 的氨基酸 2-306 具有至少 75% (例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0215] 16a. 段落 14a 的多肽,其中多肽包含与 SEQ ID NO: 2 的氨基酸 3-336 具有至少 75% (例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0216] 17a. 段落 14a 的多肽,其中多肽包含与 SEQ ID NO: 3 的氨基酸 3-305 具有至少 75% (例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0217] 18a. 段落 14a 的多肽,其中多肽包含与 SEQ ID NO: 4 的氨基酸 3-279 具有至少 75% (例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0218] 19a. 段落 14a-18a 中任一项的多肽,其中所述氨基酸序列包含 0、1 或 2 个半胱

氨酸残基。

[0219] 20a. 段落 1a 与段落 14a-19a 中任一项组合的多肽。

[0220] 21a. 段落 2a-12a 中任一项与段落 14a-19a 中任一项组合的多肽。

[0221] 22a. 段落 1a-13a 中任一项的多肽, 其中所述氨基酸序列相比于来源于原核生物体的野生型氨基酸序列包含多达 30 个修饰。

[0222] 23a. 段落 1a-13a 中任一项的多肽, 其中所述氨基酸序列相比于来源于原核生物体的成熟野生型 PAL 样酶的氨基酸序列包含以下范围内的修饰: 1-30 个修饰、1-25 个修饰、5-25 个修饰、1-20 个修饰、10-20 个修饰、1-12 个修饰、1-18 个修饰或 12-18 个修饰。

[0223] 24a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽选自:

[0224] (a) 包含相比于 SEQ ID NO: 1 的氨基酸 2-306 具有修饰的氨基酸序列或由之组成的多肽,

[0225] (b) 包含相比于 SEQ ID NO: 2 的氨基酸 3-336 具有修饰的氨基酸序列或由之组成的多肽,

[0226] (c) 包含相比于 SEQ ID NO: 3 的氨基酸 3-305 具有修饰的氨基酸序列或由之组成的多肽, 和

[0227] (d) 包含相比于 SEQ ID NO: 4 的氨基酸 3-279 具有修饰的氨基酸序列或由之组成的多肽,

[0228] 其中所述修饰为 1-30 个氨基修饰、1-25 个修饰、5-25 个修饰、1-20 个修饰、10-20 个修饰、1-12 个修饰、1-18 个修饰或 12-18 个修饰。

[0229] 25a. 段落 22a-24a 中任一项的多肽, 其中所述修饰由序列中一个或多个氨基酸的一个或多个置换、缺失、添加(包括插入)组成。

[0230] 26a. 一种分离多肽, 其选自:

[0231] (i) 一种分离多肽, 其为能够催化 α -羟基甘氨酸转化为 α -酰胺的酶, 其特征在于该酶来源于原核生物体,

[0232] (ii) 一种分离多肽, 其为能够催化 α -羟基甘氨酸转化为 α -酰胺的酶, 其中所述多肽的特征在于它在大肠杆菌中可被表达为可溶性蛋白,

[0233] (iii) 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽, 其特征在于它在大肠杆菌中被表达为可溶性蛋白,

[0234] (iv) 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽, 其特征在于它包含 0、1 或 2 个半胱氨酸残基,

[0235] (v) 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽, 其特征在于它不包含半胱氨酸残基。

[0236] 27a. 段落 26a(i) 的多肽, 其中酶相比于野生型酶包含多达 30 个修饰。

[0237] 28a. 段落 27a 的多肽, 其中所述氨基酸序列相比于来源于所述原核生物体的成熟野生型 PAL 样酶的氨基酸序列包含以下范围内的修饰: 1-30 个修饰、1-25 个修饰、5-25 个修饰、1-20 个修饰、10-20 个修饰、1-12 个修饰、1-18 个修饰或 12-18 个修饰。

[0238] 29a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽包含 SEQ ID NO: 1 的氨基酸序列 2-306 或由之组成。

[0239] 30a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽包含 SEQ

ID NO: 2 的氨基酸序列 3-336 或由之组成。

[0240] 31a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽包含 SEQ ID NO: 3 的氨基酸序列 3-305 或由之组成。

[0241] 32a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽包含 SEQ ID NO: 4 的氨基酸序列 3-279 或由之组成。

[0242] 33a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽包含氨基酸序列 SEQ ID NO: 13 或由之组成。

[0243] 34a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽, 该多肽包含氨基酸序列 SEQ ID NO: 15 或由之组成。

[0244] 35a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽, 该多肽可通过包括以下步骤的方法获得:

[0245] (i) 在适合于表达所述多肽的条件下培养非哺乳动物来源的重组表达宿主细胞, 该宿主细胞包含含有编码具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶 (PAL) 活性的多肽的核苷酸序列的核酸构建体; 和

[0246] (ii) 从 (a) 细胞破碎和离心后的上清液和 / 或 (b) 生长培养基中回收多肽,

[0247] 其中宿主细胞为非哺乳动物来源, 例如大肠杆菌菌株, 且其中当在步骤 (ii) 中回收时, 多肽是可溶的并为有活性的形式, 意指多肽不需要重折叠的步骤来获得催化活性。

[0248] 36a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽, 该多肽可通过包括以下步骤的方法获得:

[0249] (i) 在适合于表达所述多肽的条件下培养非哺乳动物来源的重组表达宿主细胞, 该宿主细胞包含含有编码具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶 (PAL) 活性的多肽的核苷酸序列的核酸构建体; 和

[0250] (ii) 从 (a) 细胞破碎和离心后的上清液和 / 或 (b) 生长培养基中回收多肽,

[0251] 其中宿主细胞为非哺乳动物来源, 例如大肠杆菌菌株。

[0252] 37a. 段落 35a 或 36a 的多肽, 其进一步的特征在于段落 1a。

[0253] 38a. 段落 35a 或 36a 的多肽, 其进一步的特征在于段落 2a-14a 中任一项。

[0254] 39a. 段落 365 或 36a 的多肽, 其中所述多肽包含 0、1 或 2 个半胱氨酸残基。

[0255] 40a. 段落 35a 或 36a 或 39a 的多肽, 该多肽来源于原核生物体。

[0256] 41a. 段落 40a 的多肽, 其中所述原核生物体选自: 赤细菌属、微小杆菌属、食土性细菌、浮霉状菌属、伯克氏菌属种、 α 变形细菌、甲烷八叠球菌属种、堆囊菌属、中慢生根瘤菌属、慢生根瘤菌属、海洋放线菌种、和土细菌种, 优选所述原核生物体为赤细菌属、微小杆菌属、食土性细菌和浮霉状菌属种。

[0257] 42a. 段落 40a 或 41a 的多肽, 其中所述多肽为野生型序列。

[0258] 43a. 段落 42a 的多肽, 其中所述多肽为不包含信号肽的成熟野生型序列。

[0259] 44a. 段落 40a-43a 中任一项的多肽, 其中所述多肽的所述氨基酸序列相比于来源于所述原核生物体的成熟野生型 PAL 样酶的氨基酸序列包含以下范围内的修饰: 1-30 个修饰、1-25 个修饰、5-25 个修饰、1-20 个修饰、10-20 个修饰、1-12 个修饰、1-18 个修饰或 12-18 个修饰。

[0260] 45a. 段落 44a 的多肽, 其中所述修饰由序列中一个或多个氨基酸的一个或多个

置换、缺失、添加（包括插入）组成。

[0261] 46a. 段落 40a-43a 中任一项的多肽，其中所述氨基酸序列与来源于所述原核生物体的成熟野生型 PAL 样酶的氨基酸序列有至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98% 同一性。

[0262] 47a. 段落 1a-46a 中任一项的多肽，其中该多肽能够将 α -羟基苯甲酰甘氨酸转化为苯甲酰胺。

[0263] 48a. 段落 1a-46a 中任一项的多肽，其中所述 α -羟基甘氨酸为式 R-Gly(OH) 的 C 端 α -羟基甘氨酸，其中 R 为肽，Gly(OH) 为与所述肽的 C 端连接的 α -羟基甘氨酸残基，且其中所述 α -酰胺具有式 R-NH₂。

[0264] 49a. 段落 1a-46a 中任一项的多肽，该多肽为肽基- α -羟基甘氨酸 α -酰胺化裂合酶 (PAL)。

[0265] 50a. 段落 1a-49a 中任一项的多肽，其中该多肽包含融合伴侣。

[0266] 51a. 段落 1a-49a 中任一项的多肽，其中该多肽包含 TAP 标签。

[0267] 52a. 段落 1a-49a 中任一项的多肽，其中该多肽包含纯化标签，纯化标签例如选自：固定化金属亲和和标签 (His6 或 His8)、谷胱甘肽转移酶标签、用抗体回收的标签 (FLAG 标签、HA 标签、MYC 标签)、生物素和链霉亲和素。

[0268] 53a. 段落 50a-52a 中任一项的多肽，其中该酶与融合伴侣通过接头序列连接。

[0269] 54a. 段落 53a 的多肽，其中接头序列具有 1-30、1-25、1-20 或 1-15 个氨基酸。

[0270] 55a. 段落 1a-54a 中任一项的多肽，其中所述多肽为实质上纯。

[0271] 56a. 段落 1a-54a 中任一项的多肽，其中该多肽制品：

[0272] (i) 包含至多 50%、或至多 10% 或至多 5% 重量的与其天然或重组缔合的其他多肽材料，或 (ii) 以重量计，相比于与其天然或重组缔合的其他多肽材料，多肽为至少 50% 纯、至少 90% 纯或至少 95% 纯。

[0273] 57a. 段落 1a-54a 中任一项的多肽，其中所述多肽相比于其他多肽材料为至少 5% 纯、至少 10% 纯、至少 25% 纯、至少 50% 纯、至少 75% 纯、至少 80% 纯、至少 90% 纯，如通过 SDS-PAGE 所测定。

[0274] 58a. 一种分离多核苷酸，其包含编码段落 1a-57a 或 73a 中任一项的多肽的核苷酸序列。

[0275] 59a. 一种分离重组核酸构建体，其包含段落 58a 的多核苷酸。

[0276] 60a. 段落 59a 的核酸构建体，其中核苷酸序列与指导多肽在表达宿主中产生的一种或多种控制序列有效连接。

[0277] 61a. 段落 60a 或 61a 的核酸构建体，其包含启动子。

[0278] 62a. 段落 59a-61a 中任一项的核酸，其中核酸还编码纯化标签和 / 或接头。

[0279] 63a. 一种重组表达宿主细胞，其包含段落 59a-62a 中任一项的核酸构建体。

[0280] 64a. 段落 63a 的宿主细胞，其中重组核酸存在于宿主细胞的基因组中或在宿主细胞中自主复制的载体中。

[0281] 65a. 段落 63a 或 64a 的宿主细胞，其中宿主细胞选自：(i) 细菌或真菌，(ii) 酵母，(iii) 哺乳动物宿主细胞和 (iv) 原核宿主细胞。

[0282] 66a. 段落 63a-65a 中任一项的宿主细胞，其中宿主细胞是大肠杆菌。

[0283] 67a. 一种用于产生能够催化 α -羟基甘氨酸转化为 α -酰胺的酶的方法,其包括在适合于产生所述酶的条件下维持段落 63a-66a 中任一项的宿主细胞。

[0284] 68a. 一种产生段落 1a-57a 或 73a 中任一项的多肽的方法,其包括:

[0285] (i) 在有利于产生多肽的条件下培养包含含有编码所述多肽的核苷酸序列的核酸构建体的宿主细胞;和(ii) 回收多肽。

[0286] 69a. 权利要求 68a 的方法,其中从细胞破碎后的上清液中以可溶性形式回收多肽。

[0287] 70a. 一种用于产生具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶(PAL)活性的酶的方法,其包括以下步骤:

[0288] i) 在适合于酶表达的条件下,培养非哺乳动物来源的重组表达宿主细胞,该宿主细胞包含含有编码酶的核苷酸序列的核酸构建体;和

[0289] ii) 从 a) 细胞破碎和离心后的上清液中和 / 或 b) 生长培养基中回收酶

[0290] 其中宿主细胞为非哺乳动物来源,例如大肠杆菌菌株,且其中当在步骤(ii)中回收时,该酶是可溶的。

[0291] 71a. 段落 70a 的方法,其中该酶具有包含 0、1 或 2 个 Cys 残基的氨基酸序列。

[0292] 72a. 段落 70a 或 71a 的方法,其中当在步骤(ii)中回收时,酶为催化活性形式,意指酶不需要重折叠的步骤来获得催化活性。

[0293] 73a. 一种分离多肽,其为可通过段落 70a-72a 中任一项的方法获得的具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的酶,且其中所述酶能够催化反应: $R-Gly(OH) \rightarrow R-NH_2$,

[0294] 其中 R 为肽, Gly(OH) 为与所述肽的 C 端连接的 α -羟基甘氨酸残基。

[0295] 74a. 段落 1a-57a 或 73a 中任一项的多肽用于催化肽中 C 端 α -羟基甘氨酸残基转化为 α -酰胺残基的用途。

[0296] 75a. 段落 74a 的用途,其中该多肽在用于制备 α -酰胺化肽的方法中使用。

[0297] 76a. 段落 74a 或 75a 的用途,其中当使用所述多肽时,存在 Cu^{2+} 离子和 / 或抗坏血酸。

[0298] 77a. 段落 74a-76a 中任一项的用途,其中具有 PHM 活性的酶在用于制备 α -酰胺化肽的方法中使用。

[0299] 78a. 一种用于产生 α -酰胺化肽的方法,其包括以下步骤:(i) 在适合于酶促活性的条件下,使得具有 C 端甘氨酸残基的靶肽与肽基甘氨酸 α -羟化单加氧酶(PHM) 和段落 1a-57a 或 73a 中任一项的多肽两者能够反应,其中用所述 PHM 和所述多肽对所述靶肽的反应在两个单独的步骤中或同时进行;(ii) 回收 C 端 α -酰胺化肽。

[0300] 79a. 一种用于产生 α -酰胺化肽的方法,其包括:

[0301] (i) 通过原核宿主中重组表达肽来产生酰胺化肽的前体(靶肽),其中所述前体具有 C 端 Glu 残基;

[0302] ii) 用 PHM 样酶处理所述前体以获得所述肽的 C 端 α -羟基甘氨酸;

[0303] iii) 用段落 1a-57a 或 73a 中任一项的多肽处理所述 C 端 α -羟基甘氨酸;

[0304] iv) 回收 C 端 α -酰胺化肽。

[0305] 80a. 段落 74a-79a 中任一项的用途或方法,其中具有 PHM 的酶的氨基酸序列来源

于哺乳动物生物体。

[0306] 81a. 段落 74a-80a 中任一项的用途或方法,其中回收的 α -酰胺化肽被用作药物。

[0307] 82a. 权利要求 81a 的用途或方法,其包括步骤 (iii):制备包含回收的 C 端 α -酰胺化肽的药物制剂。

[0308] 83a. 段落 74a-82a 中任一项的用途或方法,其中所述靶肽选自:胰淀粉样肽、神经肽 Y (NPY)、肽 YY (PYY)、PYY-3-36、胰多肽 (PP)、胰高血糖素样肽 (GLP-1)、胃泌素、降钙素、降钙素相关肽 (CGRP)、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、胰抑释素、阿黑皮质素原 (POMC)、 α -促黑素细胞激素 (α MSH)、 γ -促黑素细胞激素 (γ 1MSH) 和酰胺化铰链肽 (HP-N),或者它们的功能类似物。

[0309] 84a. 段落 74a-83a 中任一项的用途或方法,其中 α -酰胺化肽用于治疗或预防下述病况和/或疾病之一:肥胖症、高血压、高血糖症(包括应激性高血糖症)、2 型糖尿病、葡萄糖耐量减低、1 型糖尿病、烧伤、手术创伤、在治疗中需要合成代谢作用的其他疾病或损伤、心肌梗死、中风、冠心病、其他心血管疾病,用于治疗病危糖尿病和非糖尿病患者和多神经病。

[0310] 85a. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的分离多肽,其中所述多肽具有包含下述基序(命名为基序 2)的氨基酸序列:Asp Gly Tyr Xaa₁₇ Asn Xaa₁₈ Arg Xaa₁₉ Xaa₂₀ Xaa₂₁ Phe Xaa₂₂ Xaa₂₃ Xaa₂₄ Gly Xaa₂₅ Xaa₂₆ Xaa₂₇ Xaa₂₈ Xaa₂₉ Xaa₃₀ Xaa₃₁ Xaa₃₂ Xaa₃₃ Gly Xaa₃₄ Xaa₃₅ Xaa₃₆ Gly Xaa₃₇ Phe,其中 Xaa₁₇-Xaa₃₇独立选自天然存在氨基酸,条件是 Xaa₁₇不是 Cys。

[0311] 86a. 段落 85a 的分离多肽,其中该多肽来源于原核生物体。

[0312] 87a. 段落 85a 或 86a 的多肽,其中所述多肽包含 0、1 或 2 个半胱氨酸残基。

[0313] 88a. 段落 85a-87a 中任一项的多肽,其中所述 α -羟基甘氨酸是式 R-Gly(OH) 的 C 端 α -羟基甘氨酸,其中 R 是肽且 Gly(OH) 是与所述肽的 C 端连接的 α -羟基甘氨酸残基,和其中 α -酰胺具有式 R-NH₂。

[0314] 89a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含氨基酸序列 SEQ ID NO: 19

[0315]

(PNSQPNPYRTVEGWFGMPAGRTMGSTSAVFVAPGGHIWVAERCGANTCAGSDVAPL
LEFDTAGKVVSSFGAGMFQFPHGWIIEPDGSIWLTDGQGANGKGHQVFKFSPQGKVLML
TLGKAGVAGDGPDTFNQPNNAVAVSANGDVFISDGHNAGRGNARVLKYSKDGTFIKQW
GGHGSGPGQFEVPHTLAFDSKGRLFVGDRANNRIQIFDQDGKFLDEWKQFGRPSGIFI
DRNDAMYVTDSESTDRDGYGHNPGWKRGIRIGSAKDGSVTAFIPDPSPGAGATSAEAG
VAADAKGNVYGAIEVGPKDVKKYVRK

来源于土细菌属)或由之组成。

[0316] 90a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含氨基酸序列 SEQ ID NO: 20

[0317] (TVSDPAVGAIPEDRLLRPWGRLPATARIGTGITVGPTGEIALLHRAGTAFAYDAVIDTDT
VVVLNPRDGTVRQTWGAGRFRSPHSITADSEGRYWVTDVSTNKITTFDAAGROVGELG
HDYPTGLETALRVNRVLSNLPCTLDEYIFARPTDVAVSADGSIVVADGYRNSRVARFDT
HRVLTGQWGELGDQPAQFNIPHGVALDSNGAVYVADRRNARVQVFNADGSVRHVWH
SSALGRPYDVAIGPDDAVYVLDGGDLLDENNGEQRGYVCRLSTTGRVTHRVALADQR
ANPHQLAIGVRGEIYVAALAGAPLWRWAPO,

来源于海洋放线菌属)或由之组成。

[0318] 91a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含氨基酸序列 SEQ ID NO: 21

[0319] (MLTSGHGMYITAQDDLFVVTYDAHQVLHFDPNQELVKALGTFNEPHWEAPFNHPTDVA
VGAEGDIYVADGYGNAQVHRFASDGSYLNWGTGPGISAGEFSTPHAIWVLPDERVLVV
DRDNDRIQVFDNRNGAVLDEWRGLVRPMDIWADPEGQRIYVTEQAPRITCLDSAGVVIGR
ARTFGIYPHGIWGAPDGSILYVAEQGYPHQIVKYERI,

来源于伯克氏菌属)或由之组成。

[0320] 92a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含氨基酸序列 SEQ ID NO: 22

[0321] (MLPGPSVITDGQYRYHADQHWAKLPAGQQFGPISQLAVDASNRVIVVQRDTPAVLVFS
HDGTLELAWHHPKLTSVHGLCVAPDESLFVVSFDAHQVLKFSRSGELLLELGKFSSPNWI
EPFNHPTDVAVANDGEIYVTDGYGNARVHRFAADGTIYGWGWQHGNKTGEFSCPHGI
WIDEDVGRVLAVDRDNDRVQVFDRSGQYLSEWTGFRRPMDIWGTPNGTFYIVDOAPG
LSIVDRNGRLRARARMYPTYSHGIGGDANGNLFVAAQGSPSRVVRLERLDNEE,

来源于伯克氏菌属)或由之组成。

[0322] 93a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含氨基酸序列 SEQ ID NO: 23

[0323] QLGPSDAYRTTYGWEKMPAGRALGVSSGVFPDRDGKHIWILARCGGNNCAGSDADPIL
KFDMAGNLVTSFGKGVLAFFPHGFFIDAEGNVWVTEGAPVGDRRGDAGFKIGKGHQVFK
FSPEGKLLMTLGAVAGVAGDDNKHFNPGPSGVAIAPNGDIWVVDGHRGGNNRVVRFSKD
GKFIRAIGGGVGSSESADTGRFSDPHGIAIDSAGRIFVADRGNNRIQILDPEGNFLAEWKQ
FGKPSGVYIDARDRIYVGDMSTPERNPGVVPGRVGDAGTKGKVTAFIPDNEKYKQGES
GVEFLAADADGNIYAGEVTRQRFKGHIPLK

或由之组成。

[0324] 94a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含与 SEQ ID NO: 19 具有至少 75% (例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0325] 95a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含与 SEQ ID NO: 20 具有至少 75% (例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0326] 96a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含与 SEQ ID NO: 21 具有至少 75%(例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0327] 97a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含与 SEQ ID NO: 22 具有至少 75%(例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0328] 98a. 一种具有肽基- α -羟基甘氨酸 α -酰胺化裂合酶活性的分离多肽,该多肽包含与 SEQ ID NO: 23 具有至少 75%(例如至少 80%、至少 85%、至少 90%、至少 95%、至少 96% 或至少 98%) 同一性程度的氨基酸序列或由之组成。

[0329] 99a. 段落 89a-98a 中任一项的多肽,其中所述氨基酸序列包含 0、1 或 2 个半胱氨酸残基。

[0330] 100a. 段落 85a-99a 中任一项的多肽,其中该多肽包括融合伴侣。

[0331] 101a. 段落 85a-100a 中任一项的多肽,其中所述多肽能够催化 α -羟基甘氨酸转化为 α -酰胺,其中所述 α -羟基甘氨酸为式 R-Gly(OH) 的 C 端 α -羟基甘氨酸,其中 R 为肽和 Gly(OH) 为与所述肽的 C 端连接的 α -羟基甘氨酸残基,且所述 α -酰胺具有式 R-NH₂。

[0332] 102a. 段落 85a-101a 中任一项所限定的多肽在段落 74a-84a 中任一项所限定的用途或方法中的用途。

[0333] 103a. 一种分离的多核苷酸,其包含编码段落 85a-101a 中任一项的多肽的核苷酸序列。

[0334] 104a. 一种分离的重组核酸构建体,其包含段落 85a-101a 的多核苷酸。

[0335] 105a. 段落 104a 的核酸构建体,其中核苷酸序列与指导多肽在表达宿主中产生的一种或多种控制序列有效连接。

[0336] 106a. 段落 104a 或 105a 的核酸构建体,其包含启动子。

[0337] 107a. 段落 104a-106a 中任一项的核酸,其中该核酸还编码纯化标签和/或接头。

[0338] 108a. 一种重组表达宿主细胞,其包含段落 104a-107a 中任一项的核酸构建体。

[0339] 109a. 段落 108a 的宿主细胞,其中重组核酸存在于宿主细胞的基因组中或者在宿主细胞中自主复制的载体中。

[0340] 110a. 段落 107a 或 109a 的宿主细胞,其中宿主细胞选自:(i) 细菌或真菌,(ii) 酵母,(iii) 哺乳动物宿主细胞,和(iv) 原核宿主细胞。

[0341] 111a. 段落 108a-110a 中任一项的宿主细胞,其中宿主细胞是大肠杆菌。

[0342] 本发明进一步在下述段落中概述:

[0343] 1. 一种能够催化 α -羟基甘氨酸转化为 α -酰胺的酶,其特征在于该酶来源于原核生物体。

[0344] 2. 根据段落 1 的酶,其中该酶包含下述基序:Xaa₁ Val Xaa₂ Asp Arg Xaa₃ Xaa₄ Xaa₅ Arg Xaa₆ Gln Xaa₇ Xaa₈ Xaa₉ Xaa₁₀ Xaa₁₁ Gly Xaa₁₂ Xaa₁₃ Xaa₁₄ Xaa₁₅ Xaa₁₆ Trp;其中 Xaa₁和 Xaa₇可为除 Cys 外的任何天然存在氨基酸,和 Xaa₂、Xaa₃、Xaa₄、Xaa₅、Xaa₆、Xaa₈、Xaa₉、Xaa₁₀、Xaa₁₁、Xaa₁₂、Xaa₁₃、Xaa₁₄、Xaa₁₅和 Xaa₁₆可为任何天然存在氨基酸。

[0345] 3. 根据段落 2 的酶,其中 Xaa₄可为 Asn。

[0346] 4. 根据段落 2 或 3 的酶,其中 Xaa₁₄可为 Leu。

- [0347] 5. 根据段落 2-4 的酶,其中 Xaa₇可为 Val 或 Ile。
- [0348] 6. 根据段落 2-5 的酶,其中 Xaa₈可为 Phe 或 Leu。
- [0349] 7. 根据段落 2-6 的酶,其中 Xaa₉可为 Asp 或 Ser。
- [0350] 8. 根据段落 1-7 的酶,其中该酶分离自赤细菌属、微小杆菌属、食土性细菌或浮霉状菌属种。
- [0351] 9. 根据段落 1-3 的酶,其中该酶能够将 α -羟基苯甲酰甘氨酸转化为苯甲酰胺。
- [0352] 10. 根据段落 1-9 的酶,其中该酶为野生型酶。
- [0353] 11. 根据段落 1-5 的酶,其中该酶相比于野生型酶包含多达 30 个修饰。
- [0354] 12. 根据段落 1-11 的酶,其中该酶包含 1-30 个修饰。
- [0355] 13. 根据段落 1-12 的酶,其中该酶包含 5-25 个修饰。
- [0356] 14. 根据段落 1-13 的酶,其中该酶包含 10-20 个修饰。
- [0357] 15. 根据段落 1-14 的酶,其中该酶包含 12-18 个修饰。
- [0358] 16. 根据段落 1 和 11-15 的酶,其中至少一个 Cys 残基被替换或缺失。
- [0359] 17. 根据段落 1-16 的酶,其中该酶包含融合伴侣。
- [0360] 18. 根据段落 1-17 的酶,其中该酶包含 TAP 标签。
- [0361] 19. 根据段落 1-17 的酶,其中该酶包含选自以下的纯化标签:固定化金属亲和标签(His6 或 His8)、谷胱甘肽转移酶标签、用抗体回收的标签(FLAG 标签、HA 标签、MYC 标签)、生物素和链霉亲和素。
- [0362] 20. 根据段落 17-19 的酶,其中该酶通过接头序列与融合伴侣连接。
- [0363] 21. 根据段落 20 的酶,其中接头序列具有 1-30 个、1-25 个、1-20 个或 1-15 个氨基酸。
- [0364] 22. 根据段落 1-21 中任一项的酶,其中 Xaa₄可为 Asn、Xaa₁₄可为 Leu、Xaa₇可为 Val 或 Ile、Xaa₈可为 Phe 或 Leu 和 Xaa₉可为 Asp 或 Ser。
- [0365] 23. 一种包含 SEQ ID NO: 1 的酶,该酶能够催化 α -羟基甘氨酸转化为 α -酰胺。
- [0366] 24. 一种包含 SEQ ID NO: 2 的酶,该酶能够催化 α -羟基甘氨酸转化为 α -酰胺。
- [0367] 25. 一种包含 SEQ ID NO: 3 的酶,该酶能够催化 α -羟基甘氨酸转化为 α -酰胺。
- [0368] 26. 一种包含 SEQ ID NO: 4 的酶,该酶能够催化 α -羟基甘氨酸转化为 α -酰胺。
- [0369] 27. 一种包含 SEQ ID NO: 13 的酶,该酶能够催化 α -羟基甘氨酸转化为 α -酰胺。
- [0370] 28. 一种包含 SEQ ID NO: 15 的酶,该酶能够催化 α -羟基甘氨酸转化为 α -酰胺。
- [0371] 29. 根据段落 23-27 的酶,其中该酶包含 1-30 个修饰。
- [0372] 30. 根据段落 23-28 的酶,其中该酶包含 5-25 个修饰。
- [0373] 31. 根据段落 23-29 的酶,其中该酶包含 10-20 个修饰。
- [0374] 32. 根据段落 23-30 的酶,其中该酶包含 12-18 个修饰。
- [0375] 33. 根据段落 23-31 的酶,其中至少一个 Cys 残基被置换或缺失。
- [0376] 34. 根据段落 1-33 中任一项的酶用于催化肽中 C 端 α -羟基甘氨酸转化为 α -酰胺的用途。
- [0377] 35. 根据段落 34 的用途,其为在用于制备 α -酰胺化肽的方法中的用途。

- [0378] 36. 根据段落 34-35 的用途,其中存在 Cu^{2+} 离子和 / 或抗坏血酸。
- [0379] 37. 根据段落 34-36 的用途,其中具有 PHM 活性的酶在用于制备 α -酰胺化肽的方法中使用。
- [0380] 38. 根据段落 34-37 的用途,其中该肽为选自以下的靶肽:胰淀粉样肽、NPY、PYY-3-36、PP、GLP-1、胃泌素、降钙素、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、POMC、 α MSH、 γ 1MSH 和 HP-N,或者它们的功能类似物。
- [0381] 39. 根据段落 34-38 的用途,其中 α -酰胺化肽用于治疗或预防下述病况之一:肥胖症、高血压、高血糖症(包括应激性高血糖症)、2 型糖尿病、葡萄糖耐量减低、1 型糖尿病、烧伤、手术创伤、在治疗中需要合成代谢作用的其他疾病或损伤、心肌梗死、中风、冠心病、其他心血管疾病,用于治疗病危糖尿病和非糖尿病患者和多神经病。
- [0382] 40. 一种用于产生 α -酰胺化肽的方法,其包括使靶肽与段落 1-33 的酶反应和纯化 α -酰胺化肽。
- [0383] 41. 根据段落 40 的方法,其中具有 PHM 活性的酶在该方法中使用。
- [0384] 42. 根据段落 40-41 的方法,其中靶肽选自:胰淀粉样肽、NPY、PYY-3-36、PP、GLP-1、胃泌素、降钙素、胃泌素释放肽、加压素、催产素、神经激肽 A、胰泌素、POMC、 α MSH、 γ 1MSH 和 HP-N,或者它们的功能类似物。
- [0385] 43. 一种分离核酸,其编码段落 1-33 中任一项的酶。
- [0386] 44. 一种重组核酸,其包含:
- [0387] a. 启动子
- [0388] b. 段落 43 的分离核酸。
- [0389] 45. 根据段落 44 的重组核酸,其中该核酸还编码纯化标签。
- [0390] 46. 根据段落 44-45 的重组核酸,其中该核酸还编码接头。
- [0391] 47. 一种载体,其包含段落 44-46 的重组核酸。
- [0392] 48. 一种宿主细胞,其包含段落 44-46 的重组核酸。
- [0393] 49. 段落 48 的宿主细胞,其中重组核酸存在于宿主细胞的基因组中或者在宿主细胞中自主复制的载体中。
- [0394] 50. 根据段落 48-49 的宿主细胞,其中该宿主细胞是细菌、真菌(例如酵母)。
- [0395] 51. 根据段落 48-50 的宿主细胞,其中该宿主细胞是大肠杆菌。
- [0396] 52. 一种用于产生能够催化 α -羟基甘氨酸转化为 α -酰胺的酶的方法,其包括在适合于产生所述酶的条件下维持段落 47 的宿主细胞。

实施例

[0397] 通过使用数据库和下文的表征鉴定的本发明所有实施例的酶含有信号肽,其使用 Signal P (Henrik Nielsen 等, "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites (原核生物和真核生物信号肽的鉴定及其切割位点的预测)" Protein Engineering 10, 1-6 (1997)) 以高可信度预测。因为信号肽通常不是成熟功能性酶的一部分,所以根据预测从初始序列中移除这些信号肽。用 GlyPro N 端延伸设计下文所述的实施例 1-6 中的酶,以促进识别序列 GluValLeuPheGln / GlyPro 的 HRV14 3C 蛋白酶对 TAP 融合伴侣的移除。

[0398] 实施例 1

[0399] 克隆具有作为纯化标签的核糖体蛋白 L9 和来自赤细菌属 SD-21 的 PAL 样结构域的融合蛋白

[0400] 克隆、表达并评价赤细菌属 PAL 样结构域的四种序列变体。制备变体来评价不同的融合伴侣、接头或 PAL 样结构域的 N 端延伸是否影响表达和溶解度。

[0401] 蛋白 1：

[0402] SEQ ID NO: 13 编码来源于赤细菌属 SD-21 细菌的 PAL 样结构域。该蛋白的分子量计算为 33467,1 Da。当与热稳定碱性蛋白 (TAP) 纯化标签 (SEQ ID NO: 7) 通过富 Gly/Ser 的柔性接头 (SEQ ID NO: 10) 融合时,融合蛋白将具有 51231.7 Da 的分子量并代表如表 1 中所述的蛋白 1。

[0403] 为了获得编码蛋白 1 的质粒 A,对编码 SEQ ID NO: 13 的 XhoI/BamHI 基因片段进行密码子优化以便在大肠杆菌中表达,并作为具有 5' 端 XhoI 和 3' 端 BamHI 克隆位点的合成基因片段获得 (GeneScript)。在 5' 端,将编码接头 (SEQ ID NO: 10) 伴侣的短片段直接导入编码 PAL 样结构域 (SEQ ID NO: 13) 的核苷酸片段上游。

[0404] 通过制造商所述的方法,使用 LigaFast™ Rapid Ligation System (Promega) 将 XhoI/BamHI 片段连接到已编码包含在 NdeI/XhoI 片段中的纯化标签 (核糖体蛋白 L9 来自海栖热袍菌 (*T. maritima*),在以编号 WO 2006/108826 和 WO 2008/043847 公布的国际专利申请中所述,(SEQ ID NO: 7)) 的 pET11a (Novagen) 表达载体中。将连接产物用于转化 TOP10 (Invitrogen) 感受态大肠杆菌细胞,并在 LB (Luria-Bertani) 培养基上与 150 μg/ml 氨苄青霉素琼脂板孵育过夜。在液体培养基中质粒增殖和标准小量制备质粒后,从培养物阳性克隆中获得编码目标融合蛋白的质粒 A。通过使用 T7 启动子/终止子序列特异性引物的 DNA 测序,证实质粒 A 的正确核苷酸序列。

[0405] 蛋白 2：

[0406] 通过将约 110 bp 在接头区和赤细菌属 PAL 样结构域的 N 端部分中含有微小改变的 XhoI/NcoI 合成片段 (Geneart) 连接到质粒 A 的 XhoI/NcoI 位点中,获得赤细菌属 SD-21 PAL 样结构域的另一变体 (质粒 B),从而产生蛋白 2。蛋白 2 包含与对蛋白 1 所述相同的纯化标签 SEQ ID NO: 7,但具有拥有 HRV14 3C 蛋白酶切割位点 (ELTFQ) 的接头 (SEQ ID NO: 14),并在 PAL 样结构域 (SEQ ID NO: 15) 的 N 端中具有改变。该融合蛋白的理论分子量为 52092.6 Da,而 HRV14 3C 蛋白酶释放的 PAL 样结构域为 33823,5 Da。

[0407] 蛋白 3：

[0408] 通过将在接头区和赤细菌属 PAL 样结构域的 N 端部分中含有微小改变的 XhoI/NcoI 合成片段 (Geneart) 连接到质粒 B 的 XhoI/NcoI 位点中,获得赤细菌属 PAL 样结构域的另一变体 (质粒 C,在图 1 的载体图谱中例示),从而产生蛋白 3。蛋白 3 包含 SEQ ID NO: 7 纯化标签,但具有稍修饰的接头 (SEQ ID NO: 8),使得相比于蛋白 2 能够用 HRV14 3C 蛋白酶切割,且相比于蛋白 2 具有截短了 4 个氨基酸的 PAL 样赤细菌属结构域 (SEQ ID NO: 1),相比于蛋白 1 截短了 1 个氨基酸并包含 N 端 Gly 残基 (SEQ ID NO: 1)。该融合蛋白的理论分子量为 51591.1 Da,而 HRV14 3C 蛋白酶释放的 PAL 样结构域为 33324 Da。

[0409] 克隆具有 His6 纯化标签和来自赤细菌属 SD-21 的 PAL 样结构域的融合蛋白

[0410] 蛋白 4

[0411] 通过自载体 C 切除 XhoI/BamHI 片段,并将该片段连接到包含编码连续的 SEQ ID NO: 9 和 10 序列的融合伴侣的载体中以获得质粒 D,获得赤细菌属 SD-21 PAL 样结构域 SEQ ID NO: 1 的另一变体。编码的融合蛋白包含 N 端组氨酸 6 标签 (SEQ ID NO: 9) 和赤细菌属 PAL 样结构域 SEQ ID NO: 1 之前的具有 HRV14 3C 蛋白酶切割位点的接头 (SEQ ID NO: 12),并具有 35645.4 Da 的理论分子量。

[0412] 克隆真核生物 PAL 结构域用于比较分析

[0413] 为了比较本发明中公开的新的细菌 PAL 样结构域和已充分描述的 PAL 结构域,基本上如¹对蛋白 1 所述地,使用相同的融合伴侣 (SEQ ID NO: 7) 克隆褐家鼠 (编码蛋白 5 的质粒 E 和编码蛋白 6 的质粒 F 包含接头中的差异) 和非洲爪蟾 (编码蛋白 7 的质粒 G) PAL 结构域,以评价之前所述的包含两个二硫桥的典型真核生物 PAL 的表达概况 (例如 Stoffers, D. A 等 :Proc. Natl. Acad. Sci. U. S. A. 86:735-739 (1989) 和 Mizuno K 等 : Biochem. Biophys. Res. Commun. 148:546-552 (1987))。

[0414] 制备下述构建体 (表 1) :

[0415]

蛋白名	质粒号	物种	融合伴侣	接头	PAL 样序列
1	A	赤细菌属	SEQ ID NO: 7	SEQ ID NO: 10	SEQ ID NO: 13
2	B	赤细菌属	SEQ ID NO: 7	SEQ ID NO: 14	SEQ ID NO: 15
3	C	赤细菌属	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 1
4	D	赤细菌属	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 1
5	E	褐家鼠	SEQ ID NO: 7	SEQ ID NO: 9	SEQ ID NO: 5
6	F	褐家鼠	SEQ ID NO: 7	SEQ ID NO: 16	SEQ ID NO: 5
7	G	非洲爪蟾	SEQ ID NO: 7	SEQ ID NO: 9	SEQ ID NO: 6

[0416] 下文描述核苷酸序列 :

[0417] 质粒号 :A (赤细菌属 1) :SEQ ID NO: 26

[0418]

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACCTATCTGATTCGCGTGGCTTTGCGAAAAGATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAATCCTGAAAAGAACTGAAAAAACGTACCCATGTGGTGAAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACCACCG
GCCTGAAACTGGATAAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGTGGCAGCGGCGGTAGCGGTGAAGCCCCGCGGTTACCATCGATGAAAGCTGGC
CGGATATCCGGAAAGCGCCGTGTTTGGCGAACCGACCGCCATTGATGTGGATAGCCATGGCCA
TATCTTTGTGCTGCATCGTGCCGTCGCGAATGGACCCAGCCGTTTCCGAGCGATCCGATTAGCG
AACCGACCGTGTTTCATGTTCCGCGCGAACGGCAAACTGCTGAGCAAATGGGGCGCGGGCGAACT
GGTTATGCCCGATGGTCTGAGCATTGATGGTGATAACAAAGTGTGGATTACCGATGTGGCGCGT
GAACAGGTTCTGCGTTTTACCCATGAAGGTGCGGAAGAACTGGTTCTGGGCACCCGTGGCGAAA
CCGGCCAGGATGAAAGCCATTTCCGGCCGTCCGGCGGATGTTACCTTTGTGGGTGATCGTGTGCT
GGTGCCGATGCTATCTGAACCGTCTGATTATGGTGTGTTGATCGCGCGGGCAACTTCTTGAA
CAGTGGGGTAAAGAAAGGTGAAGATGCGGGCGAATTTAATCTGCCCGATGCGATCGCGGCCGATA
GCGAACGTATTTATGTGGCGGATCGTGAAAACGCGCGTGTTTCAAGTGCTGAGCCTGGATGGTGA
ACCGCTGGCCCGCTGGCGCCAGGATGGCACCGGCCATCCGTATGCCGTGAAACCGATTGGCAG
CGGCTATGTTCTGGCGATTGAAGGCCGCGATCGCGCGGGTTCGCAATACCGCCATTGGCCGCATT
TATCGTGCGGATGGTGGCCTGGAACGTGTGTTTATGTCGGGCGTTGAACCGCATACCGGCACCA
GCCTGGGCCATGATGTGGCGATTGGTCCGATGGTAGCGCGTATATGGTCGACAACAAAGCGAA
TCGTGTTATTAATTTGATCTGAGCCGCGCGCGCGTGAAGAAGCGGATGCGGAT

[0419] 质粒号 :B (赤细菌属 2) :SEQ ID NO: 27

[0420]

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACCTATCTGATTCGCGTGGCTTTGCGAAAAGATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAATCCTGAAAAGAACTGAAAAAACGTACCCATGTGGTGAAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACCACCG
GCCTGAAACTGGATAAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGCGGCGAGCGGTAGCGAAACCTGTTTCAGGCGCGTGAAAGAAGCGCGCGCGGTTA
CCATTGATGAAAGCTGGCCGGATATTCGGAAAGCGCGGTGTTTGGCGAACCGACCGCGATTGA
TGTGGATAGCCATGGCCATATTTTGTGCTGCATCGCGCGGGCCGCGAATGGACCCAGCCGTTT
CCGAGCGATCCGATTAGCGAACCGACCGTGTTTATGTTTGGCGGCAACGGCAAACTGCTGAGCA
AATGGGGCGCGGGCGAACTGGTGATGCCCGATGGCCTGAGCATTGATGGCGATAACAAAGTGT
GGATTACCGATGTGGCGCGCGAACAGGTGCTGCGCTTACCCATGAAGGCGCGGAAGAACTGGT
GCTGGGCACCCGCGGCGAAACCGGCCAGGATGAAAGCCATTTTGGCCGCCCGGCGGATGTGAC
CTTTGTGGGCGATCGCGTGCTGGTGGCGGATGGCTATCTGAACCGCCGATTATGGTGTGTTGAT
CGCGCGGGCAACTTTCTGGAACAGTGGGGCAAAGAAGGCGAAGATGCGGGCGGAATTTAACCTG
CCGCATGCGATTGCGGCGGATAGCGAACGCATTTATGTGGCGGATCGCGAAAACGCGCGCGTG
CAGGTGCTGAGCCTGGATGGCGAACCGCTGGCGCGCTGGCGCCAGGATGGCACCGGCCATCCG
TATGCGGTGAAACCGATTGGCAGCGGCTATGTGCTGGCGATTGAAGGCCGCGATCGCGCGGGC
CGCAACACCGCGATTGGCCGATTTATCGCGCGGATGGCGGCGCTGGAACGCGTGTGTTGATGCGG
GCGTGGAACCGCATACCGGCACCGAGCCTGGGCCATGATGTGGCGATTGGCCCGGATGGCAGCG
CGTATATGGTCGACAACAAAGCGAACCGCGTGATTAAATTTGATCTGAGCCGCGCGGGCGTGGA
AGAAGCGGATGCGGAT

[0421] 质粒号 :C (赤细菌属 3) :SEQ ID NO: 28

[0422]

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACTATCTGATTCCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAATCCTGAAAGAACTGAAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACCACCG
GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAATTTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGGCGGCAGCGGTAGCGAAGTGCTGTTTCAGGGCCCCGCGGTTACCATTGATGAAA
GCTGGCCGGATATTCGGAAAGCGCGGTGTTTGGCGAACCGACCGCGATTGATGTGGATAGCCA
TGCCCATATCTTTGTGCTGCATCGTGCCGGTCGCGAATGGACCCAGCCGTTTCCGAGCGATCCG
ATTAGCGAACCGACCGTGTTCATGTTCCGCGCGAACGGCAAACCTGCTGAGCAAATGGGGCGCGG
GCGAGTGGTTATGCCGATGGTCTGAGCATTGATGGTGATAACAAAGTGTTGATTACCGATGTG
GCGCGTGAACAGGTTCTGCGTTTTACCCATGAAGGTGCGGAAGAACTGGTTCTGGGCACCCGTTG
GCGAAACCGGCCAGGATGAAAGCCATTTTCGGCCGTCGCGCGGATGTTACCTTTGTGGGTGATCG
TGTGCTGGTGGCCGATGGCTATCTGAACCGTCGTATTATGGTGTGTTGATCGCGCGGGCAACTTCC
TGGAACAGTGGGGTAAAGAAGGTGAAGATGCGGGCGAATTTAATCTGCCGCATGCGATCGCGGC
CGATAGCGAACGTATTTATGTGGCGGATCGTGAAAACGCGCGTGTTCAGGTGCTGAGCCTGGAT
GGTGAACCGCTGGCCCCGCTGGCGCCAGGATGGCACCGGCCATCCGTATGCCGTGAAACCGATT
GGCAGCGGCTATGTTCTGGCGATTGAAGGCCGCGATCGCGCGGGTCGCAATACCGCCATTGGCC
GCATTTATCGTGCGGATGGTGGCCTGGAACGTGTGTTTGATGCGGGCGTTGAACCGCATACCGG
CACCAGCCTGGGCCATGATGTGGCGATTGGTCCGGATGGTAGCGCGTATATGGTCGACAACAAA
GCGAATCGTGTTATTAAATTTGATCTGAGCCGCGCCGGCGTTGAAGAAGCGGATGCGGAT

[0423] 质粒号:D (赤细菌属 4):SEQ ID NO: 29

[0424]

ATGGGCAGCAGCCATCATCATCATCACTCGAGCGGCGGCAGCGGTAGCGAAGTGCTGTTTC
AGGGCCCCGCGGTTACCATTGATGAAAGCTGGCCGGATATTCGGAAAGCGCGGTGTTTGGCGA
ACCGACCGCGATTGATGTGGATAGCCATGGCCATATCTTTGTGCTGCATCGTGCCGGTCGCGAAT
GGACCCAGCCGTTTCCGAGCGATCCGATTAGCGAACCGACCGTGTTCATGTTCCGCGCGAACGG
CAAACCTGCTGAGCAAATGGGGCGCGGGCGAACTGGTTATGCCGCATGGTCTGAGCATTGATGGT
GATAACAAAGTGTTGATTACCGATGTGGCGCGTGAACAGGTTCTGCGTTTTACCCATGAAGGTGC
GGAAGAACTGGTTCTGGGCACCCGTGGCGAAACCGGCCAGGATGAAAGCCATTTCCGGCCGTCC
GGCGGATGTTACCTTTGTGGGTGATCGTGTGCTGGTGGCCGATGGCTATCTGAACCGTCGTATTA
TGGTGTGTTGATCGCGCGGGCAACTTCTGGAACAGTGGGGTAAAGAAGGTGAAGATGCGGGCG
AATTTAATCTGCCGCATGCGATCGCGGCCGATAGCGAACGTATTTATGTGGCGGATCGTGAAAAC
GCGCGTGTTCAGGTGCTGAGCCTGGATGGTGAACCGCTGGCCCCGCTGGCGCCAGGATGGCACC
GGCCATCCGTATGCCGTGAAACCGATTGGCAGCGGCTATGTTCTGGCGATTGAAGGCCGCGATC
GCGCGGGTCGCAATACCGCCATTGGCCGCAATTATCGTGCGGATGGTGGCCTGGAACGTGTGTT
TGATGCGGGCGTTGAACCGCATACCGGCACCAGCCTGGGCCATGATGTGGCGATTGGTCCGGAT
GGTAGCGCGTATATGGTCGACAACAAAGCGAATCGTGTTATTAAATTTGATCTGAGCCGCGCCGG
CGTTGAAGAAGCGGATGCGGAT

[0425] 质粒号:E (大鼠 1):SEQ ID NO: 30

[0426]

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGAGCGACGGCTACGCGCGTAACTATCTGATTCCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCTGGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAAAGCGAAAAATCCTGAAAGAACTGAAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAGGCGGCAAAATTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACACCGGCCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGCTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAATTCGCGTGGAACGCGAAGAAAAGCTCGAGCGGGAGCGGCGAAGTTCTGTTTCAGGGCGATTTTCATGTGGAAGAGGAACTGGATTGGCCGGGTGTTTATCTGCTGCCGGGGCAGGTGAGCGGCGTTGCCCTGGATAGTAAAAACAACCTGTGATTTTCCACCGTGAGATCATGTGTGGGATGGTAACCTCATTCGATAGCAAAATTTGTCTATCAGCAGCGGCCTGGGTCTATTGAAGAAGATACCATCCTGGTTATCGATCCGAATAATGCGGAAATTCTGCAAAGTTCTGGAAAAACCTGTTCTACCTGCCTCATGGGCTGAGCATCGATACCGATGGTAACCTAGTGGTTACCGATGTAGCCCTGCATCAGGTCTTCAAACCTGGACCCACATTCAAAGAAGGTCCGCTGCTGATTCTGGGACGTTCTATGCAGCCTGGCAGCGATCAGAACCATTTTTGCCAACCAACAGATGTCCCGTGGAGCCGTGAGCCTGCTGCTGTTGTGTCGGACGGCTACTGTAACCTCCGTATCGTCCAGTTTAGCCCGTCAGGTAAATTTGTTACACAGTGGGGCGAAGAGTCGTCTGGCTCTAGTCCGCGCCCGGGCCAGTTTTCGGTGCCGCATAGCCTGGCCCTGGTGCCGCATCTGGACCAGCTGTGTGTGGCCGACCGTGAGAACGGCCGCATTCAATGCTTTAAACCTGATACAAAGGAATTTGTTGCGCAAATTAACACGCCAGCTTTGGACGTAATGTTTTGCGGATTTATACATTCCAGGCTTTCTGTTTGCAGTCAACGGCAAACCATATTTGGCGATCAGGAACCGGTACAGGGTTTTGTTATGAATTTTCATCGGGCGAAATTATTGATGTTTTTAAACCTGTGCGCAAACATTTGATATGCCACATGATATTGTGGCTAGCGAGGACGGTACGTTTTATTTGGGGATGCGCACACGAACACCGTGTGGAAATTCACCTGACCGAAAAAATGGAACACCGTTCGGTC

[0427] 质粒号:F (大鼠2):SEQ ID NO: 31

[0428]

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGAGCGACGGCTACGCGCGTAACTATCTGATTCCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCTGGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAAAGCGAAAAATCCTGAAAGAACTGAAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAGGCGGCAAAATTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACACCGGCCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGCTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAATTCGCGTGGAACGCGAAGAAAAGCTCGAGCGGGGTAGCGGCGGCTCTGGCGATTTTCATGTGGAAGAGGAACTGGATTGGCCGGGTGTTTATCTGCTGCCGGGGCAGGTGAGCGGCGTTGCCCTGGATAGTAAAAACAACCTGGTGATTTTCCACCGTGAGATCATGTGTGGGATGGTAACCTCATTCGATAGCAAAATTTGTCTATCAGCAGCGCGGCTGGGTCTATTGAAGAAGATACCATCCTGGTTATCGATCCGAATAATGCGGAAATTTCTGCAAAGTTCTGGAAAAACCTGTTCTACCTGCCTCATGGGCTGAGCATCGATACCGATGGTAACCTGTTGGTTACCGATGTAGCCCTGCATCAGGTCTTCAAACCTGGACCCACATTCAAAGAAGGTCCGCTGCTGATTCTGGGACGTTCTATGCAGCCTGGCAGCGATCAGAACCATTTTTGCCAACCAACAGATGTCGCCGTGGAGCCGTGACTGGTGCTGTGTTTGTGTCGGACGGCTACTGTAACCTCCCGTATCGTCCAGTTTAGCCCGTCAGGTAAATTTGTTACACAGTGGGGCGAAGAGTCGTCTGGCTCTAGTCCGCGCCCGGGCCAGTTTTCGGTGCCGCATAGCCTGGCCCTGGTGCCGCATCTGGACCAGCTGTGTGTGGCCGACCGTGAGAACGGCCGCATTCAATGCTTTAAACCTGATACAAAGGAATTTGTTGCGGAAATTAACACGCCAGCTTTGGACGTAATGTTTTGCGGATTTATACATTCCAGGCTTTCTGTTTGCAGTCAACGGCAAACCATATTTGGCGATCAGGAACCGGTACAGGGTTTTGTTATGAATTTTCATCGGGCGAAATTATTGATGTTTTTAAACCTGTGCGCAAACATTTGATATGCCACATGATATTGTGGCTAGCGAGGACGGTACGTTTTATTTGGGGATGCGCACACGAACACCGTGTGGAAATTCACCTGACCGAAAAAATGGAACACCGTTCGGTC

[0429] 质粒号:G (非洲爪蟾属(Xenopus)):SEQ ID NO: 32

[0430]

```
ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACCTATCTGATTCCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAATCCTGAAAGAACTGAAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACCACCG
GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGGAGCGGCGAAGTTCTGTTTCAGGGCGATGTGCATCTGGAAGAAGATACCGACTG
GCCGGGTGTGAACCTTAAAGTGGGCCAGGTGAGCGGGCTAGCCCTGGACCCGAAAAACAACCT
GGCCATTTTTCATCGTGCGGATCATGTGTGGGATGAAAACAGCTTTGATCGCAACTTTGTGTATC
AGCAGCGTGGCATTGGCCCGATTGAGGAAAGCACCATTCTGGTTGTGATCCGAGCAGCAGCAA
AGTGCTGAAAAGCACCGGCAAAAAACCTGTTTTTCTGCCGCATGGCCTGACCATTGATCSTGATG
GCAACTATTGGGTGACCGATGTGGCGCTGCATCAGGTGTTTAACTGGGTGCGGGCAAGAAAC
CCCGATGCTGGTGTGGTCTGCGTTTCAGCCGGGCAGCGATCGTAAACATTTTGGCAGCCG
ACCGATGTTGCGGTTGATCCGATTACCGGCAACTTTTTTGTGGCGGATGGCTATTGCAACAGCCG
TATTATGCAGTTTAGCCCGAACGGCATGTTTATTATGCAGTGGGGCGAAGAAACCAGCAGCAACG
TGCCGCGTCCGGGCCAGTTTCGATTCCGCACAGCCTGACCATGGTGCCGGATCAGGGCCAGCT
GTGCGTGGCGGATCGTGAAACGGCCGATTTCAGTGCTTTCATGCGGAAACCGGCAATTTTGTG
AAACAAATCAAACATCAGGAATTTGGCCGCGAAGTTTTCGCGTGAGCTATGCGCCGGGTGGTG
TGCTGTATGCGGTGAACGGCAACCGTATTATGGCTATAGCGCACCGGTGCAGGGCTTTATGCT
GAACTTTAGCAACGGCGATATTCTGGATACCTTTATTCCGGCGCGTAAAAACTTTGATATGCCGCA
TGATATTGCGGCAGCGGATGATGGCACCGTGTATGTGGGCGATGCGCATGCGAACGCGGTGTG
GAAATTTAGCCCGAGCAAAGCGGAACATCGGTCCGTG
```

[0431] 构建体的表达:

[0432] 将具有正确 DNA 序列的质粒转化到大肠杆菌表达菌株 BL21 (DE3) 中, 将其接种在 LB/ 氨苄青霉素平板上过夜。在 LB 培养基中增殖具有质粒的 BL21 (DE3) 宿主细胞, 使用 T7/ IPTG 表达系统诱导蛋白表达。

[0433] 将用编码融合构建体的质粒转化的大肠杆菌 BL21 (DE3) 细胞在 37℃ 在 LB/ 氨苄青霉素培养基中使用摇瓶培养至光密度 (OD_{600nm}) 为约 0.4 -0.6。如果测试的诱导温度为 18℃ 或 30℃, 那么相应地降低温度至 18℃ 或 30℃, 持续大约 30 分钟, 并向培养物中加入 0.5 mM IPTG, 持续 3 小时。蛋白诱导后, 沉淀培养物并在 25 mM NaPO₄ pH 7 缓冲液中通过超声处理裂解细胞。在含有诱导和未诱导的细胞以及诱导的细胞的可溶性和不溶性级分的样品缓冲液 (通过在由 25 mM NaPO₄ pH 7 组成的缓冲液中超声处理和离心而获得) 中进行裂解物样品的 SDS-PAGE 分析。进行该分析以评价变体的表达水平和溶解度。

[0434] 赤细菌属变体的表达概况以及与已充分描述的来自褐家鼠和非洲爪蟾的真核生物 PAL 结构域的比较

[0435] 当在 1L 带挡板的摇瓶中于 30℃ 自 BL21 (DE3) 中的质粒表达 3 小时时, 蛋白 3、蛋白 1 或蛋白 2 赤细菌属蛋白变体显示预期的约 50kDa 的分子量和非常类似的表达水平。在中性 pH 下观察到一致的高程度的溶解度 (约 80%) (在包含 25 mM NaPO₄ pH 7 的缓冲液中, 在图 2A 中对蛋白 2 例示)。

[0436] 使用与蛋白 1 相同的 N 端标签和接头, 对来自大鼠的 PAL (蛋白 6) 或非洲爪蟾 PAL (蛋白 7), 观察到相同的表达水平。然而, 当使用相同的表达条件时, 与赤细菌属 PAL 样结构域相比, 来自大鼠和非洲爪蟾的 PAL 结构域一致地产生不溶性蛋白 (在图 2B 中对蛋白 6 例示)。具有不同接头的大鼠 PAL (蛋白 5) 不改变表达概况, 但使用相同的表达条件仍产生不溶性蛋白。在 18℃、30℃ 或 37℃ 的表达条件下, 表达概况的比较 (蛋白 3 相对蛋白 6) 显示, 赤细菌属蛋白在所有 3 种温度下均为可溶性蛋白, 而大鼠 PAL 在所有 3 种温度下为不溶的。这显示表达原核生物 PAL 的优势, 其相比于真核生物 PAL 不包含二硫桥, 真核生物 PAL

在表达期间更可能形成聚集体。尽管在 PAL 蛋白 N 端的 5aa 区和接头区中的变化, 赤细菌属 PAL 样结构域还是可被制备为可溶性融合蛋白。

[0437] 当在 30℃ 和 37℃ 下表达时, 赤细菌属蛋白的加上 His 标签的变体 (蛋白 4) 均产生可溶性蛋白, 因此证明, 赤细菌属蛋白的高溶解度不取决于所使用的 N 端融合标签的性质, 而取决于蛋白的内在性质。

[0438] 在 SP Sepharose FF 上纯化加上 SEQ ID NO: 7 标签的赤细菌属 PAL 样融合蛋白

[0439] PAL 样酶 (蛋白 1、蛋白 2 或蛋白 3) 或大鼠 PAL (蛋白 6) 在 1-2L 带挡板的摇瓶中于 30℃ 如上文所述表达 3 小时后, 如下所述进行通过使用 SP Sepharose FF 5ml 柱的阳离子交换色谱进行的纯化:

[0440] 通过在总计 20ml 的 25mM NaPO₄ pH 7 缓冲液中超声处理, 破碎酶变体的细胞培养物沉淀 (来自 80ml 的培养物 OD₆₀₀: 约 1.6-1.8)。通过离心 (4000 rpm, 15 分钟) 旋下细胞碎片。将上清液无菌过滤 (0.45 μm 过滤器), 并用 25mM NaPO₄ pH 7 缓冲液稀释至 40ml 的总体积来获得蛋白装载物。使用 AKTA explorer 100 纯化系统 (GE Healthcare) 进行纯化。使用下述缓冲液, 以 3ml/min 的流速将 5ml 柱体积的预装填的 SP Sepharose FF HiTrap 柱 (GE Healthcare, 产品号: 17-5157-01) 用于分离:

[0441] 缓冲液 A: 50mM 磷酸钠, pH 7

[0442] 缓冲液 B: 50mM 磷酸钠, pH 7 + 1M NaCl

[0443] 首先用 7 个柱体积的缓冲液 A 平衡柱。在装载装载物后, 通过使用 7 个柱体积的缓冲液 A 洗涤, 移除未结合的蛋白。20 个柱体积的线性梯度 0-100 % 缓冲液 B 用于从柱上洗脱酶。

[0444] 通过 SDS-PAGE 分离梯度内洗脱的装载物、流通级分和代表蛋白的级分, 并在考马斯亮蓝染色或 LC-MS 分析后分析凝胶。

[0445] 以约 0.3 M NaCl 的盐浓度以单峰从柱上洗脱的蛋白 1、蛋白 2 或蛋白 3, 其根据 SDS-PAGE 具有约 80% 的高纯度 (图 3A 和 B 对蛋白 3 例示)。在流通级分中未观察到显著量的蛋白, 表明赤细菌属变体都被 SP Sepharose FF 柱有效捕获。

[0446] 与此相比, 试图使用完全相同的用于样品制备和纯化的条件纯化大鼠 (蛋白 6) 或非爪蟾 PAL (蛋白 7) 融合蛋白失败, 表明真核生物 PAL 在大肠杆菌中表达时采用不正确的折叠构象, 其使得不能容易地通过阳离子色谱纯化。

[0447] 根据制造商的说明书, 使用 Vivaspin MWC010.000 柱 (Vivaspin), 将含有洗脱的赤细菌属 PAL 样酶的级分合并在一起, 并在包含 50mM Tris pH 7.5 的缓冲液中浓缩和脱盐。加入甘油至总浓度为 10% 并在 20℃ 保存直至使用。

[0448] 纯化和 LC-MS 分析无纯化标签的成熟赤细菌属 PAL 样 (SEQ ID NO: 1)

[0449] 用加上 TAP 标签的 HRV14 3C 蛋白酶切割

[0450] 为了从蛋白 3 中移除纯化标签, 使用加上 TAP 标签的 HRV14 3C 蛋白酶 (根据 WO 2008/043847) 在 3ml 的反应体积中以 1:25 的酶 / 底物比 (根据制造商的说明书, 使用 NanoDrop2000, Thermo Scientific 测量蛋白浓度后, 估计比值), 30℃ 过夜切割存在于 50 mM Tris HCl pH 7.5, 10% 甘油缓冲液, 1 mM TCEP 中的约 1.5 mg 融合蛋白。考马斯染色的 SDS-PAGE 凝胶和 LC-MS 分析用于证实酶切割的发生, 使用下述方案: 使用制造商推荐的 MS 设置, 用 LC-MSD_TOF (Agilent technologies) 仪器来分析样品, 并在 0.3 ml/min 流速和

40℃柱温的标准 HPLC 条件下,使用分析 Poroshell 300SB-C8, Micro Bore 1.0 x 75 mm, 5 微米 (Agilent Technologies) 柱:在 20 分钟内形成梯度洗脱。使用 0.1% 甲酸水 (缓冲液 A) 和乙腈 (缓冲液 B) 中的 8.8 mM 甲酸铵如下运行:

[0451]

时间 (分钟)	% 缓冲液 B
0	22
3	22
15	75
15.1	90
20	90

[0452] HRV14 3C 蛋白酶过夜消化蛋白 3 的 LC-MS 分析,得到两条测定为 33323 Da 和 18284.78 Da 的片段,其分别对应于释放的 PAL 样结构域 (SEQ ID NO: 1, 计算的质量为 33323, 97 Da) 以及释放的纯化标签和 HRV14 3C 接头 (SEQ ID NO: 7+ SEQ ID NO: 8, 计算的质量为 18285, 18 Da)。通过 SDS-PAGE,在未切割的对照或消化液中均未观察到可见的降解产物,表明成熟 PAL 样结构域和融合蛋白都高度稳定。

[0453] 在 *Q-sepharose HP* 上纯化成熟结构域

[0454] 用 25 mM NaPO₄ pH 7 以 1:3 稀释包含释放的成熟 PAL 结构域的消化液,并将样品装载物装载在 1 ml HiTrap Q Sepharose 高效 (HP) 阴离子交换柱 (GE Healthcare 17-1153-01) 上,并使用 AKTA Explorer 100 系统 (GE Healthcare) 用下述纯化缓冲液分离:

[0455] 缓冲液 A :50 mM 磷酸钠, pH 7

[0456] 缓冲液 B :50 mM 磷酸钠, pH 7 + 1M NaCl

[0457] 首先用 7 个柱体积的缓冲液 A 平衡柱。在装载装载物后,通过使用 7 个柱体积的缓冲液 A 洗涤,移除未结合的蛋白。20 个柱体积的线性梯度 0-100 % 缓冲液 B 用于从柱上洗脱酶。

[0458] 以约 0.3 M 的 NaCl 浓度在梯度内观察到单峰。

[0459] 装载物、流通级分和涵盖洗脱蛋白的级分的 SDS-PAGE 显示,在流通中存在代表纯化标签和 HRV14 3C 接头的释放的 SEQ ID NO: 7+SEQ ID NO: 8 片段,且释放的 PAL 样结构域 (SEQ ID NO: 1) 存在于从梯度洗脱的单主峰中。洗脱蛋白的纯度估计为约 90%,表明成熟 PAL 样结构域可在仅两个层析步骤中被纯化为高纯度。

[0460] 使用 Vivaspin MWC0 10.000 Da 柱 (Vivaspin),将含有洗脱的成熟酶 (SEQ ID NO: 1) 的级分合并在一起,并在包含 50 mM Tris pH 7.5, 10% 甘油的缓冲液中浓缩和脱盐,并在 20℃ 保存直至使用。

[0461] 实施例 1 表明,即使在 PAL 蛋白 N 端的 5aa 区中具有变化,本发明的 PAL 酶也可被制备为可溶性蛋白。

[0462] 实施例 2

[0463] 克隆、表达具有作为纯化标签的核糖体蛋白 L9 和来自微小杆菌属 (菌株 ATCC BAA-1283 / AT1b) 的 PAL 样结构域的融合蛋白以及纯化融合蛋白和成熟 PAL 样结构域:

[0464] 蛋白 8 为具有 N 端纯化标签 (SEQ ID NO: 7) 和含有 HRV14 3C 蛋白酶位点的接头 (SEQ ID NO: 8) 的微小杆菌属 PAL 样结构域 (SEQ ID NO: 2)。融合蛋白具有 56.115 Da 的计算分子量,和成熟 PAL 样结构域具有 37848, 5 Da 的计算分子量。

[0465] 制备下述构建体：

[0466]

蛋白名	质粒号	物种	融合伴侣	接头	PAL 样序列
8	H	微小杆菌属	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 2

[0467] 质粒号 :H (微小杆菌属) :SEQ ID NO: 33

[0468]

```

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACCTATCTGATTCCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAAATCCTGAAAGAACTGAAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGCGGCAGCGGTAGCGAAGTGCTGTTTCAGGGTCCGCGTACCGATCCGATTTTTAA
AGATGAATATGATGAAAAAGCCAAAAGCAGCCGTTATACCAGCAGCTGGGTTTGGCCTGAAAAAG
ATAGCGTTAGCCATCGTGGTGGTGAAGGTAGCGGTGTTAGCACCAGCCCCGAGCGGTTATGTTTA
TTATCTGCATCGTGGTGATGGTAGCTATGCAAATGAAGAACTGATTACCACCCCGACCATTTACCG
TTTTTGATCCGAATACCAATGAAATTGTGGATGAATTTGGCGATAACCTGTTTCAGTCTCCGCATG
GTATTGAAGTTGATAGCCAGAATAATATTTGGGTGACCGATATTATGCTGAATAAAGTGTTTAAAC
TGGATGAACGTGGTAATGTTCTGGCCACCTTTGGTGATGATTATCGTCTGGGCACCGAAACCAGC
CTGCGTATTCGTAATGAACTGCCGAATTTCCGGTTCGATGAATGAATACACCTTTGCACGTCCG
ACCGATGTTACCGTTATGGAAGATGGCAGCTTTATTGTTGCAGATGGCTATCGCAATCATCGTATT
TGGAATTTAATCGTGATGGCAATATTCAGTGGGAAGTTGATGCATATGGTAGCTCTGATGGCGA
ATTFAATCTGCCGCATGGCATTACCCATGATCAGAGCGGCAATATTTATGTTGCCGATCGCAATAA
TGCACGTATTCAGGTGTTTGATCAGGATGGTCAGCATCTGAGCACCTGGGATGATACCGAAATTG
GCCGTCCTTATGGAATCGATGCAGGCAATGATGGTAATATTTATCTGGTGATGGTGGCGATTAT
CTGAATGGTGAACGTGAAACCCCGAAAAAGCCAGATTGTTGTTCTGAGCCCCGAAAGGTGAAGTTAT
TGAACGTTTTGGTAGCTGGGGTAACAAAATGGGTGAGCTGCGTATTCGCGATGATCTGACCGTGC
TGGAAGATGGTACAATTTTGTGGCCGAACCTGCTGAATGAACGTCTGCAGAAATTTACCATTACCG
AA

```

[0469] 通过将编码接头和 PAL 样区的合成 XhoI/BamHI 片段连接到已编码 TAP 标签 (SEQ ID NO: 7) 的 pET11a 中, 获得编码蛋白 8 的质粒 H, 并进行如实施例 1 中所述的验证步骤。如实施例 1 中所述, 在 30℃ 表达 3 小时并进行 SDS-PAGE 分析。

[0470] 蛋白 8 的表达水平与赤细菌属变体非常相似。蛋白同样高度可溶, 在细胞裂解物的不溶性级分中几乎检测不到蛋白。

[0471] 基本上如实施例 1 中所述, 进行阳离子交换捕获、用 HRV14 3C 切割和成熟微小杆菌属 PAL 样结构域的纯化。在 SP Sepharose FF 上捕获后, 以约 0.25 M 的 NaCl 浓度洗脱主峰, 基于 SDS-PAGE 分析, 捕获为高效的且纯化的融合蛋白的最初纯度为约 80%。

[0472] HRV14 3C 蛋白酶切割合并并浓缩的融合蛋白 (使用 1:25 酶/底物比过夜) 后, 如实施例 1 中所述的相同效率进行成熟微小杆菌属 PAL 样蛋白 (SEQ ID NO: 2) 的纯化。以约 0.5 M 的 NaCl 浓度从 Q Sepharose HP 柱上洗脱成熟 PAL 样结构域 (图 4A)。SDS-PAGE 分析表明, 装载物中存在对应于成熟微小杆菌属 PAL 样结构域 (SEQ ID NO: 2) (上方条带) 和释放的 TAP 标签 (SEQ ID NO: 7-SEQ ID NO: 8) (下方条带) 的两条片段 (图 4B)。阴离子交换分离后, 在从梯度洗脱的主峰 (级分 12-15) 中发现成熟微小杆菌属 PAL 样结构域, 在流通级分中存在释放的标签 (图 4B)。

[0473] 使用 Vivaspin MWC0 10.000 Da 柱 (Vivaspin), 将含有洗脱酶的级分合并在一起, 并在包含 50 mM Tris pH 7.5 的缓冲液中浓缩和脱盐, 并在 20℃ 保存直至使用。

[0474] 实施例 3

[0475] 克隆、表达和纯化具有作为纯化标签的核糖体蛋白 L9 和来自食土性细菌 (*Chthoniobacter flavus*) Ellin428 的 PAL 样结构域的融合蛋白

[0476] 蛋白 9 为具有 N 端纯化标签 (SEQ ID NO: 7) 和含有 HRV14 3C 蛋白酶位点的接头 (SEQ ID NO: 8) 的食土性细菌 Ellin428 (SEQ ID NO: 3)。融合蛋白具有 51650.4 Da 的计算分子量,和具有 Gly-Pro N 端的成熟 PAL 样结构域具有 33383,3 Da 的计算分子量。

[0477] 制备下述构建体:

[0478]

蛋白名	质粒号	物种	融合伴侣	接头	PAL 样序列
9	I	食土性细菌	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 3

[0479] 质粒号 I: (食土性细菌):SEQ ID NO: 34

[0480]

```
ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACCTATCTGATTCGCGTGCGCTTTGCGAAAGAATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
GCGGGCAAAATTTTGGCGCGGTGACCGCGGGCGACCGTGCGCGGAAGAAATTAGCAAAACCACCG
GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGCGGCGAGCGGTAGCGAAGTGCTGTTTCAGGGTCCGGAAGCCGTTATGAAGTTGT
TCCGCATTGGCCCTGTTCTGCCGGAAGGTCGTAGCCTGGGTGTTTGTGCAGGTGTTGGTGTTGATA
GCCATGGTAATGTGTTTGTGTTTCATCGTAATGAACGTAATTGGACCGCAGCATTTCGGAAGAA
CCGATTGCAGAACCGACCATTTCTGTTTTGTGTTGTCAGGTCAGAGCGGCAAACTGCTGACCGAATGGGG
TGCAGGTGAATTTATTATGCCGCATGGTCTGACCCTGGATCGTGAAGATAATGTTTGGCTGACCG
ATGTTGGTCGTCAGCAGGTTTTTAAATATGCCCATGATGGTCATCTGCTGCTGACCCCTGGGTGAA
CGTGGTGTTGCAGGTAGCGATCAGACCCATTTTAATCTGCCGACCGATGTTGCAGTGCTGCCTGA
TGGAAGCTTTTATGTGTCTGATGGCTATCGTAATACCCGTGTGGTTAAATTTGATGCCGCAGGCC
ATTATCAGTTTGAATGGGGTGGTAAAGGCACCGAACCAGGGTAAATTTCTGCTGCCTCATGGAGTT
GCAGTTGATTCTCATGGTCGTGTTTGTGTTTGCATCGTACCAATAGCCGCTGTCAGGTTTTTGAT
CCGAAAGGTAAATTTCTGCCGAATGGAAGGTCCGCAGGTTGGTCGTCCGTATGGTGTTAGCG
TTGCAGCAAATGATCATGTGTTTGTGATTGATGGTGGTGATCAGCTGCCGAATCAGCCGGAACAT
GCAAAAGCAGTTGAACTGGACCCTGAAGGTAATGTTGTTCCGCGTTTTGGTAGCTATGGTCGTGA
TCCGGGTCAGTTTCACTGGGTGATGATATTGCAGTTGCACCGGATGGTTCTGTTTATGTTGGTG
ATGCCAAAGGTAAACGCGTTCAAGAAATTTGTTCCGGTGTCATCCG
```

[0481] 通过将编码接头和 PAL 样区的合成 XhoI/BamHI 片段连接到已编码纯化标签 (SEQ ID NO: 7) 的 pET11a 中,获得编码蛋白 9 的质粒 I,并进行如实施例 1 中所述的验证步骤。如实施例 1 中所述,在 30℃表达 3 小时并进行 SDS-PAGE 分析。

[0482] 蛋白 9 的表达水平低于赤细菌属和微小杆菌属 PAL 样变体。然而,蛋白在中性 pH 下显著可溶。

[0483] 通过在 SP Sepharose FF 上阳离子交换捕获,在约 0.25 M 的 NaCl 浓度下观察到主峰。尽管基于 SDS-PAGE 分析,纯度始终为可比较高,然而捕获还是比两种之前的结构域所观察到的捕获低效。

[0484] 使用 Vivaspin MWC0 10.000 Da 柱 (Vivaspin),将含有洗脱酶的级分合并在一起,并在包含 50 mM Tris pH 7.5 的缓冲液中浓缩和脱盐,并在 20℃保存直至使用。

[0485] 实施例 4

[0486] 克隆、表达和纯化具有作为纯化标签的核糖体蛋白 L9 和来自海洋浮霉菌 (*Planctomyces mares*) DSM 8797 的 PAL 样结构域的融合蛋白

[0487] 蛋白 10 为具有 N 端纯化标签 (SEQ ID NO: 7) 和含有 HRV14 3C 蛋白酶位点的接

头 (SEQ ID NO: 8) 的来自浮霉状菌属 DSM 8797 的 PAL 样结构域 (SEQ ID NO: 4)。对于浮霉状菌属 PAL 样结构域,用丙氨酸和缬氨酸残基置换野生型序列中存在的两个 Cys 残基。

[0488] 融合蛋白具有 48678, 56 Da 的计算分子量,和带有 Gly-Pro N 端的成熟 PAL 样结构域具有 30411.4 Da 的计算分子量。

[0489] 制备下述构建体:

[0490]

蛋白名	质粒号	物种	融合伴侣	接头	PAL 序列
10	J	浮霉状菌属	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 4

[0491] 质粒号:J 浮霉状菌属:SEQ ID NO: 35

[0492]

```
ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGSCAAAAAAGGCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCGTAACATCTGATTCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAAATCCTGAAAGAACTGAAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACCAACG
GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGTGGTAGCGGTAGCGAAGTTCTGTTTCAGGGTCCGGCAGATAAAATTGATTTTGA
ACCGGCAGCCATTAATATTGAACTGCCGGAAGGTCTGGCACTGGGTCTGCAAGCGCAGTTGAT
TTTGATAGCAAAGGTCGCATGTACCTGTTTCATCGTGGTCCGCAGCCGATTCTGGTTTTTGATCAG
AGCGGTAAATTTGTTTCGTAGCTGGGGTGATAAACTGATTAGCCAGGCACATGGCTTAAGAGTTGC
ACCGGATGAAACCATTTGGGTTACCGATATTGGCAACCATGGTGTTCAGTTTAACCCGGAAG
GTAAACTGCTGCTGGCCCTGGGTGAGGCAGGTAAACCGGGTGATAGCCAGGATCAGTTTAATAA
ACCGACCGATATTGCATTTGGTCCGCAGGGCGAATTCTATATTTCTGACGGCTATGGTAATAGCC
GTGTGATGAAATTTGCAGCCAATGGTAAAAATCTGGGTGAGTGGGGTACACCGGGTAAAGGTCC
GGGTGAATTTAATCTGCCGATAGCATTTCTGGTTGATGCAAAAGGTGCTGTTCTGGTTGGTGATC
GTGAAAATGATCGCGTCGAGATCTTTGATCTGGGAAGGCAATCTGCTGGAAAATTTGGACCGGTTTT
GCACCGTATGGTATGGAATTTGATAGCCGTGGTAACTGTTTGTTCAGATGGTCTGTCAAATAA
AGTTCTGCAGCTGAATGCAAGCGGTAAAGTTGAAAATAGCTGGGGTAAACCGGCAAAAGAACCG
GGTGAGTATAACCTGCCTCACATGCTGGCAGTCGACGCAGCAGGCAACCTGTTTGTGACCGAAA
TTGGTGGTAAACGTCTGCAGAACTGCAGCGCAA
```

[0493] 通过将编码接头和 PAL 样区的合成 XhoI/BamHI 片段连接到已编码纯化标签 (SEQ ID NO: 7) 的 pET11a 中,获得编码蛋白 10 的质粒 J,并进行如实施例 1 中所述的验证步骤。在 BL21 (DE3) 中的表达产生具有预期分子量的蛋白,如通过实施例 1 中所述的 LC-MS 分析评定,且蛋白 10 的表达水平和高溶解度特性与实施例 1 中所述的赤细菌属变体类似。

[0494] 基本上如实施例 1 中所述进行阳离子交换捕获。在 SP Sepharose FF 上捕获后,在约 0.25 M 的 NaCl 浓度下观察到主峰,基于 SDS-PAGE 分析,捕获为有效的,在流通中检测到非常低量的蛋白,且来自主峰的蛋白纯度估计为约 80%。

[0495] 使用 Vivaspin MWC0 10.000 Da 柱 (Vivaspin),将含有洗脱酶的级分合并在一起,并在包含 50 mM Tris pH 7.5 的缓冲液中浓缩和脱盐,并在 20℃ 保存直至使用。

[0496] 实施例 5

[0497] 用于评价 α -羟基苯甲酰甘氨酸和苯甲酰胺转化的酶活性测定

[0498] 酶促反应:

[0499] 之前描述了用于测量 PAL 活性的酶促测定 (Katopodis AG 等: Biochemistry. (1990) 29(26):6115-20),其测量 α -羟基苯甲酰甘氨酸向苯甲酰胺的转化。为了测试 PAL 结构域转化 α -羟基苯甲酰甘氨酸为苯甲酰胺的能力,如下建立酶促反应:

[0500] 酶 (上文实施例中所述):约 0.2 mg/ml

[0501] α -羟基苯甲酰甘氨酸 (Sigma) :1 mg/ml

[0502] Zn_2SO_4 :1 mM

[0503] 缓冲液 :MES pH 5.5 或 Tris pH 7.5 :100 mM

[0504] 在 37℃ 孵育反应 3 小时, 并使用 UPLC (超高效液相色谱) 代替之前所述的 HPLC 分离后, 评价对应于 α -羟基苯甲酰甘氨酸和苯甲酰胺的峰。

[0505] UPLC 分离条件

[0506] Waters ACQ-TUV 仪器设置 :

[0507] 柱 :UPLC BEH300 C18, 1.7mm, 2.1x150mm 柱, 来自 Waters (part#186003687)

[0508] 溶剂 A :0.1% TFA,

[0509] 溶剂 B :90% MeCN, 0.1% TFA (v/v)

[0510] 梯度表

[0511]

步骤	时间 (分钟)	流速	%A	%B	曲线
1	开始	0.3	95	5	
2	5	0.3	85	15	6
3	5.50	0.3	10	90	6
4	6.50	0.3	10	90	6
5	7	0.3	95	5	6
6	8	0.3	95	55	6
7	20	0.020	95	5	11

[0512] 用酶孵育后, 分别代表 α -羟基甘氨酸和 α -酰胺的合成 α -羟基苯甲酰甘氨酸和苯甲酰胺的水平 UPLC 分析, 表明酶是否能够将 α -羟基甘氨酸转化为 α -酰胺基。与无酶的对照相比, 苯甲酰胺峰面积相对于 α -羟基苯甲酰甘氨酸峰面积的增加证实酶具有 PAL 活性。基于代表 α -羟基苯甲酰甘氨酸和苯甲酰胺的峰下面积的半定量分析, 证实所有 4 种酶 (蛋白 3、蛋白 8、蛋白 9 和蛋白 10) 能够催化 α -羟基甘氨酸转化为 α -酰胺。来自赤细菌属和微小杆菌属的 PAL 样结构域 (蛋白 3 和蛋白 8) 在较高 pH (pH 7.5) 下显著有活性, 而来自食土性细菌和浮霉菌属的 PAL 样结构域在较低 pH (pH 5.5) 下明显更具活性, 表明对酶的最优 pH 不同 (如对在图 5 中所示赤细菌属和食土性细菌 PAL 样结构域所例示)。赤细菌属 PAL 样蛋白 (蛋白 3) 和微小杆菌属 PAL 样蛋白 (蛋白 8) 的成熟变体和加上 TAP 标签的变体的分析表明, 与不具有 N 端融合伴侣 (SEQ ID NO: 7) 的成熟酶相比, 加上 TAP 标签的 PAL 样蛋白是具有活性的功能性酶 (图 5)。

[0513] 实施例 6

[0514] 测试 PAL 样结构域对用大鼠 PHM 共处理的加上 TAP 标签的 C 端 Gly 延伸的胰淀粉样肽类似物底物的活性

[0515] 为了测试细菌 PAL 样结构域是否适合用于 C 端 Gly 延伸的重组肽的 α 酰胺化, 用细菌 PAL 样结构域和来自大鼠 (褐家鼠) 的公知 PHM 结构域建立酶促反应。基本上如 Husten EJ 等 (1993) J. Biol. Chem. ;268(13):9709-17 中所述在 HEK293 细胞中瞬时重组表达后, 纯化用于促进 C 端 α -羟基甘氨酸形成的大鼠 PHM 结构域 (包含 976aa 全长大鼠 PAM 序列的氨基酸序列 36-497)。

[0516] 设计酶的相关模型肽底物, 其包含如在 SEQ ID NO: 17 中所述的序列。该肽由以下组成: 含有海栖热袍菌核糖体蛋白 L27 的 N 端 TAP 标签、具有肠激酶位点的位于中间的接头

和具有 C 端 Gly 延伸的人胰淀粉样肽类似物（包含下述氨基酸置换：Val17His、Ala25Pro、Ser28Pro 和 Ser29Pro）。优化包含该序列的基因片段的密码子以用于在大肠杆菌中表达，并连接到 pET11a 载体中，在 30℃ 在 BL21 (DE3) 中表达 3 小时。使用 SP Sepharose FF 柱，用基本上如实施例 1 中所述的缓冲液和设置进行纯化。将级分合并，并在包含 100 mM Tris pH 7.5 的缓冲液中浓缩至 0.3 mg/ml 的浓度。

[0517] 将模型肽与 PAL 样结构域和如下所述的其他相关组分孵育。

[0518] 5 μ L 大鼠 PHM-His (0.15 mg/ml)

[0519] 5 μ L PAL 样结构域（蛋白 3 (0.4 mg/ml) 或蛋白 10 (0.6 mg/ml)）

[0520] 10 μ L 100 mM Tris pH 7.5

[0521] 10 mM 抗坏血酸

[0522] 10 μ M ZnSO₄

[0523] 10 μ M CuSO₄

[0524] 30 μ g/ml 过氧化氢酶

[0525] 30 μ L 加上 TAP 标签的胰淀粉样肽类似物 ((V17H) (0.3 mg /ml))

[0526] 在 37℃ 孵育样品，用 3 μ L 100% 乙酸终止酶促反应，并在 1h、2h、4h 和 5h 后通过 LC-MS 分析，如实施例 1 中所述在 Poroshell C8 SB300 1mm X 7.5mm 反相柱上通过 LC-MSD-TOF 进行。

[0527] 计算模型蛋白 SEQ ID NO: 17 前体的预测平均同位素质量（由于 SEQ ID NO: 17 的第二位置中的丙氨酸，通过大肠杆菌甲硫氨酸氨肽酶移除起始甲硫氨酸）：

[0528] 前体形式（包含 C 端 Gly 残基）：13934.66 Da

[0529] 中间体形式（包含 C 端 α -羟基甘氨酸）：13951 Da

[0530] C 端 α 酰胺化形式：13876.62 Da

[0531] 用大鼠 PHM 单独处理模型肽 2 小时后，观察到绝大多数模型肽为包含 C 端 α -羟基甘氨酸的中间体形式，如通过 LC-MS 分析所确定（图 6B）。由于自发转化也出现了少量 α 酰胺化产物，如对实施例 5 中的合成苯甲酰基 (benzoyl) 衍生物底物所观察到的。然而，通过加入 PAL 样结构域，观察到向 α 酰胺化肽的显著转化，但在两种测试酶的特异性活性中具有差异。使用所述的孵育条件，1h-2h 孵育后已经观察到蛋白 10 完全转化为前体（图 6C）。相比之下，需要 5 小时与蛋白 3 酶孵育，才近似完全转化为 α 酰胺化形式（图 6D 和 6F）。总之这些数据表明，在具有 PHM 活性的酶存在下，细菌 PAL 样结构域能够催化重组 C 端 Gly 延伸的蛋白前体转化为 C 端完全 α 酰胺化的蛋白。

[0532] 实施例 7

[0533] 测试 PAL 样结构域对加上 TAP 标签的 C 端 Gly 延伸的胰淀粉样肽类似物（普兰林肽）底物的活性

[0534] 在被称为普兰林肽的 C 端 Gly 延伸的人胰淀粉样肽类似物（包含下述氨基酸置换：Ala25Pro、Ser28Pro 和 Ser29Pro）(SEQ ID NO: 25) 上测试细菌 PAL 样结构域，该类似物通过固相肽合成和纯化后的冻干而获得。三种形式的普兰林肽的理论平均同位素质量为：

[0535] Gly 延伸的普兰林肽：4007, 5 Da

[0536] α 羟基甘氨酸延伸的普兰林肽：4023, 5 Da

[0537] α 酰胺化的普兰林肽 3949, 4 Da

[0538] SEQ ID NO: 25 :KCNTATCATQRLANFLVHSSNFGPILPPT NVGSNTYG。

[0539] 使用大鼠 PHM 制备 C 端具有 α -羟基甘氨酸延伸的普兰林肽

[0540] 将合成的 C 端 Gly 延伸的普兰林肽溶解在 100mM Tris pH7.5 中至终浓度 1mg/ml。用（实施例 6 中使用的）大鼠 PHM 结构域孵育溶液，使用如下表所述的 1:20 (w/w) 的酶/底物比：

[0541]

PHM	普兰林肽	CuSO ₄ (uM)	过氧化氢酶 ug/ml	抗坏血酸 (mM)
0.2mg	4mg	5	20	10

[0542] 在 37℃ 进行反应 2 小时，然后通过加入 TFA 至终浓度为 0.1% 来终止反应。反应后，通过 LC-MS（基本上如实施例 6 中所述，除了使用 C18 反相柱外）分析来自反应混合物的 20ul 样品，并证实获得普兰林肽的中间体 α -羟基甘氨酸形式。

[0543] 通过反相 C18 制备 HPLC 来纯化大鼠 PHM 处理的样品，在 Agilent 1100 HPLC 仪器上使用 Agilent Zorbac C18 延伸柱 (Extend column)，使用如下表中所列的条件和乙腈梯度。

[0544] 缓冲液 A :95% H₂O 5% ACN 0.1%TFA

[0545] 缓冲液 B :70% ACN 20% 异丙醇 0.1%TFA

[0546]

时间(分钟)	B%	流速(ml/min)
0	0	0.3
4	0	0.3
4.5	10	0.3
15	100	0.3
16	100	0.3
17	10	0.3
20	10	0.3

[0547] 纯化 C 端 α -羟基甘氨酸延伸的普兰林肽中间体。收集肽并合并相关级分后，使用 Therm DNA120 Speed Vac 干燥样品，并将肽复溶在 1mL 100 mM Tris pH 7.5 缓冲液中。通过 UV280 吸收测量 α -羟基甘氨酸延伸的普兰林肽的最终量，并将浓度调整为 0.5 mg/ml 肽。

[0548] 评估细菌 PAL 样结构域对普兰林肽 α -羟基甘氨酸中间体肽的活性

[0549] 为了测试 PAL 样结构域 (SEQ ID NO: 1、2、3 和 4) 的活性，使用制备的 α -羟化 SEQ ID NO: 25 作为肽底物，在 100mM Tris pH 7.5 缓冲液中建立下述反应：

[0550]

PAL 样结构域 (μg)	普兰林肽_Gly (OH) (μg 和 mg/ml)	ZnSO_4 (μM)	酶底物比 (w/w)
2	50 μg 0.5 mg/ml	0.5	25
1	50 μg 0.5 mg/ml	0.5	50
0.5	50 μg 0.5 mg/ml	0.5	100
0.1	25 0.5 mg/ml	0.5	250
0.05	25 0.5 mg/ml	0.5	500
0.02	25 0.5 mg/ml	0.5	1250

[0551] 在 37℃ 孵育反应混合物 1 小时,并在 LC-MS 分析前通过加入 0.1% 的 TFA 终止反应,然后通过 LC/MS 分析样品。

[0552] 在去卷积谱上观察到的代表 α 羟基甘氨酸中间体和完全 α 酰胺化普兰林肽的两个主峰。

[0553] 所有 4 种细菌 PAL 样结构域可将普兰林肽的 α 羟基甘氨酸中间体转化为 α 酰胺化普兰林肽。赤细菌属 (图 7A) 和浮霉状菌属 (图 7B) PAL 样结构域在催化反应方面明显更有效,因为在测试的条件下使用 1:100 的酶/底物比,它们可几乎彻底将肽底物转化为完全 α -酰胺化肽。微小杆菌属 (图 7C) 和食土性细菌 (图 7D) PAL 不能将所有的肽底物转化为 α -酰胺化肽,但如对于 4023 (α 羟基甘氨酸延伸的普兰林肽) 和 3948 Da (α 酰胺化普兰林肽) 的相对峰强度所观察到的,可得出结论,它们催化反应,但速率较慢,需要较高浓度的酶。两组 PAL 样结构域之间的不同可能由所观察到的实施例 5 中所述的 pH 偏好不同所导致。

[0554] 结果提供证据,所有 4 种细菌 PAL 样结构域具有肽基羟基甘氨酸 α 酰胺化裂合酶活性,并可在需要 α 酰胺化来获得生物活性的典型肽底物上催化与之前关于真核生物 PAL 所述的相同反应。

[0555] 测定

[0556] 测定 (I)

[0557] 活性的评价

[0558] 可如实施例 5 中所述,测量 PAL 酶的活性。可如 Katopodis AG 等, Biochemistry. 1990, 29(26):6115-6120 中所述,通过测量 α -羟基苯甲酰甘氨酸向苯甲酰胺的转化来证明 PAL 酶的活性。代替使用所述的 HPLC 法,有利地使用修改的 UPLC 法。关于 UPLC 法的条件在实施例 5 中有述。

[0559] 本文引用的包括出版物、专利申请和专利在内的所有参考文献通过引用以其整体结合到本文中,引用的程度如同各参考文献单独并明确表示通过引用结合,并以其整体示出在本文中 (至法律允许的最大程度)。

[0560] 本文使用的所有标题和小标题仅用于方便,且不应该被认为以任何方式限制本发明。

[0561] 除非另外声明,本文提供的任何和所有实施例或例示措词(例如,“例如”)的使用仅旨在更好地阐述本发明,并不对本发明范围构成限制。说明书中的措词均不应该被认为表示任何未要求保护的要素对本发明的实施是必需的。

[0562] 本文引用和结合的专利文件仅用于方便,并不反映这样的专利文件的有效性、可专利性和/或实施性的任何观点。

[0563] 在适用法律允许下,本发明包括在本文所附的权利要求中列举的主题的所有修改和等价物。

<110> 诺沃—诺迪斯克有限公司

<120> 新的肽基 α -羟基甘氨酸 α -酰胺化裂合酶

<130> 8086.204-W0

<140> -

<141> 2010-11-26

<160> 18

<170> PatentIn 版本 3.5

<210> 1

<211> 306

<212> PRT

<213> 赤细菌属 (*Erythrobacter* sp.) SD-21

<400> 1

Gly	Pro	Pro	Val	Thr	Ile	Asp	Glu	Ser	Trp	Pro	Asp	Ile	Pro	Glu	Ser
1			5						10					15	

Ala	Val	Phe	Gly	Glu	Pro	Thr	Ala	Ile	Asp	Val	Asp	Ser	His	Gly	His
			20						25					30	

Ile	Phe	Val	Leu	His	Arg	Ala	Gly	Arg	Glu	Trp	Thr	Gln	Pro	Phe	Pro
			35						40					45	

Ser	Asp	Pro	Ile	Ser	Glu	Pro	Thr	Val	Phe	Met	Phe	Ala	Ala	Asn	Gly
			50						55					60	

Lys	Leu	Leu	Ser	Lys	Trp	Gly	Ala	Gly	Glu	Leu	Val	Met	Pro	His	Gly
			65						70					75	80

Leu Ser Ile Asp Gly Asp Asn Lys Val Trp Ile Thr Asp Val Ala Arg
85 90 95

Glu Gln Val Leu Arg Phe Thr His Glu Gly Ala Glu Glu Leu Val Leu
100 105 110

Gly Thr Arg Gly Glu Thr Gly Gln Asp Glu Ser His Phe Gly Arg Pro
115 120 125

Ala Asp Val Thr Phe Val Gly Asp Arg Val Leu Val Ala Asp Gly Tyr
130 135 140

Leu Asn Arg Arg Ile Met Val Phe Asp Arg Ala Gly Asn Phe Leu Glu
145 150 155 160

Gln Trp Gly Lys Glu Gly Glu Asp Ala Gly Glu Phe Asn Leu Pro His
165 170 175

Ala Ile Ala Ala Asp Ser Glu Arg Ile Tyr Val Ala Asp Arg Glu Asn
180 185 190

Ala Arg Val Gln Val Leu Ser Leu Asp Gly Glu Pro Leu Ala Arg Trp
195 200 205

Arg Gln Asp Gly Thr Gly His Pro Tyr Ala Val Lys Pro Ile Gly Ser
210 215 220

Gly Tyr Val Leu Ala Ile Glu Gly Arg Asp Arg Ala Gly Arg Asn Thr

225	230	235	240
-----	-----	-----	-----

Ala Ile Gly Arg Ile Tyr Arg Ala Asp Gly Gly Leu Glu Arg Val Phe
245 250 255

Asp Ala Gly Val Glu Pro His Thr Gly Thr Ser Leu Gly His Asp Val
260 265 270

Ala Ile Gly Pro Asp Gly Ser Ala Tyr Met Val Asp Asn Lys Ala Asn
275 280 285

Arg Val Ile Lys Phe Asp Leu Ser Arg Ala Gly Val Glu Glu Ala Asp
290 295 300

Ala Asp
305

<210> 2
<211> 336
<212> PRT
<213> 微小杆菌属 (Exiguobacterium sp.) (菌株 ATCC BAA-1283 / AT1b)

<400> 2

Gly Pro Arg Thr Asp Pro Ile Phe Lys Asp Glu Tyr Asp Glu Lys Ala
1 5 10 15
Lys Ser Ser Arg Tyr Thr Ser Ser Trp Val Trp Pro Glu Lys Asp Ser
20 25 30
Val Ser His Arg Gly Gly Glu Gly Ser Gly Val Ser Thr Ser Pro Ser

35	40	45
Gly Tyr Val Tyr Tyr Leu His Arg Gly Asp Gly Ser Tyr Ala Asn Glu		
50	55	60
Glu Leu Ile Thr Thr Pro Thr Ile Thr Val Phe Asp Pro Asn Thr Asn		
65	70	75
Glu Ile Val Asp Glu Phe Gly Asp Asn Leu Phe Gln Ser Pro His Gly		
85	90	95
Ile Glu Val Asp Ser Gln Asn Asn Ile Trp Val Thr Asp Ile Met Leu		
100	105	110
Asn Lys Val Phe Lys Leu Asp Glu Arg Gly Asn Val Leu Ala Thr Phe		
115	120	125
Gly Asp Asp Tyr Arg Leu Gly Thr Glu Thr Ser Leu Arg Ile Arg Asn		
130	135	140
Glu Leu Pro Asn Phe Pro Val Pro Met Asn Glu Tyr Thr Phe Ala Arg		
145	150	155
Pro Thr Asp Val Thr Val Met Glu Asp Gly Ser Phe Ile Val Ala Asp		
165	170	175
Gly Tyr Arg Asn His Arg Ile Val Lys Phe Asn Arg Asp Gly Asn Ile		
180	185	190

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

Pro His Gly Ile Thr His Asp Gln Ser Gly Asn Ile Tyr Val Ala Asp
210 215 220

Arg Asn Asn Ala Arg Ile Gln Val Phe Asp Gln Asp Gly Gln His Leu
225 230 235 240

Ser Thr Trp Asp Asp Thr Glu Ile Gly Arg Pro Tyr Gly Ile Asp Ala
245 250 255

Gly Asn Asp Gly Asn Ile Tyr Leu Val Asp Gly Gly Asp Tyr Leu Asn
260 265 270

Gly Glu Arg Glu Thr Pro Lys Ser Gln Ile Val Val Leu Ser Pro Lys
275 280 285

Gly Glu Val Ile Glu Arg Phe Gly Ser Trp Gly Asn Lys Met Gly Gln
290 295 300

Leu Arg Ile Pro His Asp Leu Thr Val Leu Glu Asp Gly Thr Ile Phe
305 310 315 320

Val Ala Glu Leu Leu Asn Glu Arg Leu Gln Lys Phe Thr Ile Thr Glu
325 330 335

<210> 3

<211> 305

<212> PRT

<213> 食土性细菌 (Chtoniobacter flavus) Ellin428

<400> 3

Gly Pro Glu Ser Arg Tyr Glu Val Val Pro His Trp Pro Val Leu Pro
1 5 10 15

Glu Gly Arg Ser Leu Gly Val Cys Ala Gly Val Gly Val Asp Ser His
20 25 30

Gly Asn Val Phe Val Phe His Arg Asn Glu Arg Asn Trp Thr Ala Ala
35 40 45

Phe Pro Glu Glu Pro Ile Ala Glu Pro Thr Ile Ser Val Phe Asp Gly
50 55 60

Gln Ser Gly Lys Leu Leu Thr Glu Trp Gly Ala Gly Glu Phe Ile Met
65 70 75 80

Pro His Gly Leu Thr Leu Asp Arg Glu Asp Asn Val Trp Leu Thr Asp
85 90 95

Val Gly Arg Gln Gln Val Phe Lys Tyr Ala His Asp Gly His Leu Leu
100 105 110

Leu Thr Leu Gly Glu Arg Gly Val Ala Gly Ser Asp Gln Thr His Phe
115 120 125

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

Ser Asp Gly Tyr Arg Asn Thr Arg Val Val Lys Phe Asp Ala Ala Gly
145 150 155 160

His Tyr Gln Phe Glu Trp Gly Gly Lys Gly Thr Glu Pro Gly Lys Phe
165 170 175

Arg Leu Pro His Gly Val Ala Val Asp Ser His Gly Arg Val Phe Val
180 185 190

Cys Asp Arg Thr Asn Ser Arg Leu Gln Val Phe Asp Pro Lys Gly Lys
195 200 205

Phe Leu Ala Glu Trp Lys Gly Pro Gln Val Gly Arg Pro Tyr Gly Val
210 215 220

Ser Val Ala Ala Asn Asp His Val Phe Val Ile Asp Gly Gly Asp Gln
225 230 235 240

Leu Pro Asn Gln Pro Glu His Ala Lys Ala Val Glu Leu Asp Pro Glu
245 250 255

Gly Asn Val Val Pro Arg Phe Gly Ser Tyr Gly Arg Asp Pro Gly Gln
260 265 270

Phe Gln Leu Gly His Asp Ile Ala Val Ala Pro Asp Gly Ser Val Tyr
275 280 285

Val Gly Asp Ala Lys Gly Lys Arg Val Gln Lys Phe Val Pro Val His
290 295 300

Pro
305

<210> 4
<211> 279
<212> PRT
<213> 海洋浮霉状菌 (*Planctomyces maris*) DSM 8797

<400> 4

Gly Pro Ala Asp Lys Ile Asp Phe Glu Pro Ala Ala Ile Asn Ile Glu
1 5 10 15

Leu Pro Glu Gly Leu Ala Leu Gly Pro Ala Ser Ala Val Asp Phe Asp
20 25 30

Ser Lys Gly Arg Met Tyr Leu Phe His Arg Gly Pro Gln Pro Ile Leu
35 40 45

Val Phe Asp Gln Ser Gly Lys Phe Val Arg Ser Trp Gly Asp Lys Leu
50 55 60

Ile Ser Gln Ala His Gly Leu Arg Val Ala Pro Asp Glu Thr Ile Trp
65 70 75 80

Val Thr Asp Ile Gly Asn His Met Val Phe Gln Phe Asn Pro Glu Gly
85 90 95

Lys Leu Leu Leu Ala Leu Gly Gln Ala Gly Lys Pro Gly Asp Ser Gln
100 105 110

Asp Gln Phe Asn Lys Pro Thr Asp Ile Ala Phe Gly Pro Gln Gly Glu
115 120 125

Phe Tyr Ile Ser Asp Gly Tyr Gly Asn Ser Arg Val Met Lys Phe Ala
130 135 140

Ala Asn Gly Lys Asn Leu Gly Gln Trp Gly Thr Pro Gly Lys Gly Pro
145 150 155 160

Gly Glu Phe Asn Leu Pro His Ser Ile Leu Val Asp Ala Lys Gly Arg
165 170 175

Val Leu Val Gly Asp Arg Glu Asn Asp Arg Val Gln Ile Phe Asp Leu
180 185 190

Glu Gly Asn Leu Leu Glu Ile Trp Thr Gly Phe Ala Pro Tyr Gly Met
195 200 205

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

Lys Val Leu Gln Leu Asn Ala Ser Gly Lys Val Glu Asn Ser Trp Gly
225 230 235 240

Lys Thr Gly Lys Glu Pro Gly Glu Tyr Asn Leu Pro His Met Leu Ala
245 250 255

Val Asp Ala Ala Gly Asn Leu Phe Val Thr Glu Ile Gly Gly Lys Arg

260

265

270

Leu Gln Lys Leu Gln Arg Lys

275

<210> 5

<211> 324

<212> PRT

<213> 褐家鼠 (Rattus norvegicus)

<400> 5

Gly Asp Phe His Val Glu Glu Glu Leu Asp Trp Pro Gly Val Tyr Leu

1

5

10

15

Leu Pro Gly Gln Val Ser Gly Val Ala Leu Asp Ser Lys Asn Asn Leu

20

25

30

Val Ile Phe His Arg Gly Asp His Val Trp Asp Gly Asn Ser Phe Asp

35

40

45

Ser Lys Phe Val Tyr Gln Gln Arg Gly Leu Gly Pro Ile Glu Glu Asp

50

55

60

Thr Ile Leu Val Ile Asp Pro Asn Asn Ala Glu Ile Leu Gln Ser Ser

65

70

75

80

Gly Lys Asn Leu Phe Tyr Leu Pro His Gly Leu Ser Ile Asp Thr Asp

85

90

95

Gly Asn Tyr Trp Val Thr Asp Val Ala Leu His Gln Val Phe Lys Leu

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

Phe Gly Asp Gln Glu Pro Val Gln Gly Phe Val Met Asn Phe Ser Ser
260 265 270

Gly Glu Ile Ile Asp Val Phe Lys Pro Val Arg Lys His Phe Asp Met
275 280 285

Pro His Asp Ile Val Ala Ser Glu Asp Gly Thr Val Tyr Ile Gly Asp
290 295 300

Ala His Thr Asn Thr Val Trp Lys Phe Thr Leu Thr Glu Lys Met Glu
305 310 315 320

His Arg Ser Val

<210> 6

<211> 323

<212> PRT

<213> 非洲爪蟾 (*Xenopus laevis*)

<400> 6

Gly Asp Val His Leu Glu Glu Asp Thr Asp Trp Pro Gly Val Asn Leu
1 5 10 15

Lys Val Gly Gln Val Ser Gly Leu Ala Leu Asp Pro Lys Asn Asn Leu
20 25 30

Ala Ile Phe His Arg Gly Asp His Val Trp Asp Glu Asn Ser Phe Asp
35 40 45

Arg Asn Phe Val Tyr Gln Gln Arg Gly Ile Gly Pro Ile Gln Glu Ser
50 55 60

Thr Ile Leu Val Val Asp Pro Ser Ser Ser Lys Val Leu Lys Ser Thr
65 70 75 80

Gly Lys Asn Leu Phe Phe Leu Pro His Gly Leu Thr Ile Asp Arg Asp
85 90 95

Gly Asn Tyr Trp Val Thr Asp Val Ala Leu His Gln Val Phe Lys Leu
100 105 110

Gly Ala Gly Lys Glu Thr Pro Leu Leu Val Leu Gly Arg Ala Phe Gln
115 120 125

Pro Gly Ser Asp Arg Lys His Phe Cys Gln Pro Thr Asp Val Ala Val
130 135 140

Asp Pro Ile Thr Gly Asn Phe Phe Val Ala Asp Gly Tyr Cys Asn Ser
145 150 155 160

Arg Ile Met Gln Phe Ser Pro Asn Gly Met Phe Ile Met Gln Trp Gly
165 170 175

Glu Glu Thr Ser Ser Asn Val Pro Arg Pro Gly Gln Phe Arg Ile Pro
180 185 190

His Ser Leu Thr Met Val Pro Asp Gln Gly Gln Leu Cys Val Ala Asp
195 200 205

Arg Glu Asn Gly Arg Ile Gln Cys Phe His Ala Glu Thr Gly Asn Phe
210 215 220

Val Lys Gln Ile Lys His Gln Glu Phe Gly Arg Glu Val Phe Ala Val
225 230 235 240

Ser Tyr Ala Pro Gly Gly Val Leu Tyr Ala Val Asn Gly Lys Pro Tyr
245 250 255

Tyr Gly Tyr Ser Ala Pro Val Gln Gly Phe Met Leu Asn Phe Ser Asn
260 265 270

Gly Asp Ile Leu Asp Thr Phe Ile Pro Ala Arg Lys Asn Phe Asp Met
275 280 285

Pro His Asp Ile Ala Ala Ala Asp Asp Gly Thr Val Tyr Val Gly Asp
290 295 300

Ala His Ala Asn Ala Val Trp Lys Phe Ser Pro Ser Lys Ala Glu His
305 310 315 320

Arg Ser Val

<210> 7

<211> 149

<212> PRT

<213> 海栖热袍菌 (Thermotoga maritima)

<400> 7

Met Lys Val Ile Leu Leu Arg Asp Val Pro Lys Ile Gly Lys Lys Gly
1 5 10 15

Glu Ile Lys Glu Val Ser Asp Gly Tyr Ala Arg Asn Tyr Leu Ile Pro
20 25 30

Arg Gly Phe Ala Lys Glu Tyr Thr Glu Gly Leu Glu Arg Ala Ile Lys
35 40 45

His Glu Lys Glu Ile Glu Lys Arg Lys Lys Glu Arg Glu Arg Glu Glu
50 55 60

Ser Glu Lys Ile Leu Lys Glu Leu Lys Lys Arg Thr His Val Val Lys
65 70 75 80

Val Lys Ala Gly Glu Gly Gly Lys Ile Phe Gly Ala Val Thr Ala Ala
85 90 95

Thr Val Ala Glu Glu Ile Ser Lys Thr Thr Gly Leu Lys Leu Asp Lys
100 105 110

Arg Trp Phe Lys Leu Asp Lys Pro Ile Lys Glu Leu Gly Glu Tyr Ser
115 120 125

Leu Glu Val Ser Leu Pro Gly Gly Val Lys Asp Thr Ile Lys Ile Arg
130 135 140

Val Glu Arg Glu Glu
145

<210> 8
<211> 13
<212> PRT
<213> 人工的

<220>
<223> 肽

<400> 8

Ser Ser Ser Gly Gly Ser Gly Ser Glu Val Leu Phe Gln
1 5 10

<210> 9
<211> 11
<212> PRT
<213> 人工的

<220>
<223> 肽

<400> 9

Ser Ser Ser Gly Ser Gly Glu Val Leu Phe Gln
1 5 10

<210> 10
<211> 10
<212> PRT
<213> 人工的

<220>
<223> 肽

<400> 10

Ser Ser Ser Gly Gly Ser Gly Gly Ser Gly
1 5 10

<210> 11

<211> 10

<212> PRT

<213> 赤细菌属 SD-21

<400> 11

Met Gly Ser Ser His His His His His His
1 5 10

<210> 12

<211> 12

<212> PRT

<213> 人工的

<220>

<223> 肽

<400> 12

Ser Ser Gly Gly Ser Gly Ser Glu Val Leu Phe Gln
1 5 10

<210> 13

<211> 307

<212> PRT

<213> 赤细菌属 SD-21

<400> 13

Glu Ala Pro Pro Val Thr Ile Asp Glu Ser Trp Pro Asp Ile Pro Glu
1 5 10 15

Ser Ala Val Phe Gly Glu Pro Thr Ala Ile Asp Val Asp Ser His Gly
20 25 30

His Ile Phe Val Leu His Arg Ala Gly Arg Glu Trp Thr Gln Pro Phe
35 40 45

Pro Ser Asp Pro Ile Ser Glu Pro Thr Val Phe Met Phe Ala Ala Asn
50 55 60

Gly Lys Leu Leu Ser Lys Trp Gly Ala Gly Glu Leu Val Met Pro His
65 70 75 80

Gly Leu Ser Ile Asp Gly Asp Asn Lys Val Trp Ile Thr Asp Val Ala
85 90 95

Arg Glu Gln Val Leu Arg Phe Thr His Glu Gly Ala Glu Glu Leu Val
100 105 110

Leu Gly Thr Arg Gly Glu Thr Gly Gln Asp Glu Ser His Phe Gly Arg
115 120 125

Pro Ala Asp Val Thr Phe Val Gly Asp Arg Val Leu Val Ala Asp Gly
130 135 140

Tyr Leu Asn Arg Arg Ile Met Val Phe Asp Arg Ala Gly Asn Phe Leu
145 150 155 160

Glu Gln Trp Gly Lys Glu Gly Glu Asp Ala Gly Glu Phe Asn Leu Pro

165	170	175
His Ala Ile Ala Ala Asp Ser Glu Arg Ile Tyr Val Ala Asp Arg Glu		
180	185	190
Asn Ala Arg Val Gln Val Leu Ser Leu Asp Gly Glu Pro Leu Ala Arg		
195	200	205
Trp Arg Gln Asp Gly Thr Gly His Pro Tyr Ala Val Lys Pro Ile Gly		
210	215	220
Ser Gly Tyr Val Leu Ala Ile Glu Gly Arg Asp Arg Ala Gly Arg Asn		
225	230	235 240
Thr Ala Ile Gly Arg Ile Tyr Arg Ala Asp Gly Gly Leu Glu Arg Val		
245	250	255
Phe Asp Ala Gly Val Glu Pro His Thr Gly Thr Ser Leu Gly His Asp		
260	265	270
Val Ala Ile Gly Pro Asp Gly Ser Ala Tyr Met Val Asp Asn Lys Ala		
275	280	285
Asn Arg Val Ile Lys Phe Asp Leu Ser Arg Ala Gly Val Glu Glu Ala		
290	295	300
Asp Ala Asp		
305		

<210> 14
<211> 13
<212> PRT
<213> 人工的

<220>
<223> 肽

<400> 14

Ser Ser Ser Gly Gly Ser Gly Ser Glu Thr Leu Phe Gln
1 5 10

<210> 15
<211> 310
<212> PRT
<213> 赤细菌属 SD-21

<400> 15

Ala Arg Glu Glu Ala Pro Pro Val Thr Ile Asp Glu Ser Trp Pro Asp
1 5 10 15

Ile Pro Glu Ser Ala Val Phe Gly Glu Pro Thr Ala Ile Asp Val Asp
20 25 30

Ser His Gly His Ile Phe Val Leu His Arg Ala Gly Arg Glu Trp Thr
35 40 45

Gln Pro Phe Pro Ser Asp Pro Ile Ser Glu Pro Thr Val Phe Met Phe
50 55 60

Ala Ala Asn Gly Lys Leu Leu Ser Lys Trp Gly Ala Gly Glu Leu Val
65 70 75 80

Met Pro His Gly Leu Ser Ile Asp Gly Asp Asn Lys Val Trp Ile Thr
85 90 95

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

Glu Leu Val Leu Gly Thr Arg Gly Glu Thr Gly Gln Asp Glu Ser His
115 120 125

Phe Gly Arg Pro Ala Asp Val Thr Phe Val Gly Asp Arg Val Leu Val
130 135 140

Ala Asp Gly Tyr Leu Asn Arg Arg Ile Met Val Phe Asp Arg Ala Gly
145 150 155 160

Asn Phe Leu Glu Gln Trp Gly Lys Glu Gly Glu Asp Ala Gly Glu Phe
165 170 175

Asn Leu Pro His Ala Ile Ala Ala Asp Ser Glu Arg Ile Tyr Val Ala
180 185 190

Asp Arg Glu Asn Ala Arg Val Gln Val Leu Ser Leu Asp Gly Glu Pro
195 200 205

Leu Ala Arg Trp Arg Gln Asp Gly Thr Gly His Pro Tyr Ala Val Lys
210 215 220

Pro Ile Gly Ser Gly Tyr Val Leu Ala Ile Glu Gly Arg Asp Arg Ala

225	230	235	240
-----	-----	-----	-----

Gly Arg Asn Thr Ala Ile Gly Arg Ile Tyr Arg Ala Asp Gly Gly Leu		
245	250	255

Glu Arg Val Phe Asp Ala Gly Val Glu Pro His Thr Gly Thr Ser Leu		
260	265	270

Gly His Asp Val Ala Ile Gly Pro Asp Gly Ser Ala Tyr Met Val Asp		
275	280	285

Asn Lys Ala Asn Arg Val Ile Lys Phe Asp Leu Ser Arg Ala Gly Val		
290	295	300

Glu Glu Ala Asp Ala Asp	
305	310

<210> 16
<211> 9
<212> PRT
<213> 人工的

<220>
<223> 肽

<400> 16

Ser Ser Ser Gly Gly Ser Gly Gly Ser	
1	5

<210> 17
<211> 129

<212> PRT

<213> 人工的

<220>

<223> 肽

<400> 17

Met Ala His Lys Lys Ser Gly Gly Val Ala Lys Asn Gly Arg Asp Ser
1 5 10 15

Leu Pro Lys Tyr Leu Gly Val Lys Val Gly Asp Gly Gln Ile Val Lys
20 25 30

Ala Gly Asn Ile Leu Val Arg Gln Arg Gly Thr Arg Phe Tyr Pro Gly
35 40 45

Lys Asn Val Gly Val Gly Arg Asp Phe Thr Leu Phe Ala Leu Lys Asp
50 55 60

Gly Arg Val Lys Phe Glu Thr Lys Asn Asn Lys Lys Tyr Val Ser Val
65 70 75 80

Tyr Glu Glu Ser Ser Ser Asp Asp Asp Asp Lys Lys Cys Asn Thr Ala
85 90 95

Thr Cys Ala Thr Gln Arg Leu Ala Asn Phe Leu His His Ser Ser Asn
100 105 110

Asn Phe Gly Pro Ile Leu Pro Pro Thr Asn Val Gly Ser Asn Thr Tyr
115 120 125

Gly

<210> 18

<211> 23

<212> PRT

<213> 人工的

<220>

<223> 肽

<220>

<221> misc_feature

<222> (1)..(1)

<223> Xaa 可为除了 Cys 外的任何天然存在氨基酸

<220>

<221> misc_feature

<222> (3)..(3)

<223> Xaa 可为任何天然存在氨基酸

<220>

<221> misc_feature

<222> (6)..(8)

<223> Xaa 可为任何天然存在氨基酸

<220>

<221> misc_feature

<222> (10)..(10)

<223> Xaa 可为任何天然存在氨基酸

<220>

<221> misc_feature

<222> (12)..(12)

<223> Xaa 可为除了 Cys 外的任何天然存在氨基酸

<220>

<221> misc_feature

<222> (13)..(16)

<223> Xaa 可为任何天然存在氨基酸

<220>

<221> misc_feature

<222> (18)..(22)

<223> Xaa 可为任何天然存在氨基酸

<400> 18

Xaa Val Xaa Asp Arg Xaa Xaa Xaa Arg Xaa Gln Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Gly Xaa Xaa Xaa Xaa Xaa Trp

20

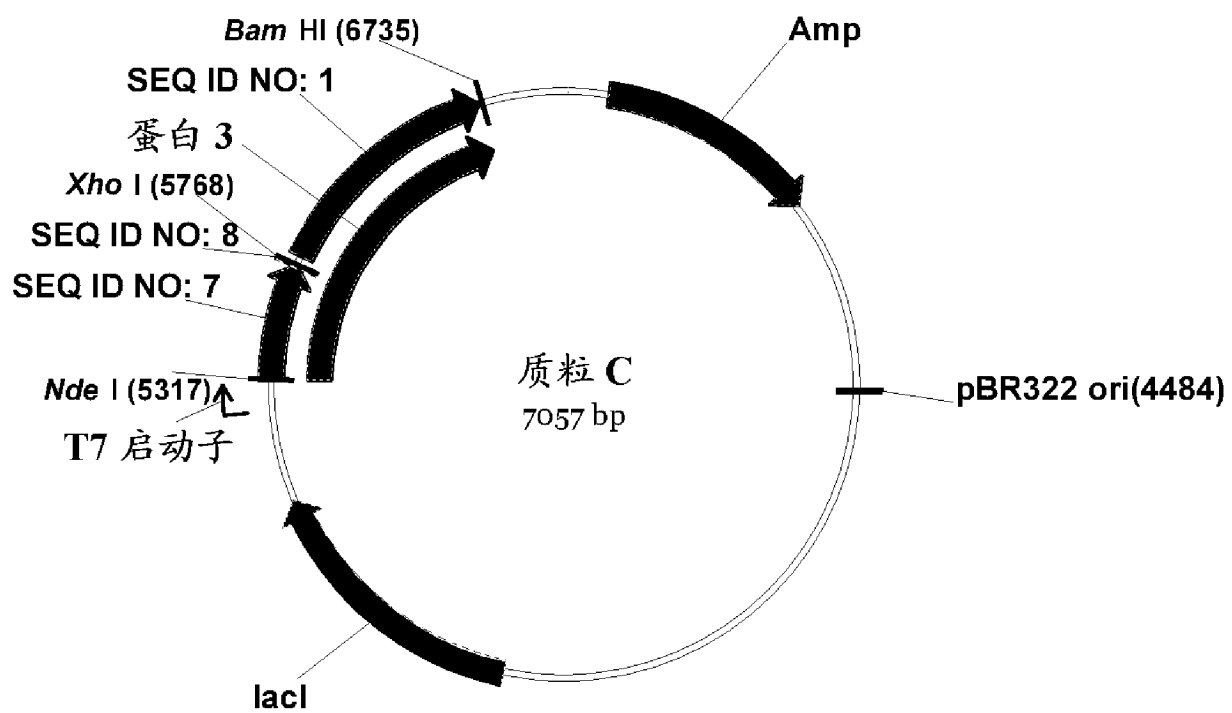


图 1/7

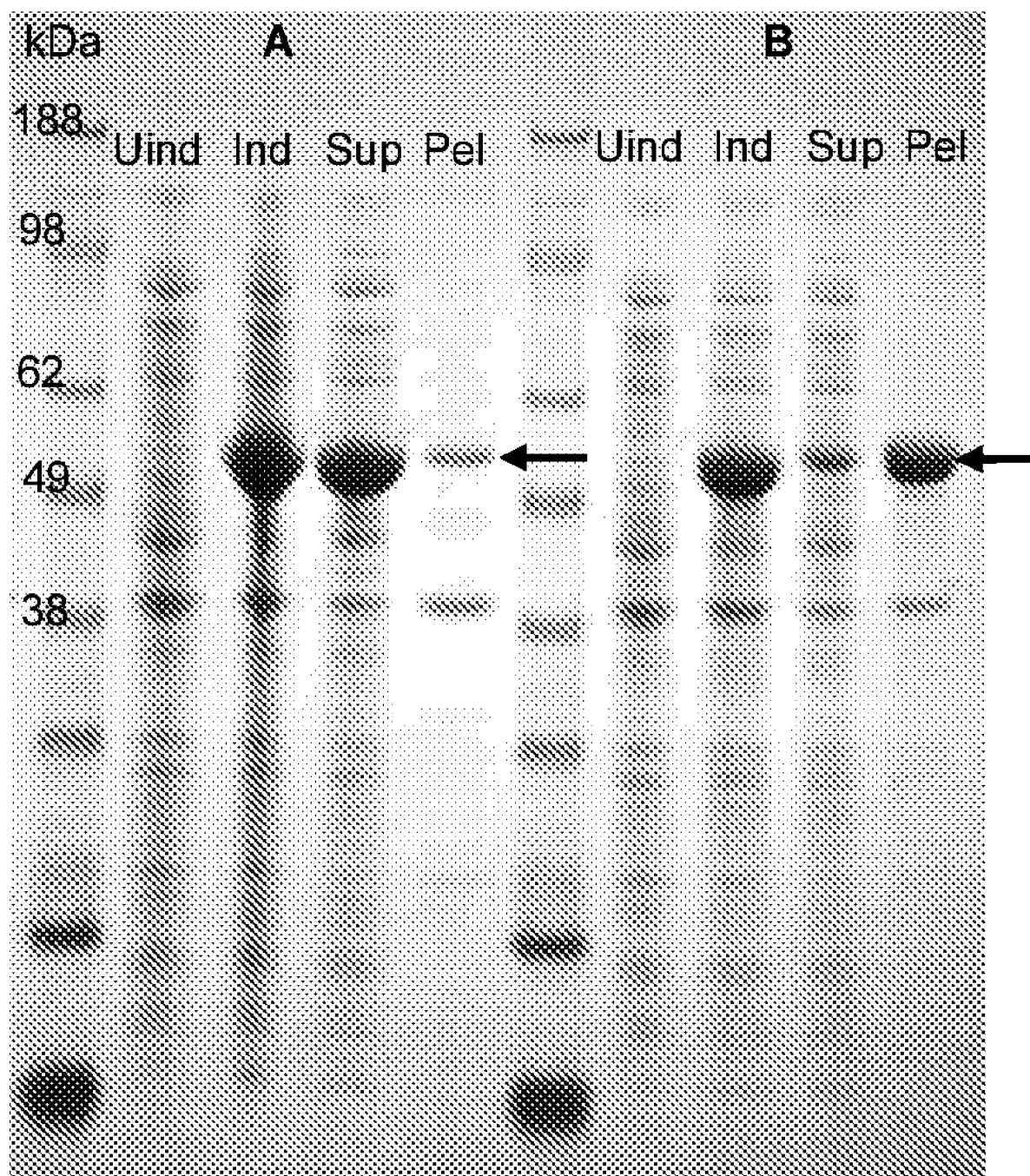


图 2/7

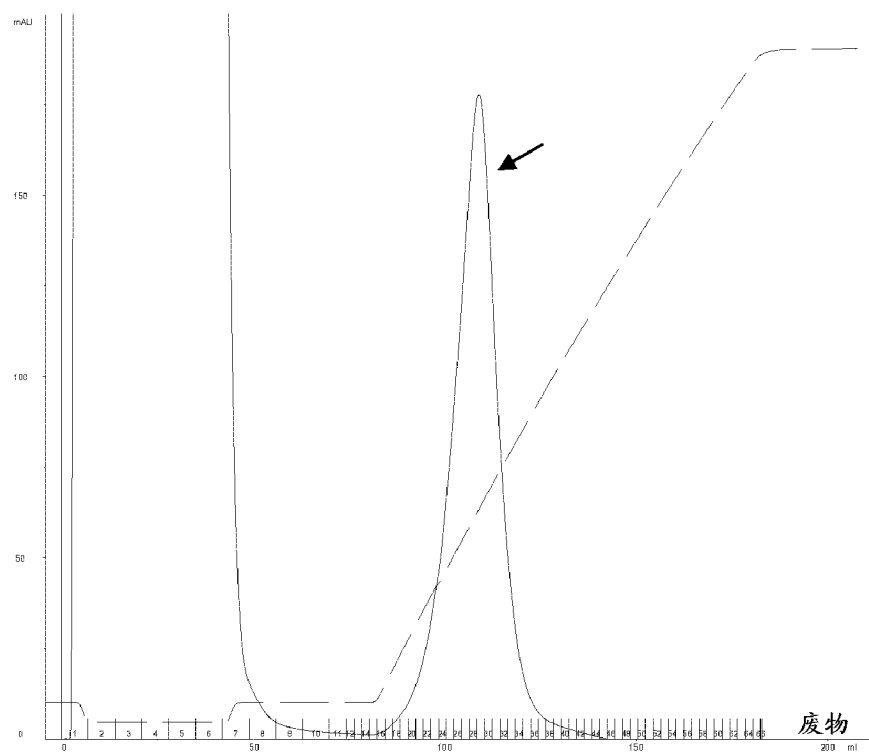
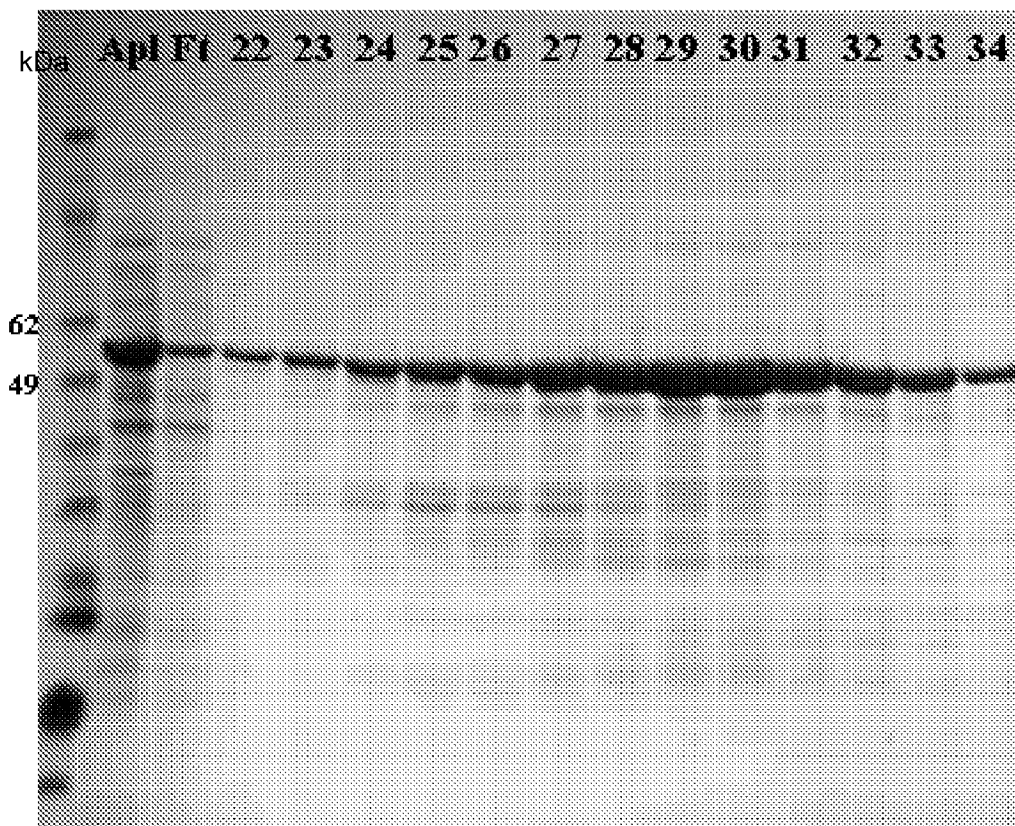
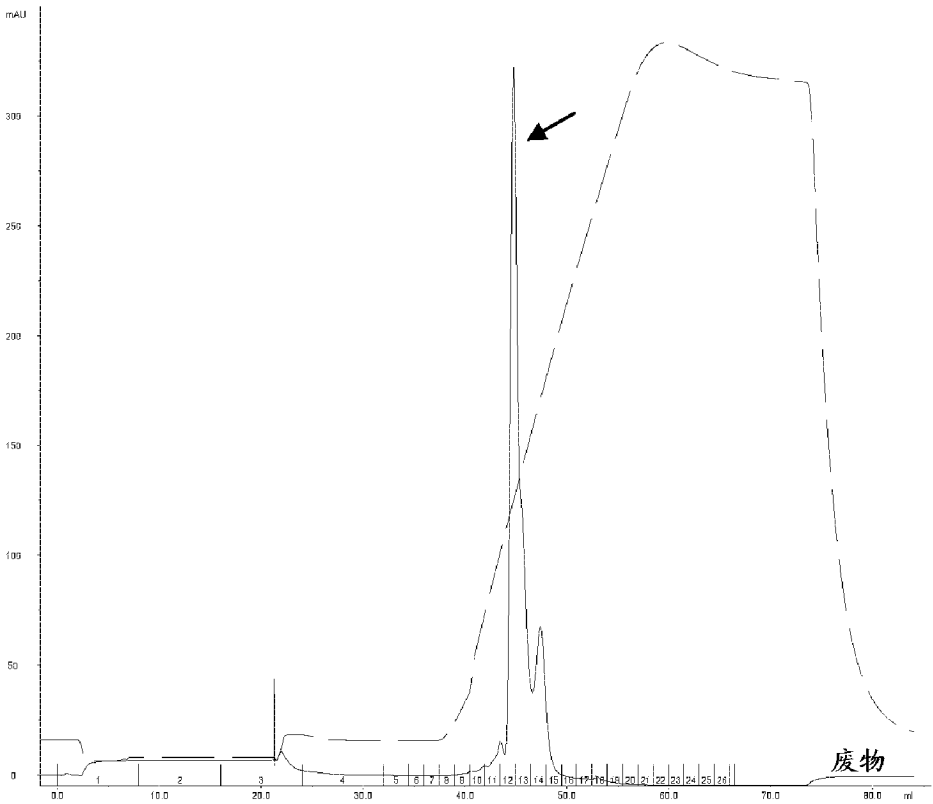
A**B**

图 3/7

A



B

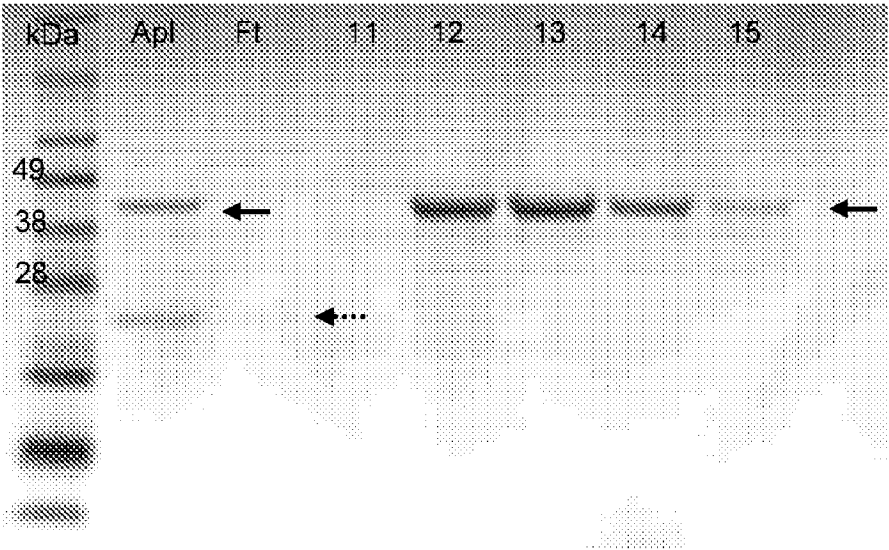


图 4/7

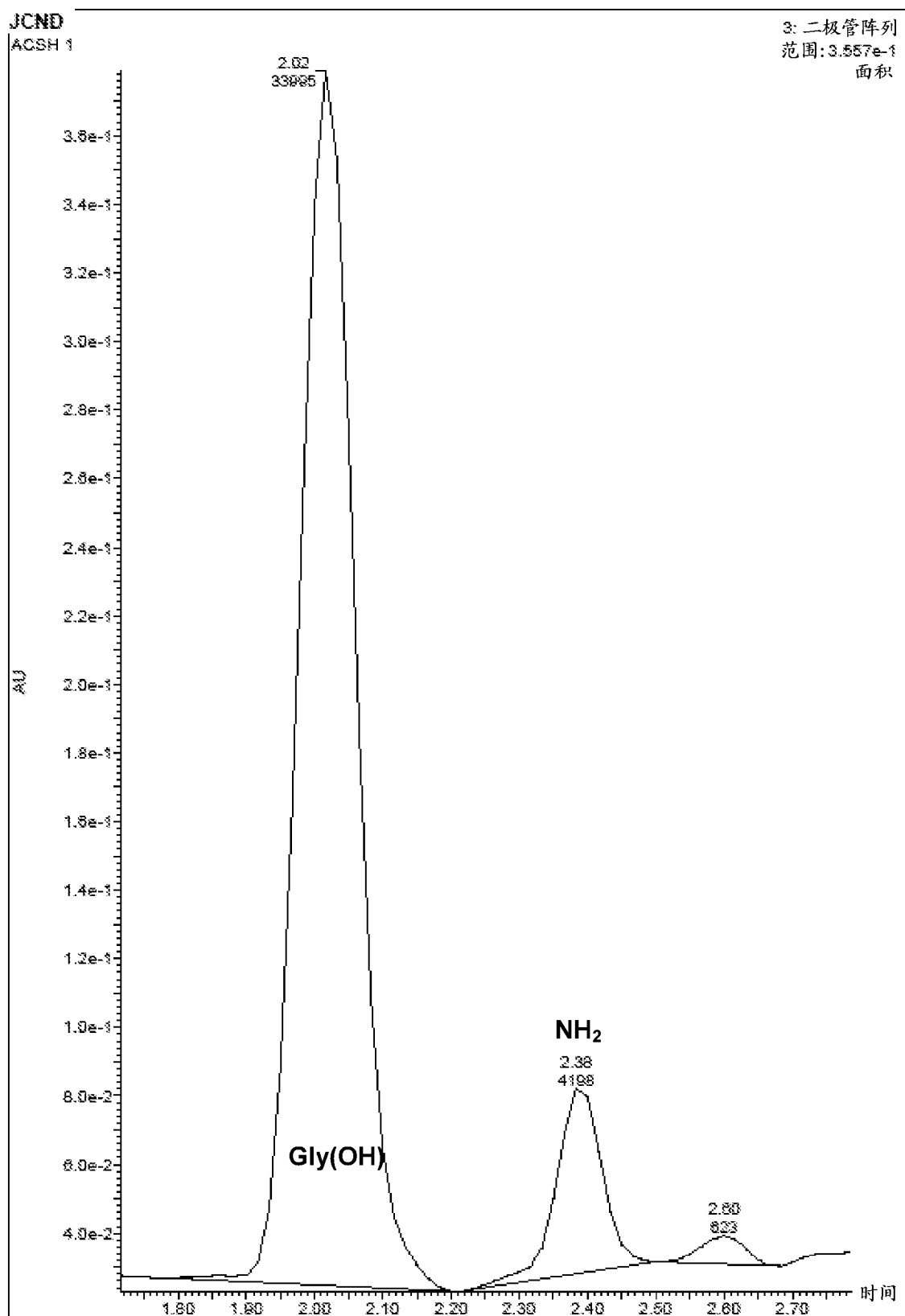
+TAP, MES pH5.5

图 5A1/7

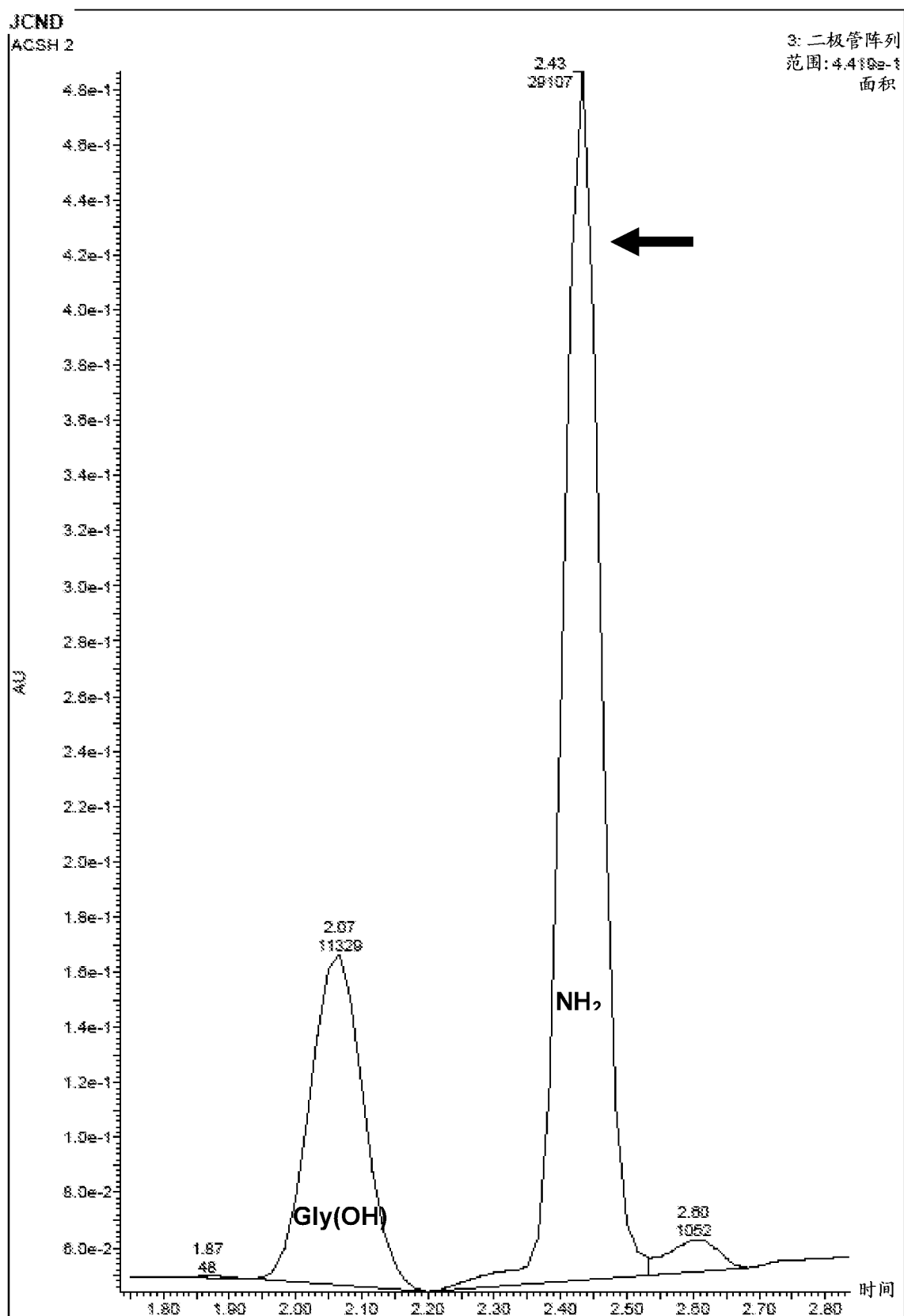
+TAP, Tris pH7.5

图 5A2/7

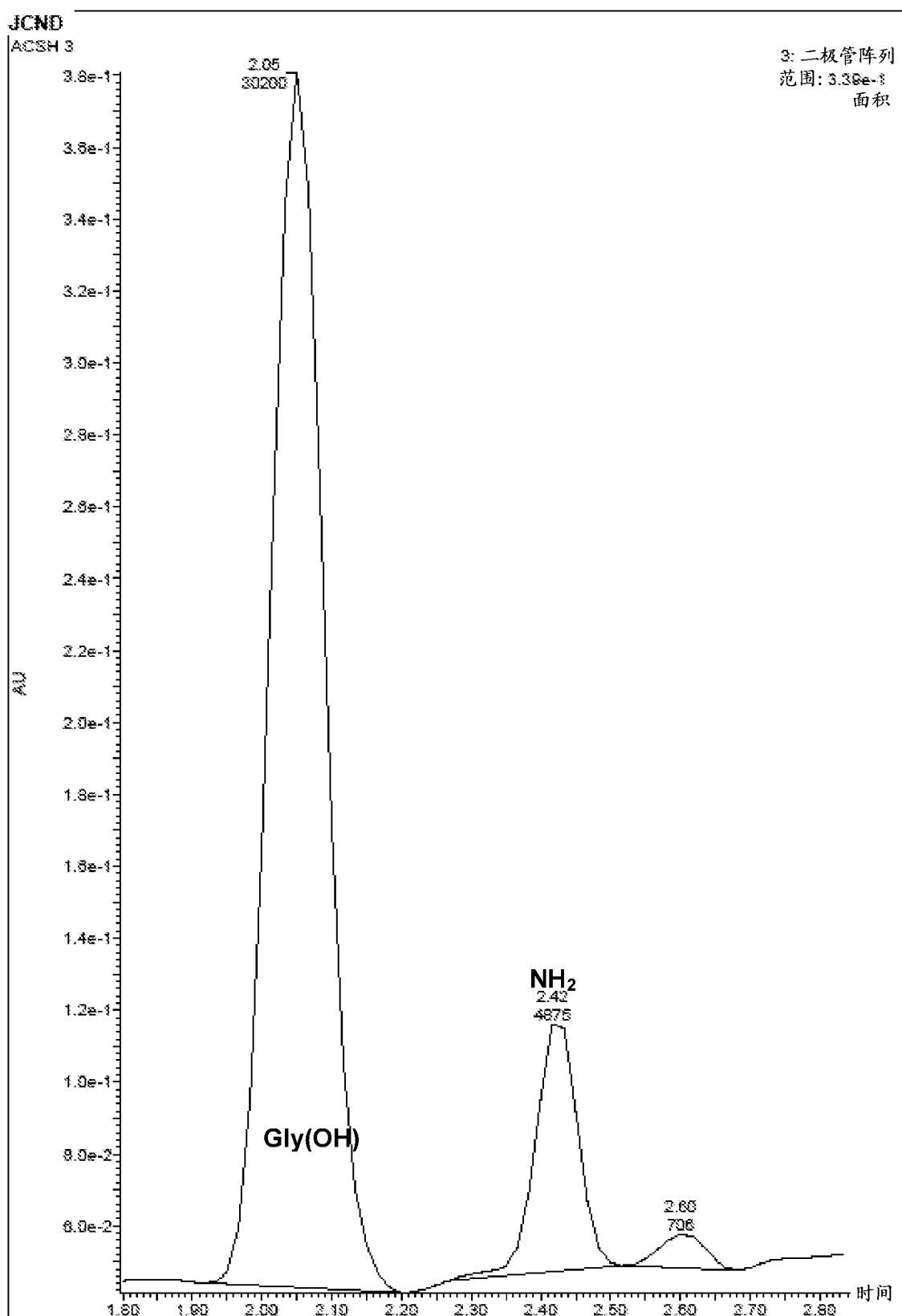
-TAP, MES pH5.5

图 5A3/7

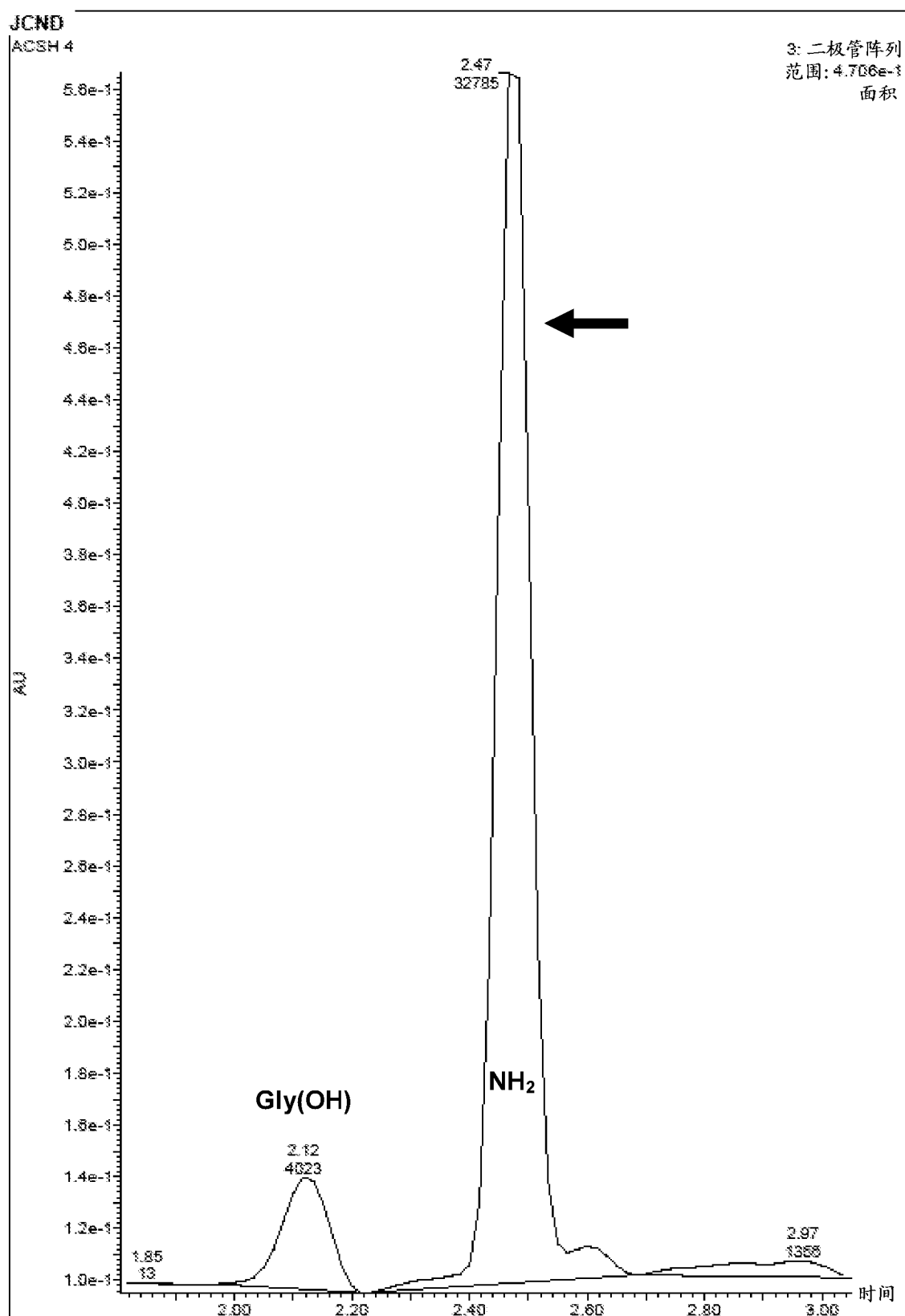
-TAP, Tris pH7.5

图 5A4/7

MES pH 5.5, neg.ctrl

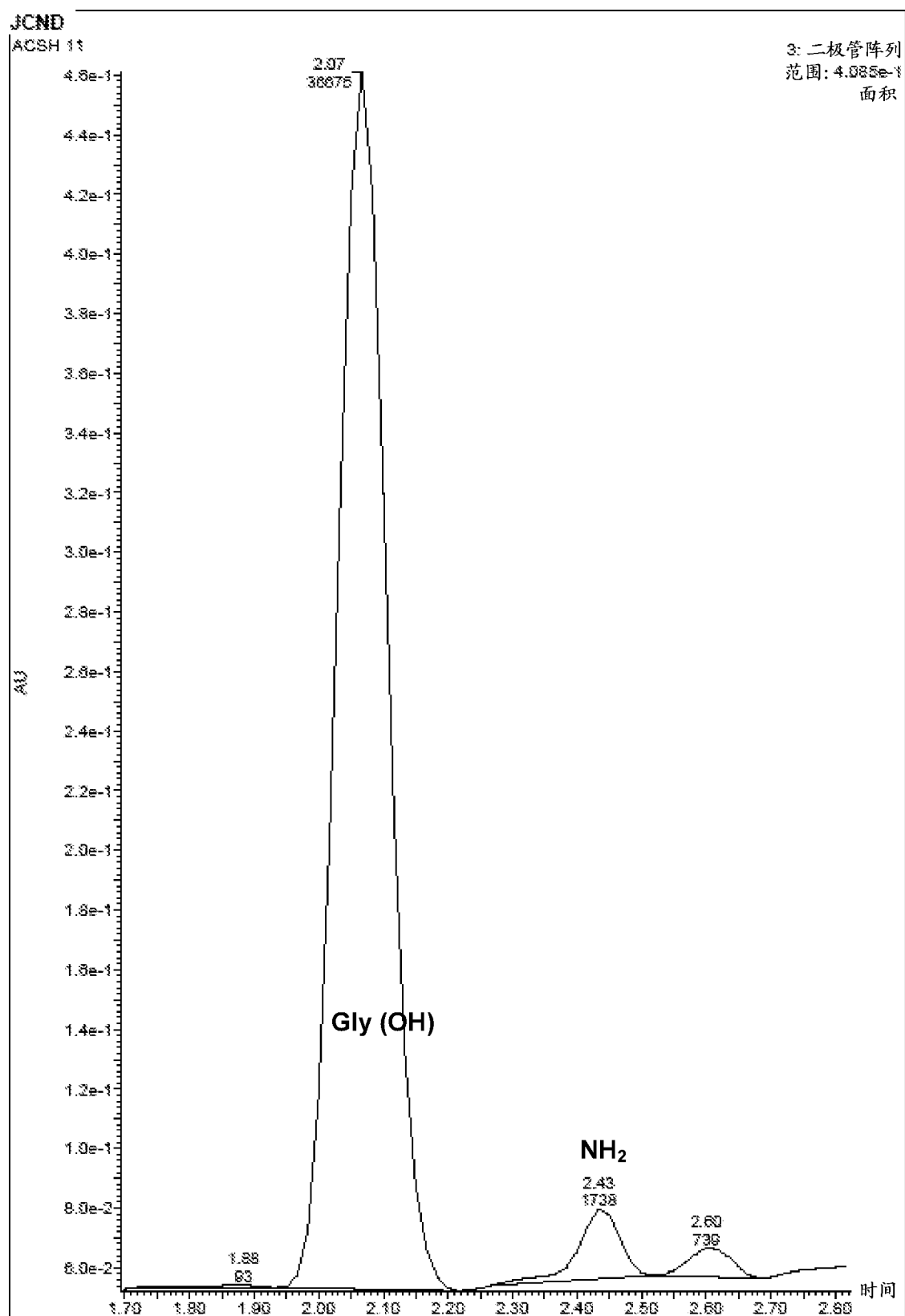


图 5A5/7

Tris pH 7.5, neg.ctrl

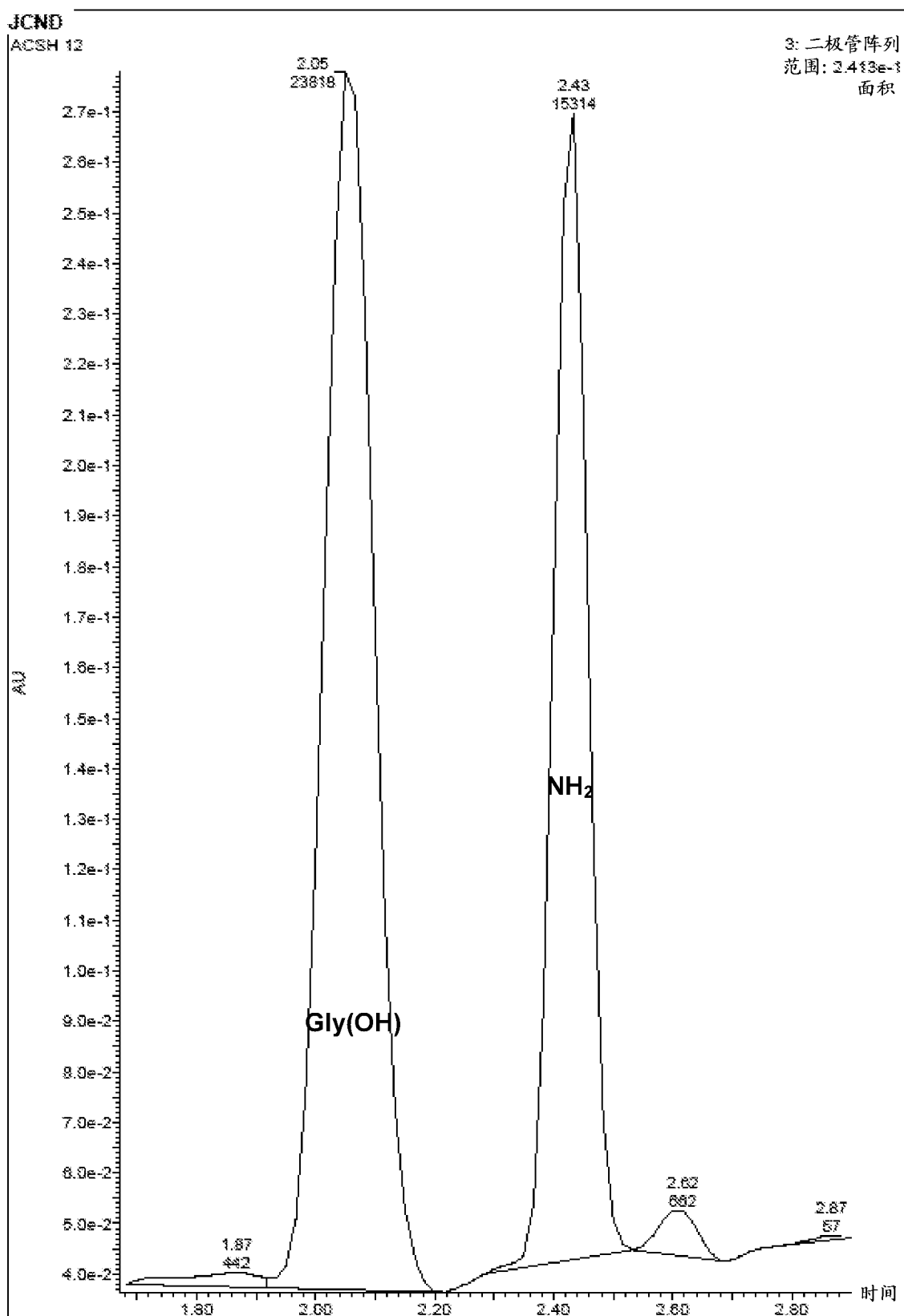


图 5A6/7

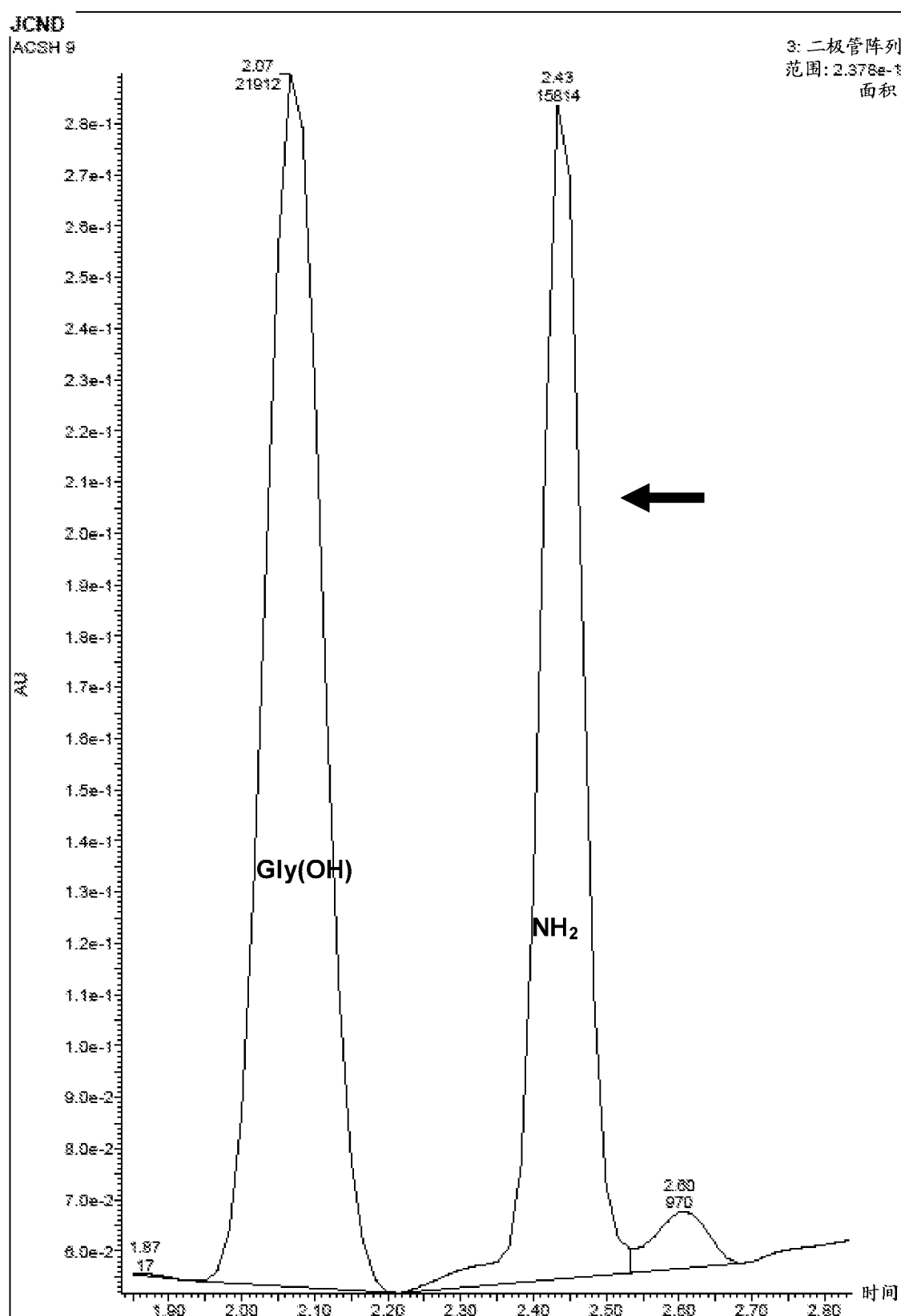
+TAP tag MES pH 5.5

图 5B1/7

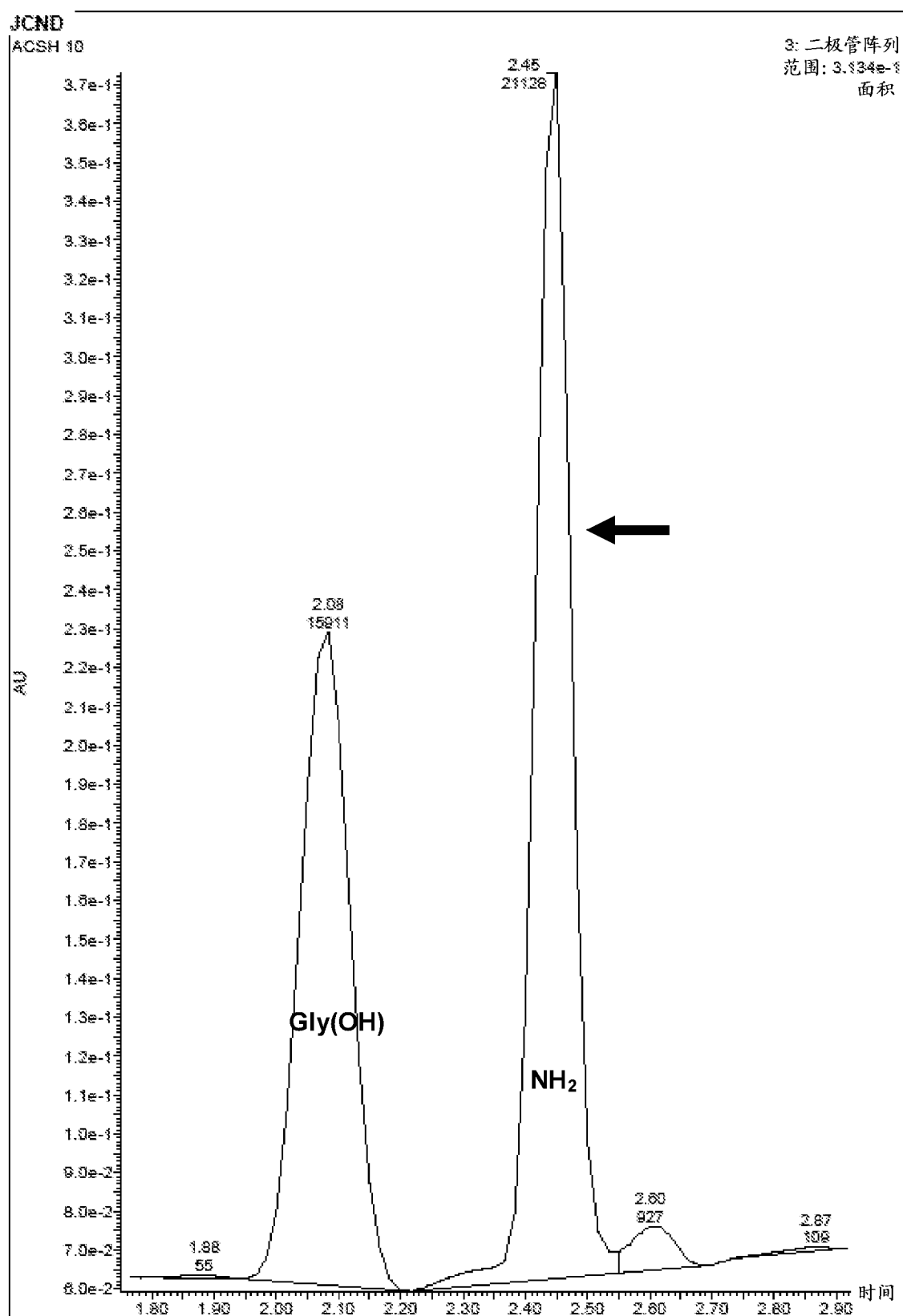
+TAP tag, Tris pH 7.5

图 5B2/7

Neg. Ctrl, MES pH 5.5

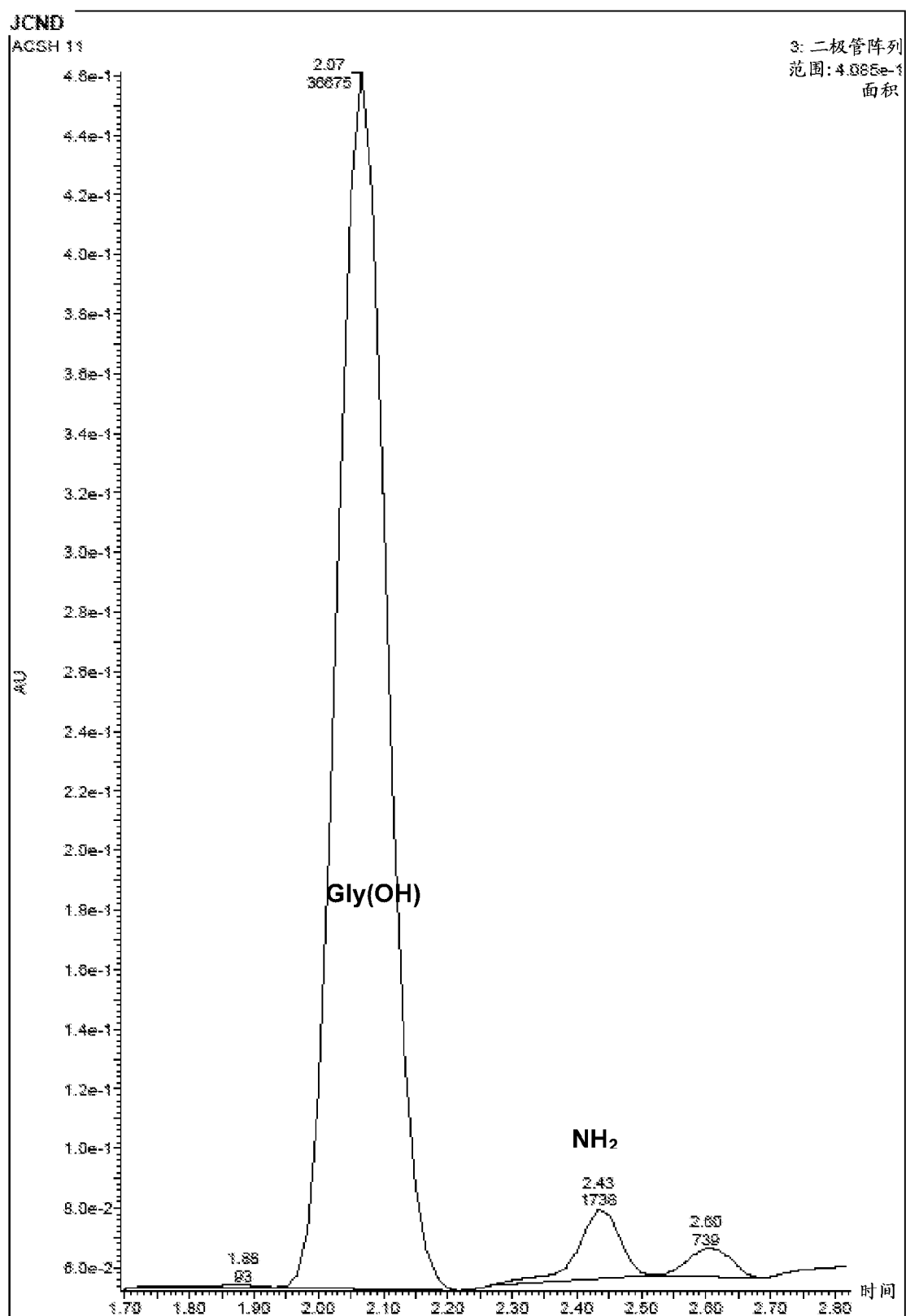


图 5B3/7

Neg. Ctrl, Tris pH 7.5

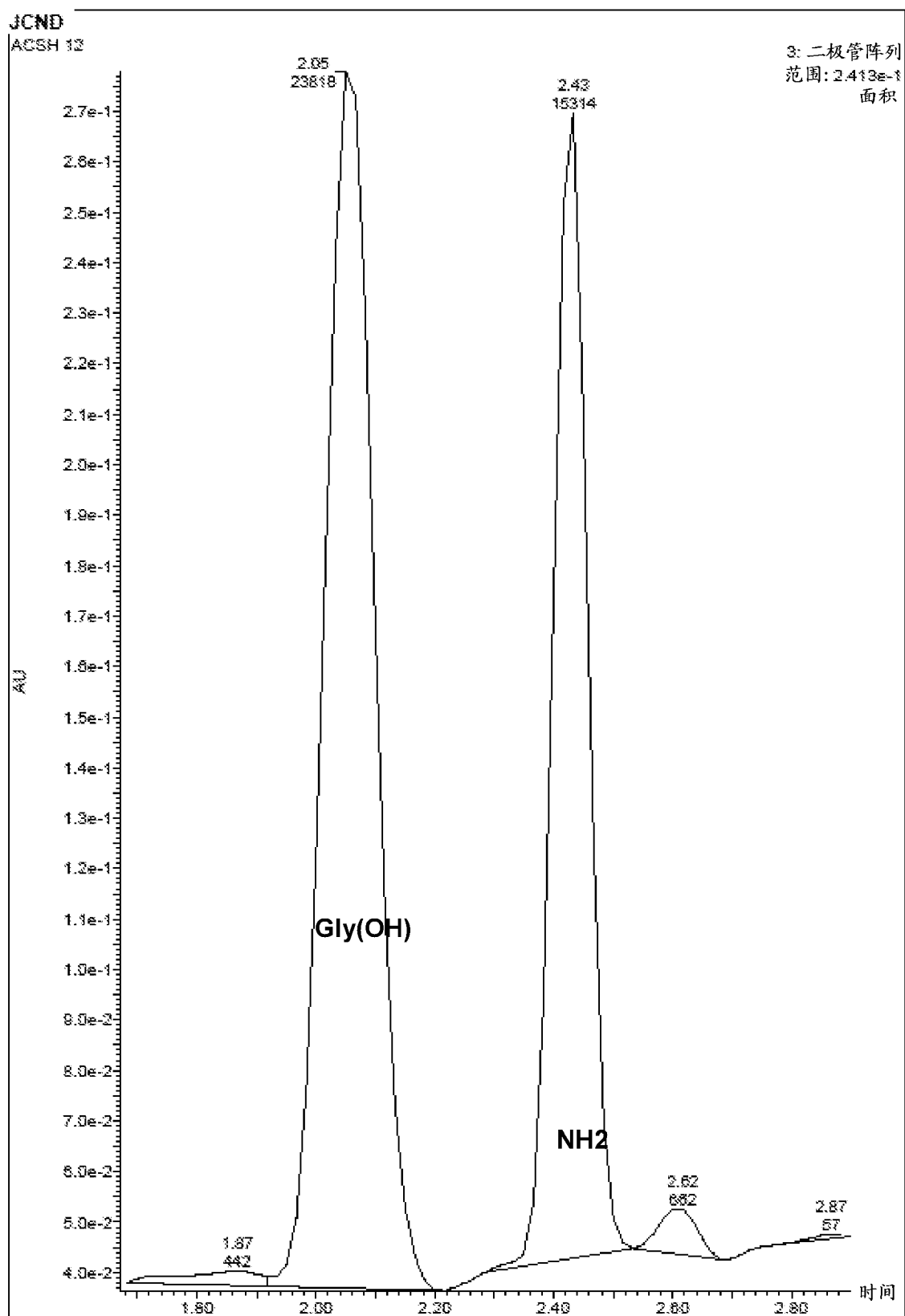


图 5B4/7

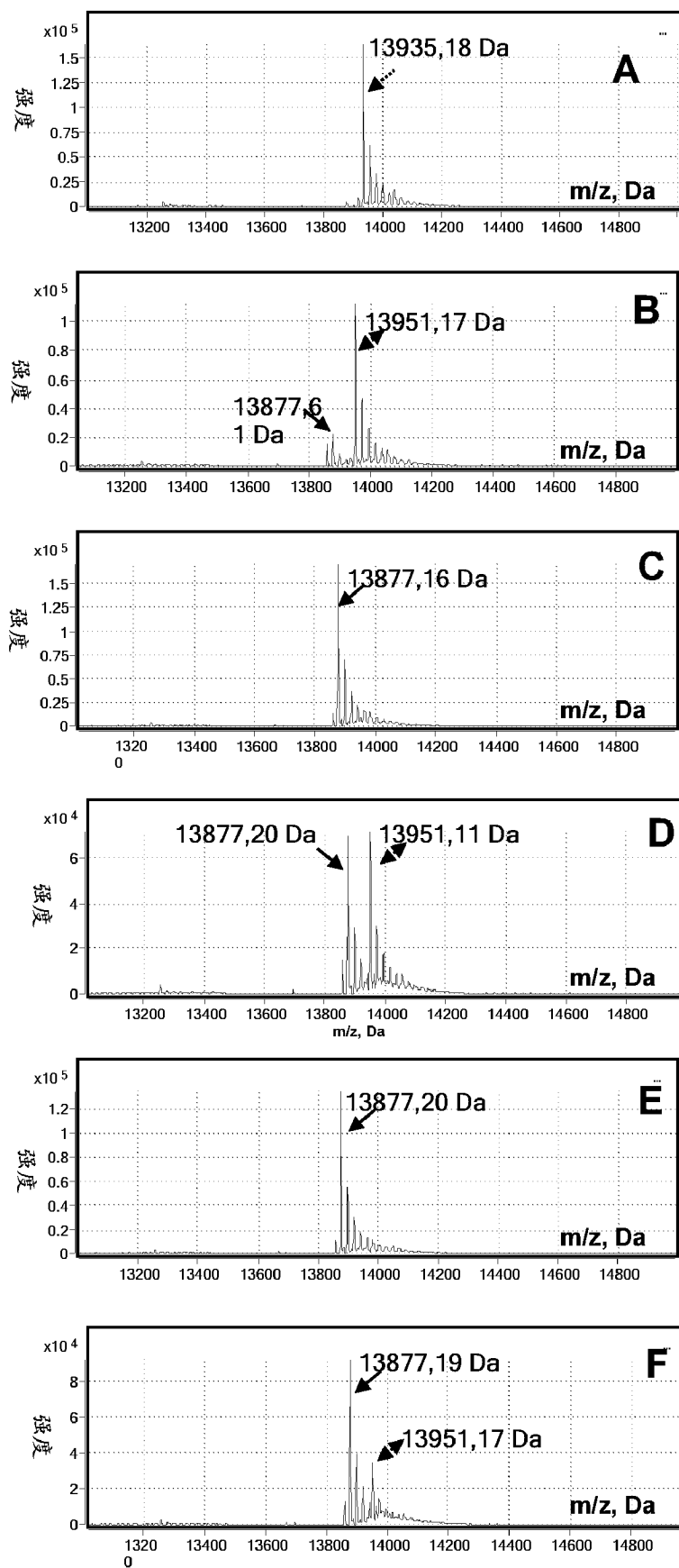
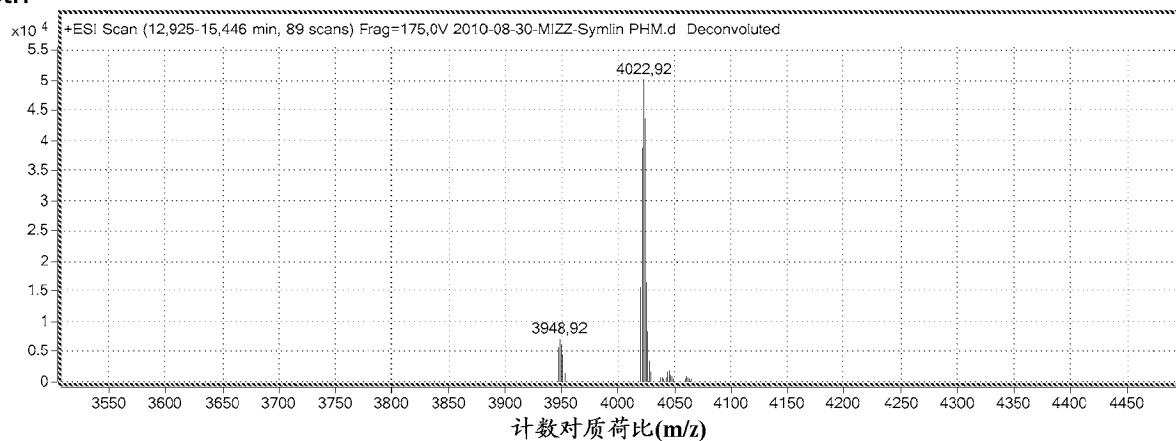


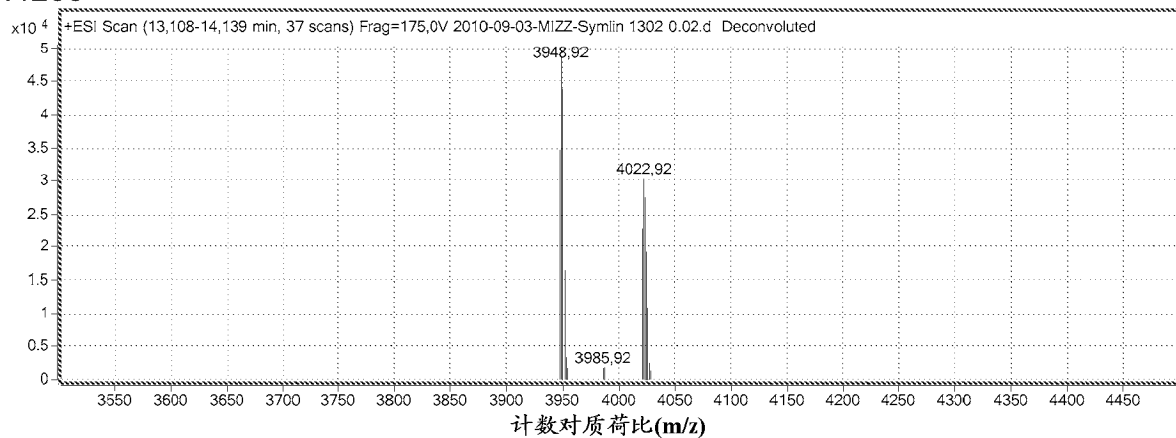
图 6/7

赤细菌属 PAL 样结构域 (SEQ ID NO: 1)

ctrl



1:1250



1:500

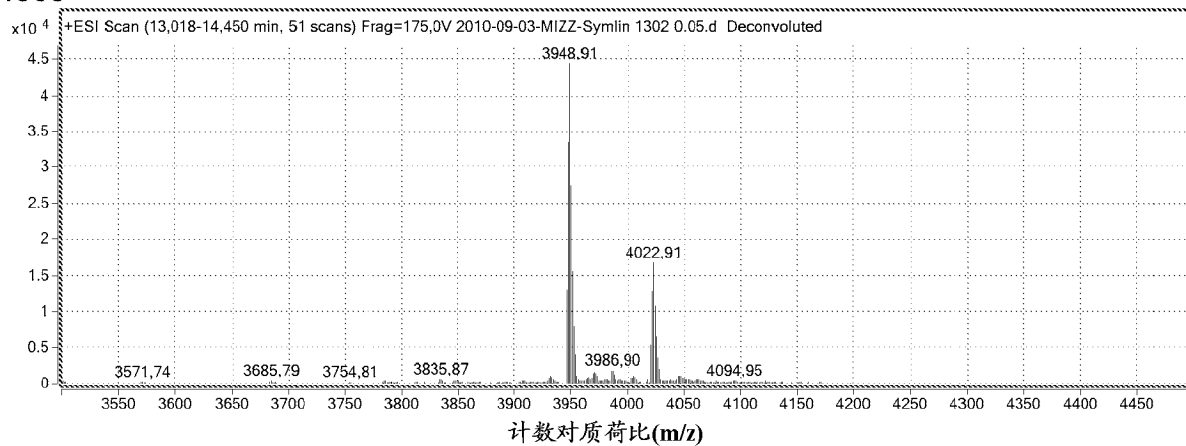


图 7A/7

1:100

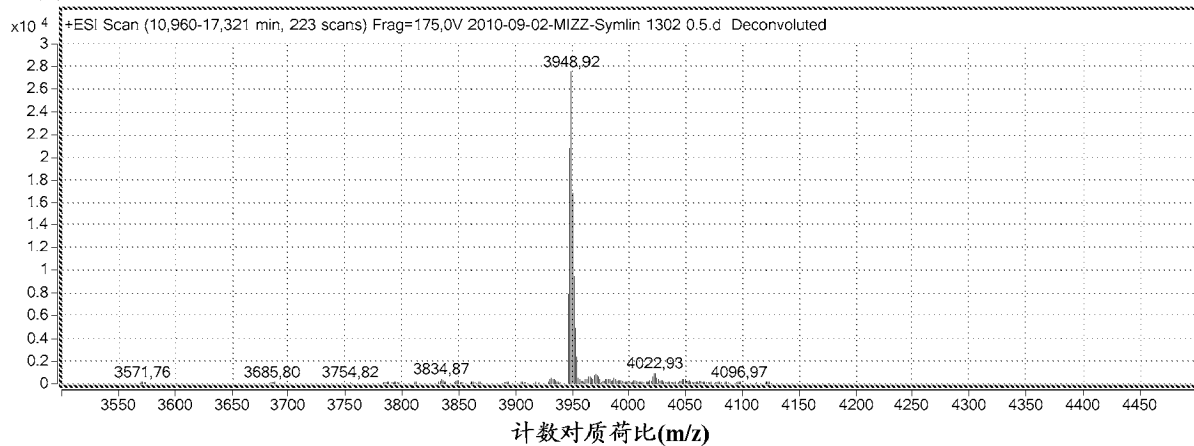
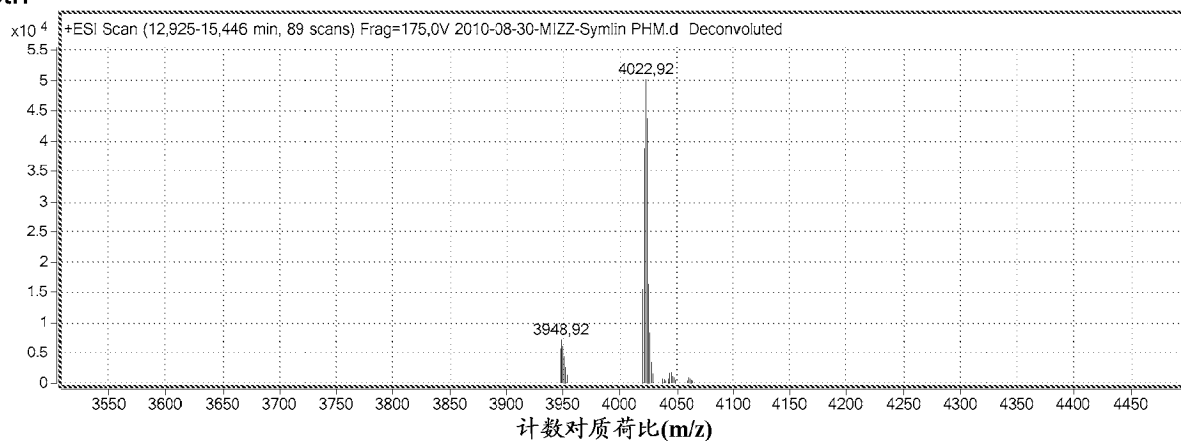


图 7A/7

浮霉菌属 PAL 样结构域 (SEQ ID NO: 4)

ctrl



1:1250

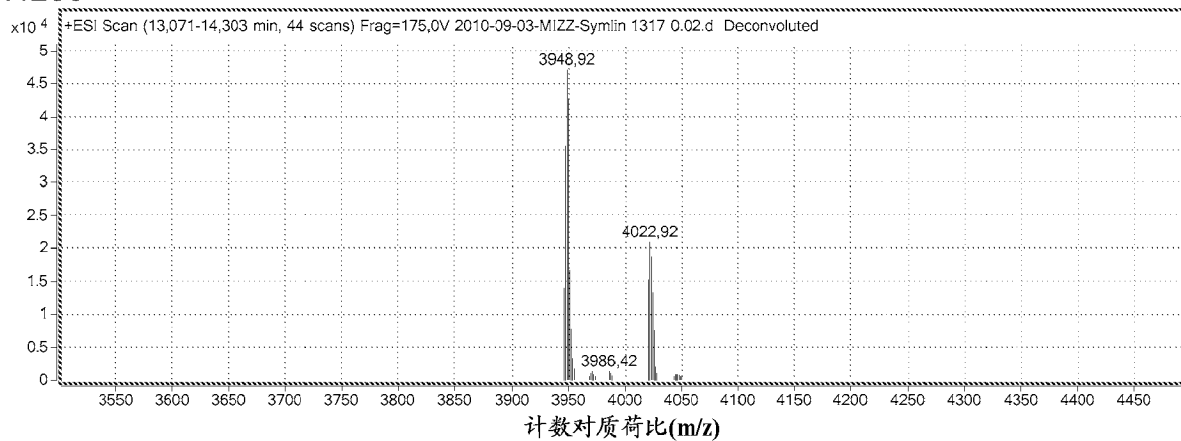
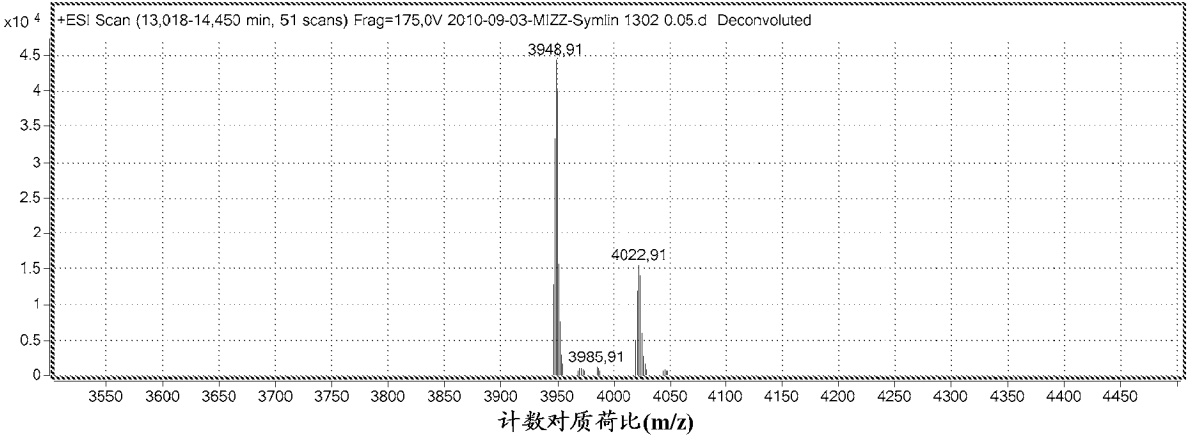


图 7B/7

1:500



1:100

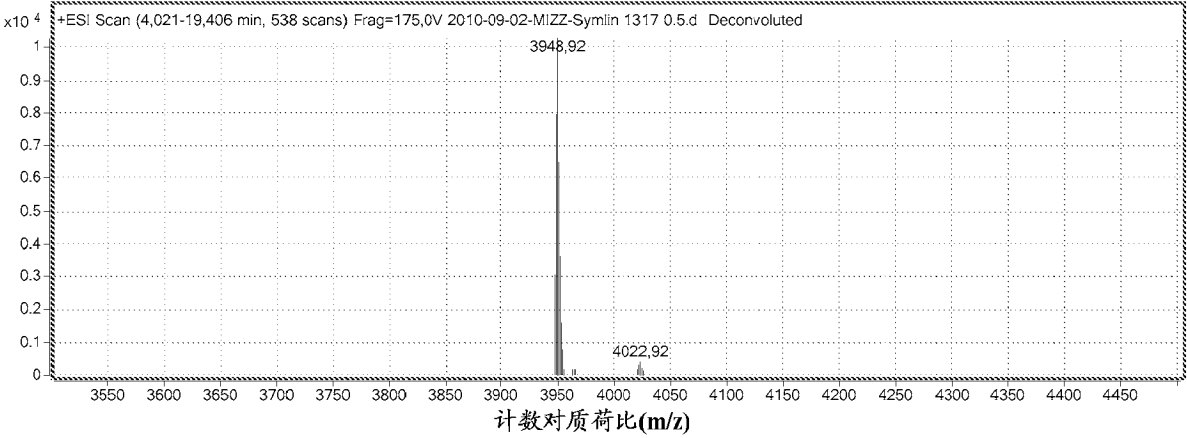


图 7B/7

微小杆菌属 PAL 样结构域 (SEQ ID NO: 2)

ctrl

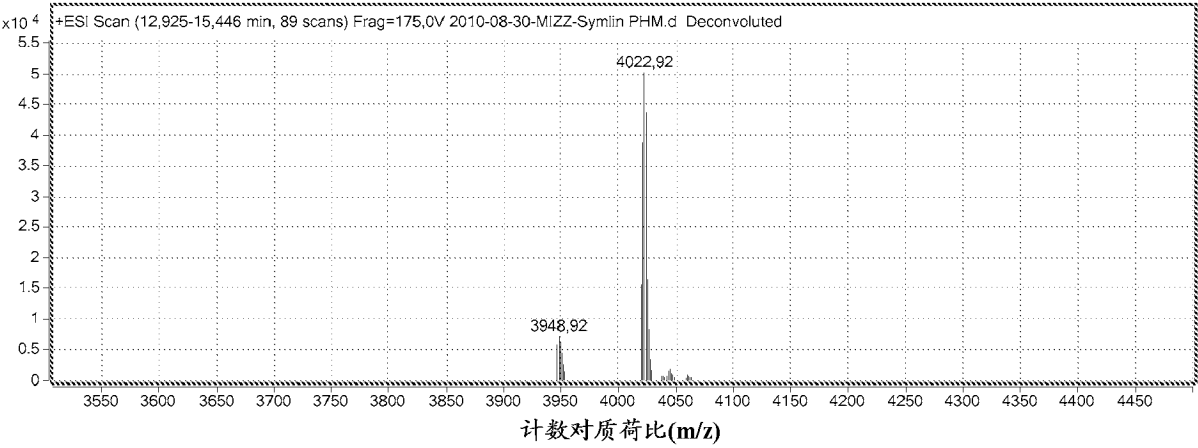
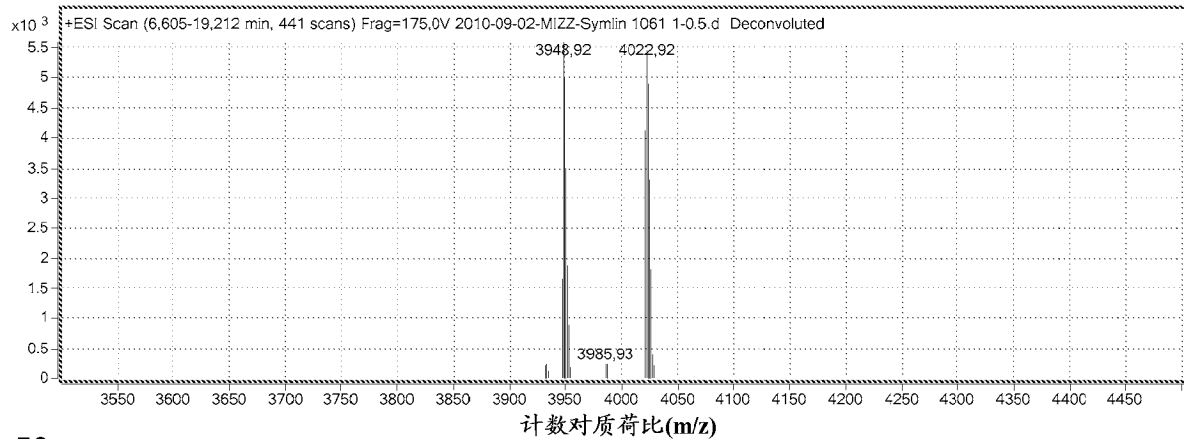
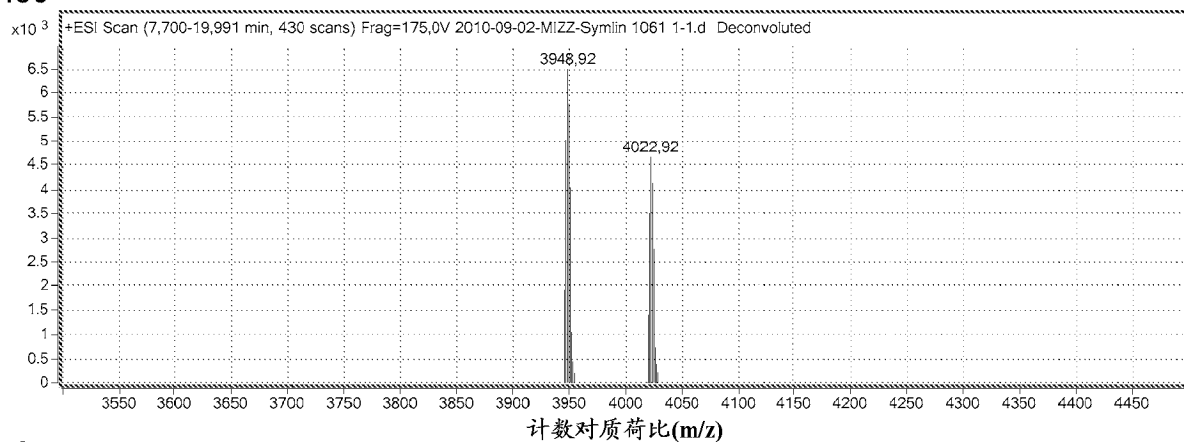


图 7C/7

1:100



1:50



1:25

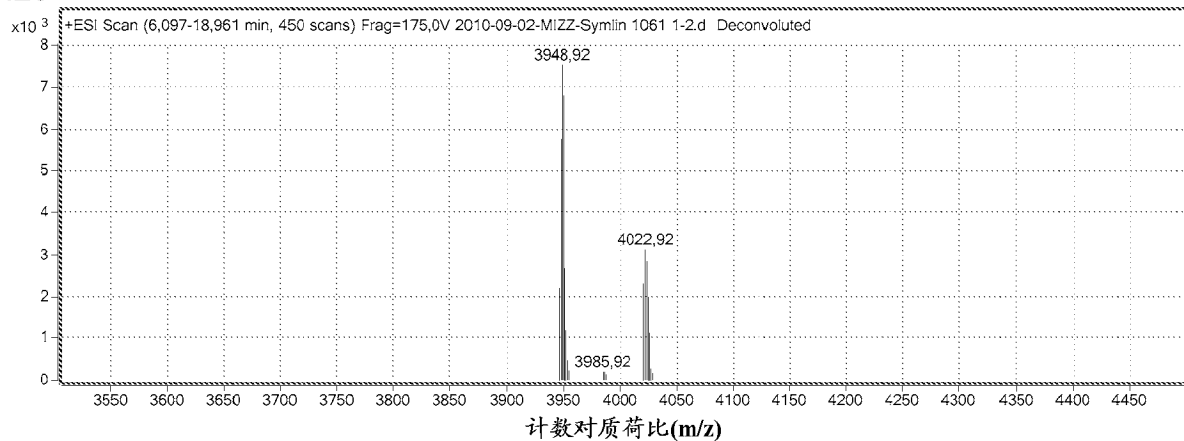
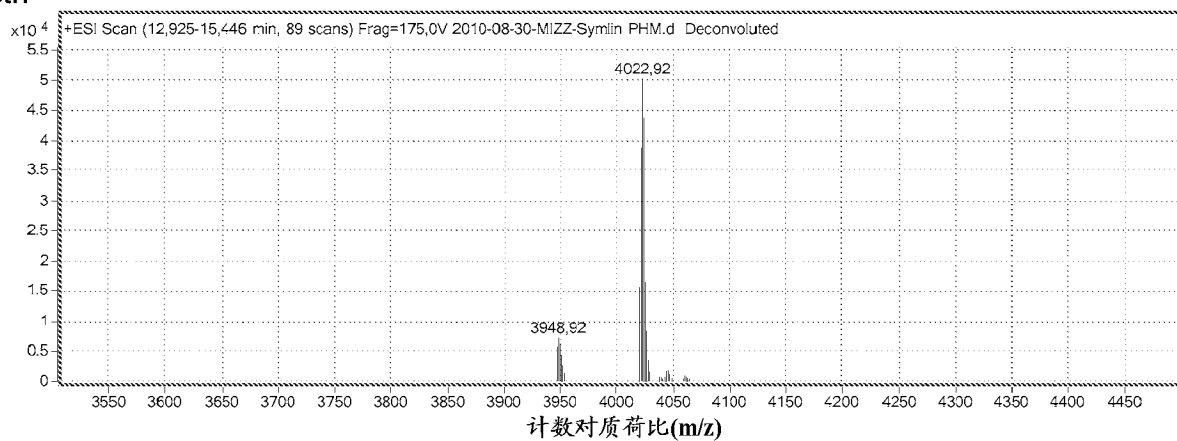


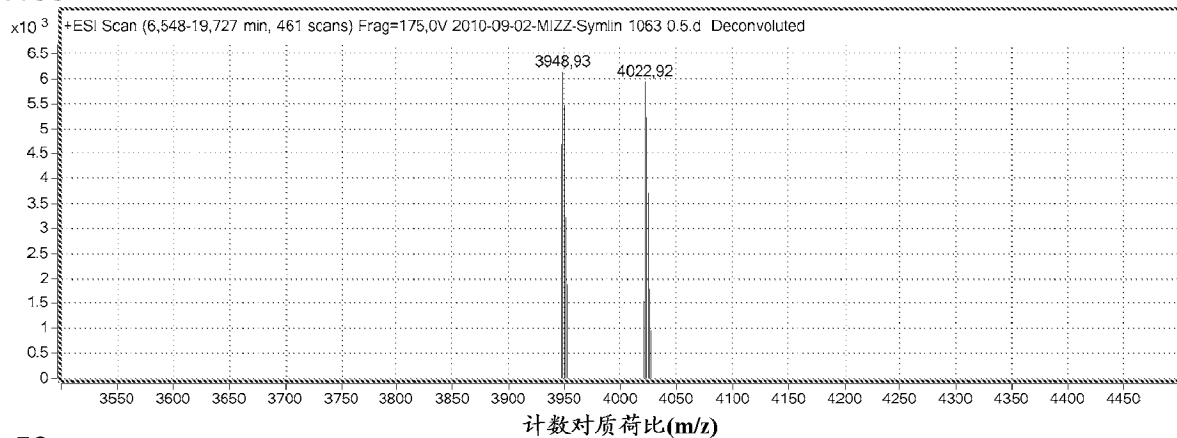
图 7C/7

食土性细菌 PAL 样结构域 (SEQ ID NO: 3)

ctrl



1:100



1:50

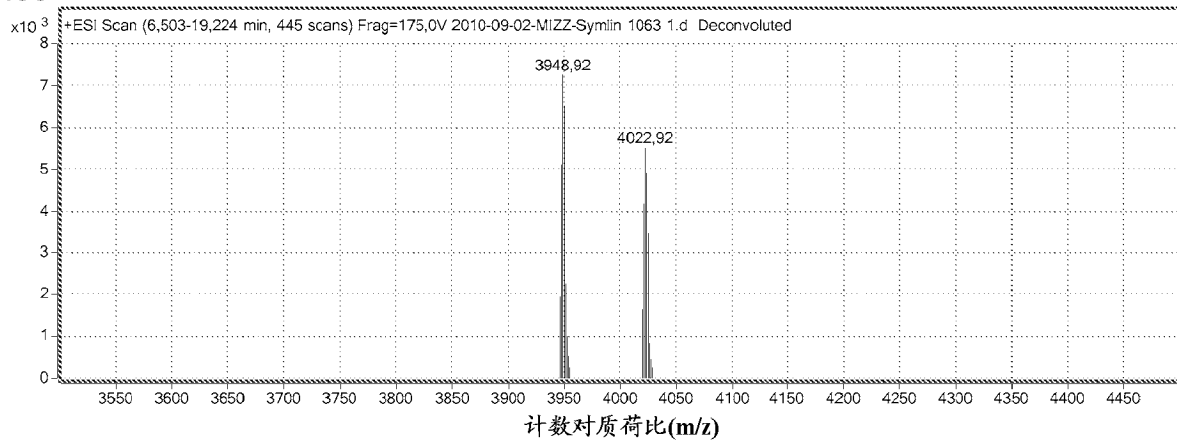


图 7D/7

1:25

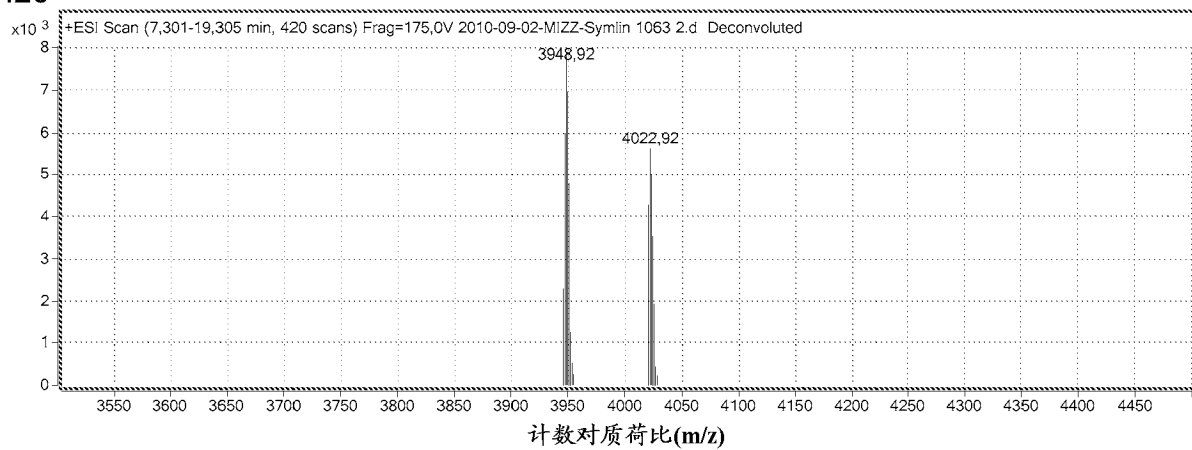


图 7D/7