



(21) 申请号 201811343013.3

(22) 申请日 2014.04.17

(65) 同一申请的已公布的文献号
申请公布号 CN 109486803 A

(43) 申请公布日 2019.03.19

(30) 优先权数据
61/813,586 2013.04.18 US
61/897,932 2013.10.31 US

(62) 分案原申请数据
201480035201.3 2014.04.17

(73) 专利权人 科德克希思公司
地址 美国加利福尼亚州

(72) 发明人 吉伽特·W·哈思曼
尼古拉斯·J·阿加德
本杰明·米杰茨 乔纳森·弗罗姆
张希云

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

专利代理师 王玮玮 郑霞

(51) Int.Cl.
C12N 9/88 (2006.01)

(56) 对比文件
CN 101842482 A, 2010.09.22
CN 102753566 A, 2012.10.24
CN 101803492 A, 2010.08.11
US 2009263369 A1, 2009.10.22
Tse Siang Kang 等. Converting an injectable protein therapeutic into an oral form: phenylalanine ammonia lyase for phenylketonuria. 《Mol Genet Metab.》. 2009, 第991卷第4-9页.

审查员 吴胜

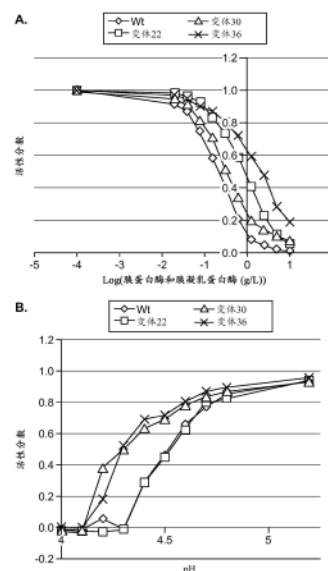
权利要求书4页 说明书103页
序列表48页 附图7页

(54) 发明名称

工程化苯丙氨酸解氨酶多肽

(57) 摘要

本申请涉及工程化苯丙氨酸解氨酶多肽。特别地, 本申请提供了工程化苯丙氨酸解氨酶 (phenylalanine ammonia-lyase, PAL) 多肽及其组合物, 以及编码工程化苯丙氨酸解氨酶 (PAL) 多肽的多核苷酸。在一些实施方案中, 工程化PLA多肽被优化以提供增强的催化活性、以及降低的对蛋白水解的敏感性和增加的对酸性pH水平的耐受性。在一些实施方案中, 工程化PLA多肽被去免疫化 (deimmunized)。本申请还涉及包含工程化PAL多肽的组合物用于治疗 and 工业目的用途。



1. 一种具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽, 所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 A39V、T54K、G59R、S73K、A91V、N290G、R305M、H307G、L407V、C503Q、Q521K、T524S、C565P; 或 A39V、T54K、G59R、S73K、A91V、N290G、R305M、H307G、D353A/N、L407V、C503Q、Q521K、T524S、C565P 与选自下列置换中的 1-5 个氨基酸残基差异的组合: R134Q、Q240W/R、Y304H、E509L 和 P564L/M/G; 或 F18H、A39V、T54K、G59R、S73K、A91V、N290G、R305M、H307G、L407V、C503Q、Q521K、T524S、C565P 与选自下列置换中的 1-5 个氨基酸残基差异的组合: L47A/E、L214Q、E308K/Q、F450Y、T460G、R467G、P528L、E540D 和 S546R; 并且所述工程化多肽与参考序列 SEQ ID NO: 4 相比表现出选自以下的改进的特性: i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、iii) 增加的对酸性 pH 的耐受性、iv) 降低的免疫原性、或 i)、ii)、iii) 和 iv) 的任何组合。

2. 如权利要求 1 所述的工程化多肽, 其中所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 A39V、T54K、G59R、S73K、A91V、Q240W/R、N290G、R305M、H307G、D353A/N、L407V、C503Q、Q521K、T524S、C565P 与选自下列置换中的 1-4 个氨基酸残基差异的组合: R134Q、Y304H、E509L 和 P564L/M/G。

3. 如权利要求 1 所述的工程化多肽, 其中所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 F18H、A39V、T54K、G59R、S73K、A91V、L214Q、N290G、R305M、H307G、L407V、C503Q、Q521K、T524S、C565P 与选自下列置换中的 1-4 个氨基酸残基差异的组合: L47A/E、E308K/Q、F450Y、T460G、R467G、E540D 和 S546R。

4. 一种具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽, 其中所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 A39V、T54K、G59R、S73K、A91V、N290G、R305M、H307G、L407V、C503Q、Q521K、T524S、C565P 与选自 F18H、L47A 和 L214Q 的 1-3 个氨基酸残基差异的组合; 并且所述工程化多肽与参考序列 SEQ ID NO: 4 相比表现出选自以下的改进的特性: i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、iii) 增加的对酸性 pH 的耐受性、iv) 降低的免疫原性、或 i)、ii)、iii) 和 iv) 的任何组合。

5. 如权利要求 4 所述的工程化多肽, 其中所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 A39V、T54K、G59R、S73K、A91V、N290G、R305M、H307G、L407V、C503Q、Q521K、T524S、C565P 与选自 F18H 和 L47A 的 1-2 个氨基酸残基差异的组合。

6. 一种具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽, 所述工程化多肽由以下变体中的任一种的氨基酸序列组成: 说明书表 2-3 中的变体号 43-47、说明书表 2-4 中的变体号 61、说明书表 2-5 中的变体号 1011-1090、说明书表 9-1 中的变体号 100-354、说明书表 9-2 中的变体号 400-416 和 418-420、说明书表 9-3 中的变体号 500-565、说明书表 9-4 中的变体号 600-615、说明书表 9-5 中的变体号 700-774、说明书表 9-6 中的变体号 800-845 以及说明书表 9-7、说明书表 9-8 和说明书表 9-9 中的变体号 900-1010。

7. 一种具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽, 所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 A91V、Y158H 和 S180A; 或所述工程化多肽的氨基酸序列与 SEQ ID NO: 4 相比的氨基酸残基差异为 A39V、A91V、Y158H、S180A、S399N 与选自下列置换的 4-7 个氨基酸残基差异的组合: K195E、T243L、I245L、A256G、N290G、A394V、L407V、Y459F、T463N、N474Q、K522Y/F/N、T524S; 并且所述工程化多肽与参考序列 SEQ ID NO: 4 相比表现出选自以下的改进的特性: i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、

iii) 增加的对酸性pH的耐受性、iv) 降低的免疫原性、或i)、ii)、iii) 和iv) 的任何组合。

8. 一种具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽, 所述工程化多肽的氨基酸序列与SEQ ID NO: 4相比的氨基酸残基差异选自:

- 1) A39V、A91V、Y158H、S180A、N290G、A394V、S399N、N474Q、K522Y和T524S;
- 2) A39V、A91V、Y158H、A256G、A394V、P404A和N474Q;
- 3) A39V、A91V、S180A、A394V、K522F和T524S;
- 4) A39V、A91V、Y158H、T243I、A256G、S399N、P404A、L407V和N474Q;
- 5) E308Q;
- 6) N400K;
- 7) L364Q;
- 8) A256G、N290G、P404A、L407V、N474Q和K522F;
- 9) Y158H、S180A、A394V、T463N、N474Q和T524S;
- 10) A39V、A91V、Y158H、S180A、K195E、A256G、S399N、L407V、Y459F、T463N、K522N和T524S;
- 11) A39V、A91V、Y158H、S180A、K195E、A256G、N290G、S399N、Y459F和T463N;
- 12) A39V、A91V、Y158H、S180A、K195E、T243I和A394V;
- 13) A39V、A91V、Y158H、K195E、T243I、A256G、A394V、S399N、N474Q、K522Y和T524S;
- 14) A39V、A91V、Y158H、S180A、K195E、I245L、S399N、L407V、Y459F、T463N、N474Q和T524S;
- 15) A39V、A91V、Y158H、S180A、T243L、A394V、S399N、T463N和K522F;
- 16) A91V和N474Q;
- 17) S180A和K195E;
- 18) A91V、Y158H、T243I、A256G、N290G、S399N、L407V、Y459F、T463N、N474Q、K522N和T524S;
- 19) A39V、Y158H、S180A和S399N;
- 20) A91V和N270K;
- 21) N453G;
- 22) L257W;
- 23) A91V、Y158H、I245L、A256G、S399N、Y459F、T463N、K522Y和T524S;
- 24) F443H;
- 25) A91V、Y158H和S180A;
- 26) A91V、Y158H、K195E和S399N;
- 27) A39V、A91V、A256G、N290G、A394V、S399N、P404A、L407V、K522Y和T524S;
- 28) L257A;
- 29) Y158H、K195E、T243L、A256G、A394V、S399N、L407V、N474Q、K522F和T524S;
- 30) L349M;
- 31) I326F;
- 32) T460G和P528L;
- 33) A39V、A91V、A256G、N290G、S399N、P404A、L407V和K522Y。

9. 如前述权利要求中任一项所述的工程化多肽, 其中所述改进的特性选自降低的对蛋白水解的敏感性和/或增加的对酸性pH的耐受性。

10. 如权利要求1-8中任一项所述的工程化多肽, 其中所述工程化多肽是多变鱼腥藻酶的变体。

11. 如权利要求1-8中任一项所述的工程化多肽, 其中所述工程化多肽耐受蛋白水解。

12. 如权利要求11所述的工程化多肽, 其中所述工程化多肽耐受通过至少一种消化道酶的蛋白水解。

13. 如权利要求12所述的工程化多肽, 其中所述工程化多肽耐受通过胰凝乳蛋白酶、胰蛋白酶、羧肽酶和/或弹性蛋白酶的蛋白水解。

14. 如权利要求1-8中任一项所述的工程化多肽, 其中所述工程化多肽是酸稳定的。

15. 如权利要求1-8中任一项所述的工程化多肽, 其中所述工程化多肽被去免疫化。

16. 如权利要求15所述的工程化多肽, 其中所述多肽具有与SEQ ID NO: 8、10、12、14、16、18、20、22、24和/或26至少95%的序列同一性。

17. 如权利要求1-8中任一项所述的工程化多肽, 其中所述多肽是纯化的。

18. 一种多核苷酸, 所述多核苷酸编码至少一种如前述权利要求1-17中任一项所述的工程化多肽。

19. 如权利要求18所述的多核苷酸, 其中所述多核苷酸被可操作地连接到控制序列。

20. 如权利要求18或19所述的多核苷酸, 其中所述多核苷酸是密码子优化的。

21. 一种表达载体, 所述表达载体包含至少一种如权利要求18-20中任一项所述的多核苷酸, 和至少一种控制序列。

22. 如权利要求21所述的表达载体, 其中所述控制序列是启动子。

23. 如权利要求22所述的表达载体, 其中所述启动子是异源启动子。

24. 一种宿主细胞, 所述宿主细胞用至少一种如权利要求18-20中任一项所述的多核苷酸, 和/或权利要求21-23中任一项所述的载体来转化, 其中所述宿主细胞不是胚胎干细胞、生殖细胞或植物细胞。

25. 如权利要求24所述的宿主细胞, 其中所述宿主细胞是大肠杆菌(*E. coli*)。

26. 一种在宿主细胞中产生工程化PAL多肽的方法, 所述方法包括在适合的培养条件下培养包含编码至少一种如权利要求1至17中任一项所述的工程化多肽的至少一种多核苷酸、和/或至少一种如权利要求18-20中任一项所述的多核苷酸、和/或至少一种如权利要求21-23中任一项所述的载体的宿主细胞, 使得产生至少一种工程化PAL多肽, 其中所述宿主细胞不是人胚胎干细胞。

27. 如权利要求26所述的方法, 所述方法还包括从培养物和/或所述宿主细胞回收至少一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽。

28. 如权利要求27所述的方法, 所述方法还包括纯化产生的所述至少一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽的步骤。

29. 一种组合物, 所述组合物包含至少一种如权利要求1至17中任一项所述的工程化多肽。

30. 如权利要求29所述的组合物, 其中所述组合物是药物组合物。

31. 如权利要求30所述的组合物, 所述组合物还包含至少一种药学上可接受的赋形剂

和/或载体。

32. 如权利要求29-31中任一项所述的组合物,其中所述组合物适合于苯丙酮尿症的治疗。

33. 如权利要求29-31中任一项所述的组合物,其中所述组合物适合于口服施用至人。

34. 如权利要求30-31中任一项所述的组合物,其中所述组合物呈丸剂、片剂、胶囊、囊形片、液体或乳剂的形式。

35. 如权利要求34所述的组合物,其中所述丸剂、片剂、胶囊、或囊形片还包含肠溶包衣。

36. 如权利要求30-31中任一项所述的组合物,其中所述组合物适合于肠胃外注射入人中。

37. 如权利要求30-31中任一项所述的组合物,其中所述组合物与至少一种另外的治疗有效化合物共施用。

38. 如权利要求37所述的组合物,其中所述组合物包含至少一种另外的治疗有效化合物。

39. 权利要求29-38中任一项所述的组合物在制备用于治疗 and/或预防受试者中苯丙酮尿症的症状的药物中的用途。

40. 如权利要求39所述的用途,其中所述苯丙酮尿症的症状被改善。

41. 如权利要求39或40所述的用途,其中所述受试者能够食用在其甲硫氨酸、苯丙氨酸和/或酪氨酸含量方面比由还未被提供包含至少一种如权利要求1-17中所阐述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽的至少一种药物组合物的受试者所需要的饮食较不受限的饮食。

42. 如权利要求39或40所述的用途,其中所述受试者是婴儿或儿童。

43. 如权利要求41所述的用途,其中所述受试者是婴儿或儿童。

44. 如权利要求39或40所述的用途,其中所述受试者是成人或青年。

45. 如权利要求41所述的用途,其中所述受试者是成人或青年。

工程化苯丙氨酸解氨酶多肽

[0001] 本申请是申请日为2014年4月17日,申请号为201480035201.3,发明名称为“工程化苯丙氨酸解氨酶多肽”的申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求2013年4月18日提交的美国临时专利申请系列号61/813,586和2013年10月31日提交的美国临时专利申请系列号61/897,932的优先权,两者都通过引用以其整体被并入用于全部目的。

[0004] 对作为ASCII文本文件提交的“序列表”、表格,或计算机程序列表附件的引用

[0005] 写入文件CX7-131W02_ST25.TXT中,创建日期为2014年4月15日,127,412个字节,机器格式IBM-PC,MS-Windows操作系统的序列表在此通过引用并入。

发明领域

[0006] 本发明提供了工程化苯丙氨酸解氨酶(PAL)多肽及其组合物,以及编码工程化苯丙氨酸解氨酶(PAL)多肽的多核苷酸。在一些实施方案中,工程化PAL多肽被优化以提供增强的催化活性,以及降低的对蛋白水解的敏感性和增加的对酸性pH水平的耐受性。在一些实施方案中,工程化PAL多肽被去免疫化。本发明还涉及包含工程化PAL多肽的组合物用于治疗 and 工业目的的用途。

[0007] 发明背景

[0008] 苯丙氨酸解氨酶(PAL)连同组氨酸解氨酶(HAL)和酪氨酸解氨酶(TAL)一起为芳香族氨基酸裂解酶家族(EC 4.3.1.23-1.25和4.3.1.3)的成员。更具体地,具有PAL活性的酶(EC 4.3.1.23-1.25和先前分类为EC4.3.1.5)催化L-苯丙氨酸的非氧化性脱氨基为(E)-肉桂酸。PAL为非哺乳动物酶,其广泛分布于植物中并且还在真菌和有限数目的细菌中被鉴定出。PAL酶可以被用于用于治疗代谢紊乱苯丙酮尿症(PKU)的治疗性蛋白。PKU是常染色体代谢遗传紊乱,其中肝酶苯丙氨酸羟化酶(PAH)或参与辅因子四氢生物蝶呤的合成或再循环的一种或更多种酶,由于在相应基因的一个中的突变而无功能。这种功能的缺乏导致血流中高水平的苯丙氨酸。苯丙氨酸被转化为苯丙酮酸盐(苯基酮)和其他衍生物。在人中,如果PKU未及早治疗,高水平的苯丙氨酸和其分解产物的部分可引起显著医学问题,所述医学问题包括智力障碍、小头畸形和癫痫发作。许多研究都集中在通过酶替代将PAL用于治疗PKU(Ambrus等人,Science 201:837-839[1978];Bourget等人,Appl.Biochem.Biotechnol., 10:57-59[1984];和Sarkissian等人,Proc.Natl.Acad.Sci.USA 96:2339-2344[1999])。

[0009] 除去血流中苯丙氨酸的一种方法是使用可注射的重组PAL和通过聚乙二醇化修饰的PAL变体(PEG-PAL)。聚乙二醇化已被证明改进酶的半衰期并降低受试者抗原反应(参见例如,WO 2008/153776、WO 2011/097335和美国专利号7,531,341)。在PEG-PAL组合物中有用的PAL变体已经被描述为野生型点状念珠藻(Nostoc punctiforme)(NpPAL);多变鱼腥藻(Anabaena variabilis)(AvPAL)和圆红冬孢酵母(Rhodospiridium toruloides)(RtPAL)的变体。特别地,已描述了野生型AvPAL的变体,其中在位置64、318、503和565处的半胱氨酸残基已用丝氨酸置换(参见例如,美国专利号:7,790,433;7,560,263;和7,537,923)。

[0010] PAL施用作为减少PKU受试者中的L-苯丙氨酸的血浆浓度的方法的替代方案是非侵入性制剂,诸如口服制剂(Sarkissian等人,Proc.Natl.Acad.Sci.USA 96:2339-2344 [1999])。口服递送PAL的关键优势是酶对免疫系统的减少的暴露,从而最小化用可注射PEG-PAL观察到的免疫反应。然而,关于PAL的口服制剂的主要限制是酶活性在胃和肠腔中的损失。为了有效而实用,PAL必须耐受酸性pH和蛋白酶的降解,所述蛋白酶诸如通常降解蛋白食物为寡肽和氨基酸的胰蛋白酶、胰凝乳蛋白酶、羧肽酶和胃蛋白酶。在一些以前的研究中(Sarkissian,同上),部分由于通过蛋白酶的酶降解且部分由于在pH 7.0下相对低的比活性,为了实现口服施用PAL的显著效应,需要大量的酶。已经开发多种方法以抑制消化后的PAL降解(Kim等人,Molec.Therap.,10:220-224 [2004];和Shah等人,Int.J.Pharmaceut,356:61-68 [2008])。

[0011] 提高PAL在消化道的苛刻条件下的效力的一种方法是提供耐受固有的苛刻条件的工程化PAL多肽。Kang等人使用胰凝乳蛋白酶裂解位点的定点诱变和AvPAL的表面赖氨酸的聚乙二醇化以减少蛋白水解失活(参见,Kang等人,Mol.Gen.Metabol.,99:4-9 [2010])。在这些研究中,特异性突变了十个裂解位点,并且这些所得突变体除了两个(F18A和R94G)以外的所有都损失多于50%的原始酶活性。没有突变体显示出增加的活性,并且F18A突变体显示胰蛋白酶耐受性的轻微增加(Kang等人,同上)。用PAL的另外的研究,虽然有效,但通常没有导致较长存活的酶。因此,先前描述的PAL突变体及其衍生物的口服施用不会导致有效治疗PKU。

[0012] 尽管PAL的多种制剂取得了进展,对具有改进的特性的用于口服施用的PAL多肽仍存在需要。这些改进的特性包括但不限于较长的半衰期、增加的催化活性、改进的对消化道中条件的稳定性和减少的聚集。

[0013] 除了治疗应用以外,PAL酶还可用于工业合成L-苯丙氨酸和其他取代的L-苯丙氨酸衍生物。然后,这些衍生物可用作药物前体(Gloge等人,6:3386-3390 [2000];Bartsch等人,Prot.Eng.Des.Sel.,23:929-933 [2010];和Turner,Curr.Opin.Chem.Biol.,234-240 [2011])。

[0014] PAL酶还可用于农业应用。PAL在植物、真菌和细菌中的苯丙素(phenylpropanoid)(诸如类黄酮和木质素)生物合成中起着重要作用,并且可以用作防御相关酶(Bate等人,Proc.Natl.Acad.Sci.USA 91:7608-7612 [1994])。通过使用具有PAL活性的重组多肽的PAL活性的调节可潜在地导致有效的除草剂。

[0015] 发明概述

[0016] 本发明提供了工程化苯丙氨酸解氨酶(PAL)多肽及其组合物,以及编码工程化苯丙氨酸解氨酶(PAL)多肽的多核苷酸。在一些实施方案中,工程化PAL多肽被优化以提供增强的催化活性,以及降低的对蛋白水解的敏感性和增加的对酸性pH水平的耐受性。在一些实施方案中,工程化PAL多肽被去免疫化。本发明还涉及包含工程化PAL多肽的组合物用于治疗 and 工业目的的用途。在一些实施方案中,本发明涉及具有改进的特性的工程化苯丙氨酸解氨酶(PAL)多肽及其生物活性片段和类似物,所述改进的特性诸如增加的对酸性pH的耐受性和/或降低的对蛋白水解的敏感性。

[0017] 本发明涉及当在基本相同的条件下与野生型PAL酶或参考PAL多肽相比时具有改进的特性的工程化PAL多肽及其生物活性片段和类似物。本发明还涉及在治疗性和/或工业

组合物中使用工程化PAL多肽及其生物活性片段和类似物的方法以及使用此类组合物用于治疗 and/或工业目的的方法。

[0018] 在第一个方面,本发明提供了工程化苯丙氨酸解氨酶(PAL)多肽,其中当在基本相同的条件下测量时,所述工程化苯丙氨酸解氨酶(PAL)多肽与参考序列相比具有选自以下的组的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、或i)、ii)、iii)、或iv)的任何组合。在一些具体实施方案中,工程化PAL多肽具有两种改进的特性。在其他具体实施方案中,改进的特性是降低的对蛋白水解的敏感性,并且在又其他具体实施方案中,改进的特性是增加的对酸性pH的耐受性。

[0019] 在第二个方面,工程化PAL多肽包括蛋白,所述蛋白包含与SEQ ID NO:4或其功能片段至少85%的氨基酸序列同一性,和当与SEQ ID NO:4的多肽最佳比对时在对应于以下位置的位置处的氨基酸残基差异:X39;X91;X158;X180;X195;X243;X245;X256;X257;X270;X290;X307;X308;X326;X349;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X522;X524;和X528。

[0020] 在第一个和第二个方面的一些具体实施方案中,当与SEQ ID NO:4的多肽最佳比对时,工程化PAL多肽包含对应于以下的一个或多个氨基酸残基位置处的至少一个氨基酸残基差异:A39;A91;Y158;S180;K195;T243;I245;A256;L257;N270;N290;H307;E308;I326;L349;L364;A394;S399;N400;P404;L407;F443;N453;Y459;T460;T463;N474;K522;T524;和P528。在一些具体实施方案中,工程化PAL多肽包含与包含SEQ ID NO:4的氨基酸序列的参考多肽的至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少15个和至少20个氨基酸残基差异。

[0021] 在第一个和第二个方面的其他具体实施方案中,工程化PAL多肽包含与SEQ ID NO:4至少90%(至少91%、92%、93%、94%、95%、96%、97%、98%和99%)的氨基酸序列同一性。在又另外的具体实施方案中,工程化PAL多肽包含与SEQ ID NO:4至少90%(至少91%、92%、93%、94%、95%、96%、97%、98%和99%)的氨基酸序列同一性,并且包含以下置换中的一个或多个:A39V;A91V;Y158H;S180A;K195E;T243I/L;I245L;A256G;L257W/A;N270K;N290G;H307G/Q/M;E308Q;I326F;L349M;L364Q;A394V;S399N;N400K;P404A;L407V;F443H;N453G;Y459F;T460G;T463N;N474Q;K522Y/F/N;T524S;和P528L。

[0022] 在其他具体实施方案中,工程化PAL多肽源自野生型多变鱼腥藻PAL(诸如ATCC29413;NCBI蛋白参考序列YP_324488.1;SEQ ID NO:4)。

[0023] 在第三个方面,被本发明包括的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:10或其功能片段至少99%的序列同一性的氨基酸序列。

[0024] 在第四个方面,被本发明包括的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:10或其功能片段至少95%的序列同一性且还包含与SEQ ID NO:10相比在一个、两个、三个、四个、五个或六个或多个氨基酸位置处的氨基酸残基差异的氨基酸序列。

[0025] 在第五个方面,本发明提供了多核苷酸序列,所述多核苷酸序列编码如本文描述的工程化PAL多肽中的任何一种。

[0026] 在第六个方面,本发明提供了药物组合物或工业组合物,所述药物组合物或工业

组合物包含如本文描述的工程化PAL多肽中的任何一种。

[0027] 在一些实施方案中,本发明提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:4或其功能片段至少85%的序列同一性;b)与SEQ ID NO:4或其功能片段相比在一个或多个氨基酸位置处的氨基酸残基差异;并且c)所述工程化多肽与参考序列相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、或i)、ii)、iii)、或iv)的任何组合。在一些实施方案中,当与SEQ ID NO:4的氨基酸序列最佳比对时,一个或多个氨基酸位置选自:X39;X54;X59;X73;X91;X158;X112,X134,X180;X195;X240;X243;X245;X256;X257;X270;X290;X304,X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;X564;和/或其组合。在一些另外的实施方案中,改进的特性选自减少的对蛋白水解的敏感性和/或增加的对酸性pH的耐受性。在又另外的实施方案中,参考序列是源自多变鱼腥藻的野生型PAL。在一些另外的实施方案中,参考序列SEQ ID NO:4的氨基酸残基对应于A39;T54;G59,S73;A91;Y158;S180;K195;A112;R134;Q240;T243;I245;A256;L257;N270;N290;Y304;R305;H307;E308;I326;L349;D353;L364;A394;S399;N400;P404;L407;F443;N453;Y459;T460;T463;N474;E509;Q521;K522;T524;P528;S546;和/或P564。在一些实施方案中,当与SEQ ID NO:4的多肽最佳比对时,与SEQ ID NO:4相比的氨基酸残基差异选自以下置换中的一个或多个:A39V;T54K;G59R;S73K;A112C;R134Q;A91V;Y158H;S180A;K195E;Q240R/W;T243I/L;I245L;A256G;L257W/A;N270K;N290G;Y304H;R305M;H307G/Q/M;E308Q;I326F;L349M;D353A/N;L364Q;A394V;S399N;N400K;P404A;L407V;F443H;N453G;Y459F;T460G;T463N;N474Q;E509L;Q521K/S;K522Y/F/N;T524S;P528L;S546R;和P564G/L/M。在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%、或约100%的序列同一性。在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%的序列同一性。在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少约90%的序列同一性。在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少约95%的序列同一性。在一些另外的实施方案中,工程化多肽具有与SEQ ID NO:4至少约90%的序列同一性;和在位置H307处的氨基酸残基差异。在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少90%的序列同一性。在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少95%的序列同一性。在一些另外的实施方案中,工程化多肽具有与SEQ ID NO:4至少90%的序列同一性;和在位置H307处的氨基酸残基差异。在一些另外的实施方案中,氨基酸残基差异为H307G/Q/M。在一些另外的实施方案中,氨基酸残基差异选自以下的一个或多个的组合:A39;A91;Q240;A256;N290;Y304;R305;H307;D353A394;S399;P404;L407;Q521;K522;和T524。

[0028] 本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约

99%、或至少约100%的序列同一性的氨基酸序列。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少约95%的序列同一性的氨基酸序列。

[0029] 在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少约99%的序列同一性的氨基酸序列。

[0030] 本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少约至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或至少100%的序列同一性的氨基酸序列。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少约95%的序列同一性的氨基酸序列。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少约99%的序列同一性的氨基酸序列。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少95%的序列同一性的氨基酸序列。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14或其功能片段至少99%的序列同一性的氨基酸序列。

[0031] 本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与SEQ ID NO:4或其功能片段至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或至少约100%的序列同一性的氨基酸序列,其中该工程化多肽被去免疫化。本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与SEQ ID NO:4或其功能片段至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%的序列同一性的氨基酸序列,其中该工程化多肽被去免疫化。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包含具有与SEQ ID NO:4或其功能片段至少95%的序列同一性的氨基酸序列,其中该工程化多肽被去免疫化。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽是在表9-1至9-7的任一个中提供的变体PAL。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的去免疫化的工程化多肽包含具有与SEQ ID NO:6、8、10、12和/或14至少95%序列同一性的氨基酸序列。

[0032] 本发明还提供了多核苷酸序列,所述多核苷酸序列编码本文提供的具有PAL活性的工程化多肽。在一些实施方案中,多核苷酸序列被可操作地连接到控制序列。本发明还提供了载体,所述载体包含编码具有PAL活性的至少一种工程化多肽的至少一种多核苷酸序列。本发明还提供了宿主细胞,所述宿主细胞用编码如本文所提供的具有PAL活性的工程化多肽的至少一种多核苷酸序列转化。

[0033] 本发明还提供了在宿主细胞中产生工程化PAL多肽的方法,所述方法包括在适合的培养条件下培养包含编码至少一种工程化PAL多肽的至少一种多核苷酸的宿主细胞。本发明还提供了在宿主细胞中产生工程化PAL多肽的方法,所述方法包括在适合的培养条件下培养包含编码工程化PAL多肽的多核苷酸的宿主细胞。在一些实施方案中,该方法还包括从培养物和/或宿主细胞中回收工程化PAL多肽。

[0034] 本发明还提供了组合物,所述组合物包含至少一种如本文所提供的具有PAL活性

的工程化多肽。在一些实施方案中,组合物为药物组合物,所述药物组合物还包含药学上可接收的载体。本发明还提供了这些组合物的用途。

[0035] 本发明还提供了具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽,所述工程化多肽包含:a) 氨基酸序列,所述氨基酸序列具有与具有苯丙氨酸解氨酶 (PAL) 活性的参考序列或其功能片段至少约85%、至少约86%、至少约87%、至少约88%、至少约89%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或更大的序列同一性;b) 多肽序列,所述多肽序列包含与具有苯丙氨酸解氨酶 (PAL) 活性的参考序列或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c) 所述工程化多肽与具有苯丙氨酸解氨酶 (PAL) 活性的参考序列相比表现出选自以下的改进的特性:i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、iii) 增加的对酸性pH的耐受性、iv) 减少的聚集、v) 降低的免疫原性、或i)、ii)、iii)、iv) 或v) 的任何组合。在一些实施方案中,参考序列为原核PAL,而在一些其他实施方案中,参考序列为真核PAL。在一些实施方案中,参考序列为细菌PAL(例如,多变鱼腥藻PAL),而在一些其他实施方案中,它是人或其他PAL。在一些另外的实施方案中,参考序列为野生型序列(例如,野生型多变鱼腥藻PAL),而在一些可选的实施方案中,参考序列为变体酶(例如,具有PAL活性的工程化多肽)。

[0036] 在一些实施方案中,本发明的具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽包含:a) 氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:4或其功能片段至少约85%、至少约86%、至少约87%、至少约88%、至少约89%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或更大的序列同一性;b) 多肽序列,所述多肽序列包含与SEQ ID NO:4或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c) 所述工程化多肽与参考序列SEQ ID NO:4相比表现出选自以下的改进的特性:i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、iii) 增加的对酸性pH的耐受性、iv) 减少的聚集、v) 降低的免疫原性、或i)、ii)、iii)、iv) 或v) 的任何组合。

[0037] 在一些另外的实施方案中,具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽包含:a) 氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:4或其功能片段至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的序列同一性;b) 多肽序列,所述多肽序列包含与SEQ ID NO:4或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c) 所述工程化多肽与参考序列SEQ ID NO:4相比表现出选自以下的改进的特性:i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、iii) 增加的对酸性pH的耐受性、iv) 减少的聚集、v) 降低的免疫原性、或i)、ii)、iii)、iv) 或v) 的任何组合。

[0038] 在一些实施方案中,具有苯丙氨酸解氨酶活性 (PAL) 的工程化多肽包含在以下氨基酸位置的一个或更多个处的至少一个置换:20、24、27、39、43、45、47、54、58、59、62、70、73、80、82、91、94、98、104、105、110、112、115、117、118、119、121、123、124、125、126、127、128、129、130、131、133、134、135、139、140、141、142、143、144、145、146、147、149、150、151、153、154、156、157、158、159、172、174、175、176、177、178、180、187、191、195、199、205、206、210、212、213、214、232、240、243、245、247、248、250、256、257、266、270、275、278、279、285、

286、289、290、292、304、305、307、308、309、319、321、326、331、332、334、349、353、355、364、365、369、370、371、372、374、375、377、378、379、381、382、383、384、385、387、389、394、396、399、400、403、404、407、417、418、425、431、432、433、434、435、436、437、438、439、443、446、447、453、456、459、460、461、463、471、472、473、474、475、476、477、478、479、482、483、503、507、509、521、522、524、525、528、538、546、547、551、558、560、564、565和/或其任何组合,其中氨基酸位置参考SEQ ID NO:4来编号。在一些实施方案中,参考序列SEQ ID NO:4的氨基酸残基对应于A39、T54、G59、S73、A91、Y158、S180、K195、A112、R134、Q240、T243、I245、A256、L257、N270、N290、Y304、R305、H307、E308、I326、L349、D353、L364、A394、S399、N400、P404、L407、F443、N453、Y459、T460、T463、N474、E509、Q521、K522、T524、P528、S546和/或P564。在一些另外的实施方案中,当与SEQ ID NO:4的多肽最佳比对时,与SEQ ID NO:4相比的氨基酸残基差异选自以下置换中的一个或更多个:A39V、T54K、G59R、S73K、A112C、R134Q、A91V、Y158H、S180A、K195E、Q240R/W、T243I/L、I245L、A256G、L257W/A、N270K、N290G、Y304H、R305M、H307G/Q/M、E308Q、I326F、L349M、D353A/N、L364Q、A394V、S399N、N400K、P404A、L407V、F443H、N453G、Y459F、T460G、T463N、N474Q、E509L、Q521K/S、K522Y/F/N、T524S、P528L、S546R和P564G/L/M。在一些另外的实施方案中,工程化多肽具有与SEQ ID NO:4至少约90%的序列同一性;和在位置H307处的氨基酸残基差异。在一些实施方案中,氨基酸残基差异为H307G/Q/M。在一些仍然另外的实施方案中,氨基酸残基差异选自以下的一个或更多个的组合:A39、A91、Q240、A256、N290、Y304、R305、H307、D353、A394、S399、P404、L407、Q521、K522和T524。在一些另外的实施方案中,具有苯丙氨酸解氨酶活性(PAL)的工程化多肽的改进的特性选自降低的对蛋白水解的敏感性和/或增加的对酸性pH的耐受性。

[0039] 本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:10或其功能片段至少约85%、至少约86%、至少约87%、至少约88%、至少约89%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或更大的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:10或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与参考序列SEQ ID NO:10相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0040] 在一些实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:10或其功能片段至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:10或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与参考序列SEQ ID NO:10相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0041] 在一些实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多

肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:10或其功能片段至少85%的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:10或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与参考序列SEQ ID NO:10相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0042] 在一些实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与参考序列SEQ ID NO:10至少85%的序列同一性的氨基酸序列,和与SEQ ID NO:10相比的至少一个氨基酸残基差异,且所述工程化多肽与SEQ ID NO:10相比表现出选自以下的至少一种改进的特性:增强的催化活性、降低的对蛋白水解的敏感性、增加的对酸性pH的耐受性、减少的聚集、和/或降低的免疫原性。

[0043] 在一些实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或更多个:

[0044] I27E/V39A;I27E/V39A/R43L/V105C/A153R/L214E/P266H/L278D/C503Q;

[0045] I27E/V39A/R43L/L214E/A547D;

[0046] I27E/V39A/V105C/A112C/R134Q/L214E/L278D/C503Q/A547D/C565N;

[0047] I27E/V39A/V105C/A112C/R134Q/A153R/Q205T/L214E/P266H/L278D/C503Q/A551D;

[0048] I27E/V39A/V105C/A112C/Q205T/P210C/P266H/C503Q/A547D;

[0049] I27E/V39A/V105C/A112C/Q205T/P266H/I285E/C503Q/A551D;

[0050] I27E/V39A/V105C/A112C/L214E/I285E/C503Q/A547D;

[0051] I27E/V39A/V105C/S131N/R134Q/Q205T/L214E/C503Q/A547D/C565N;

[0052] I27E/V39A/V105C/R134Q/A153R/P210C/L278D/I285E/C503Q/A547D/A551D;

[0053] I27E/V39A/V105C/R134Q/Q205T/P210C/L278D/C503Q/A547D;

[0054] I27E/V39A/V105C/R134Q/Q205T/L214E;

[0055] I27E/V39A/V105C/R134Q/Q205T/L214E/A551D/C565N;

[0056] I27E/V39A/V105C/R134Q/Q205T/L278D/I285E/C503Q/A547D/A551D/C565N;

[0057] I27E/V39A/V105C/R134Q/P210C;I27E/V39A/V105C/R134Q/P210C/L214E;

[0058] I27E/V39A/V105C/R134Q/P210C/L214E/I285E/A547D;

[0059] I27E/V39A/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N;

[0060] I27E/V39A/V105C/R134Q/L214E/L278D/A547D/A551D;

[0061] I27E/V39A/V105C/R134Q/L214E/I285E/C503Q/A547D/A551D;

[0062] I27E/V39A/V105C/R134Q/P266H/C503Q;

[0063] I27E/V39A/V105C/R134Q/P266H/C503Q/A547D/A551D;

[0064] I27E/V39A/V105C/R134Q/L278D/C503Q/C565N;

[0065] I27E/V39A/V105C/R134Q/L278D/I285E/C503Q;I27E/V39A/V105C/R134Q/L278D/A551D;

[0066] I27E/V39A/V105C/R134Q/I285E/A547D/A551D;I27E/V39A/V105C/R134Q/C503Q/

A551D;

[0067] I27E/V39A/V105C/A153R/Q205T/L278D/C503Q/A547D/A551D;

[0068] I27E/V39A/V105C/A153R/L214E; I27E/V39A/V105C/A153R/I285E;

[0069] I27E/V39A/V105C/A153R/C503Q/A547D/C565N; I27E/V39A/V105C/A153R/A551D/
C565N;

[0070] I27E/V39A/V105C/Q205T/P210C/L214E/L278D/A547D;

[0071] I27E/V39A/V105C/Q205T/P210C/L278D/C503Q;

[0072] I27E/V39A/V105C/Q205T/P210C/L278D/A547D;

[0073] I27E/V39A/V105C/Q205T/L214E/L278D/C503Q/A547D;

[0074] I27E/V39A/V105C/Q205T/L278D/C503Q/A547D;

[0075] I27E/V39A/V105C/P210C/I285E/C503Q/A547D/A551D/C565N;

[0076] I27E/V39A/V105C/P210C/L214E/P266H/L278D;

[0077] I27E/V39A/V105C/L214E/P266H/C503Q/A547D/C565N;

[0078] I27E/V39A/V105C/L214E/L278D/L309P/C503Q/A547D/A551D;

[0079] I27E/V39A/V105C/L278D/C503Q/A547D/C565N; I27E/V39A/V105C/I285E/A547D;

[0080] I27E/V39A/V105C/C503Q/A551D; I27E/V39A/V105C/C503Q/A547D/A551D/C565N;

[0081] I27E/V39A/A112C/R134Q/Q205T/P210C/L214E/A551D/C565N;

[0082] I27E/V39A/A112C/R134Q/L214E/P266H/A551D;

[0083] I27E/V39A/A112C/R134Q/L214E/C503Q/A547D; I27E/V39A/A112C/R134Q/P266H/
I285E;

[0084] I27E/V39A/A112C/Q205T/L214E/P266H/C503Q/A551D/C565N;

[0085] I27E/V39A/A112C/Q205T/L278D/I285E; I27E/V39A/A112C/L214E;

[0086] I27E/V39A/A112C/L214E/L278D/C503Q/A547D/A551D; I27E/V39A/A112C/I285E;

[0087] I27E/V39A/A112C/A547D; I27E/V39A/R134Q;

[0088] I27E/V39A/R134Q/A153R/Q205T/L214E/P266H/C503Q;

[0089] I27E/V39A/R134Q/A153R/P210C/L214E/L278D/I285E/A547D/C565N;

[0090] I27E/V39A/R134Q/A153R/L214E/P266H/L278D/C503Q/A547D/C565N;

[0091] I27E/V39A/R134Q/A153G/L214E/P266H/I285E/C503Q/A551D/C565N;

[0092] I27E/V39A/R134Q/A153R/L214E/C503Q/A547D; I27E/V39A/R134Q/A153R/L278D;

[0093] I27E/V39A/R134Q/A153R/L278D/A547D/A551D; I27E/V39A/R134Q/A153R/A547D;

[0094] I27E/V39A/R134Q/Q205T/L214E/P266H/I285E/C503Q/A551D/C565N;

[0095] I27E/V39A/R134Q/Q205T/P266H/C503Q/A551D/C565N;

[0096] I27E/V39A/R134Q/P210C/L214E/C503Q; I27E/V39A/R134Q/P210C/C503Q/A551D;

[0097] I27E/V39A/R134Q/L214E/P266H/A551D;

[0098] I27E/V39A/R134Q/L278D/I285E/C503Q/A547D/A551D;

[0099] I27E/V39A/R134Q/L278D/C503Q/A547D; I27E/V39A/R134Q/C503Q/A547D;

[0100] I27E/V39A/R134Q/A547D/C565N; I27E/V39A/Q205T/L214E/C503Q/C565N;

[0101] I27E/V39A/Q205T/P266H/I285E/A547D/A551D/C565N; I27E/V39A/Q205T/P266H/
A551D;

- [0102] I27E/V39A/Q205T/L278D/C503Q/A551D/C565N; I27E/V39A/Q205T/L278D/C503Q/C565N;
- [0103] I27E/V39A/Q205T/C503Q/A547D/C565N; I27E/V39A/P210C/T212S;
- [0104] I27E/V39A/P210C/L214E/L278D/C503Q/A551D; I27E/V39A/P210C/L214E/I285E/C503Q/A551D;
- [0105] I27E/V39A/P210C/P266H/I285E/C503Q/A547D; I27E/V39A/P210C/P266H/C503Q/A551D;
- [0106] I27E/V39A/L214E; I27E/V39A/L214E/P266H/L278D/C503Q/A547D/A551D/C565N;
- [0107] I27E/V39A/L214E/L278D/C503Q; I27E/V39A/L214E/L278D/C503Q/A547D/C565N;
- [0108] I27E/V39A/L214E/C503Q/A551D; I27E/V39A/P266H; I27E/V39A/P266H/L278D;
- [0109] I27E/V39A/L278D; I27E/V39A/L278D/A547D; I27E/V39A/L278D/I285E/C503Q/A547D;
- [0110] I27E/V39A/L278D/C503Q/C565N; I27E/V39A/C503Q; I27E/G45D/Q205T/P266H/C565N;
- [0111] I27E/V105C; I27E/V105C/R134Q/A153R/P210C/L214E/C503Q/A547D;
- [0112] I27E/V105C/R134Q/A153R/I285E/A547D; I27E/V105C/R134Q/A153R/C503Q;
- [0113] I27E/V105C/R134Q/Q205T/P210C/C503Q;
- [0114] I27E/V105C/R134Q/Q205T/L214E/P266H/L278D/C503Q/C565N;
- [0115] I27E/V105C/Q205T/P266H/C503Q;
- [0116] I27E/V105C/R134Q/P210C/L214E/P266H/L278D/A551D/C565N;
- [0117] I27E/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N;
- [0118] I27E/V105C/R134Q/P210C/P266H/L278D/I285E/C503Q/A551D/C565N;
- [0119] I27E/V105C/R134Q/L214E/L278D/C503Q/A547D;
- [0120] I27E/V105C/R134Q/L214E/L278D/C503Q/A547D/A551D/C565N; I27E/V105C/Q205T;
- [0121] I27E/V105C/Q205T/L214E/P266H; I27E/V105C/Q205T/L214E/P266H/A551D/C565N;
- [0122] I27E/V105C/Q205T/L214E/L278D/I285E/C503Q/A547D/A551D/C565N;
- [0123] I27E/V105C/Q205T/C503Q/A547D/A551D/C565N; I27E/V105C/L214E;
- [0124] I27E/V105C/L214E/P266H/C503Q; I27E/V105C/L214E/I285E/A551D/C565N;
- [0125] I27E/V105C/L214E/A547D/A551D/C565N; I27E/V105C/L214E/A551D/C565N;
- [0126] I27E/V105C/P266H; I27E/V105C/P266H/I285E/C503Q/A547D/C565N;
- [0127] I27E/V105C/L278D/A547D; I27E/V105C/I285E/C503Q/A547D/A551D/C565N;
- [0128] I27E/V105C/C503Q/A547D/C565N; I27E/V105C/C503Q/A547D/A551D/C565N;
- [0129] I27E/A112C/R134Q/A153R/L214E/P266H/C503Q;
- [0130] I27E/A112C/R134Q/L278D/I285E/C503Q/A551D/C565N;
- [0131] I27E/A112C/R134Q/Q205T/L278D/C503Q; I27E/A112C/R134Q/Q205T/I285E/C503Q;
- [0132] I27E/A112C/Q205T/P266H/L278D/I285E/C503Q; I27E/A112C/P210C/L214E/

C503Q/A547D;

[0133] I27E/R134Q;I27E/R134Q/A153R/I285E/C503Q/A547D;

[0134] I27E/R134Q/Q205T/I285E/C503Q/A551D;I27E/R134Q/Q205T/P266H/L278D/A547D;

[0135] I27E/R134Q/P210C;I27E/R134Q/L214E/C503Q;I27E/R134Q/L214E/C503Q/A547D;

[0136] I27E/R134Q/L214E/C503Q/A547D/A551D;I27E/R134Q/L214E/C503Q/C565N;

[0137] I27E/R134Q/L278D/I285E/A551D/C565N;I27E/R134Q/I285E/C503Q;

[0138] I27E/A153R/L214E/L278D/I285E/A551D/C565N;I27E/A153R/L214E/L278D/A551D;

[0139] I27E/Q205T;I27E/Q205T/L214E/L278D/I285E/C503Q/C565N;

[0140] I27E/Q205T/L214E/C503Q/A547D/C565N;I27E/Q205T/P266H/L278D/I285E/A551D/C565N;

[0141] I27E/Q205T/L278D/A551D;I27E/P210C;I27E/P210C/L214E/C503Q/A547D;

[0142] I27E/P210C/L278D/C503Q;I27E/P210C/C503Q;I27E/P210C/C503Q/C565N;I27E/P210C/A551D;

[0143] I27E/L214E;I27E/L214E/P266H/L278D/I285E/A551D;I27E/L214E/L278D;

[0144] I27E/L214E/L278D/C503Q;I27E/L214E/C503Q;I27E/L214E/C503Q/A547D;

[0145] I27E/L214E/C503Q/A547D/C565N;I27E/L214E/A551D;I27E/P266H/L278D/C503Q;

[0146] I27E/P266H/A547D/A551D;I27E/L278D/C503Q/A551D;I27E/L278D/C503Q/A551D/C565N;

[0147] I27E/A547D/C565N;V39A/G45S/L278D/C503Q/A551D;

[0148] V39A/V105C/R134Q/A153R/Q205T/A551D;V39A/V105C/R134Q/P210C/L214E/A551D;

[0149] V39A/V105C/R134Q/L214E/C503Q/A547D/A551D;V39A/V105C/A153R/P266H/A547D/A551D;

[0150] V39A/V105C/Q205T/C503Q;V39A/V105C/Q205T/A551D;V39A/V105C/P210C/A547D;

[0151] V39A/V105C/L214E/P266H/A547D/C565N;V39A/V105C/L214E/I285E/C503Q/A551D/C565N;

[0152] V39A/A112C/R134Q/Q205T/L214E/L278D;V39A/A112C/R134Q/L214E/C503Q/A547D/A551D;

[0153] V39A/A112C/A153R/Q205T/L278D/C503Q/A547D;V39A/R134Q;

[0154] V39A/R134Q/Q205T/L214E/C503Q/C565N;V39A/R134Q/P210C/L214E/A547D/C565N;

[0155] V39A/A153R/C503Q/A547D;V39A/Q205T/L278D/A547D/A551D;

[0156] V39A/P210C/L214E/L278D/I285E/C503Q/A551D;V39A/P266H;

[0157] V39A/P275R/L278D/C503Q/A551D;V39A/C503Q;V39A/C503Q/A551D/C565N;V105C;

[0158] V105C/A112C/R134Q/Q205T/L214E/Y492H/C503Q/A547D;

[0159] V105C/R134Q/A153R/Q205T/L214E/C503Q;V105C/R134Q/Q205T/L214E/A547D;

[0160] V105C/R134Q/Q205T/P266H/L278D;V105C/R134Q/L214E/P266H/I285E/C503Q/

A551D/C565N;

[0161] V105C/R134Q/L214E/L278D/C565N;V105C/R134Q/L214E/C503Q/A547D;

[0162] V105C/R134Q/L214E/C503Q/A547D/A551D;V105C/R134Q/C503Q;

[0163] V105C/R134Q/C503Q/A547D;V105C/R134Q/C503Q/A547D/C565N;

[0164] V105C/A153R/Q205T/L214E/P266H/C503Q/A547D;

[0165] V105C/A153R/Q205T/P266H/I285E/A547D/C565N;V105C/Q205T/P210C/L214E/C503Q/A547D;

[0166] V105C/Q205T/L214E/L278D;V105C/Q205T/L214E/C503Q/A547D/A551D/C565N;

[0167] V105C/Q205T/C503Q/A551D;V105C/L214E/P266H/L278D/A547D;

[0168] V105C/L214E/L278D/C503Q/A547D/A551D;V105C/L214E/I285E;

[0169] V105C/L214E/I285E/C503Q/A547D/A551D/C565N;V105C/L214E/I285E/A547D/C565N;

[0170] V105C/L278D/C503Q/A551D;V105C/I285E;V105C/I285E/A547D;V105C/C503Q;

[0171] V105C/A547D/A551D;A112C/R134Q/A153R/L214E/L278D/I285E/C503Q/A547D/A551D/C565N;

[0172] A112C/R134Q/L214E/C503Q/A547D/A551D/C565N;A112C/L214E/L278D;

[0173] A112C/L278D/C503Q/A547D;R134Q/Q205T/L214E/I285E/C503Q/A551D/C565N;

[0174] R134Q/Q205T/C503Q;R134Q/P210C/L214E/L278D/C503Q/A547D/C565N;

[0175] R134Q/P210C/L214E/C503Q/A547D/A551D;R134Q/L214E;R134Q/L214E/L278D/C503Q;

[0176] R134Q/L214E/L278D/C503Q/A551D;R134Q/L214E/I285E/C503Q;R134Q/C503Q;

[0177] R134Q/C503Q/A547D/A551D;A153R;Q205T/L214E/I285E/C503Q/A551D;

[0178] Q205T/L214E/I285E/C503Q/C565N;Q205T/L214E/C503Q/A547D/C565N;

[0179] Q205T/L278D/I285E/A547D/A551D;P210C/L214E;P210C/L214E/P266H;L214E/P266H;

[0180] L214E/P266H/C503Q/A547D/A551D/C565N;L214E/C503Q/A547D;L214E/A547D;

[0181] P266H/L278D/C503Q;P266H/C565N;L278D/A547D;C503Q;C503Q/A547D;

[0182] C503Q/A547D/A551D/C565N;C503Q/A547D/C565N;C503Q/A551D;C503Q/A551D/C565N;

[0183] A547D;和/或C565N。

[0184] 在一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个:V80I/R134C/P564Q;V121C;A123G;A124G;M125L;L126I/T;L126M/R134L;L127A;A129G/L;N130Q;N130C/M370I R134W;M133R;R134I;R134N/G307C;G135C/S;和/或G135A/A394E。

[0185] 在一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个:G20S/I144L;R43S;L47M/I144L;L47M/R146E;L47M/M147G/A383E;L47M/P157C;Q58H/L143V;Q58K/P157D/G369C;A62S/M147V;S82I/G135C/P157F/W279L;R94C/I149E;

T110I/I139R;L118M/L141H;A119E/T156H/A289D;I139M/V;R140D/G/M;R140N/A199E;R140E/A334S/A551D;L141K/Q/P/T;E142H/P/V;E142D/G371D;L143F/M;I144L/N/V;K145N/Q/R;K145G/P157T;R146H/L;R146W/D191Y;M147A;I149L/R;F150K/L/M;L151M;A153C/G;A153S/H250N;G154R;G154Y/L174M/Q321K/S456I/G483C;T156K/G483C;P157D/F/H/Y;Y158E;V159C/H/L/M;M247I;L319M;和/或Q389K。

[0186] 在仍然一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个: P117T/Y176Q;V172I/C/L;L174M;S175G;Y176E/I/M/R/V;I177M/V;T178L/A477S;和/或S180C/T。

[0187] 在一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个: R43S/H374K;R43S/H374R;A112S/M370A/A507E;M147I/H374S;S187R/L381V;D191Y/H385N;A232S;Q240K/H374R;A256S/L381N;P275Q/M370S;P275T/H374R;Q332K/Y377M;A334S/H374V;L349M;Q355K/H374S;M370G/I/S;G371H/N/Q/S;M372A/V;H374A/D/G/L/N/R/S/T;H374Q/P396Q;H374R/G417C;L375I;L375M;Y377C/I/N;Y378C/D/E/I/L/N/S;Y378F/P404Q;I379C/H/L/M/N;L381G/V;L381M/Q560K;L382C/H/I/M/S;A383S/V;K384R;H385C/G/N;H385M/P403H;H385S/P403H;D387S;L418M;G425V;A447S;S461G;和/或S525L。

[0188] 在一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个: A24S/F434M;A62S/T433N;S98I;L213M/S438L;Q240K/T433Y;S286R/Y435T;A289S/L431E;S331I;L431C/E/G/P/S/V;L432C/V;T433A/I/L/N/P/Q/R/S/V/W;F434C;Y435L;Y435Q/H446N;G436M;G436D/T;N437E/G/Q;N437T/L538M;S438C/F/M/R/T;I439C/F/L/V;和/或A477S。

[0189] 在一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个: A24E;Q58R/Y475H;A70S/N474E;L104M/V476L;A119E/G365A;L206M;P275Q;G276V;Q292H/A479G;Q355H/I478C;P404T/A477V;I471F/G/K/M/N/R/V/W;F472G;Q473H/K/M/R/S;Q473H/A507S;N474A/H/R/W;N474D/R490H;Y475C/F/L/Q;V476C/I/L;I478N/S;A479G/S;F482C/L;G483C/H/S;G483A/S524I;G483R/G537C;和/或A558S。

[0190] 在一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或多个: V39A/K115E/M133R/C565N;V39A/M133R/F472G/C503Q/C565N;V39A/M133R/F472G/C565N;V39A/M133R/C503Q;V39A/M133R/C503Q/C565N;V39A/M147A/Y378E/C503Q/C565N;V39A/M147A/Y378E/C565N;V39A/M147A/L381G/F472G/C503Q/C565N;V39A/M147A/L381G/C503Q/C565N;V39A/M147A/F472G/C503Q/C565N;V39A/M147A/F472G/C565N;V39A/M147A/C565N;V39A/G248C/L381G/F472G/C503Q/C565N;V39A/Y378E/C503Q/C565N;V39A/Y378E/C565N;V39A/L381G;V39A/F472G/C503Q/C565N;V39A/C503Q/C565N;M133R/L381G/C565N;M133R/C503Q;Y378D/C503Q;Y378E/F472G/C503Q/C565N;L381G/F472G/C503Q/C565N;和/或F472G/C503Q/C565N。

[0191] 在仍然一些另外的实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中与SEQ ID NO:10相比的氨基酸残基差异选自以下置换或置换集合中的一个或更多个:

- [0192] I27E/V39A;I27E/V39A//R43L/V105C/A153R/L214E/P266H/L278D/C503Q;
- [0193] I27E/V39A/R43L/L214E/A547D;
- [0194] I27E/V39A/V105C/A112C/R134Q/L214E/L278D/C503Q/A547D/C565N;
- [0195] I27E/V39A/V105C/A112C/R134Q/A153R/Q205T/L214E/P266H/L278D/C503Q/A551D;
- [0196] I27E/V39A/V105C/A112C/Q205T/P210C/P266H/C503Q/A547D;
- [0197] I27E/V39A/V105C/A112C/Q205T/P266H/I285E/C503Q/A551D;
- [0198] I27E/V39A/V105C/A112C/L214E/I285E/C503Q/A547D;
- [0199] I27E/V39A/V105C/S131N/R134Q/Q205T/L214E/C503Q/A547D/C565N;
- [0200] I27E/V39A/V105C/R134Q/A153R/P210C/L278D/I285E/C503Q/A547D/A551D;
- [0201] I27E/V39A/V105C/R134Q/Q205T/P210C/L278D/C503Q/A547D;
- [0202] I27E/V39A/V105C/R134Q/Q205T/L214E;
- [0203] I27E/V39A/V105C/R134Q/Q205T/L214E/A551D/C565N;
- [0204] I27E/V39A/V105C/R134Q/Q205T/L278D/I285E/C503Q/A547D/A551D/C565N;
- [0205] I27E/V39A/V105C/R134Q/P210C;I27E/V39A/V105C/R134Q/P210C/L214E;
- [0206] I27E/V39A/V105C/R134Q/P210C/L214E/I285E/A547D;
- [0207] I27E/V39A/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N;
- [0208] I27E/V39A/V105C/R134Q/L214E/L278D/A547D/A551D;
- [0209] I27E/V39A/V105C/R134Q/L214E/I285E/C503Q/A547D/A551D;
- [0210] I27E/V39A/V105C/R134Q/P266H/C503Q;
- [0211] I27E/V39A/V105C/R134Q/P266H/C503Q/A547D/A551D;
- [0212] I27E/V39A/V105C/R134Q/L278D/C503Q/C565N;
- [0213] I27E/V39A/V105C/R134Q/L278D/I285E/C503Q;I27E/V39A/V105C/R134Q/L278D/A551D;
- [0214] I27E/V39A/V105C/R134Q/I285E/A547D/A551D;I27E/V39A/V105C/R134Q/C503Q/A551D;
- [0215] I27E/V39A/V105C/A153R/Q205T/L278D/C503Q/A547D/A551D;
- [0216] I27E/V39A/V105C/A153R/L214E;I27E/V39A/V105C/A153R/I285E;
- [0217] I27E/V39A/V105C/A153R/C503Q/A547D/C565N;I27E/V39A/V105C/A153R/A551D/C565N;
- [0218] I27E/V39A/V105C/Q205T/P210C/L214E/L278D/A547D;
- [0219] I27E/V39A/V105C/Q205T/P210C/L278D/C503Q;
- [0220] I27E/V39A/V105C/Q205T/P210C/L278D/A547D;
- [0221] I27E/V39A/V105C/Q205T/L214E/L278D/C503Q/A547D;
- [0222] I27E/V39A/V105C/Q205T/L278D/C503Q/A547D;
- [0223] I27E/V39A/V105C/P210C/I285E/C503Q/A547D/A551D/C565N;

- [0224] I27E/V39A/V105C/P210C/L214E/P266H/L278D;
[0225] I27E/V39A/V105C/L214E/P266H/C503Q/A547D/C565N;
[0226] I27E/V39A/V105C/L214E/L278D/L309P/C503Q/A547D/A551D;
[0227] I27E/V39A/V105C/L278D/C503Q/A547D/C565N; I27E/V39A/V105C/I285E/A547D;
[0228] I27E/V39A/V105C/C503Q/A551D; I27E/V39A/V105C/C503Q/A547D/A551D/C565N;
[0229] I27E/V39A/A112C/R134Q/Q205T/P210C/L214E/A551D/C565N;
[0230] I27E/V39A/A112C/R134Q/L214E/P266H/A551D;
[0231] I27E/V39A/A112C/R134Q/L214E/C503Q/A547D; I27E/V39A/A112C/R134Q/P266H/
I285E;
[0232] I27E/V39A/A112C/Q205T/L214E/P266H/C503Q/A551D/C565N;
[0233] I27E/V39A/A112C/Q205T/L278D/I285E; I27E/V39A/A112C/L214E;
[0234] I27E/V39A/A112C/L214E/L278D/C503Q/A547D/A551D; I27E/V39A/A112C/I285E;
[0235] I27E/V39A/A112C/A547D; I27E/V39A/R134Q;
[0236] I27E/V39A/R134Q/A153R/Q205T/L214E/P266H/C503Q;
[0237] I27E/V39A/R134Q/A153R/P210C/L214E/L278D/I285E/A547D/C565N;
[0238] I27E/V39A/R134Q/A153R/L214E/P266H/L278D/C503Q/A547D/C565N;
[0239] I27E/V39A/R134Q/A153G/L214E/P266H/I285E/C503Q/A551D/C565N;
[0240] I27E/V39A/R134Q/A153R/L214E/C503Q/A547D; I27E/V39A/R134Q/A153R/L278D;
[0241] I27E/V39A/R134Q/A153R/L278D/A547D/A551D; I27E/V39A/R134Q/A153R/A547D;
[0242] I27E/V39A/R134Q/Q205T/L214E/P266H/I285E/C503Q/A551D/C565N;
[0243] I27E/V39A/R134Q/Q205T/P266H/C503Q/A551D/C565N;
[0244] I27E/V39A/R134Q/P210C/L214E/C503Q; I27E/V39A/R134Q/P210C/C503Q/A551D;
[0245] I27E/V39A/R134Q/L214E/P266H/A551D;
[0246] I27E/V39A/R134Q/L278D/I285E/C503Q/A547D/A551D;
[0247] I27E/V39A/R134Q/L278D/C503Q/A547D; I27E/V39A/R134Q/C503Q/A547D;
[0248] I27E/V39A/R134Q/A547D/C565N; I27E/V39A/Q205T/L214E/C503Q/C565N;
[0249] I27E/V39A/Q205T/P266H/I285E/A547D/A551D/C565N; I27E/V39A/Q205T/P266H/
A551D;
[0250] I27E/V39A/Q205T/L278D/C503Q/A551D/C565N; I27E/V39A/Q205T/L278D/C503Q/
C565N;
[0251] I27E/V39A/Q205T/C503Q/A547D/C565N; I27E/V39A/P210C/T212S;
[0252] I27E/V39A/P210C/L214E/L278D/C503Q/A551D; I27E/V39A/P210C/L214E/I285E/
C503Q/A551D;
[0253] I27E/V39A/P210C/P266H/I285E/C503Q/A547D; I27E/V39A/P210C/P266H/C503Q/
A551D;
[0254] I27E/V39A/L214E; I27E/V39A/L214E/P266H/L278D/C503Q/A547D/A551D/C565N;
[0255] I27E/V39A/L214E/L278D/C503Q; I27E/V39A/L214E/L278D/C503Q/A547D/C565N;
[0256] I27E/V39A/L214E/C503Q/A551D; I27E/V39A/P266H; I27E/V39A/P266H/L278D;
[0257] I27E/V39A/L278D; I27E/V39A/L278D/A547D; I27E/V39A/L278D/I285E/C503Q/

A547D;

[0258] I27E/V39A/L278D/C503Q/C565N; I27E/V39A/C503Q; I27E/G45D/Q205T/P266H/C565N;

[0259] I27E/V105C; I27E/V105C/R134Q/A153R/P210C/L214E/C503Q/A547D;

[0260] I27E/V105C/R134Q/A153R/I285E/A547D; I27E/V105C/R134Q/A153R/C503Q;

[0261] I27E/V105C/R134Q/Q205T/P210C/C503Q;

[0262] I27E/V105C/R134Q/Q205T/L214E/P266H/L278D/C503Q/C565N;

[0263] I27E/V105C/Q205T/P266H/C503Q;

[0264] I27E/V105C/R134Q/P210C/L214E/P266H/L278D/A551D/C565N;

[0265] I27E/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N;

[0266] I27E/V105C/R134Q/P210C/P266H/L278D/I285E/C503Q/A551D/C565N;

[0267] I27E/V105C/R134Q/L214E/L278D/C503Q/A547D;

[0268] I27E/V105C/R134Q/L214E/L278D/C503Q/A547D/A551D/C565N; I27E/V105C/Q205T;

[0269] I27E/V105C/Q205T/L214E/P266H; I27E/V105C/Q205T/L214E/P266H/A551D/C565N;

[0270] I27E/V105C/Q205T/L214E/L278D/I285E/C503Q/A547D/A551D/C565N;

[0271] I27E/V105C/Q205T/C503Q/A547D/A551D/C565N; I27E/V105C/L214E;

[0272] I27E/V105C/L214E/P266H/C503Q; I27E/V105C/L214E/I285E/A551D/C565N;

[0273] I27E/V105C/L214E/A547D/A551D/C565N; I27E/V105C/L214E/A551D/C565N;

[0274] I27E/V105C/P266H; I27E/V105C/P266H/I285E/C503Q/A547D/C565N;

[0275] I27E/V105C/L278D/A547D; I27E/V105C/I285E/C503Q/A547D/A551D/C565N;

[0276] I27E/V105C/C503Q/A547D/C565N; I27E/V105C/C503Q/A547D/A551D/C565N;

[0277] I27E/A112C/R134Q/A153R/L214E/P266H/C503Q;

[0278] I27E/A112C/R134Q/L278D/I285E/C503Q/A551D/C565N;

[0279] I27E/A112C/R134Q/Q205T/L278D/C503Q; I27E/A112C/R134Q/Q205T/I285E/C503Q;

[0280] I27E/A112C/Q205T/P266H/L278D/I285E/C503Q; I27E/A112C/P210C/L214E/C503Q/A547D;

[0281] I27E/R134Q; I27E/R134Q/A153R/I285E/C503Q/A547D;

[0282] I27E/R134Q/Q205T/I285E/C503Q/A551D; I27E/R134Q/Q205T/P266H/L278D/A547D;

[0283] I27E/R134Q/P210C; I27E/R134Q/L214E/C503Q; I27E/R134Q/L214E/C503Q/A547D;

[0284] I27E/R134Q/L214E/C503Q/A547D/A551D; I27E/R134Q/L214E/C503Q/C565N;

[0285] I27E/R134Q/L278D/I285E/A551D/C565N; I27E/R134Q/I285E/C503Q;

[0286] I27E/A153R/L214E/L278D/I285E/A551D/C565N; I27E/A153R/L214E/L278D/A551D;

[0287] I27E/Q205T; I27E/Q205T/L214E/L278D/I285E/C503Q/C565N;

[0288] I27E/Q205T/L214E/C503Q/A547D/C565N; I27E/Q205T/P266H/L278D/I285E/

A551D/C565N;

[0289] I27E/Q205T/L278D/A551D; I27E/P210C; I27E/P210C/L214E/C503Q/A547D;

[0290] I27E/P210C/L278D/C503Q; I27E/P210C/C503Q; I27E/P210C/C503Q/C565N; I27E/P210C/A551D;

[0291] I27E/L214E; I27E/L214E/P266H/L278D/I285E/A551D; I27E/L214E/L278D;

[0292] I27E/L214E/L278D/C503Q; I27E/L214E/C503Q; I27E/L214E/C503Q/A547D;

[0293] I27E/L214E/C503Q/A547D/C565N; I27E/L214E/A551D; I27E/P266H/L278D/C503Q;

[0294] I27E/P266H/A547D/A551D; I27E/L278D/C503Q/A551D; I27E/L278D/C503Q/A551D/C565N;

[0295] I27E/A547D/C565N; V39A/G45S/L278D/C503Q/A551D;

[0296] V39A/V105C/R134Q/A153R/Q205T/A551D; V39A/V105C/R134Q/P210C/L214E/A551D;

[0297] V39A/V105C/R134Q/L214E/C503Q/A547D/A551D; V39A/V105C/A153R/P266H/A547D/A551D;

[0298] V39A/V105C/Q205T/C503Q; V39A/V105C/Q205T/A551D; V39A/V105C/P210C/A547D;

[0299] V39A/V105C/L214E/P266H/A547D/C565N; V39A/V105C/L214E/I285E/C503Q/A551D/C565N;

[0300] V39A/A112C/R134Q/Q205T/L214E/L278D; V39A/A112C/R134Q/L214E/C503Q/A547D/A551D;

[0301] V39A/A112C/A153R/Q205T/L278D/C503Q/A547D; V39A/R134Q;

[0302] V39A/R134Q/Q205T/L214E/C503Q/C565N; V39A/R134Q/P210C/L214E/A547D/C565N;

[0303] V39A/A153R/C503Q/A547D; V39A/Q205T/L278D/A547D/A551D;

[0304] V39A/P210C/L214E/L278D/I285E/C503Q/A551D; V39A/P266H;

[0305] V39A/P275R/L278D/C503Q/A551D; V39A/C503Q; V39A/C503Q/A551D/C565N; V105C;

[0306] V105C/A112C/R134Q/Q205T/L214E/Y492H/C503Q/A547D;

[0307] V105C/R134Q/A153R/Q205T/L214E/C503Q; V105C/R134Q/Q205T/L214E/A547D;

[0308] V105C/R134Q/Q205T/P266H/L278D; V105C/R134Q/L214E/P266H/I285E/C503Q/A551D/C565N;

[0309] V105C/R134Q/L214E/L278D/C565N; V105C/R134Q/L214E/C503Q/A547D;

[0310] V105C/R134Q/L214E/C503Q/A547D/A551D; V105C/R134Q/C503Q;

[0311] V105C/R134Q/C503Q/A547D; V105C/R134Q/C503Q/A547D/C565N;

[0312] V105C/A153R/Q205T/L214E/P266H/C503Q/A547D;

[0313] V105C/A153R/Q205T/P266H/I285E/A547D/C565N; V105C/Q205T/P210C/L214E/C503Q/A547D;

[0314] V105C/Q205T/L214E/L278D; V105C/Q205T/L214E/C503Q/A547D/A551D/C565N;

[0315] V105C/Q205T/C503Q/A551D; V105C/L214E/P266H/L278D/A547D;

[0316] V105C/L214E/L278D/C503Q/A547D/A551D; V105C/L214E/I285E;

[0317] V105C/L214E/I285E/C503Q/A547D/A551D/C565N; V105C/L214E/I285E/A547D/

C565N;

[0318] V105C/L278D/C503Q/A551D;V105C/I285E;V105C/I285E/A547D;V105C/C503Q;

[0319] V105C/A547D/A551D;A112C/R134Q/A153R/L214E/L278D/I285E/C503Q/A547D/A551D/C565N;

[0320] A112C/R134Q/L214E/C503Q/A547D/A551D/C565N;A112C/L214E/L278D;

[0321] A112C/L278D/C503Q/A547D;R134Q/Q205T/L214E/I285E/C503Q/A551D/C565N;

[0322] R134Q/Q205T/C503Q;R134Q/P210C/L214E/L278D/C503Q/A547D/C565N;

[0323] R134Q/P210C/L214E/C503Q/A547D/A551D;R134Q/L214E;R134Q/L214E/L278D/C503Q;

[0324] R134Q/L214E/L278D/C503Q/A551D;R134Q/L214E/I285E/C503Q;R134Q/C503Q;

[0325] R134Q/C503Q/A547D/A551D;A153R;Q205T/L214E/I285E/C503Q/A551D;

[0326] Q205T/L214E/I285E/C503Q/C565N;Q205T/L214E/C503Q/A547D/C565N;

[0327] Q205T/L278D/I285E/A547D/A551D;P210C/L214E;P210C/L214E/P266H;L214E/P266H;

[0328] L214E/P266H/C503Q/A547D/A551D/C565N;L214E/C503Q/A547D;L214E/A547D;

[0329] P266H/L278D/C503Q;P266H/C565N;L278D/A547D;C503Q;C503Q/A547D;

[0330] C503Q/A547D/A551D/C565N;C503Q/A547D/C565N;C503Q/A551D;C503Q/A551D/C565N;

[0331] A547D;C565N;V80I/R134C/P564Q;V121C;A123G;A124G;M125L;L126I/T;L126M/R134L;

[0332] L127A;A129G/L;N130Q;N130C/M370I R134W;M133R;R134I;R134N/G307C;G135C/S;

[0333] G135A/A394E;G20S/I144L;R43S;L47M/I144L;L47M/R146E;L47M/M147G/A383E;

[0334] L47M/P157C;Q58H/L143V;Q58K/P157D/G369C;A62S/M147V;S82I/G135C/P157F/W279L;

[0335] R94C/I149E;T110I/I139R;L118M/L141H;A119E/T156H/A289D;I139M/V;R140D/G/M;

[0336] R140N/A199E;R140E/A334S/A551D;L141K/Q/P/T;E142H/P/V;E142D/G371D;L143F/M;

[0337] I144L/N/V;K145N/Q/R;K145G/P157T;R146H/L;R146W/D191Y;M147A;I149L/R;F150K/L/M;

[0338] L151M;A153C/G;A153S/H250N;G154R;G154Y/L174M/Q321K/S456I/G483C;T156K/G483C;

[0339] P157D/F/H/Y;Y158E;V159C/H/L/M;M247I;L319M;Q389K;P117T/Y176Q;V172I/C/L;

[0340] L174M;S175G;Y176E/I/M/R/V;I177M/V;T178L/A477S;S180C/T;R43S/H374K;R43S/H374R;

[0341] A112S/M370A/A507E;M147I/H374S;S187R/L381V;D191Y/H385N;A232S;Q240K/H374R;

- [0342] A256S/L381N;P275Q/M370S;P275T/H374R;Q332K/Y377M;A334S/H374V;L349M;
- [0343] Q355K/H374S;M370G/I/S;G371H/N/Q/S;M372A/V;H374A/D/G/L/N/R/S/T;H374Q/P396Q;
- [0344] H374R/G417C;L375I;L375M;Y377C/I/N;Y378C/D/E/I/L/N/S;Y378F/P404Q;I379C/H/L/M/N;
- [0345] L381G/V;L381M/Q560K;L382C/H/I/M/S;A383S/V;K384R;H385C/G/N;H385M/P403H;
- [0346] H385S/P403H;D387S;L418M;G425V;A447S;S461G;S525L;A24S/F434M;A62S/T433N;S98I;
- [0347] L213M/S438L;Q240K/T433Y;S286R/Y435T;A289S/L431E;S331I;L431C/E/G/P/S/V;L432C/V;
- [0348] T433A/I/L/N/P/Q/R/S/V/W;F434C;Y435L;Y435Q/H446N;G436M;G436D/T;N437E/G/Q;
- [0349] N437T/L538M;S438C/F/M/R/T;I439C/F/L/V;A477S;A24E;Q58R/Y475H;A70S/N474E;
- [0350] L104M/V476L;A119E/G365A;L206M;P275Q;G276V;Q292H/A479G;Q355H/I478C;
- [0351] P404T/A477V;I471F/G/K/M/N/R/V/W;F472G;Q473H/K/M/R/S;Q473H/A507S;N474A/H/R/W;
- [0352] N474D/R490H;Y475C/F/L/Q;V476C/I/L;I478N/S;A479G/S;
- [0353] F482C/L;G483C/H/S;G483A/S524I;G483R/G537C;A558S;V39A/K115E/M133R/C565N;
- [0354] V39A/M133R/F472G/C503Q/C565N;V39A/M133R/F472G/C565N;V39A/M133R/C503Q;
- [0355] V39A/M133R/C503Q/C565N;V39A/M147A/Y378E/C503Q/C565N;V39A/M147A/Y378E/C565N;
- [0356] V39A/M147A/L381G/F472G/C503Q/C565N;V39A/M147A/L381G/C503Q/C565N;
- [0357] V39A/M147A/F472G/C503Q/C565N;V39A/M147A/F472G/C565N;V39A/M147A/C565N;
- [0358] V39A/G248C/L381G/F472G/C503Q/C565N;V39A/Y378E/C503Q/C565N;V39A/Y378E/C565N;
- [0359] V39A/L381G;V39A/F472G/C503Q/C565N;V39A/C503Q/C565N;M133R/L381G/C565N;
- [0360] M133R/C503Q;Y378D/C503Q;Y378E/F472G/C503Q/C565N;L381G/F472G/C503Q/C565N;
- [0361] 和/或F472G/C503Q/C565N。
- [0362] 本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:26或其功能片段至少约85%、至少约86%、至少约87%、至少约88%、至少约89%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或更大的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:26或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与参考序列SEQ ID NO:26相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的

对蛋白水解的敏感性、iii) 增加的对酸性pH的耐受性、iv) 减少的聚集、v) 降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0363] 在一些实施方案中,本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:26或其功能片段至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:26或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与参考序列SEQ ID NO:26相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0364] 在一些另外的实施方案中,本发明提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:26或其功能片段至少85%的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:26或其功能片段相比在一个或多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与参考序列SEQ ID NO:26相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,包含具有与参考序列SEQ ID NO:26至少85%的序列同一性的氨基酸序列,和与SEQ ID NO:126相比的至少一个氨基酸残基差异,并且所述工程化多肽与SEQ ID NO:26相比表现出选自以下的至少一种改进的特性:增强的催化活性、降低的对蛋白水解的敏感性、增加的对酸性pH的耐受性、减少的聚集和/或降低的免疫原性。在工程化多肽的一些实施方案中,与SEQ ID NO:26相比的氨基酸残基差异选自以下置换或置换集合中的一个或更多个:A24E/G381L;L127V;A129I/V;S131C/T;H132L/S;R134C/F/H/K;R134H/Y378E/G381L;R134H/Y378E/G381L/V388T;R134H/V388T;A136K;A289S;M372L;H374G/M/Q;G381A/C/F/I/L/M/N/Q/S/T;A383C/M;V388C/T;L431M;和/或L563M。

[0365] 在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽具有与参考序列SEQ ID NO:4至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的序列同一性。在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽具有与参考序列SEQ ID NO:4至少约90%的序列同一性,而在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少约95%的序列同一性。在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽具有与参考序列SEQ ID NO:4至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的序列同一性。在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽具有与参考序列SEQ ID NO:4至少90%的序列同一性,而在一些另外的实施方案中,工程化多肽具有与参考序列SEQ ID NO:4至少95%的序列同一性。在一些另外的实施方案中,工程化多肽包含本发明的具有苯丙氨酸解氨酶(PAL)活性的多肽(例如,本文在表中提供的任何变体)的功能片段。

[0366] 在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽具有

与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的序列同一性。在一些实施方案中,本发明的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的序列同一性。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽包括具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少约90%序列同一性的氨基酸序列。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)的酶活性的工程化多肽包括具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26或其功能片段至少约99%的序列同一性的氨基酸序列。在一些另外的实施方案中,工程化多肽包含本发明的具有苯丙氨酸解氨酶(PAL)活性的多肽的功能片段(例如,SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26,以及本文中表中提供的任何变体的功能片段)。

[0367] 本发明还提供了具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中该工程化多肽是表2-1至2-5和/或表9-1至9-7中的任一个提供的变体PAL。

[0368] 在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽是多变鱼腥藻酶。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽是热稳定的。在一些实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽耐受蛋白水解。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽耐受通过至少一种消化道酶的蛋白水解。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽耐受通过胰凝乳蛋白酶、胰蛋白酶、羧肽酶和/或弹性蛋白酶的蛋白水解。在一些另外的实施方案中,具有苯丙氨酸解氨酶(PAL)活性的工程化多肽是酸稳定的。

[0369] 本发明还提供了去免疫化的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽。在一些实施方案中,去免疫化的工程化多肽包含具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或更大的序列同一性的氨基酸序列。在一些另外的实施方案中,去免疫化的工程化多肽包含具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或更大的序列同一性的氨基酸序列。在一些实施方案中,去免疫化的工程化多肽包含具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少约95%序列同一性的氨基酸序列。在一些实施方案中,去免疫化的工程化多肽包含与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26具有95%的序列同一性的氨基酸序列。

[0370] 在仍然一些另外的实施方案中,本发明提供了纯化的工程化多肽,所述工程化多肽具有苯丙氨酸解氨酶(PAL)活性。

[0371] 本发明还提供了多核苷酸序列,所述多核苷酸序列编码至少一种如本文所述的具有苯丙氨酸解氨酶(PAL)的工程化多肽。在一些实施方案中,多核苷酸序列被可操作地连接到控制序列。在一些另外的实施方案中,多核苷酸序列是密码子优化的。

[0372] 本发明还提供了表达载体,所述表达载体包含编码至少一种如本文所提供的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽的至少一种多核苷酸序列。在一些实施方案中,表

达载体还包含至少一种控制序列。在一些实施方案中,控制序列是启动子。在一些另外的实施方案中,启动子是异源启动子。

[0373] 本发明还提供了宿主细胞,所述宿主细胞用编码具有苯丙氨酸解氨酶(PAL)活性的至少一种工程化多肽的至少一种多核苷酸序列来转化,和/或用至少一种表达载体来转化,所述表达载体包含编码具有苯丙氨酸解氨酶(PAL)活性的至少一种工程化多肽的至少一种多核苷酸序列和至少一种控制序列。在一些实施方案中,宿主细胞包含具有苯丙氨酸解氨酶(PAL)活性的至少一种工程化多肽,所述工程化多肽是密码子优化的。在一些实施方案中,宿主细胞是大肠杆菌(E.coli)。

[0374] 本发明还提供了在宿主细胞中产生至少一种工程化PAL多肽的方法,所述方法包括在适合的培养条件下,培养包含编码具有苯丙氨酸解氨酶(PAL)活性的至少一种工程化多肽的至少一种多核苷酸、和/或包含编码具有苯丙氨酸解氨酶(PAL)活性的至少一种工程化多肽的至少一种多核苷酸序列的至少一种表达载体的宿主细胞,使得该工程化PAL多肽被产生。在一些实施方案中,该方法还包括从培养物和/或宿主细胞中回收具有苯丙氨酸解氨酶(PAL)的至少一种工程化多肽的步骤。在一些另外的实施方案中,该方法还包括纯化具有苯丙氨酸解氨酶(PAL)的至少一种工程化多肽的步骤。

[0375] 本发明还提供了组合物,所述组合物包含至少一种如本文所提供的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽。在一些实施方案中,组合物是药物组合物。在一些实施方案中,组合物是膳食和/或营养补充物。在一些另外的实施方案中,药物组合物还包含至少一种药学上可接受的赋形剂和/或载体。在一些另外的实施方案中,组合物适合于苯丙酮尿症的治疗。在一些另外的实施方案中,组合物适合于口服施用至人。在一些实施方案中,组合物呈丸剂、片剂、胶囊、囊形片(gelcap)、液体或乳剂的形式。在又一些另外的实施方案中,丸剂、片剂、胶囊、或囊形片还包含肠溶包衣。在一些另外的实施方案中,药物组合物适合于肠胃外注射到人中。在一些实施方案中,药物组合物与至少一种另外的治疗有效化合物共施用。在一些另外的实施方案中,药物组合物包含至少一种另外的治疗有效化合物。在一些另外的实施方案中,药物组合物存在于膳食和/或营养补充物中。

[0376] 本发明还提供了用于治疗和/或预防受试者中苯丙酮尿症的症状的方法,所述方法包括:提供患有苯丙酮尿症的受试者,以及向受试者提供本文所提供的至少一种组合物。在一些实施方案中,组合物包含药物组合物,而在一些可选的实施方案中,组合物包含膳食/营养补充物。在该方法的一些实施方案中,苯丙酮尿症的症状被改善。在一些另外的实施方案中,被治疗的受试者能够食用在其甲硫氨酸、苯丙氨酸和/或酪氨酸含量方面比由表现出苯丙氨酸症状的受试者所需要的饮食较不受限的饮食。在一些实施方案中,被治疗的受试者(即,已经被提供了包含至少一种如本文所提供的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽的至少一种组合物的受试者)能够食用在其甲硫氨酸、苯丙氨酸和/或酪氨酸含量方面比由还未被提供至少一种如本文所提供的组合物的受试者所需要的饮食较不受限的饮食。在一些实施方案中,向受试者提供的组合物包含药物组合物,而在一些可选的实施方案中,组合物包含膳食/营养补充物。本发明还提供了被治疗的受试者,其中该受试者已被施用了至少一种组合物和/或药物组合物,所述至少一种组合物和/或药物组合物包含至少一种如本文所提供的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽。在一些实施方案中,受试者是选自以下的动物:灵长类、啮齿类动物和兔类动物。在一些另外的实施方案中,受

试者是小鼠。在一些另外的实施方案中,受试者是人。在仍然一些另外的实施方案中,受试者是人类婴儿或儿童,而在一些可选的实施方案中,受试者是人类成人或青年。I

[0377] 本发明还提供了包含至少一种本文所提供的具有苯丙氨酸解氨酶 (PAL) 活性的工程化多肽的组合物的用途。

[0378] 附图简述

[0379] 图1提供了野生型PAL蛋白序列的比对:多变鱼腥藻PAL (NCBI YP_324488.1 (SEQ ID NO:4));点状念珠藻苯丙氨酸/组氨酸解氨酶“NpPHAL” (NCBI YP_001865631.1 (SEQ ID NO:30));胶须藻属 (*Rivularia* sp.) 组氨酸解氨酶“RspHAL” (NCBI YP_007056096.1 (SEQ ID NO:31));颤藻属 (*Oscillatoria* sp.) 组氨酸解氨酶“Osp HAL” (NCBI YP_07108482.1 (SEQ ID NO:32));和粘球藻属 (*Gloeocapsa* sp.) 组氨酸解氨酶“GspHAL” (NCBI YP_007127054.1) (SEQ ID NO:33)。

[0380] 图2的A图显示了与野生型AvPAL相比,在pH 7.0下测试的变体22号 (SEQ ID NO:8)、变体30号 (SEQ ID NO:6) 和变体36号 (SEQ ID NO:10) 的降低的对蛋白水解的敏感性 (表示为胰凝乳蛋白酶和胰蛋白酶耐受性),如实施例4中进一步描述的。

[0381] 图2的B图提供了显示与野生型AvPAL相比,在pH 4.0至5.2下测试的变体22号、30号和36号的增加的对酸性pH的耐受性的图,如实施例4中进一步描述的。

[0382] 图3提供了野生型PAL和变体36号的 K_M 结果。

[0383] 图4提供了显示野生型PAL和变体36号的氨基酸特异性的数据。

[0384] 图5提供了显示暴露于人胰凝乳蛋白酶和胰蛋白酶的野生型PAL和变体36号的相对稳定性的结果。

[0385] 图6提供了显示暴露于猪胰腺提取物的野生型PAL和变体36号、42号和43号的相对稳定性的结果。

[0386] 图7提供了显示实施例5获得的结果的图。

[0387] 图8提供了显示实施例8获得的结果的图。

[0388] 发明描述

[0389] 本发明提供了工程化PAL多肽、突变体、生物活性片段及其类似物,以及包含工程化PAL多肽、突变体、生物活性片段及其类似物的药物组合物和工业组合物。

[0390] 本发明提供了工程化苯丙氨酸解氨酶 (PAL) 多肽及其组合物,以及编码工程化苯丙氨酸解氨酶 (PAL) 多肽的多核苷酸。在一些实施方案中,工程化PAL多肽被优化以提供增强的催化活性,以及降低的对蛋白水解的敏感性和增加的对酸性pH水平的耐受性。在一些实施方案中,工程化PAL多肽被去免疫化。本发明还涉及包含工程化PAL多肽的组合物用于治疗 and 工业目的的用途。

[0391] 缩写和定义:

[0392] 除非另外定义,本文所使用的所有技术和科学术语通常具有与本发明所属领域普通技术人员通常理解的相同的含义。通常,本文中所使用的命名法和下面所述的细胞培养、分子遗传学、微生物学、有机化学、分析化学和核酸化学的实验室程序是本领域熟知和通常采用的那些。此类技术在本领域技术人员熟知的许多文本和参考作品中被熟知和描述。标准技术或其修改用于化学合成和化学分析。本文中上文和下文均提及的所有专利、专利申请、文章和出版物在此通过引用明确并入本文。

[0393] 尽管发现在本发明的实践中使用类似或等同于本文所述那些的任何适合的方法和材料,一些方法和材料被本文描述。应当理解,本发明并不限于所述的具体方法学、方案和试剂,因为这些取决于本领域技术人员使用其的技术范围而变化。因此,下面即将定义的术语通过参考本申请作为整体而被更充分地描述。本文上文和下文均提及的所有专利、专利申请、文章和出版物,在此通过引用明确并入本文。

[0394] 此外,除非上下文另外清楚地指示,如本文所用的单数“一(a)”、“一(an)”和“该(the)”包括复数指示物。

[0395] 数值范围包括限定该范围的数字。因此,本文公开的每个数值范围意在包括落在此类较宽数值范围内的每一较窄数值范围,如同此类较窄数值范围在本文被全部清楚地写出。它还意在本文公开的每一最大(或最小)数值极限包括每一较低(或较高)的数值极限,如同此类较低(或较高)数值极限在本文被清楚地写出。

[0396] 术语“约”意指特定值的可接受误差。在某些情况下,“约”意指在给定值范围的0.05%、0.5%、1.0%或2.0%内。在某些情况下,“约”意指在给定值的1、2、3或4个标准差内。

[0397] 此外,本文所提供的标题不是可通过参考本申请作为整体而被具有的本发明的各个方面或实施方案的限制。因此,下面即将定义的术语通过参考本申请作为整体而被更充分地描述。尽管如此,为了便于理解本发明,许多术语定义如下。

[0398] 除非另外指示,核酸以5'至3'方向从左到右书写;氨基酸序列分别以氨基至羧基方向从左至右书写。

[0399] 如本文所用的术语“包含(comprising)”及其同源词以其包含的意义被使用(即,等同于术语“包括(including)”和其相应的同源词)。

[0400] “EC”数字是指Enzyme Nomenclature of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology(生物化学和分子生物学国际联合命名委员会)(NC-IUBMB)的酶命名法。该IUBMB生化分类是基于其催化的化学反应的酶数值分类系统。

[0401] “ATCC”是指American Type Culture Collection(美国典型培养物保藏中心),其生物保藏收集包括基因和菌株。

[0402] “NCBI”是指National Center for Biological Information(国家生物信息中心)和在其中所提供的序列数据库。

[0403] 如本文所用,术语“苯丙氨酸解氨酶(PAL)多肽”是指在还包括组氨酸解氨酶和酪氨酸解氨酶的芳香族氨基酸裂解酶家族(EC 4.3.1.23, EC 4.3.1.24和EC4.3.1.25)内的一类酶。因为一些PAL酶可使用酪氨酸和苯丙氨酸作为底物, PAL多肽有时还被称为苯丙氨酸/酪氨酸解氨酶。然而,本文公开和要求保护的AvPAL和变体不使用酪氨酸作为底物。PAL多肽催化L-苯丙氨酸转化为反式肉桂酸和氨。PAL活性指PAL多肽的酶活性。在一些优选的实施方案中, PAL酶还包含辅因子3,5-二氢-5-亚甲基-4H-咪唑-4-酮(MIO)。该辅因子可能为催化活性所需要,并通过保守的活性位点Ala167-Ser168-Gly169三肽区段的环化和脱水来形成。

[0404] “蛋白”、“多肽”和“肽”在本文可互换使用,以表示不管长度或翻译后修饰(例如,糖基化或磷酸化)通过酰胺键共价连接的至少两个氨基酸聚合物。

[0405] “氨基酸”通过其通常已知的三字母符号或通过IUPAC-IUB生物化学命名委员会推荐的单字母符号在本文被提及。同样地,核苷酸可通过其通常可接受的单字母代码被提及。

[0406] 当关于细胞、多核苷酸或多肽使用时,术语“工程化”、“重组”、“非天然存在的”和“变体”指如下材料或与该材料的天然或自然形式相应的材料:已经以天然本来不存在的方式被修饰或与其相同但由合成材料产生或衍生和/或通过使用重组技术操作产生。

[0407] 如本文所用,“野生型”和“天然存在的”是指在自然界中发现的形式。例如野生型多肽或多核苷酸序列是生物体中存在的序列,其可从天然来源分离且未通过人为操作而被有意识地修饰。

[0408] 如本文所用的“去免疫化”是指操作蛋白创造不如野生型或参考蛋白那样是免疫原性的变体。在一些实施方案中,去免疫化被完成,其中变体蛋白不刺激该变体蛋白被施用至其的患者中的免疫应答。该应答可通过多种方法来测量,所述多种方法包括但不限于,施用蛋白后中和(即,耐药物抗体)的存在或丰度、过敏性应答的存在或细胞因子释放的普遍程度(prevalence)或强度。在一些实施方案中,变体蛋白比野生型蛋白或参考蛋白的免疫原性更小。在一些实施方案中,去免疫化包括对被T细胞受体识别的蛋白(例如,表位)的修饰。在一些实施方案中,T细胞表位从野生型蛋白或参考蛋白中除去,以便产生去免疫化的变体蛋白。在一些实施方案中,去免疫化的蛋白显示了对人体免疫应答的生物化学和细胞生物学的预测因子的较低水平的应答,该应答包括树突状细胞T细胞活化测定或人白细胞抗原(HLA)肽结合测定。

[0409] “编码序列”指编码蛋白的氨基酸序列的核酸的部分(例如,基因)。

[0410] 术语“序列同一性百分比(%)”在本文中使用以指多核苷酸和多肽之间的比较,并通过跨比较窗比较两条最佳比对的序列来确定,其中多核苷酸或多肽序列在比较窗中的部分与参考序列相比可包括添加或缺失(即,缺口),以用于两个序列的最佳比对。百分比可如下计算:通过确定两个序列中出现相同的核酸碱基或氨基酸残基的位置的数目,以产生匹配位置的数目,将匹配位置的数目除以比较窗中位置的总数目,并将结果乘以100以得到序列同一性的百分比。可选地,百分比可如下计算:通过确定两个序列中出现相同的核酸碱基或氨基酸残基或核酸碱基或氨基酸残基与缺口对齐的位置的数目,以产生匹配位置的数目,将匹配位置的数目除以比较窗中位置的总数目,并将结果乘以100以得到序列同一性的百分比。本领域技术人员理解,存在许多可用于比对两个序列的已建立的算法。用于比较的最佳序列比对如下进行,例如,如本领域已知的,通过Smith和Waterman的局部同源性算法(Smith和Waterman,Adv. Appl. Math., 2:482[1981])、通过Needleman和Wunsch的同源性比对算法(Needleman和Wunsch,J. Mol. Biol., 48:443[1970])、通过Pearson和Lipman的相似度检索方法(Pearson和Lipman,Proc. Natl. Acad. Sci. USA 85:2444[1988])、通过这些算法的计算机实现(例如,在GCG Wisconsin软件包中的GAP、BESTFIT、FASTA和TFASTA)或通过目测。适合于确定序列同一性百分比和序列相似性的算法的实例包括但不限于BLAST和BLAST 2.0算法(参见例如,Altschul等,J. Mol. Biol., 215:403-410[1990];和Altschul等,Nucleic Acids Res., 3389-3402[1977])。用于进行BLAST分析的软件为通过美国国家生物技术信息中心网站公共可获得的。该算法包括首先通过识别查询序列(query sequence)中具有长度“W”的短字来确定高评分序列对(HSP),当其与数据库序列中相同长度的字比对

时,所述短字匹配或满足某个正值阈值评分“T”。T被称为邻近字评分阈值(参见,Altschul等,同上)。这些最初的邻近字击中(word hit)用作启动检索的种子以找到更长的包括它们的HSP。然后字击中沿着每个序列的两个方向延伸到累积比对评分不能增加的程度。对于核苷酸序列,累积评分使用参数“M”(用于一对匹配残基的奖励评分;总是>0)和“N”(用于错配残基的惩罚评分;总是<0)来计算。对于氨基酸序列,使用评分矩阵以计算累积评分。当发生以下情况时字击中在每个方向的延伸停止:累积比对评分从其最大获得的值下降了量“X”时;由于一个或多个负评分残基比对的累积,累积评分变成零或以下时;或达到任一序列的末端。BLAST算法参数W、T和X决定比对的灵敏度和速度。BLASTN程序(对于核苷酸序列)使用11的字长(W)、10的期望值(E)、M=5、N=-4、以及两个链的比较作为默认参数。对于氨基酸序列,BLASTP程序使用3的字长(W)、10的期望值(E),和BLOSUM62评分矩阵作为缺省(参见例如,Henikoff和Henikoff,Proc.Natl.Acad.Sci.USA 89:10915[1989])。序列比对与%序列同一性的示例性确定可使用GCG Wisconsin软件包(Accelrys,Madison WI)中的BESTFIT或GAP程序,使用提供的默认参数。

[0411] “参考序列”是指用作序列比较的基础的确定序列。参考序列可以是更大序列的子集,例如,全长基因或多肽序列的片段。通常,参考序列是至少20个核苷酸或氨基酸残基的长度、至少25个残基的长度、至少50个残基的长度、至少100个残基的长度或核酸或多肽的全长。由于两种多核苷酸或多肽可以各自(1)包括两个序列之间相似的序列(即,完整序列的一部分),和(2)还可包括两个序列之间不同的序列,两种(或更多种)多核苷酸或多肽之间的序列比较通常通过在“比较窗口”内比较两种多核苷酸或多肽的序列来进行,以确定和比较具有序列相似性的局部区域。在一些实施方案中,“参考序列”可基于基本氨基酸序列,其中参考序列是可在基本序列中具有一个或更多变化的序列。例如,短语“基于SEQ ID NO:4的在相应于X39的残基上具有缬氨酸的参考序列”指在SEQ ID NO:4中的位置X39上相应的残基(例如,丙氨酸)已被改变为缬氨酸的参考序列。

[0412] “比较窗”是指至少约20个连续核苷酸位置或氨基酸残基的概念性区段,其中序列可以与至少20个连续核苷酸或氨基酸的参考序列进行比较,并且其中在比较窗口中序列的一部分与用于两个序列的最佳比对的参考序列(其不包含添加或缺失)相比,可包括20%或更少的添加或缺失(即,缺口)。比较窗口可以比20个连续残基更长,并任选地包括30、40、50、100或更长的窗。

[0413] 当在给定的氨基酸或多核苷酸序列编号的上下文中使用时,“对应于”、“关于”和“相对于”是指当给定的氨基酸或多核苷酸序列与指定的参考序列相比时,参考序列的残基编号。换言之,给定的聚合物的残基数目或残基位置关于参考序列被指定,而不是通过给定的氨基酸或多核苷酸序列内的残基的实际数字位置被指定。例如,给定的氨基酸序列,诸如工程化PAL的氨基酸序列可通过引入空位以优化两条序列之间的残基匹配与参考序列比对。在这些情况中,尽管存在空位,在给定的氨基酸或多核苷酸序列中的残基的编号关于与其比对的参考序列作出。

[0414] “氨基酸差异”和“残基差异”是指多肽序列的位置处的氨基酸残基相对于参考序列中在相应位置处的氨基酸残基的差异。氨基酸差异的位置通常在本文中称为“X_n”,其中n指残基差异基于其的参考序列中的相应位置。例如,“与SEQ ID NO:4相比在位置X₉₁上的残基差异”是指相应于SEQ ID NO:4的位置91的多肽位置处的氨基酸残基的差异。因此,如果

SEQ ID NO:4的参考多肽在位置91上具有丙氨酸,那么“与SEQ ID NO:4相比在位置X91上的残基差异”是指相应于SEQ ID NO:4的位置的91的多肽位置处的丙氨酸以外的任何残基的氨基酸置换。在本文的大多数情况下,位置处的特定氨基酸残基差异表示为“X_nY”,其中“X_n”是指定参考多肽的相应残基和位置(如上所述),并且“Y”是工程化多肽中发现的氨基酸的单字母标识符(即,与参考多肽相比的不同残基)。在一些情况下(例如,在实施例的表中),本公开内容还提供了由常规表示法“AnB”表示的特定氨基酸差异,其中A是参考序列中残基的单字母标识符,“n”是参考序列中的残基位置的数目,且B是工程化多肽的序列中的残基置换的单字母标识符。在一些情况下,本公开内容的多肽相对于参考序列可包含一个或更多个氨基酸残基,其由相对于参考序列存在残基差异的一列特定位置指示。在一些实施方案中,当多于一个氨基酸可在多肽的特定残基位置中使用时,可使用的各种氨基酸残基由“/”分离(例如,X307G/X307Q或X307G/Q)。本公开内容包括包含保守和非保守的氨基酸置换的任一个/或两者的一个或更多个氨基酸差异的工程化多肽序列置换。

[0415] 术语“氨基酸置换集合”和“置换集合”是指多肽序列内的一组氨基酸置换。在一些实施方案中,置换集合包括1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个或更多个氨基酸置换。在一些实施方案中,置换集合指存在于在实施例的任一表中列出的变体AvPAL多肽的任何一个中的氨基酸置换的集合。例如,在变体36中存在的置换集合为A39V/A91V/N290G/H307G/L407V/T524S,其中所述氨基酸位置是相对于SEQ ID NO:4。

[0416] “保守氨基酸置换”是指用具有相似侧链的不同残基置换残基,并且因此通常涉及在相同或相似定义类别的氨基酸内的氨基酸置换多肽中的氨基酸。通过示例的方式而非限制,具有脂族侧链的氨基酸可被另一个脂肪族氨基酸置换(例如,丙氨酸、缬氨酸、亮氨酸和异亮氨酸);具有羟基侧链的氨基酸被具有羟基侧链的另一个氨基酸置换(例如,丝氨酸和苏氨酸);具有芳香族侧链的氨基酸被具有芳香族侧链的另一个氨基酸置换(例如,苯丙氨酸、酪氨酸、色氨酸和组氨酸);具有碱性侧链的氨基酸被具有碱性侧链的另一个氨基酸置换(例如,赖氨酸和精氨酸);具有酸性侧链的氨基酸被具有酸性侧链的另一个氨基酸置换(例如,天冬氨酸或谷氨酸);且疏水性氨基酸或亲水性氨基酸分别被另一个疏水性氨基酸或亲水性氨基酸置换。

[0417] “非保守置换”是指用具有显著差异侧链性质的氨基酸置换多肽中的氨基酸置换。非保守置换可以使用限定组之间,而不是它们之内的氨基酸,并影响:(a)在置换的区域中肽骨架的结构(例如,脯氨酸置换甘氨酸);(b)电荷或疏水性;和/或(c)侧链的体积。通过示例的方式而非限制,示例性非保守性置换包括酸性氨基酸被碱性或脂肪族氨基酸置换;芳香族氨基酸被小氨基酸置换;和亲水性氨基酸被疏水性氨基酸置换。

[0418] “缺失”是指通过从参考多肽中除去一个或更多个氨基酸而对多肽进行的修饰。缺失可包括除去1个或更多个氨基酸、2个或更多个氨基酸、5个或更多个氨基酸、10个或更多个氨基酸、15个或更多个氨基酸或20个或更多个氨基酸、多至组成参考酶的氨基酸总数的10%、或多至氨基酸总数的20%,同时保留酶活性和/或保留工程化转氨酶的改进的特性。缺失可涉及多肽的内部和/或端部。在各个实施方案中,缺失可包括连续的区段或可以是不连续的。

[0419] “插入”是指通过向参考多肽添加一个或更多个氨基酸而对多肽进行的修饰。插入

可以在多肽的内部或到羧基或氨基末端。如本文所用的插入包括如本领域已知的融合蛋白。插入可以是氨基酸的连续区段或由天然存在的多肽中的一个或多个氨基酸分隔。

[0420] 术语“功能片段”和“生物活性片段”在本文可互换使用,是指如下多肽:所述多肽具有氨基末端和/或羧基末端缺失和/或内部缺失,但其中剩余的氨基酸序列与它与其相比的序列(例如,本发明的全长工程化PAL)中的相应位置相同,并且其保留全长多肽的基本上所有的活性。

[0421] “分离的多肽”是指如下多肽:所述多肽与其天然伴随的其他污染物(例如,蛋白、脂质和多核苷酸)基本上分离。术语包括已从它们天然存在的环境或表达系统(例如,宿主细胞或体外合成)中取出或纯化的多肽。重组PAL多肽可存在于细胞内、存在于细胞培养基中,或以各种形式制备,诸如裂解物或分离的制剂。因此,在一些实施方案中,本文提供的重组PAL多肽是分离的多肽。

[0422] “实际上纯化的多肽”指如下组合物,在所述组合物中多肽物种是存在的优势物种(即,在摩尔或重量基础上,它比在该组合物中的任何其他单独的大分子物质更丰富),并且当目标物种构成存在的大分子物种的按摩尔或%重量计至少约50%时,通常是基本上纯化的组合物。通常,基本上纯化的PAL组合物将构成该组合物中存在的所有大分子物种的按摩尔或%重量计约60%或更多、约70%或更多、约80%或更多、约90%或更多、约95%或更多以及约98%或更多。在一些实施方案中,将目标物种纯化至基本的均一性(即,通过常规检测方法不能在组合物中检测出污染物物种),其中该组合物基本上由单一大分子物种组成。溶剂物种、小分子(<500道尔顿)和元素离子物种不被认为是大分子物种。在一些实施方案中,分离的重组PAL多肽是基本上纯化的多肽组合物。

[0423] “改进的酶特性”是指与参考PAL多肽诸如野生型PAL多肽(例如,具有SEQ ID NO:4的AvPAL野生型)或另一种工程化PAL多肽相比表现出在任何酶特性中的改进的工程化的PAL多肽。改进的特性包括但不限于这样的特性,如增加的蛋白表达、增加的热活性(thermoactivity)、增加的热稳定性、增加的pH活性、增加的稳定性、增加的酶活性、增加的底物特异性和/或亲和力、增加的比活性、增加的对底物和/或终产物抑制的抑制性、增加的化学稳定性、改进的化学选择性、改进的溶剂稳定性、增加的对酸性pH耐受性、增加的对蛋白水解活性耐受性(即,降低的对蛋白水解的敏感性)、降低聚集、增加的溶解度、降低的免疫原性和改变的温度曲线。

[0424] “增加的酶活性”和“增强的催化活性”是指工程化PAL多肽的改进的特性,其可被表示为与参考PAL酶(例如,野生型AvPAL和/或另一种工程化AvPAL)相比,比活性的增加(例如,产生的产物/时间/重量蛋白)和/或底物至产物的转化率百分比的增加(例如,在指定的时间段使用指定量的PAL,起始量的底物至产物的转化率百分比)。确定酶活性的示例性方法在实施例中被提供。可影响与酶活性相关的任何特性,包括经典酶特性 K_m 、 V_{max} 或 k_{cat} ,它们的改变可导致增加的酶活性。酶活性的改进可以是相应野生型酶的酶活性的约1.1倍到相比于天然存在的PAL或从其衍生PAL多肽的另一种工程化PAL的多达2倍、5倍、10倍、20倍、25倍、50倍、75倍、100倍、150倍、200倍或更多的酶活性。

[0425] 在一些实施方案中,工程化PAL多肽具有以下的 k_{cat} :至少0.1/秒、至少0.2/秒、至少0.3/秒、至少0.5/秒、至少1.0/秒的,且在一些优选的实施方案中大于1.0/秒。在一些实施方案中, K_m 在约1 μ M至约5mM的范围内;在约5 μ M至约2mM的范围内;在约10 μ M至约2mM的范

围内;或在约10 μ m至约1mM的范围内。在一些特别的实施方案中,工程化PAL酶表现出比参考PAL酶的酶活性大1.5至10倍、1.5至25倍、1.5至50倍、1.5至100倍或更大的范围内的改进的酶活性。PAL活性可通过本领域已知的任何标准测定来测量(例如,通过监测反应物或产物的分光光度法性质中的变化)。在一些实施方案中,产生的产物的量通过高效液相色谱法(HPLC)分离结合o-酞二醛(OPA)衍生化后的UV吸光度或荧光检测来测量。在一些实施方案中,使用确定的酶制剂、在设定条件下的确定的测定、和一种或多种确定的底物进行酶活性的比较,如本文进一步详细地描述的。通常,当比较裂解物时,细胞的数目和测定的蛋白的量被确定,并使用相同的表达系统和相同的宿主细胞以将由宿主细胞产生的和裂解物中存在的酶的量的变化最小化。

[0426] 术语“改进的对酸性pH的耐受性”意指与参考PAL相比,根据本发明的重组PAL表现出增加的稳定性(即,在暴露于酸性pH持续指定的时间段[1小时,多达24小时]后,在约pH 7.0下的更高的保留的活性)。

[0427] 如本文所用的“生理pH”意指通常在受试者(例如,人)的小肠内发现的pH范围。通常从幽门瓣到大肠存在约6.0至7.5的范围内梯度pH。

[0428] 关于改进的对酸性pH条件的稳定性或增加的对酸性pH的耐受性所用的术语“酸性pH”意指约1.5至6.8的pH范围。

[0429] 本文可互换使用的术语“蛋白水解活性”和“蛋白水解”是指蛋白分解为较小的多肽或氨基酸。蛋白的分解通常是肽键通过蛋白酶(蛋白酶)的酶水解的结果。蛋白酶包括但不限于,胃蛋白酶、胰蛋白酶、胰凝乳蛋白酶、弹性蛋白酶;羧肽酶A和B和肽酶(例如,氨基肽酶、二肽酶和肠肽酶)。

[0430] 短语“降低的对蛋白水解敏感性”和“降低的蛋白水解敏感性”在本文中可互换使用,意指在标准测定(例如,如实施例中公开的)中,在用一种或更多种蛋白酶上理之后,根据本发明的工程化PAL多肽与参考PAL相比将具有较高的酶活性。

[0431] “聚集”意指PAL多肽的凝结或沉淀。聚集可导致酶失活。术语“减少的聚集”意指工程化PAL多肽与参考PAL相比将更不易聚集。用于评价聚集的方法是本领域已知的,包括但不限于用适当染料(例如,硫磺素T或尼罗红)的荧光显微镜检查术、动态光散射、用适当染料(例如,氟硼荧(Bodipy))的流式细胞术、通过SDS-PAGE的过滤和分析、和/或蛋白印迹、荧光相关光谱法和电子显微镜的使用。存在评价聚集的商购可得的试剂盒(例如,ProteoStat[®]蛋白聚集测定试剂盒[Enzo])。

[0432] “转化率”指底物至相应的产物的酶促转化(或生物转化)。“百分比转化率”是指在指定条件下一段时间内被转化为产物的底物的百分比。因此,PAL多肽的“酶活性”或“活性”可表示为在指定的一段时间内底物至产物的“转化率百分比”。

[0433] “杂交严格性”涉及核酸杂交中的杂交条件,诸如洗涤的条件。通常,杂交反应在较低严格性的条件下进行,随后是不同的但较高严格性的洗涤。术语“中度严格杂交”是指允许靶-DNA结合以下互补的核酸的条件,所述互补的核酸具有与靶DNA约60%的同一性、优选地约75%的同一性、约85%的同一性,与靶-多核苷酸大于约90%的同一性。示例性中度严格条件是等同在于42 $^{\circ}$ C下于50%甲酰胺、5 $\times$ Denhart溶液、5 $\times$ SSPE、0.2%SDS中杂交,随后在42 $^{\circ}$ C下于0.2 $\times$ SSPE、0.2%SDS中洗涤的条件。“高严格度杂交”通常是指以下的条件:偏离对于定义的多核苷酸序列在溶液条件下确定的热融化温度 T_m 约10 $^{\circ}$ C或更少。在一些实施

方案中,高严格度条件指以下的条件,其仅允许在65℃下于0.018M NaCl中形成稳定的杂交体的那些核酸序列的杂交(即,如果杂交体在65℃下于0.018M NaCl中是不稳定的,它在如本文考虑的在高严格度条件下将是不稳定的)。例如,可通过在等同于在42℃下于50%甲酰胺、5×Denhart溶液、5×SSPE、0.2%SDS的条件中杂交,随后在65℃下于0.1×SSPE和0.1%SDS中洗涤,而提供高度严格条件。另一种高严格度条件是在与以下等同的条件中杂交:在65℃下于含0.1% (w:v) SDS的5×SSC中杂交,和在65℃下于含0.1%SDS的0.1×SSC中洗涤。其他高严格度杂交条件,以及中度严格条件描述于以上引用的参考文献中。

[0434] “密码子优化的”是指编码蛋白的多核苷酸的密码子变为特定生物体中优先使用的那些密码子,使得所编码的蛋白在该生物体中更有效地表达。尽管遗传密码由于大多数氨基酸被称为“同义密码子”或“同义”密码子的几种密码子代表而为简并的,但众所周知具体生物体的密码子使用是非随机的且偏向特定的密码子三联体。就给定的基因、具有共同功能或祖先起源的基因、高表达的蛋白相对于低拷贝数蛋白和生物体基因组的聚集蛋白编码区而言,这种密码子使用偏向可能更高。在一些实施方案中,对编码PAL酶的多核苷酸进行优化,以用于从被选择用于表达的宿主生物体优化生产。

[0435] 本文中“控制序列”是指包括对本公开内容的多核苷酸和/或多肽的表达是必要的或有利的的所有组分。每个控制序列对编码多肽的核酸序列来说可以是天然的或外源的。此类控制序列包括但不限于,前导序列、多腺苷酸序列、前肽序列、启动子序列、信号肽序列、起始序列和转录终止子。在最小程度上,控制序列包括启动子和转录及翻译终止信号。在一些实施方案中,控制序列设置有连接体,用于引入特异性限制位点的目的,促进控制序列与编码多肽的核酸序列的编码区的连接。

[0436] “可操作地连接的”在本文中定义为如下一种配制:在所述配置中控制序列相对于感兴趣的多核苷酸被适当放置(即,以功能关系)在一位置中,使得控制序列指导或调整编码感兴趣的多肽的多核苷酸的表达。

[0437] “启动子序列”指被宿主细胞识别用于感兴趣的多核苷酸的表达的核酸序列,诸如编码序列。启动子序列包含介导感兴趣的多核苷酸的表达的转录控制序列。启动子可以是在选择的宿主细胞中显示转录活性的任何核酸序列,包括突变、截短的和杂合启动子,的,并且可以从编码与宿主细胞同源或异源的细胞外或细胞内多肽的基因来获得。

[0438] “适当的反应条件”是指在酶促转化反应溶液中的那些条件(例如,酶载量、底物载量、温度、pH、缓冲液、助溶剂等的范围),在上述条件下本公开内容的PAL多肽能够将底物转化为期望的产物化合物。示例性的“适当的反应条件”被提供在本文中(参见,实施例)。

[0439] “载量”,诸如在“化合物载量”或“酶载量”中是指在反应开始时反应混合物中组分的浓度或量。在酶促转化反应过程的上下文中的“底物”是指由PAL多肽作用产生的化合物或分子。在酶促转化过程的上下文中的“产物”是指从PAL多肽对底物的作用产生的化合物或分子。

[0440] 如本文所用的术语“培养”是指微生物细胞群使用任何适合的培养基(例如,液体、凝胶或固体)在适当的条件下适合的生长。

[0441] 重组多肽(例如,PAL酶变体)可使用本领域已知的任何适合的方法来产生。例如,存在本领域技术人员熟知的很多不同的诱变技术。另外,诱变试剂盒还可购自许多商业分子生物学供应商。方法可用于做出在限定的氨基酸(定点)上的特定置换、在基因的局部区

域中的特异性(区域-特异性)或随机突变、或在整个基因内的随机诱变(例如,饱和诱变)。许多产生酶变体的适合的方法是本领域技术人员已知的,包括但不限于使用PCR、盒式诱变、基因合成、易错PCR、改组和化学饱和诱变或本领域已知的任何其他适合的方法进行单链DNA或双链DNA的定点诱变。用于DNA和蛋白工程化的方法的非限制性实例在以下专利中被提供:美国专利号6,117,679;美国专利号6,420,175;美国专利号6,376,246;美国专利号6,586,182;美国专利号7,747,391;美国专利号7,747,393;美国专利号7,783,428;和美国专利号8,383,346。变体产生之后,它们可被筛选任何期望的特性(例如,高或增加的活性或低或降低的活性、增加的热活性、增加的热稳定性和/或酸性pH稳定性等)。在一些实施方案中,“重组PAL多肽”(在本文中还称为“工程化PAL多肽”、“变体PAL酶”和“变体PAL体”)发现用途。

[0442] 如本文所用,“载体”是用于将DNA序列导入细胞的DNA构建体。在一些实施方案中,载体为可操作地被连接到能够影响编码多肽的DNA序列在适合的宿主中表达的适合的控制序列的表达载体。在一些实施方案中,“表达载体”具有可操作地连接到DNA序列(例如,转基因)以驱动宿主细胞中表达的启动子序列,并且在一些实施方案中,还包括转录终止子序列。

[0443] 如本文所用的,术语“表达”包括参与多肽产生的任何步骤,包括但不限于,转录、转录后修饰、翻译和翻译后修饰。在一些实施方案中,该术语还包括从细胞中分泌多肽。

[0444] 如本文所用的,术语“产生”是指蛋白和/或其他化合物通过细胞的产生。意图是,术语包括参与多肽产生的任何步骤,包括但不限于,转录、转录后修饰、翻译和翻译后修饰。在一些实施方案中,该术语还包括从细胞中分泌多肽。

[0445] 如本文所用的,如果两个序列在自然界中未缔合,氨基酸或核苷酸序列(例如,启动子序列、信号肽、终止子序列等)与它被可操作地连接至其的另一个序列是“异源的”。

[0446] 如本文所用的,术语“宿主细胞”和“宿主菌株”是指用于表达包含本文提供的DNA(例如,编码至少一种AvPAL变体的多核苷酸序列)的载体的适合的宿主。在一些实施方案中,宿主细胞是已用使用如本领域已知的重组DNA技术构建的载体转化或转染的原核或真核细胞。

[0447] 术语“类似物”意指具有与参考多肽大于70%的序列同一性,但小于100%的序列同一性的多肽(例如,大于75%、78%、80%、83%、85%、88%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的序列同一性)的多肽。在一些实施方案中,类似物包括非天然存在的氨基酸残基以及天然存在的氨基酸,所述非天然存在的氨基酸残基包括但不限于高精氨酸、鸟氨酸和正缬氨酸。在一些实施方案中,类似物还包括一种或更多种D-氨基酸残基和两个或更多个氨基酸残基之间的非肽键。

[0448] 术语“治疗的”是指向显示病理学迹象或症状的受试者施用具有有益或期望的医疗效应的化合物。

[0449] 术语“药物组合物”是指适合于哺乳动物受试者(例如,人)中药物用途的包含由本发明所包括的药学有效量的工程化PAL多肽和可接受的载体的组合物。

[0450] 术语“有效量”意指足以产生期望的结果的量。本领域一般技术人员可通过使用常规实验确定有效量是多少。

[0451] 使用术语“分离的”和“纯化的”指从它与其天然缔合的至少一种其他组分中取出

的分子(例如,分离的核酸、多肽等)或其他组分。术语“纯化的”不要求绝对纯度,而它的目的是作为相对的定义。

[0452] 术语“受试者”包括哺乳动物,诸如人、非人灵长类动物、家畜、宠物和实验动物(例如,啮齿动物和兔类动物)。它的目的是,术语包括女性以及男性。

[0453] 如本文所用的,术语“患者”意指被评价、治疗或正经历疾病的任何受试者。

[0454] 术语“婴儿”是指在出生之后第一个月至约一(1)岁的时期内的儿童。如本文所用的,术语“新生儿”是指在从出生至第生命的28天的时期内的儿童。术语“早产儿”是指完成妊娠二十周后但妊娠期满之前出生的婴儿,通常在出生时称重~500-~2499克。“极低出生体重儿”是在出生时称重低于1500g的婴儿。

[0455] 如本文所用的,术语“儿童”是指对于同意治疗或研究程序未达到法定年龄的人。在一些实施方案中,该术语是指在出生和青春期的时期之间的人。

[0456] 如本文所用的,术语“成人”是指对于相关司法权已达到法定年龄的人(例如,在美国为18岁)。在一些实施方案中,该术语是指任何完全发育成熟的生物体。在一些实施方案中,术语“青年”是指小于18岁但已经达到性成熟的人。

[0457] 如本文所用,“组合物”和“制剂”包括目的用于任何适合的用途的包含本发明的至少一种工程化PAL的产物(例如,药物组合物、膳食/营养补充物、饲料等)。

[0458] 术语“施用”和“施用”组合物意指向受试者(例如,患有PKU效应的人)提供本发明的组合物。

[0459] 当关于药物组合物使用时,术语“载体”意指标准药物载体,缓冲液和赋形剂诸如稳定剂、防腐剂和佐剂中的任一种。

[0460] 术语“药学上可接受的”意指可被施用至受试者而不引起任何不良生物效应或以有害的方式与在其中它被包含的组分的任何一个相互作用且拥有期望的生物活性的材料。

[0461] 如本文所用的,术语“赋形剂”是指任何药学上可接受的添加剂、载体、稀释剂、佐剂或其他成分,而不是活性药物成分(API;例如,本发明的工程化PAL多肽)。赋形剂通常包括配制和/或施用目的。

[0462] 当关于疾病/状况的症状使用时,术语“治疗有效量”是指改善、减弱或消除疾病/状况的一个或更多个症状或预防或延缓症状(例如,PKU)的发作的化合物(例如,工程化PAL多肽)的量 and/或浓度。在一些实施方案中,该术语关于组合物的量被使用,所述组合物的量引发被研究者、医师、兽医或其他临床医师寻找的组织、系统或动物受试者的生物学(例如,医学)应答。。

[0463] 当关于疾病/状况使用时,术语“治疗有效量”是指改善、减弱或消除该疾病/状况的组合物的量和/或浓度。

[0464] 它的目的是,术语“治疗(treating)”、“治疗(treat)”和“治疗(treatment)”包括预防的治疗(例如,预防剂),以及舒减治疗。

[0465] 工程化PAL多肽:

[0466] 本发明的工程化PAL多肽源自其的亲本PAL多肽包括细菌菌株诸如鱼腥藻属(*Anabaena*) (例如,多变鱼腥藻(*A. variabilis*))、念珠藻属(*Nostoc*) (例如,点状念珠藻(*N. punctiforme*))、红冬孢酵母属(*Rhodospiridium*) (例如,圆红冬孢酵母(*R. toruloides*))、链霉菌属(*Streptomyces*) (例如*S. maritimus*或*S. verticillatus*)、颤菌

属(*Oscillatoria* sp.)、粘球藻属(*Gloeocapsa* sp.)和胶须藻属(*Rivularia* sp.)。来自这些菌株的PAL酶已被鉴定,并是熟知的。在图1中提供了来自鱼腥藻属(多变鱼腥藻) ATCC 29413和NCBI YP_324488.1;念珠藻属(点状念珠藻) ATCC 29133和NCBI YP_00186563.1;颤藻属(*Oscillatoria* sp.) PCC 6506和NCBI ZP_07108482.1和粘球藻属(*Gloeocapsa* sp.) PCC 7428和NCBI YP_007127054.1的同源酶序列。点状念珠藻苯丙氨酸/组氨酸解氨酶“NpPHAL”(NCBI YP_001865631.1 (SEQ ID NO:30);胶须藻属(*Rivularia* sp.)组氨酸解氨酶“RspHAL”(NCBI YP_007056096.1 (SEQ ID NO:31);颤藻属(*Oscillatoria* sp.)组氨酸解氨酶“Osp HAL”(NCBI YP_07108482.1 (SEQ ID NO:32);和粘球藻属(*Gloeocapsa* sp.)组氨酸解氨酶“GspHAL”(NCBI YP_007127054.1) (SEQ ID NO:33)与AvPAL (SEQ ID NO:4)具有多于70%的同源性。

[0467] 此外,当特定的PAL变体(即,工程化PAL多肽)通过参考野生型PAL或参考PAL的序列中的特定氨基酸残基的修饰被提及时,应该理解的是,本文包括了在等同位置修饰的另一种PAL的变体(如从各自的氨基酸序列之间的任选的氨基酸序列比对中所确定的)。在一些实施方案中,工程化PAL多肽源自从以上细菌菌株(即,念珠藻属[点状念珠藻]、红冬孢酵母属[圆红冬孢酵母]、链霉菌属[*S.maritimus*或*S.verticillatus*]、颤藻属(*Oscillatoria* sp.)、粘球藻属(*Gloeocapsa* sp.)和胶须藻属(*Rivularia* sp.))列出的多肽中的任何一种。在一些另外的实施方案中,本发明的工程化PAL多肽包含保守的活性位点Ala167-Ser168-Gly169,并且包含与SEQ ID NO:4至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的序列同一性。在一些实施方案中,工程化PAL多肽不仅包含PAL活性,还具有对酪氨酸和/或组氨酸底物的活性。

[0468] 在一些实施方案中,工程化PAL多肽通过在有利于产生工程化PAL多肽的条件下培养包含编码至少一种工程化PAL多肽的至少一种多核苷酸序列的微生物来产生。在一些实施方案中,工程化PAL多肽随后从所得的培养基和/或细胞中回收。

[0469] 本发明提供了具有PAL活性的示例性工程化PAL多肽。实施例提供了显示与工程化PAL多肽的功能活性相关的特定氨基酸序列特征的序列结构信息的表。该结构-功能相关信息以相对于SEQ ID NO:4的参考工程化多肽的特定氨基酸残基差异、以及与示例性工程化PAL多肽的相关的实验上确定的活性数据的形式来提供。

[0470] 在一些实施方案中,本发明的具有PAL活性的工程化PAL多肽包含a) 氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:4至少85%的序列同一性;b) 与SEQ ID NO:4相比在一个或更多个氨基酸位置处的氨基酸残基差异;并且c) 所述工程化PAL多肽与参考序列相比表现出选自以下的改进的特性:i) 增强的催化活性、ii) 降低的蛋白水解敏感性、iii) 增加的对酸性pH的耐受性、iv) 减少的聚集、或i)、ii)、iii)、或iv)的任何组合。

[0471] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:4相比在一个或更多个氨基酸位置处(诸如与SEQ ID NO:4或具有与SEQ ID NO:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在1、2、3、4、5、6、7、8、9、

10、11、12、14、15、20或更多个氨基酸位置处)的氨基酸残基差异。在一些实施方案中,与SEQ ID NO:4相比在一个或更多个位置处的残基差异包括至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个保守氨基酸置换。在一些实施方案中,工程化PAL多肽是在实施例中提供的表中所列的多肽。

[0472] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,以及当与SEQ ID NO:4的氨基酸序列最佳比对时,与SEQ ID NO:4相比在选自以下的一个或更多个氨基酸位置处的氨基酸残基差异:X39;X54;X59;X73;X91;X158;X112,X134,X180;X195;X240;X243;X245;X256;X257;X270;X290;X304;X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;X564或其任何组合。在一些实施方案中,氨基酸差异是1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、15个或20个或更多个氨基酸位置。

[0473] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:4至少85% (至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%)的序列同一性,并且包含在位置H307处的氨基酸残基差异和任选地在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置处的氨基酸残基差异。在一些实施方案中,在位置307处的氨基酸残基差异为H307/G/Q/M。

[0474] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:4至少85% (至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%)的序列同一性,并且包含选自以下的一个或更多个的组合的至少一个氨基酸残基差异:A39;T54;G59;S73;A91;Y158;S180;K195;A112;R134;Q240;T243;I245;A256;L257;N270;N290;Y304;R305;H307;E308;I326;L349;D353;L364;A394;S399;N400;P404;L407;F443;N453;Y459;T460;T463;N474;E509;Q521;K522;T524;P528;S546;和/或P564。在一些另外的实施方案中,存在在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置处的氨基酸残基差异。

[0475] 在一些实施方案中,表现出改进的特性的工程化PAL多肽具有与SEQ ID NO:4至少85% (至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%)的序列同一性,并且包含当与SEQ ID NO:4最佳比对时选自以下的一个或更多个的组合的氨基酸残基差异:A39V;T54K;G59R;S73K;A112C;R134Q;A91V;Y158H;S180A;K195E;Q240R/W;T243I/L;I245L;A256G;L257W/A;N270K;N290G;Y304H;R305M;H307G/Q/M;E308Q;I326F;L349M;D353A/N;L364Q;A394V;S399N;N400K;P404A;L407V;F443H;N453G;Y459F;T460G;T463N;N474Q;E509L;Q521K/S;K522Y/F/N;T524S;P528L;S546R;和P564G/L/M。

[0476] 在一些实施方案中,氨基酸残基差异选自当与SEQ ID NO:4最佳比对时以下的一个或更多个的组合:A39V;A91V;A256G;N290G;A394V;S399N;P404A;L407V;K522Y/F/N;和/或T524S。

[0477] 在一些实施方案中,本发明提供了工程化PAL多肽的功能片段。在一些实施方案中,功能片段包括其所源自的工程化PAL多肽(即,亲本工程化PAL)的至少约90%、至少约

95%、至少约96%、至少约97%、至少约98%或至少约99%的活性。在一些实施方案中,功能片段包含至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的工程化PAL的亲本序列。在一些实施方案中,功能片段将被截短小于5个、小于10个、小于15个、小于10个、小于25个、小于30个、小于35个、小于40个、小于45个和小于50个氨基酸。

[0478] 在一些实施方案中,本发明提供了工程化PAL多肽的功能片段。在一些实施方案中,功能片段包括其所源自的工程化PAL多肽(即,亲本工程化PAL)的至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的活性。在一些实施方案中,功能片段包含至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的工程化PAL的亲本序列。在一些实施方案中,功能片段将被截短小于5个、小于10个、小于15个、小于10个、小于25个、小于30个、小于35个、小于40个、小于45个、小于50个、小于55个、小于60个、小于65个或小于70个氨基酸。

[0479] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:6至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:6相比,与SEQ ID NO:6或具有与SEQ ID NO:6至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或多个氨基酸位置)处的氨基酸残基差异。在一些实施方案中,工程化PAL包含与SEQ ID NO:6至少90%的序列同一性,并且包含与SEQ ID NO:6相比在至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或多个氨基酸位置处的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID NO:6的序列组成。

[0480] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:10或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:10相比,与SEQ ID NO:10或具有与SEQ ID NO:10至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或多个氨基酸位置)处的氨基酸残基的差异。在一些实施方案中,工程化PAL包含与SEQ ID NO:10至少95%的序列同一性,并且包含与SEQ ID NO:10相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID NO:10的序列组成。

[0481] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:12或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:12相比,与SEQ ID NO:12或具有与SEQ ID NO:12至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或多个氨基酸位置(比如在1

个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或更多个氨基酸位置)处的氨基酸残基的差异。在一些实施方案中,工程化PAL包含与SEQ ID NO:12至少95%的序列同一性,并且包含与SEQ ID NO:12相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID NO:12的序列组成。

[0482] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:14或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:14相比,与SEQ ID NO:14或具有与SEQ ID NO:14至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或更多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或更多个氨基酸位置)处的氨基酸残基的差异。在一些实施方案中,工程化PAL包含与SEQ ID NO:14至少95%的序列同一性,并且包含与SEQ ID NO:14相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID NO:14的序列组成。

[0483] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:16或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:16相比,与SEQ ID NO:16或具有与SEQ ID NO:16至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或更多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或更多个氨基酸位置)处的氨基酸残基差异。在一些实施方案中,工程化PAL包含与SEQ ID NO:16至少95%的序列同一性,并且包含与SEQ ID NO:16相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID NO:16的序列组成。

[0484] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID NO:18或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:18相比,与SEQ ID NO:18或具有与SEQ ID NO:18至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或更多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或多个氨基酸位置)处的氨基酸残基差异。在一些实施方案中,工程化PAL包含与SEQ ID NO:18至少95%的序列同一性,并且包含与SEQ ID NO:18相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID NO:18的序列组成。

[0485] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID

N0:20或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID N0:20相比,与SEQ ID N0:20或具有与SEQ ID N0:20至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或多个氨基酸位置)处的氨基酸残基的差异。在一些实施方案中,工程化PAL包含与SEQ ID N0:20至少95%的序列同一性,并且包含与SEQ ID N0:20相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID N0:20的序列组成。

[0486] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID N0:22或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID N0:22相比,与SEQ ID N0:22或具有与SEQ ID N0:22至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或多个氨基酸位置)处的氨基酸残基差异。在一些实施方案中,工程化PAL包含与SEQ ID N0:22至少95%的序列同一性,并且包含与SEQ ID N0:22相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID N0:22的序列组成。

[0487] 在一些实施方案中,表现出至少一种改进的特性的工程化PAL多肽具有与SEQ ID N0:24或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID N0:24相比,与SEQ ID N0:24或具有与SEQ ID N0:24至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比,在一个或多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或多个氨基酸位置)处的氨基酸残基的差异。在一些实施方案中,工程化PAL包含与SEQ ID N0:24至少95%的序列同一性,并且包含与SEQ ID N0:24相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,工程化PAL多肽由SEQ ID N0:24的序列组成。

[0488] 具有降低的对蛋白水解敏感性的变体:

[0489] 在一些实施方案中,本发明的工程化PAL多肽具有PAL活性,表现出降低的对蛋白水解的敏感性,并包含:a)具有与参考序列SEQ ID N0:4至少85%的序列同一性的氨基酸序列;b)与SEQ ID N0:4相比在一个或多个氨基酸位置处的氨基酸残基差异。

[0490] 在一些实施方案中,表现出降低的对蛋白水解的敏感性的工程化PAL多肽具有与SEQ ID N0:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID

N0:4相比在一个或多个氨基酸位置(诸如与SEQ ID N0:4或具有与SEQ ID N0:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个、20个或更多个氨基酸位置)处的氨基酸残基差异。

[0491] 在一些实施方案中,表现出降低的对蛋白水解的敏感性的工程化PAL多肽具有与SEQ ID N0:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,以及当与SEQ ID N0:4的氨基酸序列最佳比对时,与SEQ ID N0:4相比在选自以下的一个或多个氨基酸位置处的氨基酸残基差异:X39;X54;X59;X73;X91;X158;X112、X134、X180;X195;X240;X243;X245;X256;X257;X270;X290;X304、X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;X564;或其任何组合。在一些实施方案中,氨基酸差异为1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、15个或20个或更多的氨基酸位置。

[0492] 在一些实施方案中,表现出降低的对蛋白水解的敏感性的工程化PAL多肽具有与SEQ ID N0:4至少85%、至少88%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性,并包含在位置X307;X326;X460;X307;和/或X528处的氨基酸残基差异,和任选地在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置处的氨基酸残基差异。在一些实施方案中,当与SEQ ID N0:4比对时,氨基酸残基差异选自Y304H/W;R305L/M;H307G/M/Q;I326F;Q240W;T460G;P528L;和这些置换的任何组合。

[0493] 在一些实施方案中,表现出降低的对蛋白水解的敏感性的工程化PAL多肽具有与SEQ ID N0:10、12、14、16、18、20、22和/或24中的任一个或其功能片段至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID N0:10、12、14、16、18、20、22和/或24相比的氨基酸残基差异,所述氨基酸残基差异在与SEQ ID N0:10、12、14、16、18、20、22和/或24或具有与SEQ ID N0:10、12、14、16、18、20、22和/或24至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比的一个或多个氨基酸位置(诸如在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个或更多个氨基酸位置)处。在一些实施方案中,工程化PAL包含与SEQ ID N0:10、12、14、16、18、20、22和/或24至少95%的序列同一性,并且包含与SEQ ID N0:10、12、14、16、18、20、22和/或24相比至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置的氨基酸差异。在一些实施方案中,PAL包含SEQ ID N0:10、12、14、16、18、20、22和/或24的序列或由SEQ ID N0:10、12、14、16、18、20、22和/或24的序列组成。

[0494] 在一些实施方案中,在基本相同的条件下,工程化PAL多肽的蛋白水解敏感性比野生型PAL(例如,具有SEQ ID N0:4的AvPAL)的蛋白水解敏感性或与参考PAL多肽相比降低了至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少85%、至少90%或至少95%。蛋白水解活性可使用本领域已知的任何适合方法来测量,包括但不限于在实施例中所描述的那些。

[0495] 在一些实施方案中,当参考PAL和具有降低的敏感性的工程化PAL在基本相同的条件下被比较并暴露于基本上相同量和种类的蛋白酶时,具有降低的对蛋白水解的敏感性的工程化PAL多肽具有降低的对包含一种或更多种蛋白酶的组合物敏感性,所述蛋白酶包括但不限于胃蛋白酶、胰蛋白酶、胰凝乳蛋白酶、羧肽酶A和B、肽酶(例如,氨基肽酶、二肽酶和肠肽酶)。

[0496] 在一些实施方案中,具有降低的对蛋白水解的敏感性的工程化PAL多肽具有是参考PAL(例如,AvPAL)的酶活性的约1.0倍、2倍、5倍、10倍、20倍、25倍、50倍、75倍、100倍、150倍、200倍或更高倍数的酶活性水平。在一些实施方案中,当在4.5到7.5的pH范围下测量活性时,当在4.5到6.5的pH范围下测量活性时;当在5.0到7.5的pH范围下测量活性时;当在5.0至6.5的pH范围下测量活性时;当在5.5到7.5的pH范围下测量活性时;和/或还当在5.5到6.5的pH范围下测量活性时;工程化多肽具有与参考PAL相比更大的酶活性。在一些其他实施方案中,工程化PAL多肽具有在1 μ M至5mM范围内的 K_m 值。

[0497] 具有增加的对酸性pH耐受性的变体:

[0498] 在一些实施方案中,本发明的工程化PAL多肽具有PAL活性,耐受酸性pH水平,并包含:a)具有与参考序列SEQ ID NO:4或其片段至少85%的序列同一性的氨基酸序列;b)与SEQ ID NO:4相比在一个或更多个氨基酸位置处的氨基酸残基差异。

[0499] 在一些实施方案中,与野生型AvPAL和/或另一种参考多肽相比表现出增加的对酸性pH耐受性的工程化PAL多肽具有与SEQ ID NO:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:4相比,在一个或更多个氨基酸位置(诸如与SEQ ID NO:4或具有与SEQ ID NO:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性的序列相比在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、14个、15个、20个或更多个氨基酸位置)处的氨基酸残基差异。

[0500] 在一些实施方案中,与野生型AvPAL和/或另一种参考多肽相比表现出增加的对酸性pH的耐受性的工程化PAL多肽具有与SEQ ID NO:4至少85%、至少88%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或更大的氨基酸序列同一性,和与SEQ ID NO:4的氨基酸序列最佳比对时,与SEQ ID NO:4相比在一个或更多个氨基酸位置处的氨基酸残基差异,所述一个或更多个氨基酸位置选自X39;X54;X59;X73;X91;X158;X112,X134,X180;X195;X240;X243;X245;X256;X257;X270;X290;X304;X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;X564;或其任何组合。在一些实施方案中,氨基酸差异为1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、15个或20个或更多个氨基酸位置。

[0501] 在一些实施方案中,与野生型AvPAL和/或另一种参考多肽相比表现出增加的对酸性pH耐受性的工程化PAL多肽具有与SEQ ID NO:4至少85%、至少88%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性,且包含选自在以下位置处的氨基酸残基差异:X39;X54;X59;X73;X91;X158;X112,X134,X180;X195;X240;X243;X245;X256;X257;X270;X290;X304,X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;

X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;X564或其任何组合;和任选地在1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸位置处的氨基酸残基差异。在一些实施方案中,当与SEQ ID NO:4比对时,氨基酸残基差异为A39;T54;G59;S73;A91;Y158;S180;K195;A112;R134;Q240;T243;I245;A256;L257;N270;N290;Y304;R305;H307;E308;I326;L349;D353;L364;A394;S399;N400;P404;L407;F443;N453;Y459;T460;T463;N474;E509;Q521;K522;T524;P528;S546;和/或P564。在一些实施方案中,表现出增加的对酸性pH耐受性的工程化PAL多肽当与SEQ ID NO:4比对时具有与SEQ ID NO:4至少85%、至少88%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性,且包含在以下一个或更多个位置处的氨基酸残基差异:A39V;T54K;G59R;S73K;A112C;R134Q;A91V;Y158H;S180A;K195E;Q240R/W;T243I/L;I245L;A256G;L257W/A;N270K;N290G;Y304H;R305M;H307G/Q/M;E308Q;I326F;L349M;D353A/N;L364Q;A394V;S399N;N400K;P404A;L407V;F443H;N453G;Y459F;T460G;T463N;N474Q;E509L;Q521K/S;K522Y/F/N;T524S;P528L;S546R;和/或P564G/L/M。

[0502] 在一些实施方案中,当所有其他测定条件基本相同时,与参考PAL多肽相比具有增加的对酸性pH的耐受性的工程化PAL多肽具有在以下pH范围下的增加的耐受性:在1.5至6.5之间;在1.5和5.0之间;在2.0至5.5之间;在3.0和6.8之间;在3.0和5.5之间;在4.0和6.5之间;在4.0和4.5之间;在4.5和5.0之间;在4.5和5.5之间;在4.5和6.0之间;在4.5和6.5之间;在5.0和6.5之间;在5.0和6.0之间;在5.0和5.5之间;在5.5和6.0之间;在6.0和6.5之间;和/或在6.5和7.0之间。在一些实施方案中,增加的对酸性pH的耐受性表现在约3.5、4.0、4.5、5.0、5.5、6.0和/或6.5的pH下。

[0503] 在一些实施方案中,与参考PAL相比具有增加的对酸性pH的耐受性的工程化PAL多肽当通过标准测定测量时还表现出更大的PAL活性。任何适合的测定在本发明中具备实用性,包括但不限于本文所提供的那些。

[0504] 进一步设想,任何例示的工程化多肽(即,变体1号-变体1010号),例如通过随后的多轮进化通过添加来自本文描述的其他多肽和其他残基位置的多种氨基酸差异的新组合,具备作为用于合成其他工程化PAL多肽的起始氨基酸序列的用途。在一些实施方案中,另外的改进通过包括在贯穿前几轮的进化中已被保持不变的残基位置处的氨基酸差异而产生。这不意在使本发明限于任何特定的产生工程化PAL多肽的方法,因为任何适合的方法在本发明中具备实用性,包括但不限于本文所提供的那些方法。

[0505] 编码工程化多肽的多核苷酸、表达载体和宿主细胞:

[0506] 本发明提供了编码本文描述的工程化PAL多肽的多核苷酸。在一些实施方案中,多核苷酸被可操作地连接至控制基因表达的一个或更多个异源调控序列,以创建能够表达多肽的重组多核苷酸。在一些实施方案中,将包含编码工程化PAL多肽的至少一个异源的多核苷酸的表达构建体引入到适当的宿主细胞以表达相应的PAL多肽。

[0507] 如对本领域技术人员将是明显的,蛋白序列的可用性和对应各种氨基酸的密码子的知识提供了对能够编码该主题多肽的所有多核苷酸的描述。遗传密码子的简并性,其中相同氨基酸由替代的或同义的密码子编码,允许制备极大数目的核酸被,所有这些核酸编码工程化PAL多肽。因此,通过选择基于可能的密码子选择的组合,本发明提供了用于产生可制备的编码本文描述的PAL多肽的PAL多核苷酸的每种和每一种可能的变化的方法和组

合物,并且所有这些变化将被认为针对本文描述的任何多肽,包括在实施例(例如,在多个表中)中呈现的氨基酸序列被明确地公开。

[0508] 在一些实施方案中,密码子优选地被优化以被用于产生蛋白所选择的宿主细胞利用。例如,细菌中使用的优选的密码子通常被用于在细菌中的表达。因此,编码工程化PAL多肽的密码子优化的多核苷酸在全长编码区的约40%、50%、60%、70%、80%、90%或大于90%的密码子位置处包含优选的密码子。

[0509] 在一些实施方案中,PAL多核苷酸编码具有PAL活性、具有本文公开的特性的工程化多肽,其中所述多肽包含具有与选自SEQ ID NO:3、5、7、9、11、13、15、17、19、21、和/或23的参考序列或任何变体(例如,在实施例中提供的变体)的氨基酸序列至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的同一性的氨基酸序列,,和与SEQ ID NO:3、5、7、9、11、13、15、17、19、21和/或23的参考多核苷酸或实施例中公开的任何变体的氨基酸序列相比的一个或更多个残基差异(例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个氨基酸残基位置)。在一些实施方案中,参考序列选自SEQ ID NO:3、5、7、9、11、13、15、17、19、21和/或23。

[0510] 在一些实施方案中,PAL多核苷酸编码具有PAL活性的、具有本文公开的特性的工程化多肽,其中所述多肽包含具有与参考序列SEQ ID NO:4至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多的序列同一性的氨基酸序列,和当与SEQ ID NO:4的多肽最佳比对时,与SEQ ID NO:4相比在来自以下残基位置处的一个或更多个残基差异:X39;X54;X59;X73;X91;X158;X112,X134,X180;X195;X240;X243;X245;X256;X257;X270;X290;X304、X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;和/或X564。

[0511] 在一些实施方案中,编码工程化PAL多肽的多核苷酸包含选自多核苷酸序列SEQ ID NO:3、5、7、9、11、13、15、17、19、21和/或23的多核苷酸序列。在一些实施方案中,编码工程化PAL多肽的多核苷酸具有与SEQ ID NO:2、3、5、7、9、11、13、15、17、19、21和/或23至少80%、85%、90%、93%、95%、96%、97%、98%、99%的核苷酸残基同一性。

[0512] 在一些实施方案中,多核苷酸能够在高度严格条件下与选自SEQ ID NO:2、3、5、7、9、11、13、15、17、19、21和/或23的参考多核苷酸序列或其互补序列或编码本文提供的任何变体PAL多肽的多核苷酸序列杂交。在一些实施方案中,能够在高度严格条件下杂交的多核苷酸编码包含氨基酸序列的PAL多肽,所述氨基酸序列具有与SEQ ID NO:4相比在选自以下的残基位置处的一个或更多个残基差异:X39;X54;X59;X73;X91;X158;X112、X134、X180;X195;X240;X243;X245;X256;X257;X270;X290;X304、X305;X307;X308;X326;X349;X353;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X509;X521;X522;X524;X528;X546;和/或X564。

[0513] 在一些实施方案中,编码本文的工程化PAL多肽中的任何一个的分离的多核苷酸以多种方式被操作,以促进PAL多肽的表达。在一些实施方案中,编码PAL多肽的多核苷酸包含存在一个或更多个控制序列以调控PAL多核苷酸和/或多肽的表达的表达载体。取决于所利用的表达载体,所分离的多核苷酸在其插入载体中之前的操作可以是令人期望的或必要的。利用重组DNA方法修饰多核苷酸和核酸序列的技术是本领域熟知的。在一些实施方案

中,控制序列包括,尤其是,启动子、前导序列、多腺苷酸化序列、前肽序列、信号肽序列和转录终止子。在一些实施方案中,适合的启动子根据宿主细胞的选择来选择。对于细菌宿主细胞,用于指导本公开内容的核酸构建体转录的合适启动子包括但不限于从以下获得的启动子:大肠杆菌 (*E.coli*) lac操纵子、天蓝色链霉菌 (*Streptomyces coelicolor*) 琼脂酶基因 (*dagA*)、枯草芽孢杆菌 (*Bacillus subtilis*) 果聚糖蔗糖酶基因 (*sacB*)、地衣芽孢杆菌 (*Bacillus licheniformis*) α -淀粉酶基因 (*amyL*)、嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*) 麦芽糖淀粉酶基因 (*amyM*)、解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*) α 解淀粉酶基因 (*amyQ*)、地衣芽孢杆菌 (*Bacillus licheniformis*) 青霉素酶基因 (*penP*)、枯草芽孢杆菌 (*Bacillus subtilis*) *xylA*和*xylB*基因和原核 β 内酰胺酶基因(参见例如,Villa-Kamaroff等,Proc.Natl Acad.Sci.USA 75:3727-3731[1978]),以及tac启动子(参见例如,DeBoer等,Proc.Natl Acad.Sci.USA 80:21-25[1983])。用于丝状真菌宿主细胞的示例性的启动子包括,但不限于从针对以下的基因获得的启动子:米曲霉 (*Aspergillus oryzae*) TAKA淀粉酶、米黑根毛霉 (*Rhizomucor miehei*) 天冬氨酸蛋白酶、黑曲霉 (*Aspergillus niger*) 中性 α 粉淀粉酶、黑曲霉酸稳定性 α -淀粉酶、黑曲霉或泡盛曲霉 (*Aspergillus awamori*) 葡糖淀粉酶 (*glaA*)、米黑根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉磷酸丙糖异构酶、构巢曲霉 (*Aspergillus nidulans*) 乙酰胺酶和尖孢镰刀菌 (*Fusarium oxysporum*) 胰蛋白酶样蛋白酶(参见例如,WO 96/00787),以及NA2-tpi启动子(来自对于黑曲霉中性 α -淀粉酶和米曲霉磷酸丙糖异构酶启动子的基因的启动子的杂合体),和其突变的、截短的和杂合的启动子。示例性酵母细胞启动子可来自对于以下的基因:酿酒酵母 (*Saccharomyces cerevisiae*) 烯醇化酶 (*ENO-1*)、酿酒酵母半乳糖激酶 (*GAL1*)、酿酒酵母乙醇脱氢酶/甘油醛-3-磷酸脱氢酶 (*ADH2/GAP*) 和酿酒酵母3-磷酸甘油酸激酶。用于酵母宿主细胞的其他有用的启动子是本领域已知的(参见例如,Romanos等,Yeast 8:423-488 [1992])。

[0514] 在一些实施方案中,控制序列还是适合的转录终止子序列(即,由宿主细胞识别以终止转录的序列)。在一些实施方案中,终止子序列可操作地连接到编码该PAL多肽的核酸序列的3'末端。在选择的宿主细胞中有功能的任何合适的终止子在本发明中具备实用性。用于丝状真菌宿主细胞的示例性转录终止子可从对于以下的基因获得:米曲霉TAKA淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、黑曲霉 α -葡糖苷酶和尖孢镰刀菌胰蛋白酶样蛋白酶。用于酵母宿主细胞的示例性终止子可从对于以下的基因获得:酿酒酵母烯醇化酶、酿酒酵母细胞色素C (*CYC1*) 和酿酒酵母甘油醛-3-磷酸脱氢酶。用于酵母宿主细胞的其他有用的终止子是本领域已知的(参见例如,Romanos等,同上)。

[0515] 在一些实施方案中,控制序列还是适合的前导序列(即,对宿主细胞的翻译重要的mRNA的未翻译区)。在一些实施方案中,前导序列可操作地连接到编码PAL多肽的核酸序列的5'末端。在选择的宿主细胞中有功能的任何适合的前导序列在本发明中具备实用性。用于丝状真菌宿主细胞的示例性前导序列从以下基因获得:米曲霉TAKA淀粉酶和构巢曲霉磷酸丙糖异构酶。用于酵母宿主细胞的适合的前导序列从对于以下的基因获得:酿酒酵母烯醇化酶 (*ENO-1*)、酿酒酵母3-磷酸甘油酸激酶、酿酒酵母 α 因子和酿酒酵母乙醇脱氢酶/甘油醛-3-磷酸脱氢酶 (*ADH2/GAP*)。

[0516] 在一些实施方案中,控制序列还是多聚腺苷酸化序列(即,可操作地连接到核酸序

列的3'末端的序列,并且当转录时,其被宿主细胞识别为将多聚腺苷残基添加到转录的mRNA的信号)。在选择的主细胞中有功能的任何适合的多聚腺苷酸序列在本发明中具备实用性。用于丝状真菌宿主细胞的示例性多腺苷酸序列包括,但不限于以下的基因:米曲霉TAKA淀粉酶、黑曲霉葡萄糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、尖孢镰刀菌胰蛋白酶样蛋白酶和黑曲霉 α 粉葡萄糖苷酶。对酵母宿主细胞有用的多聚腺苷酸化序列是已知的(参见例如,,Guo和Sherman,Mol.Cell.Bio.,15:5983-5990[1995])。

[0517] 在一些实施方案中,控制序列还是信号肽(即,编码连接到多肽的氨基末端的氨基酸序列并将编码的多肽引导入细胞的分泌通路的编码区)。在一些实施方案中,核酸序列的编码序列的5'末端固有地包含信号肽编码区,所述信号肽编码区与编码分泌的多肽的编码区的区段符合翻译读码框地天然连接。可选地,在一些实施方案中,编码序列的5'端包含对于编码序列而言是外源性的信号肽编码区。将表达的多肽引导入选择的宿主细胞的分泌通路的任何合适的信号肽编码区具备用于工程化多肽的表达的用途。对于细菌宿主细胞有效的信号肽编码区包括但不限于从对于以下的基因获得的那些:芽孢杆菌(*Bacillus*) NC1B 11837麦芽糖淀粉酶、嗜热脂肪芽孢杆菌 α -芽淀粉酶、地衣芽孢杆菌枯草蛋白酶、地衣芽孢杆菌 β -内酰胺酶、嗜热脂肪芽孢杆菌中性蛋白酶(nprT、nprS、nprM)和枯草芽孢杆菌prSA。另外的信号肽是本领域已知的(参见例如,Simonen和Palva,Microbiol.Rev.,57:109-137[1993])。在一些实施方案中,对于丝状真菌宿主细胞有效的信号肽编码区包括但不限于从对于以下的基因获得的信号肽编码区:米曲霉TAKA淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡萄糖淀粉酶、米黑根毛霉天冬氨酸蛋白酶、特异腐质霉(*Humicola insolens*)纤维素酶和胎毛腐质霉(*Humicola lanuginosa*)脂肪酶。用于酵母宿主细胞的有用的信号肽包括但不限于来自对于以下的基因的那些:酿酒酵母 α 因子和酿酒酵母转化酶。

[0518] 在一些实施方案中,控制序列还是前肽编码区,其编码定位在多肽的氨基末端的氨基酸序列。所得多肽称为“前酶(proenzyme)”、“前多肽(propolypeptide)”或“酶原”。前多肽可被转化为成熟活性多肽,通过催化或自动催化裂解来自前多肽的前肽。前肽编码区可从任何适合的来源来获得,所述来源包括但不限于对于以下的基因:枯草芽孢杆菌碱性蛋白酶(aprE)、枯草芽孢杆菌中性蛋白酶(nprT)、酿酒酵母 α 因子、米黑根毛霉天冬氨酸蛋白酶和嗜热毁丝霉(*Myceliophthora thermophila*)乳糖酶参见例如,WO 95/33836)。当信号肽和前肽区均存在于多肽的氨基末端时,前肽区紧邻多肽的氨基末端定位并且信号肽区紧邻前肽区的氨基末端定位。

[0519] 在一些实施方案中,还利用了调节序列。这些序列促进多肽相对于宿主细胞的生长的表达的调节。调节系统的实例是引起基因表达响应于化学或物理刺激,包括调节化合物的存在而开启或关闭的那些调节系统。在原核宿主细胞中,适合的调节序列包括但不限于lac、tac和trp操纵子系统。在酵母宿主细胞中,适合的调节系统包括但不限于ADH2系统或GAL1系统。在丝状真菌中,适合的调控序列包括但不限于TAKA α -淀粉酶启动子、黑曲霉葡萄糖淀粉酶启动子和米曲霉葡萄糖淀粉酶启动子。

[0520] 在另一个方面,本发明涉及重组表达载体,所述重组表达载体包含编码工程化PAL多肽的多核苷酸和取决于其被引入的宿主细胞的类型的一个或多个表达调节区,诸如启动子和终止子、复制起点等。在一些实施方案中,将本文描述的各种核酸和控制序列连接在一起以产生重组表达载体,其包含一个或多个方便的限制位点以允许编码PAL多肽的核

酸序列在这些位点插入或置换。可选地,在一些实施方案中,本发明的核酸序列通过将核酸序列或包含该序列的核酸构建体插入合适的表达载体来表达。在涉及创建表达载体的一些实施方案中,编码序列位于载体中,以使编码序列与用于表达的适当的控制序列可操作地连接。

[0521] 重组表达载体可以是任何适合的载体(例如,质粒或病毒),其可方便地经历重组DNA程序,并致使PAL多核苷酸序列的表达。载体的选择通常取决于载体与载体将被引入的宿主细胞的相容性。载体可以是线性的或闭合的环状质粒。

[0522] 在一些实施方案中,表达载体是自主复制载体(即,作为染色体外的实体存在的载体,其复制独立于染色体复制,诸如质粒、染色体外元件、微型染色体或人工染色体)。载体可包含用于确保自我复制的任何工具(means)。在一些替代性实施方案中,载体是当被引入宿主细胞时,被整合进入基因组并与其被整合进的染色体一起复制的载体。此外,在一些实施方案中,利用了单一载体或质粒,或共同包含待引入宿主细胞的基因组的全部DNA的两种或更多种载体或质粒,和/或转座子。

[0523] 在一些实施方案中,表达载体包含一个或更多个可选择的标志物,其允许容易选择经转化的细胞。“可选择的标志物”是基因,其产物提供杀生物剂或病毒抗性、对重金属的抗性、对营养缺陷型的原养型等等。细菌的可选择的标志物的实例包括但不限于,来自枯草芽孢杆菌或地衣芽孢杆菌的 dal 基因,或赋予抗生素抗性诸如氨苄青霉素、卡那霉素、氯霉素或四环素抗性的标志物。用于酵母宿主细胞的适合的标志物包括但不限于 $ADE2$ 、 $HIS3$ 、 $LEU2$ 、 $LYS2$ 、 $MET3$ 、 $TRP1$ 和 $URA3$ 。用于丝状真菌宿主细胞的可选择的标志物包括但不限于 $amdS$ (乙酰胺酶;例如,来自构巢曲霉(*A.nidulans*)或米曲霉(*A.oryzae*)的乙酰胺酶)、 $argB$ (鸟氨酸氨甲酰基转移酶)、 bar (草胺膦乙酰基转移酶;例如,来自吸水链霉菌(*S.hygrosopicus*)的草胺膦乙酰基转移酶)、 hph (潮霉素磷酸转移酶)、 $niaD$ (硝酸还原酶)、 $pyrG$ (乳清酸核苷-5'-磷酸脱羧酶;例如,来自构巢曲霉或米曲霉的乳清酸核苷-5'-磷酸脱羧酶)、 sC (硫酸腺苷酰转移酶)和 $trpC$ (邻氨基苯甲酸合酶)以及其等同物。在另一个方面,本发明提供了宿主细胞,其包含编码至少一种本发明的工程化PAL多肽的至少一种多核苷酸,所述多核苷酸可操作地连接到用于在宿主细胞中表达工程化PAL酶的一个或更多个控制序列。适合于在表达由本发明的表达载体编码的多肽中使用的宿主细胞是本领域熟知的,并且包括但不限于细菌细胞,诸如大肠杆菌、河流弧菌(*Vibrio fluvialis*)、链霉菌属和鼠伤寒沙门氏菌(*Salmonella typhimurium*)细胞;真菌细胞,诸如酵母细胞(例如,酿酒酵母或毕赤酵母(*Pichia pastoris*)(ATCC保藏号201178));昆虫细胞,诸如果蝇属(*Drosophila*)S2和夜蛾属(*Spodoptera*)Sf9细胞;动物细胞,诸如CHO、COS、BHK、293和Bowes黑素瘤细胞;和植物细胞。示例性的宿主细胞还包括各种大肠杆菌(*Escherichia coli*)菌株(例如,W3110($\Delta fhuA$)和BL21)。

[0524] 因此,在另一个方面,本发明提供了产生工程化PAL多肽的方法,其中所述方法包括在适合于多肽表达的条件下,培养能够表达编码工程化PAL多肽的多核苷酸的宿主细胞。在一些实施方案中,该方法还包括分离和/或纯化如本文所描述的PAL多肽的步骤。

[0525] 用于宿主细胞的适当的培养基和生长条件是本领域已知的。可以预期的是,将用于将表达PAL多肽的多核苷酸引入细胞的任何适合的方法用于本发明。适合的技术包括但不限于:电穿孔、生物射弹粒子轰击、脂质体介导的转染、氯化钙转染和原生质体融合。

[0526] 具有本文公开的特性的工程化PAL多肽可通过使编码天然存在的或工程化的PAL多肽的多核苷酸经历任何适合的诱变和/或本领域中已知的和/或如本文所描述的定向进化方法来获得。示例性的定向进化技术是诱变和/或DNA改组(参见例如, Stemmer, Proc.Natl.Acad.Sci.USA 91:10747-10751[1994]; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767和美国专利6,537,746)。可使用的其他定向进化程序包括交错延伸过程(StEP)、体外重组(参见例如, Zhao等, Nat.Biotechnol., 16: 258-261[1998])、诱变PCR(参见例如, Caldwell等, PCR Methods Appl., 3:S136-S140[1994])和盒式诱变(参见例如, Black等, Proc.Natl.Acad.Sci.USA 93:3525-3529[1996])。

[0527] 诱变和定向进化的方法可容易地应用到编码PAL-的多核苷酸中, 以产生可被表达、筛选和测定的变体库。将任何适合的诱变和定向进化方法应用于本发明且是本领域熟知的(参见例如, 美国专利号5,605,793、5,830,721、6,132,970、6,420,175、6,277,638、6,365,408、6,602,986、7,288,375、6,287,861、6,297,053、6,576,467、6,444,468、5,811,238、6,117,679、6,165,793、6,180,406、6,291,242、6,995,017、6,395,547、6,506,602、6,519,065、6,506,603、6,413,774、6,573,098、6,323,030、6,344,356、6,372,497、7,868,138、5,834,252、5,928,905、6,489,146、6,096,548、6,387,702、6,391,552、6,358,742、6,482,647、6,335,160、6,653,072、6,355,484、6,03,344、6,319,713、6,613,514、6,455,253、6,579,678、6,586,182、6,406,855、6,946,296、7,534,564、7,776,598、5,837,458、6,391,640、6,309,883、7,105,297、7,795,030、6,326,204、6,251,674、6,716,631、6,528,311、6,287,862、6,335,198、6,352,859、6,379,964、7,148,054、7,629,170、7,620,500、6,365,377、6,358,740、6,406,910、6,413,745、6,436,675、6,961,664、7,430,477、7,873,499、7,702,464、7,783,428、7,747,391、7,747,393、7,751,986、6,376,246、6,426,224、6,423,542、6,479,652、6,319,714、6,521,453、6,368,861、7,421,347、7,058,515、7,024,312、7,620,502、7,853,410、7,957,912、7,904,249,和所有相关的非美国对应申请; Ling等, Anal.Biochem, 254(2):157-78[1997]; Dale等, Meth.Mol.Biol., 57:369-74[1996]; Smith, Ann.Rev.Genet, 19:423-462[1985]; Botstein等, Science, 229:1193-1201[1985]; Carter, Biochem.J., 237:1-7[1986]; Kramer等, Cell, 38:879-887[1984]; Wells等, Gene, 34:315-323[1985]; Minshull等, Curr.Op.Chem.Biol., 3:284-290[1999]; Christians等, Nat.Biotechnol., 17:259-264[1999]; Cramer等, Nature, 391:288-291[1998]; Cramer等, Nat.Biotechnol., 15:436-438[1997]; Zhang等, Proc.Nat.Acad.Sci.U.S.A., 94:4504-4509[1997]; Cramer等, Nat.Biotechnol., 14:315-319[1996]; Stemmer, Nature, 370:389-391[1994]; Stemmer, Proc.Nat.Acad.Sci.USA, 91:10747-10751[1994]; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767; WO 2009/152336和美国专利号6,537,746。其所有通过引用并入本文)。

[0528] 在一些实施方案中, 诱变处理后获得的酶克隆通过使酶制剂经历定义的温度(或其他测定条件), 并测量热处理或其他适合的测定条件之后剩余的酶活性的量进行筛选。然后将包含编码PAL多肽的多核苷酸的克隆从基因中分离、测序以鉴定核苷酸序列变化(如果有的话), 并且用于在宿主细胞中表达酶。测量来自表达文库的酶活性可使用本领域已知的任何适合的方法(例如, 标准生物化学技术, 诸如HPLC分析)来进行。

[0529] 对于已知序列的工程化多肽,编码酶的多核苷酸可根据已知的合成方法通过标准的固相方法制备。在一些实施方案中,多至约100个碱基的片段可被单独地合成,然后连接(例如,通过酶或化学连接方法(chemical ligation method)或聚合酶介导的方法)以形成任何期望的连续序列。例如,使用经典的亚磷酰胺法(参见例如,Beaucage等,Tet.Lett., 22:1859-69[1981];和Matthes等,EMBO J., 3:801-05[1984]),如通常在自动合成方法中所实践的,本文公开的多核苷酸和寡核苷酸可通过化学合成来制备。根据亚磷酰胺方法,寡核苷酸被合成(例如,在自动的DNA合成仪中,纯化、退火、连接并克隆入适合的载体)。

[0530] 因此,在一些实施方案中,用于制备工程化PAL多肽的方法可包括:(a)合成编码包含选自如本文所描述的任何变体的氨基酸序列的氨基酸序列的多肽的多核苷酸,和(b)表达由多核苷酸编码的PAL多肽。在该方法的一些实施方案中,由多核苷酸编码的氨基酸序列可任选地具有一个或几个(例如,多至3个、4个、5个或多至10个)氨基酸残基缺失、插入和/或置换。在一些实施方案中,氨基酸序列任选地具有1-2个、1-3个、1-4个、1-5个、1-6个、1-7个、1-8个、1-9个、1-10个、1-15个、1-20个、1-21个、1-22个、1-23个、1-24个、1-25个、1-30个、1-35个、1-40个、1-45个或1-50个氨基酸残基缺失、插入和/或置换。在一些实施方案中,氨基酸序列任选地具有1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、30个、30个、35个、40个、45个或50个氨基酸残基缺失、插入和/或置换。在一些实施方案中,氨基酸序列任选地具有1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、18个、20个、21个、22个、23个、24个或25个氨基酸残基缺失、插入和/或置换。在一些实施方案中,置换为保守置换或非保守置换。

[0531] 使用本领域中已知的任何适合的测定,包括但不限于本文所描述的测定和条件,评价所表达的工程化PAL多肽的任何期望的改进的特性或特性的组合(例如,活性、选择性、稳定性、酸耐受性、蛋白酶敏感性等)。

[0532] 在一些实施方案中,使用用于蛋白纯化的熟知技术中的任何一种或更多种,将在宿主细胞中表达的工程化PAL多肽的任一种从细胞和/或培养基中回收,用于蛋白纯化的熟知技术包括,溶菌酶处理、超声处理、过滤、盐析、超离心和色谱法以及其他。

[0533] 用于分离PAL多肽的色谱技术包括,除了其他的以外,反相色谱法、高效液相色谱法、离子交换色谱法、疏水相互作用色谱法、尺寸排阻色谱法、凝胶电泳和亲和色谱法。用于纯化特定酶的条件部分取决于诸如净电荷、疏水性、亲水性、分子量、分子形状等因素,并且对本领域技术人员将是明显的。在一些实施方案中,亲和技术可用于分离改进的PAL酶。对于亲和色谱纯化,可使用特异性结合感兴趣的PAL多肽的任何抗体。为了产生抗体,通过用PAL多肽或其片段注射免疫多种宿主动物,包括但不限于兔、小鼠、大鼠等。在一些实施方案中,PAL多肽或片段借助于侧链官能团或被附连至侧链官能团的衔接子被附连至适合的载体,诸如BSA。

[0534] 在一些实施方案中,工程化PAL多肽通过包括以下的方法在宿主细胞中产生:在有利于工程化PAL多肽产生的条件下培养包含编码如本文描述的工程化PAL多肽的多核苷酸序列的宿主细胞(例如,大肠杆菌菌株),并从细胞和/或培养物中回收工程化PAL多肽。在一些实施方案中,宿主细胞产生多于一种工程化PAL多肽。

[0535] 在一些实施方案中,本发明提供了产生工程化PAL多肽的方法,所述方法包括在适

合的培养条件下培养包含编码工程化PAL多肽的多核苷酸序列的重组细菌细胞以允许产生工程化PAL多肽并任选地从培养物和/或培养的细菌细胞回收工程化PAL多肽,所述工程化PAL多肽具有与参考序列SEQ ID NO:4至少85%、90%、95%、96%、97%、98%或99%的序列同一性,和当与SEQ ID NO:4的氨基酸序列最佳比对时与SEQ ID NO:4相比选自以下的一个或多个氨基酸残基差异:X39;X91;X158;X180;X195;X243;X245;X256;X257;X270;X290;X307;X308;X326;X349;X364;X394;X399;X400;X404;X407;X443;X453;X459;X460;X463;X474;X522;X524;和X528或其组合。在一些实施方案中,宿主细胞产生多于一种工程化PAL多肽。

[0536] 在一些实施方案中,工程化PAL多肽从重组宿主细胞和/或培养物中回收后,它们可通过本领域已知的任何适合的方法进一步纯化。在一些另外的实施方案中,纯化的PAL多肽与其他成分和化合物组合以提供适当包含工程化PAL多肽用于不同应用和用途的组合物和制剂(例如,药物组合物)。

[0537] 组合物:

[0538] 本发明提供了适合于在多种组合物中使用的工程化PAL多肽。这些组合物用于许多领域,包括但不限于医药、膳食/营养补充物、食品、饲料、和精细化工生产。例如,在一些实施方案中,本发明提供了食品和/或饲料,所述食品和/或饲料包含至少一种工程化PAL变体和/或编码至少一种PAL变体的至少一种多核苷酸序列。在一些实施方案中,本发明提供了饮品,所述饮品包含至少一种工程化PAL变体。

[0539] 在一些实施方案中,在食品、饲料和/或营养/膳食补充物中的工程化PAL变体是糖基化的。此外,工程化PAL变体应用于任何适合的可食用的酶递送基质。在一些实施方案中,工程化PAL变体存在于可食用的酶递送基质中,所述可食用的酶递送基质被设计用于摄取变体后PAL变体在动物的消化道内的快速扩散。

[0540] 本发明还提供了工程化PAL多肽,所述工程化PAL多肽适合于在精细化学品和其他工业上重要的化合物的产生中使用(参见例如,美国专利申请号2013/0340119、2013/0005012和2005/0260724,和WO 2012/122333)。

[0541] 药物组合物和其他组合物

[0542] 本发明提供了工程化PAL多肽,所述工程化PAL多肽适合于在药物组合物和其他组合物诸如膳食/营养补充物中使用。

[0543] 取决于施用方式,包含治疗有效量的根据本发明的工程化PAL的这些组合物呈固体、半固体或液体的形式。在一些实施方案中,组合物包括其他药学上可接受的组分,诸如稀释剂、缓冲液、赋形剂、盐、乳化剂、防腐剂、稳定剂、填料和其他成分。用于配制和施用的技术的细节是本领域熟知的,并在文献中被描述。

[0544] 在一些实施方案中,配制工程化PAL多肽用于在口服药物组合物中使用。在递送工程化PAL多肽中使用的任何适合的形式应用于本发明,所述形式包括但不限于丸剂、片剂、凝胶片(gel tabs)、胶囊、锭剂、糖衣丸、粉末、软凝胶、溶胶凝胶、凝胶、乳剂、植入物、贴剂、喷雾剂、软膏剂、搽剂、霜剂、糊剂、凝胶剂、涂剂、气雾剂、口香糖、缓和剂、条状物、悬液(包括但不限于油基悬液、水包油乳剂等)、药浆、糖浆剂、控释制剂、栓剂等。在一些实施方案中,工程化PAL多肽以适合于注射的形式提供(即,以可注射的制剂的形式)。在一些实施方案中,工程化PAL多肽以生物相容性基质诸如溶胶凝胶,包括基于二氧化硅(例如,氧基硅烷

(oxysilane))的溶胶凝胶来提供。在一些实施方案中,封装工程化PAL多肽。在一些可选的实施方案中,将工程化PAL多肽封装在纳米结构(例如,纳米管、纳米通道(nanotubule)、纳米胶囊或微米胶囊、微球体、脂质体等)中。实际上,不预期将本发明限于任何特定的递送制剂和/或递送装置。预期工程化PAL多肽通过本领域已知的任何适合的方式来施用,包括但不限于:肠胃外、口服、局部、经皮、鼻内、眼内、鞘内、经由植入物等等。

[0545] 在一些实施方案中,将工程化PAL多肽通过糖基化、聚乙二醇化(即,用聚乙二醇[PEG]或活化的PEG等修饰)或其他化合物化学修饰(参见例如,Ikeda,Amino Acids 29:283-287[2005];美国专利号7,531,341、7,534,595、7,560,263和7,53,653;美国专利申请公布号2013/0039898,2012/0177722等)。实际上,不预期使本发明限于任何特定的递送方法和/或机制。

[0546] 在一些另外的实施方案中,工程化PAL多肽被提供在包含基质稳定的酶晶体的制剂中。在一些实施方案中,该制剂包含交联的结晶工程化PAL酶和具有附连至酶晶体的反应部分的聚合物。本发明还以聚合物提供工程化PAL多肽。

[0547] 在一些实施方案中,包含本发明的工程化PAL多肽的组合物包含一种或更多种常用的载体化合物,其包括但不限于糖(例如,乳糖、蔗糖、甘露醇和/或山梨糖醇)、淀粉(例如,玉米、小麦、水稻、马铃薯或其他植物淀粉)、纤维素(例如,甲基纤维素、羟丙基甲基纤维素、羧甲基纤维素钠)、树胶(例如,阿拉伯、黄芪胶、瓜尔胶等)和/或蛋白(例如,明胶、胶原等)。在口服制剂中的另外的组分可包括着色剂和或甜味剂(例如,葡萄糖、蔗糖和甘露糖醇)和润滑剂(例如,硬脂酸镁),以及肠溶包衣(例如,甲基丙烯酸酯聚合物、羟丙基甲基纤维素邻苯二甲酸酯和/或本领域已知的任何其他适合的肠溶包衣)。在一些实施方案中,崩解剂或增溶剂被包括(例如,交联的聚乙烯吡咯烷酮、琼脂、藻酸或其盐,诸如藻酸钠)。在一些实施方案中,特别是在液体制剂中,工程化PAL多肽与多种另外的组分组合,所述另外的组分包括但不限于防腐剂、助悬剂、增稠剂、湿润剂、醇类、脂肪酸和/或乳化剂。

[0548] 在一些实施方案中,特别是在液体制剂中,工程化PAL多肽与多种另外的组分组合,所述另外的组分包括但不限于防腐剂、助悬剂、增稠剂、湿润剂、醇类、脂肪酸和/或乳化剂。在一些实施方案中,将工程化PAL多肽与用于治疗PKU的其他化合物组合施用到受试者,所述用于治疗PKU的其他化合物包括但不限于 **KUVAN[®]** 四氢生物蝶呤(BioMarin Pharmaceutical, Inc., Novato, CA)、抗酸剂(例如奥美拉唑、埃索美拉唑和其他吡唑类),以及任何其他适合的化合物。

[0549] 在一些实施方案中,本发明提供了工程化PAL多肽,所述工程化PAL多肽适合用于降低在流体诸如血液、脑脊髓液等中的苯丙氨酸的浓度。施用到动物中的工程化PAL多肽的剂量取决于状况或疾病、动物的一般情况、和本领域技术人员已知的其他因素。在一些实施方案中,组合物用于单次或多次施用至动物。在一些实施方案中,预期的是,工程化PAL多肽在施用到动物(例如,患有PKU的人)的组合物中的浓度足以有效治疗、改善和/或预防疾病(例如PKU和/或PKU相关的状况、疾病和/或症状)。在一些实施方案中,工程化PAL多肽与其他药物和/或膳食组合物组合施用。

[0550] 工业组合物

[0551] 预期的是,将本发明的工程化PAL多肽应用于工业组合物中。在一些实施方案中,配制工程化PAL多肽用于在食品和/或饲料工业中使用。在一些实施方案中,将工程化PAL多

肽配制为粒状或丸状产品,其与动物饲料组分诸如另外的酶(例如,纤维素酶、漆酶和淀粉酶)混合。在一些可选的实施方案中,工程化PAL多肽被用于液体动物饲料组合物(例如,水基或油基的浆液)。因此,在一些实施方案中,本发明的工程化PAL变体是足够耐热和热稳定的以承受用于产生丸剂和其他处理的饲料/食品的处理。

[0552] 还将本发明的工程化PAL变体用于制备苯丙氨酸和/或苯丙氨酸衍生物。

[0553] 本申请提供了以下内容:

[0554] 项目1.一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:
a) 氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:4或其功能片段至少85%的序列同一性;b) 多肽序列,所述多肽序列包含与SEQ ID NO:4或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c) 所述工程化多肽与所述参考序列SEQ ID NO:4相比表现出选自以下的改进的特性:i) 增强的催化活性、ii) 降低的对蛋白水解的敏感性、iii) 增加的对酸性pH的耐受性、iv) 减少的聚集、v) 降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0555] 项目2.如项目1所述的工程化多肽,其中所述一个或更多个氨基酸位置选自20、24、27、39、43、45、47、54、58、59、62、70、73、80、82、91、94、98、104、105、110、112、115、117、118、119、121、123、124、125、126、127、128、129、130、131、133、134、135、139、140、141、142、143、144、145、146、147、149、150、151、153、154、156、157、158、159、172、174、175、176、177、178、180、187、191、195、199、205、206、210、212、213、214、232、240、243、245、247、248、250、256、257、266、270、275、278、279、285、286、289、290、292、304、305、307、308、309、319、321、326、331、332、334、349、353、355、364、365、369、370、371、372、374、375、377、378、379、381、382、383、384、385、387、389、394、396、399、400、403、404、407、417、418、425、431、432、433、434、435、436、437、438、439、443、446、447、453、456、459、460、461、463、471、472、473、474、475、476、477、478、479、482、483、503、507、509、521、522、524、525、528、538、546、547、551、558、560、564、565和/或其任何组合,其中所述氨基酸位置参考SEQ ID NO:4来编号。

[0556] 项目3.如前述项目中任一项所述的工程化多肽,其中所述参考序列SEQ ID NO:4的氨基酸残基对应于A39、T54、G59、S73、A91、Y158、S180、K195、A112、R134、Q240、T243、I245、A256、L257、N270、N290、Y304、R305、H307、E308、I326、L349、D353、L364、A394、S399、N400、P404、L407、F443、N453、Y459、T460、T463、N474、E509、Q521、K522、T524、P528、S546和/或P564。

[0557] 项目4.如前述项目中任一项所述的工程化多肽,其中当与SEQ ID NO:4的多肽最佳比对时,与SEQ ID NO:4相比的所述氨基酸残基差异选自以下置换中的一个或更多个:A39V、T54K、G59R、S73K、A112C、R134Q、A91V、Y158H、S180A、K195E、Q240R/W、T243I/L、I245L、A256G、L257W/A、N270K、N290G、Y304H、R305M、H307G/Q/M、E308Q、I326F、L349M、D353A/N、L364Q、A394V、S399N、N400K、P404A、L407V、F443H、N453G、Y459F、T460G、T463N、N474Q、E509L、Q521K/S、K522Y/F/N、T524S、P528L、S546R和P564G/L/M。

[0558] 项目5.如项目1-4中任一项所述的工程化多肽,其中所述氨基酸残基差异选自以下的一个或更多个的组合:A39、A91、Q240、A256、N290、Y304、R305、H307、D353、A394、S399、P404、L407、Q521、K522和T524。

[0559] 项目6.如前述项目中任一项所述的工程化多肽,其中所述工程化多肽具有与SEQ

ID NO:4至少约90%的序列同一性;和在位置H307处的氨基酸残基差异。

[0560] 项目7.如项目6所述的工程化多肽,其中所述氨基酸残基差异为H307G/Q/M。

[0561] 项目8.如前述项目中任一项所述的工程化多肽,其中所述改进的特性选自降低的对蛋白水解的敏感性和/或增加的对酸性pH的耐受性。

[0562] 项目9.如前述项目中任一项所述的工程化多肽,其中所述参考序列是源自多变鱼腥藻(*Anabaena variabilis*)的野生型PAL。

[0563] 项目10.一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:10或其功能片段至少85%的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:10或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与所述参考序列SEQ ID NO:10相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0564] 项目11.一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与参考序列SEQ ID NO:10至少85%的序列同一性的氨基酸序列,和与SEQ ID NO:10相比的至少一个氨基酸残基差异,并且所述工程化多肽与SEQ ID NO:10相比表现出选自以下的至少一种改进的特性:增强的催化活性、降低的对蛋白水解的敏感性、增加的对酸性pH的耐受性、减少的聚集、和/或降低的免疫原性。

[0565] 项目12.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或更多个:

[0566] I27E/V39A;I27E/V39A/R43L/V105C/A153R/L214E/P266H/L278D/C503Q;

[0567] I27E/V39A/R43L/L214E/A547D;

[0568] I27E/V39A/V105C/A112C/R134Q/L214E/L278D/C503Q/A547D/C565N;

[0569] I27E/V39A/V105C/A112C/R134Q/A153R/Q205T/L214E/P266H/L278D/C503Q/A551D;

[0570] I27E/V39A/V105C/A112C/Q205T/P210C/P266H/C503Q/A547D;

[0571] I27E/V39A/V105C/A112C/Q205T/P266H/I285E/C503Q/A551D;

[0572] I27E/V39A/V105C/A112C/L214E/I285E/C503Q/A547D;

[0573] I27E/V39A/V105C/S131N/R134Q/Q205T/L214E/C503Q/A547D/C565N;

[0574] I27E/V39A/V105C/R134Q/A153R/P210C/L278D/I285E/C503Q/A547D/A551D;

[0575] I27E/V39A/V105C/R134Q/Q205T/P210C/L278D/C503Q/A547D;

[0576] I27E/V39A/V105C/R134Q/Q205T/L214E;

[0577] I27E/V39A/V105C/R134Q/Q205T/L214E/A551D/C565N;

[0578] I27E/V39A/V105C/R134Q/Q205T/L278D/I285E/C503Q/A547D/A551D/C565N;

[0579] I27E/V39A/V105C/R134Q/P210C;I27E/V39A/V105C/R134Q/P210C/L214E;

[0580] I27E/V39A/V105C/R134Q/P210C/L214E/I285E/A547D;

[0581] I27E/V39A/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N;

[0582] I27E/V39A/V105C/R134Q/L214E/L278D/A547D/A551D;

[0583] I27E/V39A/V105C/R134Q/L214E/I285E/C503Q/A547D/A551D;

- [0584] I27E/V39A/V105C/R134Q/P266H/C503Q;
- [0585] I27E/V39A/V105C/R134Q/P266H/C503Q/A547D/A551D;
- [0586] I27E/V39A/V105C/R134Q/L278D/C503Q/C565N;
- [0587] I27E/V39A/V105C/R134Q/L278D/I285E/C503Q; I27E/V39A/V105C/R134Q/L278D/A551D;
- [0588] I27E/V39A/V105C/R134Q/I285E/A547D/A551D; I27E/V39A/V105C/R134Q/C503Q/A551D;
- [0589] I27E/V39A/V105C/A153R/Q205T/L278D/C503Q/A547D/A551D;
- [0590] I27E/V39A/V105C/A153R/L214E; I27E/V39A/V105C/A153R/I285E;
- [0591] I27E/V39A/V105C/A153R/C503Q/A547D/C565N; I27E/V39A/V105C/A153R/A551D/C565N;
- [0592] I27E/V39A/V105C/Q205T/P210C/L214E/L278D/A547D;
- [0593] I27E/V39A/V105C/Q205T/P210C/L278D/C503Q;
- [0594] I27E/V39A/V105C/Q205T/P210C/L278D/A547D;
- [0595] I27E/V39A/V105C/Q205T/L214E/L278D/C503Q/A547D;
- [0596] I27E/V39A/V105C/Q205T/L278D/C503Q/A547D;
- [0597] I27E/V39A/V105C/P210C/I285E/C503Q/A547D/A551D/C565N;
- [0598] I27E/V39A/V105C/P210C/L214E/P266H/L278D;
- [0599] I27E/V39A/V105C/L214E/P266H/C503Q/A547D/C565N;
- [0600] I27E/V39A/V105C/L214E/L278D/L309P/C503Q/A547D/A551D;
- [0601] I27E/V39A/V105C/L278D/C503Q/A547D/C565N; I27E/V39A/V105C/I285E/A547D;
- [0602] I27E/V39A/V105C/C503Q/A551D; I27E/V39A/V105C/C503Q/A547D/A551D/C565N;
- [0603] I27E/V39A/A112C/R134Q/Q205T/P210C/L214E/A551D/C565N;
- [0604] I27E/V39A/A112C/R134Q/L214E/P266H/A551D;
- [0605] I27E/V39A/A112C/R134Q/L214E/C503Q/A547D; I27E/V39A/A112C/R134Q/P266H/I285E;
- [0606] I27E/V39A/A112C/Q205T/L214E/P266H/C503Q/A551D/C565N;
- [0607] I27E/V39A/A112C/Q205T/L278D/I285E; I27E/V39A/A112C/L214E;
- [0608] I27E/V39A/A112C/L214E/L278D/C503Q/A547D/A551D; I27E/V39A/A112C/I285E;
- [0609] I27E/V39A/A112C/A547D; I27E/V39A/R134Q;
- [0610] I27E/V39A/R134Q/A153R/Q205T/L214E/P266H/C503Q;
- [0611] I27E/V39A/R134Q/A153R/P210C/L214E/L278D/I285E/A547D/C565N;
- [0612] I27E/V39A/R134Q/A153R/L214E/P266H/L278D/C503Q/A547D/C565N;
- [0613] I27E/V39A/R134Q/A153G/L214E/P266H/I285E/C503Q/A551D/C565N;
- [0614] I27E/V39A/R134Q/A153R/L214E/C503Q/A547D; I27E/V39A/R134Q/A153R/L278D;
- [0615] I27E/V39A/R134Q/A153R/L278D/A547D/A551D; I27E/V39A/R134Q/A153R/A547D;
- [0616] I27E/V39A/R134Q/Q205T/L214E/P266H/I285E/C503Q/A551D/C565N;
- [0617] I27E/V39A/R134Q/Q205T/P266H/C503Q/A551D/C565N;
- [0618] I27E/V39A/R134Q/P210C/L214E/C503Q; I27E/V39A/R134Q/P210C/C503Q/A551D;

- [0619] I27E/V39A/R134Q/L214E/P266H/A551D;
- [0620] I27E/V39A/R134Q/L278D/I285E/C503Q/A547D/A551D;
- [0621] I27E/V39A/R134Q/L278D/C503Q/A547D; I27E/V39A/R134Q/C503Q/A547D;
- [0622] I27E/V39A/R134Q/A547D/C565N; I27E/V39A/Q205T/L214E/C503Q/C565N;
- [0623] I27E/V39A/Q205T/P266H/I285E/A547D/A551D/C565N; I27E/V39A/Q205T/P266H/A551D;
- [0624] I27E/V39A/Q205T/L278D/C503Q/A551D/C565N; I27E/V39A/Q205T/L278D/C503Q/C565N;
- [0625] I27E/V39A/Q205T/C503Q/A547D/C565N; I27E/V39A/P210C/T212S;
- [0626] I27E/V39A/P210C/L214E/L278D/C503Q/A551D; I27E/V39A/P210C/L214E/I285E/C503Q/A551D;
- [0627] I27E/V39A/P210C/P266H/I285E/C503Q/A547D; I27E/V39A/P210C/P266H/C503Q/A551D;
- [0628] I27E/V39A/L214E; I27E/V39A/L214E/P266H/L278D/C503Q/A547D/A551D/C565N;
- [0629] I27E/V39A/L214E/L278D/C503Q; I27E/V39A/L214E/L278D/C503Q/A547D/C565N;
- [0630] I27E/V39A/L214E/C503Q/A551D; I27E/V39A/P266H; I27E/V39A/P266H/L278D;
- [0631] I27E/V39A/L278D; I27E/V39A/L278D/A547D; I27E/V39A/L278D/I285E/C503Q/A547D;
- [0632] I27E/V39A/L278D/C503Q/C565N; I27E/V39A/C503Q; I27E/G45D/Q205T/P266H/C565N;
- [0633] I27E/V105C; I27E/V105C/R134Q/A153R/P210C/L214E/C503Q/A547D;
- [0634] I27E/V105C/R134Q/A153R/I285E/A547D; I27E/V105C/R134Q/A153R/C503Q;
- [0635] I27E/V105C/R134Q/Q205T/P210C/C503Q;
- [0636] I27E/V105C/R134Q/Q205T/L214E/P266H/L278D/C503Q/C565N;
- [0637] I27E/V105C/Q205T/P266H/C503Q;
- [0638] I27E/V105C/R134Q/P210C/L214E/P266H/L278D/A551D/C565N;
- [0639] I27E/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N;
- [0640] I27E/V105C/R134Q/P210C/P266H/L278D/I285E/C503Q/A551D/C565N;
- [0641] I27E/V105C/R134Q/L214E/L278D/C503Q/A547D;
- [0642] I27E/V105C/R134Q/L214E/L278D/C503Q/A547D/A551D/C565N; I27E/V105C/Q205T;
- [0643] I27E/V105C/Q205T/L214E/P266H; I27E/V105C/Q205T/L214E/P266H/A551D/C565N;
- [0644] I27E/V105C/Q205T/L214E/L278D/I285E/C503Q/A547D/A551D/C565N;
- [0645] I27E/V105C/Q205T/C503Q/A547D/A551D/C565N; I27E/V105C/L214E;
- [0646] I27E/V105C/L214E/P266H/C503Q; I27E/V105C/L214E/I285E/A551D/C565N;
- [0647] I27E/V105C/L214E/A547D/A551D/C565N; I27E/V105C/L214E/A551D/C565N;
- [0648] I27E/V105C/P266H; I27E/V105C/P266H/I285E/C503Q/A547D/C565N;
- [0649] I27E/V105C/L278D/A547D; I27E/V105C/I285E/C503Q/A547D/A551D/C565N;

- [0650] I27E/V105C/C503Q/A547D/C565N; I27E/V105C/C503Q/A547D/A551D/C565N;
- [0651] I27E/A112C/R134Q/A153R/L214E/P266H/C503Q;
- [0652] I27E/A112C/R134Q/L278D/I285E/C503Q/A551D/C565N;
- [0653] I27E/A112C/R134Q/Q205T/L278D/C503Q; I27E/A112C/R134Q/Q205T/I285E/C503Q;
- [0654] I27E/A112C/Q205T/P266H/L278D/I285E/C503Q; I27E/A112C/P210C/L214E/C503Q/A547D;
- [0655] I27E/R134Q; I27E/R134Q/A153R/I285E/C503Q/A547D;
- [0656] I27E/R134Q/Q205T/I285E/C503Q/A551D; I27E/R134Q/Q205T/P266H/L278D/A547D;
- [0657] I27E/R134Q/P210C; I27E/R134Q/L214E/C503Q; I27E/R134Q/L214E/C503Q/A547D;
- [0658] I27E/R134Q/L214E/C503Q/A547D/A551D; I27E/R134Q/L214E/C503Q/C565N;
- [0659] I27E/R134Q/L278D/I285E/A551D/C565N; I27E/R134Q/I285E/C503Q;
- [0660] I27E/A153R/L214E/L278D/I285E/A551D/C565N; I27E/A153R/L214E/L278D/A551D;
- [0661] I27E/Q205T; I27E/Q205T/L214E/L278D/I285E/C503Q/C565N;
- [0662] I27E/Q205T/L214E/C503Q/A547D/C565N; I27E/Q205T/P266H/L278D/I285E/A551D/C565N;
- [0663] I27E/Q205T/L278D/A551D; I27E/P210C; I27E/P210C/L214E/C503Q/A547D;
- [0664] I27E/P210C/L278D/C503Q; I27E/P210C/C503Q; I27E/P210C/C503Q/C565N; I27E/P210C/A551D;
- [0665] I27E/L214E; I27E/L214E/P266H/L278D/I285E/A551D; I27E/L214E/L278D;
- [0666] I27E/L214E/L278D/C503Q; I27E/L214E/C503Q; I27E/L214E/C503Q/A547D;
- [0667] I27E/L214E/C503Q/A547D/C565N; I27E/L214E/A551D; I27E/P266H/L278D/C503Q;
- [0668] I27E/P266H/A547D/A551D; I27E/L278D/C503Q/A551D; I27E/L278D/C503Q/A551D/C565N;
- [0669] I27E/A547D/C565N; V39A/G45S/L278D/C503Q/A551D;
- [0670] V39A/V105C/R134Q/A153R/Q205T/A551D; V39A/V105C/R134Q/P210C/L214E/A551D;
- [0671] V39A/V105C/R134Q/L214E/C503Q/A547D/A551D; V39A/V105C/A153R/P266H/A547D/A551D;
- [0672] V39A/V105C/Q205T/C503Q; V39A/V105C/Q205T/A551D; V39A/V105C/P210C/A547D;
- [0673] V39A/V105C/L214E/P266H/A547D/C565N; V39A/V105C/L214E/I285E/C503Q/A551D/C565N;
- [0674] V39A/A112C/R134Q/Q205T/L214E/L278D; V39A/A112C/R134Q/L214E/C503Q/A547D/A551D;
- [0675] V39A/A112C/A153R/Q205T/L278D/C503Q/A547D; V39A/R134Q;
- [0676] V39A/R134Q/Q205T/L214E/C503Q/C565N; V39A/R134Q/P210C/L214E/A547D/C565N;

- [0677] V39A/A153R/C503Q/A547D;V39A/Q205T/L278D/A547D/A551D;
- [0678] V39A/P210C/L214E/L278D/I285E/C503Q/A551D;V39A/P266H;
- [0679] V39A/P275R/L278D/C503Q/A551D;V39A/C503Q;V39A/C503Q/A551D/C565N;V105C;
- [0680] V105C/A112C/R134Q/Q205T/L214E/Y492H/C503Q/A547D;
- [0681] V105C/R134Q/A153R/Q205T/L214E/C503Q;V105C/R134Q/Q205T/L214E/A547D;
- [0682] V105C/R134Q/Q205T/P266H/L278D;V105C/R134Q/L214E/P266H/I285E/C503Q/A551D/C565N;
- [0683] V105C/R134Q/L214E/L278D/C565N;V105C/R134Q/L214E/C503Q/A547D;
- [0684] V105C/R134Q/L214E/C503Q/A547D/A551D;V105C/R134Q/C503Q;
- [0685] V105C/R134Q/C503Q/A547D;V105C/R134Q/C503Q/A547D/C565N;
- [0686] V105C/A153R/Q205T/L214E/P266H/C503Q/A547D;
- [0687] V105C/A153R/Q205T/P266H/I285E/A547D/C565N;V105C/Q205T/P210C/L214E/C503Q/A547D;
- [0688] V105C/Q205T/L214E/L278D;V105C/Q205T/L214E/C503Q/A547D/A551D/C565N;
- [0689] V105C/Q205T/C503Q/A551D;V105C/L214E/P266H/L278D/A547D;
- [0690] V105C/L214E/L278D/C503Q/A547D/A551D;V105C/L214E/I285E;
- [0691] V105C/L214E/I285E/C503Q/A547D/A551D/C565N;V105C/L214E/I285E/A547D/C565N;
- [0692] V105C/L278D/C503Q/A551D;V105C/I285E;V105C/I285E/A547D;V105C/C503Q;
- [0693] V105C/A547D/A551D;A112C/R134Q/A153R/L214E/L278D/I285E/C503Q/A547D/A551D/C565N;
- [0694] A112C/R134Q/L214E/C503Q/A547D/A551D/C565N;A112C/L214E/L278D;
- [0695] A112C/L278D/C503Q/A547D;R134Q/Q205T/L214E/I285E/C503Q/A551D/C565N;
- [0696] R134Q/Q205T/C503Q;R134Q/P210C/L214E/L278D/C503Q/A547D/C565N;
- [0697] R134Q/P210C/L214E/C503Q/A547D/A551D;R134Q/L214E;R134Q/L214E/L278D/C503Q;
- [0698] R134Q/L214E/L278D/C503Q/A551D;R134Q/L214E/I285E/C503Q;R134Q/C503Q;
- [0699] R134Q/C503Q/A547D/A551D;A153R;Q205T/L214E/I285E/C503Q/A551D;
- [0700] Q205T/L214E/I285E/C503Q/C565N;Q205T/L214E/C503Q/A547D/C565N;
- [0701] Q205T/L278D/I285E/A547D/A551D;P210C/L214E;P210C/L214E/P266H;L214E/P266H;
- [0702] L214E/P266H/C503Q/A547D/A551D/C565N;L214E/C503Q/A547D;L214E/A547D;
- [0703] P266H/L278D/C503Q;P266H/C565N;L278D/A547D;C503Q;C503Q/A547D;
- [0704] C503Q/A547D/A551D/C565N;C503Q/A547D/C565N;C503Q/A551D;C503Q/A551D/C565N;
- [0705] A547D;和/或C565N。
- [0706] 项目13.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个:V80I/R134C/P564Q;V121C;A123G;A124G;M125L;L126I/T;L126M/R134L;L127A;A129G/L;N130Q;N130C/M370I R134W;

M133R;R134I;R134N/G307C;G135C/S;和/或G135A/A394E。

[0707] 项目14.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个个:G20S/I144L;R43S;L47M/I144L;L47M/R146E;L47M/M147G/A383E;L47M/P157C;Q58H/L143V;Q58K/P157D/G369C;A62S/M147V;S82I/G135C/P157F/W279L;R94C/I149E;T110I/I139R;L118M/L141H;A119E/T156H/A289D;I139M/V;R140D/G/M;R140N/A199E;R140E/A334S/A551D;L141K/Q/P/T;E142H/P/V;E142D/G371D;L143F/M;I144L/N/V;K145N/Q/R;K145G/P157T;R146H/L;R146W/D191Y;M147A;I149L/R;F150K/L/M;L151M;A153C/G;A153S/H250N;G154R;G154Y/L174M/Q321K/S456I/G483C;T156K/G483C;P157D/F/H/Y;Y158E;V159C/H/L/M;M247I;L319M;和/或Q389K。

[0708] 项目15.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个个:P117T/Y176Q;V172I/C/L;L174M;S175G;Y176E/I/M/R/V;I177M/V;T178L/A477S;和/或S180C/T。

[0709] 项目16.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个个:R43S/H374K;R43S/H374R;A112S/M370A/A507E;M147I/H374S;S187R/L381V;D191Y/H385N;A232S;Q240K/H374R;A256S/L381N;P275Q/M370S;P275T/H374R;Q332K/Y377M;A334S/H374V;L349M;Q355K/H374S;M370G/I/S;G371H/N/Q/S;M372A/V;H374A/D/G/L/N/R/S/T;H374Q/P396Q;H374R/G417C;L375I;L375M;Y377C/I/N;Y378C/D/E/I/L/N/S;Y378F/P404Q;I379C/H/L/M/N;L381G/V;L381M/Q560K;L382C/H/I/M/S;A383S/V;K384R;H385C/G/N;H385M/P403H;H385S/P403H;D387S;L418M;G425V;A447S;S461G;和/或S525L。

[0710] 项目17.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个个:A24S/F434M;A62S/T433N;S98I;L213M/S438L;Q240K/T433Y;S286R/Y435T;A289S/L431E;S331I;L431C/E/G/P/S/V;L432C/V;T433A/I/L/N/P/Q/R/S/V/W;F434C;Y435L;Y435Q/H446N;G436M;G436D/T;N437E/G/Q;N437T/L538M;S438C/F/M/R/T;I439C/F/L/V;和/或A477S。

[0711] 项目18.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个个:A24E;Q58R/Y475H;A70S/N474E;L104M/V476L;A119E/G365A;L206M;P275Q;G276V;Q292H/A479G;Q355H/I478C;P404T/A477V;I471F/G/K/M/N/R/V/W;F472G;Q473H/K/M/R/S;Q473H/A507S;N474A/H/R/W;N474D/R490H;Y475C/F/L/Q;V476C/I/L;I478N/S;A479G/S;F482C/L;G483C/H/S;G483A/S524I;G483R/G537C;和/或A558S。

[0712] 项目19.如项目10和/或11所述的工程化多肽,其中与SEQ ID NO:10相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或多个个:V39A/K115E/M133R/C565N;V39A/M133R/F472G/C503Q/C565N;V39A/M133R/F472G/C565N;V39A/M133R/C503Q;V39A/M133R/C503Q/C565N;V39A/M147A/Y378E/C503Q/C565N;V39A/M147A/Y378E/C565N;V39A/M147A/L381G/F472G/C503Q/C565N;V39A/M147A/L381G/C503Q/C565N;V39A/M147A/F472G/C503Q/C565N;V39A/M147A/F472G/C565N;V39A/M147A/C565N;V39A/G248C/L381G/F472G/C503Q/C565N;V39A/Y378E/C503Q/C565N;V39A/Y378E/C565N;V39A/L381G;V39A/F472G/

C503Q/C565N;V39A/C503Q/C565N;M133R/L381G/C565N;M133R/C503Q;Y378D/C503Q;Y378E/F472G/C503Q/C565N;L381G/F472G/C503Q/C565N;和/或F472G/C503Q/C565N。

[0713] 项目20.一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含:a)氨基酸序列,所述氨基酸序列具有与参考序列SEQ ID NO:26或其功能片段至少85%的序列同一性;b)多肽序列,所述多肽序列包含与SEQ ID NO:26或其功能片段相比在一个或更多个氨基酸位置处的至少一个氨基酸残基差异;并且c)所述工程化多肽与所述参考序列SEQ ID NO:26相比表现出选自以下的改进的特性:i)增强的催化活性、ii)降低的对蛋白水解的敏感性、iii)增加的对酸性pH的耐受性、iv)减少的聚集、v)降低的免疫原性、或i)、ii)、iii)、iv)或v)的任何组合。

[0714] 项目21.一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与参考序列SEQ ID NO:26至少85%的序列同一性和与SEQ ID NO:26相比的至少一个氨基酸残基差异的氨基酸序列,并且所述工程化多肽与SEQ ID NO:26相比表现出选自以下的至少一种改进的特性:增强的催化活性、降低的对蛋白水解的敏感性、增加的对酸性pH的耐受性、减少的聚集、和/或降低的免疫原性。

[0715] 项目22.如项目20和/或21所述的工程化多肽,其中与SEQ ID NO:26相比的所述氨基酸残基差异选自以下置换或置换集合中的一个或更多个:A24E/G381L;L127V;A129I/V;S131C/T;H132L/S;R134C/F/H/K;R134H/Y378E/G381L;R134H/Y378E/G381L/V388T;R134H/V388T;A136K;A289S;M372L;H374G/M/Q;G381A/C/F/I/L/M/N/Q/S/T;A383C/M;V388C/T;L431M;和/或L563M。

[0716] 项目23.如前述项目中任一项所述的工程化多肽,其中所述工程化多肽具有与参考序列SEQ ID NO:4至少约90%的序列同一性。

[0717] 项目24.如项目23所述的工程化多肽,其中所述工程化多肽具有与参考序列SEQ ID NO:4至少约95%的序列同一性。

[0718] 项目25.一种具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,所述工程化多肽包含具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24、和/或26或其功能片段至少约90%的序列同一性的氨基酸序列。

[0719] 项目26.如项目25所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,包含具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24、和/或26或其功能片段至少约99%的序列同一性的氨基酸序列。

[0720] 项目27.具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽是表2-1至2-5和/或表9-1至9-7中的任一个中提供的变体PAL。

[0721] 项目28.如项目1-27中任一项所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽是多变鱼腥藻酶。

[0722] 项目29.如项目1-28中任一项所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽是热稳定的。

[0723] 项目30.如项目1-29中任一项所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽耐受蛋白水解。

[0724] 项目31.如项目30所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽耐受通过至少一种消化道酶的蛋白水解。

[0725] 项目32.如项目31所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽耐受通过胰凝乳蛋白酶、胰蛋白酶、羧肽酶和/或弹性蛋白酶的蛋白水解。

[0726] 项目33.如项目1-32中任一项所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽,其中所述工程化多肽是酸稳定的。

[0727] 项目34.如项目1-33中任一项所述的工程化多肽,其中所述工程化多肽被去免疫化。

[0728] 项目35.如项目34所述的去免疫化的工程化多肽,其中所述多肽包含具有与SEQ ID NO:6、8、10、12、14、16、18、20、22、24和/或26至少95%的序列同一性的氨基酸序列。

[0729] 项目36.如项目1-35中任一项所述的工程化多肽,其中所述多肽是纯化的。

[0730] 项目37.一种多核苷酸序列,所述多核苷酸序列编码至少一种如前述项目1-36中任一项所述的工程化多肽。

[0731] 项目38.如项目37所述的多核苷酸序列,其中所述序列被可操作地连接到控制序列。

[0732] 项目39.如项目37和/或38所述的多核苷酸序列,其中所述多核苷酸是密码子优化的。

[0733] 项目40.一种表达载体,所述表达载体包含至少一种如项目37-39中任一项所述的多核苷酸序列,和至少一种控制序列。

[0734] 项目41.如项目40所述的表达载体,其中所述控制序列是启动子。

[0735] 项目42.如项目41所述的表达载体,其中所述启动子是异源启动子。

[0736] 项目43.一种宿主细胞,所述宿主细胞用至少一种如项目35-37中任一项所述的多核苷酸序列,和/或项目40-42中任一项所述的载体来转化。

[0737] 项目44.如项目43所述的宿主细胞,其中所述宿主细胞是大肠杆菌(E.coli)。

[0738] 项目45.一种在宿主细胞中产生工程化PAL多肽的方法,所述方法包括在适合的培养条件下培养包含编码至少一种如项目1至36中任一项所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽的至少一种多核苷酸、和/或至少一种如项目37和/或39所述的多核苷酸序列、和/或至少一种如项目40-42中任一项所述的载体的宿主细胞,使得产生至少一种工程化PAL多肽。

[0739] 项目46.如项目45所述的方法,所述方法还包括从所述培养物和/或宿主细胞回收至少一种具有苯丙氨酸解氨酶(PAL)的工程化多肽。

[0740] 项目47.如项目46所述的方法,所述方法还包括纯化产生的所述至少一种具有苯丙氨酸解氨酶(PAL)的工程化多肽的步骤。

[0741] 项目48.一种组合物,所述组合物包含至少一种如项目1至36中任一项所述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽。

[0742] 项目49.如项目48所述的组合物,其中所述组合物是药物组合物。

[0743] 项目50.如项目49所述的药物组合物,所述药物组合物还包含至少一种药学上可接受的赋形剂和/或载体。

[0744] 项目51.如项目48-50中任一项所述的组合物,其中所述组合物适合于苯丙酮尿症的治疗。

[0745] 项目52.如项目49-51中任一项所述的药物组合物,其中所述组合物适合于口服施

用至人。

[0746] 项目53.如项目49-52中任一项所述的药物组合物,其中所述组合物呈丸剂、片剂、胶囊、囊形片、液体或乳剂的形式。

[0747] 项目54.如项目53所述的药物组合物,其中所述丸剂、片剂、胶囊、或囊形片还包含肠溶包衣。

[0748] 项目55.如项目49-51中任一项所述的药物组合物,其中所述组合物适合于肠胃外注射入人中。

[0749] 项目56.如项目49-55中任一项所述的药物组合物,其中所述组合物与至少一种另外的治疗有效化合物共施用。

[0750] 项目57.如项目56所述的药物组合物,其中所述组合物包含至少一种另外的治疗有效化合物。

[0751] 项目58.一种用于治疗和/或预防受试者中苯丙酮尿症的症状的方法,所述方法包括:提供患有苯丙酮尿症的受试者,以及向所述受试者提供如项目49-57中任一项所述的药物组合物。

[0752] 项目59.如项目58所述的方法,其中所述苯丙酮尿症的症状被改善。

[0753] 项目60.如项目58和/或59所述的方法,其中所述受试者能够食用在其甲硫氨酸、苯丙氨酸和/或酪氨酸含量方面比由还未被提供包含至少一种如项目1-36中所阐述的具有苯丙氨酸解氨酶(PAL)活性的工程化多肽的至少一种药物组合物的受试者所需要的饮食较不受限的饮食。

[0754] 项目61.如项目58-60中任一项所述的方法,其中所述受试者是婴儿或儿童。

[0755] 项目62.如项目58-60中任一项所述的方法,其中所述受试者是成人或青年。

[0756] 项目63.项目48-57中任一项所提供的组合物的用途。

[0757] 可以结合以下非限制性实施例更好地理解本发明的前述和其他方面。提供实施例仅用于说明目的并且不意图以任何方式限制本发明的范围。

[0758] 实验

[0759] 提供以下实施例,包括实验和获得的结果,仅用于说明的目的,而不应被解释为限制本发明。

[0760] 在下面的实验公开内容中,应用以下缩写:ppm(parts per million);M(摩尔);mM(毫摩尔);uM和 μ M(微摩尔);nM(纳摩尔);mol(摩尔);gm和g(克);mg(毫克);ug和 μ g(微克);L和l(升);ml和mL(毫升);cm(厘米);mm(毫米);uM和 μ M(微米);sec.(秒);min(s)(分钟);h(s)和hr(s)(小时);U(单位);MW(分子量);rpm(每分钟转数);psi和PSI(每平方英寸磅数);°C(摄氏温度);RT和rt(室温);CDS(编码序列);DNA(脱氧核糖核酸);RNA(核糖核酸);大肠杆菌W3110(常用的实验室大肠杆菌菌株,可从Coli Genetic Stock Center [CGSC], New Haven, CT获得);HTP(高通量);HPLC(高压液相色谱);CFSE(羧基荧光素琥珀酰亚胺酯);IPTG(异丙基 β -D-1-硫代半乳糖苷);PES(聚醚砜);PHE和phe(苯丙氨酸);BSA(牛血清白蛋白);PBMC(外周血单核细胞);PKU(苯丙酮尿症);MHC(主要组织相容性复合物);HLA(人白细胞抗原);HLA-DR(由6号染色体上的HLA复合物编码的MHC II类细胞表面受体);FIOPC(相比于阳性对照的改进倍数);LB(Luria肉汤);AthensResearch(Athens Research Technology, Athens, GA);ProSpec(ProSpec TANY Technogene, East Brunswick, NJ);

Sigma-Aldrich (Sigma-Aldrich, St. Louis, MO); Ram Scientific (Ram Scientific, Inc., Yonkers, NY); Pall Corp. (Pall Corp., Pt. Washington, NY); Millipore (Millipore Corp., Billerica MA); Difco (Difco Laboratories, BD Diagnostic Systems, Detroit, MI); Molecular Devices (Molecular Devices, LLC, Sunnyvale, CA); Kuhner (Adolf Kuhner, AG, Basel, Switzerland); Cambridge Isotope Laboratories, (Cambridge Isotope Laboratories, Inc., Tewksbury, MA); Applied Biosystems (Applied Biosystems, part of Life Technologies, Corp., Grand Island, NY), Agilent (Agilent Technologies, Inc., Santa Clara, CA); Thermo Scientific (Thermo Fisher Scientific 的一部分, Waltham, MA); Corning (Corning, Inc., Palo Alto, CA); Constant Systems (Constant Systems Ltd., Daventry, United Kingdom); Megazyme (Megazyme International, Wicklow, Ireland); Enzo (Enzo Life Sciences, Inc., Farmingdale, NY); GE Healthcare (GE Healthcare Bio-Sciences, Piscataway, NJ); Harlan (Harlan Laboratories, Indianapolis, IN); AB Sciex (AB Sciex, Framingham, MA); 和 Bio-Rad (Bio-Rad Laboratories, Hercules, CA)。

[0761] 以下多核苷酸和多肽序列应用于本发明。在某些情况下(如下所示),多核苷酸序列之后是所编码的多肽。

[0762] pET16b-AvPAL表达载体的多核苷酸序列(SEQ ID NO:1):

TCTCATGTTTGACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTATCACAGTTAAATT
GCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAAATGCGCTCATCGTCATCCTCGGCA
CCGTCACCCTGGATGCTGTAGGCATAGGCTTGTTATGCCGGTACTGCCGGGCGCTCTTGC
GGGATATCCGGATATAGTTCCTCTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGGC
CCCAAGGGGTTATGCTAGTTATTGCTCAGCGGTGGCAGCAGCCAACCTCAGCTTCCTTTTCG
GGCTTTGTTAGCAGCCGGATCCTTAATGCAGACACGGCAGAATGTCCTGAACGGCCTGA
ACAATAACACCACCGGCTGCAATATCTGCACTAATACGTGCAATATGTTTCATCCAGACCC
TGTTTCATTATCATTCCAAATATACGGACGATCTGAGGTCGGTTTCTGACCAACAACATGA
CGAACTGCGCTATACAGACGTTTCGGTTGCCGGTGACAGACAGGCACGTGCATCATAATG
ACCGGTTTTTTTTGTAGGTACGCAGATCAACTGCCTGAACACCAAACATCAGGGCAATGGC
AACATAATTCTGAAAAATATCAACGCTACGACGTGCCAGGGTTGCGCTGGTATAACCCCTG
GCTGTTAATATTCTGGTTAACTGTTCCGGCATGGGTCGGAACGATCTGCAATACTATT
ACCATAAAAGGTCAGCAGCGGCATAATGCTATTACCGCAAATCTGCAGACCTTTCAGAC
CCATATTAACTTTACGTTACGATTACCCAGCAGACTCGGAGGCAGACCATTGCTAAATT
CCGGTGATGCCAGCAGTGCAATCTGAACATCCAGATGTTTTGCCAGCAGACCGATATAAT
AGCGCAGATGATCCATACCCATACCAACATACTGACCCAGAAAATTACCACCATGATAG
CTTGCCTGATTATCAACATCAATCAGCGGGTTATCGGTAACGCTGTTAATCTCAATTTTCG
ATTTGTTTGGCAATCTGGCTAATACCATCAACAATCGGACCCAGATACTGCGGCAGACAA
CGCAGGCTATAACGATCCTGGATCAGTTCATGATCACGATAATCATGTTTACCATCCAGT
TCATCACGAACCAGCTGGCTATTGGCCAGCAGGCTAATCATCTGATCTGCTGCCACAGC

[0764]

TGACCCGGATGCGGTTTGTCTGTTATGGATAAACGGATGAAAGCTCTGATTTGTACCATTC
AGTGCCTGAATATCCAGTGCATGAACACCCATTGCAATTGCGGTCAGAATCTGGGTATCA
TAAACACAATTTGCTGCAATACCGGTCATAACGCTGGTGCCATTCATCATTGCCAGACCT
TCTTTCGGCAGCAGGGTCAGCGGACTCAGATTCAGCTGACGCAGTGCGGTGCGTGCGTCC
ATTTCTTTGCCATTAAAATCAACTTTAAAGCTCGGGTCCAGGCCAATCAGGCTACCGGTA
ATATAGCTCAGCGGAACCAGATCACCGCTGGCACCAATGCTACCAAATTCATAAACATA
CGGGGTAACACCGGCATTGAGAAAGATTTCCATGCGTTTAATCAGTTCCAGACGAATACC
GCTTGCAACCACGCATGTGGCTATTTGCACGCAGCAGCATTGCTGCACGAACATCTGCCAG
CGGCAGTTTATTACCTGCACCGGTTTTTCAGAAACCAAACCAGATTGGTCTGCAGTTCGCT
TGCCTGTTACGGCTAATTGCAACATTTGCCATACCACCAAAAACCGCTGGTAACACCATA
AATCGGTTACCGCTTTCAACTGCATTATTGATATAATCACAGCTGGCCTGAATACCCTG
CAGAATATCGGTATTATTGGTCAGGCTAACCAGGGTGCCATTACGGGCAACACGTGCAA
CATCATTGATGGTCAGTTTCTGATTACCAATAATCACATTTGCGCTGCTATTGCCGGTAAA
GCTAAACTGCTGGCTGCTGGTTTTGCTCTGTGCCTGGCTCAGGGTTTTTCATATGACGACCT
TCGATATGGCCGCTGCTGTGATGATGATGATGATGATGATGATGATGATGGCCCATGGTATAT
CTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACAATTCC
CCTATAGTGAGTCGTATTAATTTGCGGGGATCGAGATCTCGATCCTCTACGCCGGACGCA
TCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCA
CCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTA
TGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGATGCACCATTC
TTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGT
CGCATAAGGGAGAGCGTCGAGATCCCGGACACCATCGAATGGCGCAAAACCTTTGCGGG
TATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAA
CGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTTCCCGCGTGGTGA
ACCAGGCCAGCCACGTTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGA
GCTGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGA
TTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTGCGGGCGATTA
AATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGC
GTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGAT
CATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGT
TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCAT
GAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTACCAGCAAATCGC
GCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAA
ATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCA
TGTCGGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGC
TGTTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTG
CGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTAT
ATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGAC
CGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCA
CTGGTGAAAAGAAAAACCACCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTT
GGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAG
CGAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACCGGGATCTCGACCGATGCCC
TTGAGAGCCTTCAACCCAGTCAGTCTCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCC
GCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGG
GTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTGCTTGCG
GTATTCGGAATCTTGACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGT
TTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTT
GCTGGCGTTTCGCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGG

[0765]

CGGCATCGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACCATC
AGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCACTGGACCGC
TGATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCATGGATTG
TAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGTGCATGGAGCCGGG
CCACCTCGACCTGAATGGAAGCCGGCGGCACCTCGCTAACGGATTACCACTCCAAGAA
TTGGAGCCAATCAATTCTTGCAGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAGAAC
ATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCAGCGTTGGG
TCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTCTGTTGAGGACCCGGCTAGGCTGGCGG
GGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTG
CTGCTGCAAAACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTC
GTAAAGTCTGGAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATCTGCATCGC
AGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATTG
ACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAA
CGTTCCAGTAACCGGGCATGTTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTT
TCATCGGTATCATTACCCCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGACCA
AACAGGAAAAAACCGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTT
CTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCA
CGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAA
CCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGA
GCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATG
ACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACATATGCGGCATCAGAGCAG
ATTGTACTGAGAGTGCACCATATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGA
AAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTC
GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA
GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTA
AAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
CCCCCTGGAAGCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGT
CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAG
TTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGA
CCGCTGCGCCTTATCCGGTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTA
CAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCT
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAA
AAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTT
TAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACA
GTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCAT
AGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCC
CAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAA
ACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC
CAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGC
AACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCAT
TCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAG
CGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTTATCAC
TCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTGATGCCATCCGTAAGATGCTTTTC
TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTG

- [0766] CTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCT
CATCATTGGAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATC
CAGTTCGATGTAACCCACTCGTGCACCCAACCTGATCTTCAGCATCTTTTACTTTACCCAGC
GTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGA
CACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGG
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGT
TCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGAC
ATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAAT(SEQ ID NO:1)
- [0767] AvPAL ORF的多核苷酸序列(SEQ ID NO:2):
ATGAAAACCCTGAGCCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGG
CAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCAC
GTGTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTA
TTCAGGCCAGCTGTGATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTG
TTACCAGCGGTTTTGGTGGTATGGCAAATGTTGCAATTAGCCGTGAACAGGCAAGCGAA
CTGCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCA
GATGTTCTGTGCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGTATT
CGTCTGGAACCTGATTAAACGCATGGAAATCTTCTGAATGCCGGTGTTACCCCGTATGTT
TATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCCGCTGAGCTATATTACCGGT
AGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGACGC
ACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCCTGCTGCCGAAAGAAGGTCT
GGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGATAC
CCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGGCACTGAATGG
TACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTG
GGCAGCAGATCAGATGATTAGCCTGCTGGCCAATAGCCAGCTGGTTTCGTGATGAACTGG
ATGGTAAACATGATTATCGTGATCATGAACTGATCCAGGATCGTTATAGCCTGCGTTGTC
TGCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAA
TTGAGATTAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATC
ATGGTGGTAATTTTCTGGGTGAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTATAT
CGGTCTGCTGGCAAAACATCTGGATGTTTCAGATTGCACTGCTGGCATCACCGGAATTTAG
CAATGGTCTGCCTCCGAGTCTGCTGGGTAAATCGTGAACGTAAAGTTAATATGGGTCTGAA
AGGTCTGCAGATTTGCGGTAAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAGTAT
TGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATATTAACAGCCAGGGTTA
TACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTTTTCAGAATTATGTTGCCATTGC
CCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTGATTATGA
TGCACGTGCCTGTCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTTCATGTTGT
TGGTCAGAAACCGACCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGG
ATGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTTCAGGCCG
TTCAGGACATTCTGCCGTGTCTGCAT (SEQ ID NO:2)
- [0768]
- [0769] WT AvPAL的多核苷酸序列(SEQ ID NO:3):
- [0770] ATGAAAACCCTGAGCCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGG
CAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCAC

- GTGTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTA
TTCAGGCCAGCTGTGATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTATGGTG
TTACCAGCGGTTTTGGTGGTATGGCAAATGTTGCAATTAGCCGTGAACAGGCAAGCGAA
CTGCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCA
GATGTTCTGTGCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGTATT
CGTCTGGAAC TGATTAAACGCATGGAAATCTTTCTGAATGCCGGTGTTACCCCGTATGTT
TATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCCGCTGAGCTATATTACCGGT
AGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGACGC
ACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCGCTGCGCGAAAGAAGGTCT
GGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGATAC
CCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGGCACTGAATGG
TACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTCAGCTGTG
GGCAGCAGATCAGATGATTAGCCTGCTGGCCAATAGCCAGCTGGTTCGTGATGAATGG
[0771] ATGGTAAACATGATTATCGTGATCATGAACTGATCCAGGATCGTTATAGCCTGCGTTGTC
TGCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAA
TTGAGATTAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATC
ATGGTGGAATTTTCTGGGTGAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTATAT
CGGTCTGCTGGCAAAACATCTGGATGTTTCAGATTGCACTGCTGGCATCACCGGAATTTAG
CAATGGTCTGCCTCCGAGTCTGCTGGGTAAATCGTGAACGTAAAGTTAATATGGGTCTGAA
AGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAGTAT
TGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATATTAACAGCCAGGGTTA
TACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTTTTCAGAATTATGTTGCCATTGC
CCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTGATTATGA
TGCACGTGCCTGTCTGTACCGGCAACCGAACGTCTGTATAGCGCAGTTTCGTGATGTTGT
TGGTCAGAAACCGACCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGG
ATGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTCAAGCCG
TTCAGGACATTCTGCCGTGTCTGCAT (SEQ ID NO:3)
- [0772] WT AvPAL的多肽序列 (SEQ ID NO:4) :
- MKTLSQAQSKTSSQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGIQAS
CDYINNAVESGEPIYGVTSFGGGMANVAISREQASELQTNLVWFLKTGAGNKLPLADVRAA
MLLRANSHMRGASGIRLELIKRMIEFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDP
KVDENGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYVDYQILTAIAMGV
HALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLANSQVLRDEL DKGHDYRDHELIQ
[0773] DRYSLRCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDH
LRYIIGLLAKHLDVQIALLASPEFSNGLPPSLLGNRERKVNMLKGLQICGNSIMPLLTIFYGN
SIADRFPTHAEQFNQININSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKKTGHYD
ARACLS PATERLYSAVRHVVGQKPTSDRPYIWN DNEQGLDEHIARISADIAAGGVIVQAVQDI
LPCLH (SEQ ID NO:4)
- [0774] AvPAL变体30号的多核苷酸序列 (SEQ ID NO:5) :
- ATGAAAACCCTGAGTCAGGCACAGAGCAAAACCGCAGCCAGCAGTTTAGCTTTACCGG
CAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGTACG
[0775] TGTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTAT
TCAGGCCAGCTGTGATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTATGGTGT
TACCAGCGGTTTTGGTGGTATGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACT

- GCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCAG
ATGTTTCGTGCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGCGTGAAGCGGTATTC
GTCTGGAAGTGAATTAACGCATGGAAATCTTTCTGAATGCCGGTGTTACCCCGTATGTTT
ATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCCGCTGAGCTATATTACCGGTA
GCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGACGCAC
CGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAAAGAAGGTCTG
GCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGATACC
CAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTACGGGACTGAATGGT
ACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTGG
GCAGCAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTTCGTGATGAACTGGA
TGGTAAACATGATTATCGTGATCATGAACTGATCCAGGATCGTTATAGCCTGCGTTGTCT
GCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAAT
TGAGATTAAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCA
TGGTGGTAATTTTCTGGGTGAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTATATC
GGTCTGCTGGCAAAACATCTGGATGTTTCAGATTGCACTGCTGGTATCACCGGAATTTAAC
AATGGTCTGCGTGCAGTCTGGTGGGTAATCGTGAACGTAAAGTTAATATGGGTCTGAAA
GGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAGTATT
GCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATATTAACAGCCAGGGTTAT
ACCAGCGCAACCCCTGGCACGTCGTAGCGTTGATATTTTTCAGAATTATGTTGCCATTGCC
CTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCATTATGAT
GCACGTGCCTGTCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTATGTTGTT
GGTCAGTATCCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGGAT
GAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTTATTGTTTCAGGCCGTT
CAGGACATTCTGCCGTGTCTGCAT (SEQ ID NO:5)
- [0776]
- [0777] AvPAL变体30号的多肽序列 (SEQ ID NO:6) :
- MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVVRVARNGTLVSLTNNTDILQGIQAS
CDYINNAVESGEPIYGVTSFGFGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVRAA
MLLRANSHMRGASGIRLELIKRMIFLNAGVTPYVYEFSGISGDLVPLSYITGSLIGLDPSF
KVDFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCVYDTQILTAIAMGV
HALDIQGLNGTNQSFHPFIHNSKPHPGQLWAADQMISLLAGSQLVRDELKGHDYRDHELIQ
[0778] DRYSLRCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDH
LRYIIGLLAKHLDVQIALLVSPFNGLPASLVGNRERKVNMGKGLQICGNSIMPLLTIFYG
NSIADRFPTHAEQFNQININSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKKTGHY
DARACLSPATERLYSAVRHVVGQYPSSDRPYIWNENEQGLDEHIARISADIAAGGVIVQAVQ
DILPCLH (SEQ ID NO:6)
- [0779] AvPAL变体22号的多核苷酸序列 (SEQ ID NO:7) :
- ATGAAAACCCCTGAGCCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGG
CAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCAC
[0780] GTGTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTA
TTCAGGCCAGCTGTGATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTATGGTG
TTACCAGCGGTTTTGGTGGTATGGCAAATGTTGCAATTAGCCGTGAACAGGCAAGCGAA

- [0781] CTGCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCA
GATGTTCTGTCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGTATT
CGTCTGGAACCTGATTAAACGCATGGAAATCTTTCTGAATGCCGGTGTTACCCCGTATGTT
TATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCGCTGAGCTATATTACCGGT
AGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGACGC
ACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAAAGAAGGTCT
GGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGATAC
CCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGGCACTGAATGG
TACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTCAGCTGTG
GGCAGCAGATCAGATGATTAGCCTGCTGGCCAATAGCCAGCTGGTTCGTGATGAAGTGG
ATGGTAAACATGATTATCGTGATGGTGAAGTATCCAGGATCGTTATAGCCTGCGTTGTC
TGCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAA
TTGAGATTAAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATC
ATGGTGGTAATTTTCTGGGTCAAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTATAT
CGGTCTGCTGGCAAAACATCTGGATGTTTCAGATTGCACTGCTGGCATCACCGGAATTTAG
CAATGGTCTGCCTCCGAGTCTGCTGGGTAAATCGTGAACGTAAAGTTAATATGGGTCTGAA
AGGTCTGCAGATTTGCGGTAAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAGTAT
TGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAAACCAGAATATTAACAGCCAGGGTTA
TACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTTTTCAGAATTATGTTGCCATTGC
CCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCAATTATGA
TGCACGTGCCTGTCTGTACCCGGCAACCGAACGTCTGTATAGCGCAGTTTCGTATGTTGT
TGGTCAGAAACCGACCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGG
ATGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTTCAGGCCG
TTCAGGACATTCTGCCGTGTCTGCAT (SEQ ID NO:7)
- [0782] AvPAL变体22号的多肽序列 (SEQ ID NO:8) :
- [0783] MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGIQAS
CDYINNAVESGEPIYGVTSFGFGMANVAISREQASELQTNLVWFLKTGAGNKLPLADVRAA
MLLRANSHMRGASGIRLELIKRMIFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDPSF
KVDENGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYVDQTILTAIAMGV
HALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLANSQLVRLDELGKHHDYRDGELIQ
DRYSLRCLPQYLGPIVDGISQIAKQIEIINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDH
LRYIIGLLAKHLDVQIALLASPEFSNGLPPSLLGNRERKVNMLGKGLQICGNSIMPLLTIFYGN
SIADRFPTHAEQFNQININSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKKTGHYD
ARACLSPATERLYSAVRHVVGQKPTSDRPYIWNNDNEQGLDEHIARISADIAAGGVIVQAVQDI
LPCLH (SEQ ID NO:8)
- [0784] AvPAL变体36号的多核苷酸序列 (SEQ ID NO:9) :
- [0785] ATGAAAACCCTGAGTCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGG
CAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGTACG
TGTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTAT
TCAGGCCAGCTGTGATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTGT
TACCAGCGGTTTTGGTGGTATGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACT
GCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCAG
ATGTTTCGTGCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGTATTC
GTCTGGAAGTATTAAACGCATGGAAATCTTTCTGAATGCCGGTGTTACCCCGTATGTTT
ATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCGCTGAGCTATATTACCGGTA

- [0786] GCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGACGCAC
CGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAAAGAAGGTCTG
GCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTTGTGTTTATGATACC
CAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGGCACTGAATGGT
ACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTGG
GCAGCAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACTGGA
TGGTAAACATGATTATCGTGATGGTGAACCTGATCCAGGATCGTTATAGCCTGCGTTGTCT
GCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAAT
TGAGATTAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCA
TGGTGGTAATTTTCTGGGTGAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTATATC
GGTCTGCTGGCAAACATCTGGATGTTTCAGATTGCACTGCTGGCATCACCGGAATTTAGC
AATGGTCTGCCTCCGAGTCTGGTGGGTAATCGTGAACGTAAAGTTAATATGGGTCTGAAA
GGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAGTATT
GCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATATTAACAGCCAGGGTTAT
ACCAGCGCAACCTGGCACGTCGTAGCGTTGATATTTTTCAGAATTATGTTGCCATTGCC
CTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCATTATGAT
GCACGTGCCTGTCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTATGTTGTT
GGTCAGAAACCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGGA
TGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTTCAGGCCGT
TCAGGACATTCTGCCGTGTCTGCAT (SEQ ID NO:9)
- [0787] AvPAL变体36号的多肽序列 (SEQ ID NO:10) :
- [0788] MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVVRVARNGTLVSLTNNTDILQGIQAS
CDYINNAVESGEPIYGVTSGFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVRAA
MLLRANSHMRGASGIRLELIKRMIFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDPSF
KVDFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYVDYQILTAIAMGV
HALDIQALNGTNQSFHPIHNSKPHPGQLWAADQMISLLAGSQLVRDELDDGKHHDYRDGELIQ
DRYSLRCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDH
LRYIIGLLAKHLDVQIALLASPEFSNGLPPSLVGNRERKVNMLKGLQICGNSIMPLLTIFYGN
SIADRFPTHAEQFNQININSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKKTGHYD
ARACLS PATERLYSAVRHVVGQKPSSDRPYIWN DNEQGLDEHIARISADIAAGGVIVQAVQDI
LPCLH (SEQ ID NO:10)
- [0789] AvPAL变体42号的多核苷酸序列 (SEQ ID NO:11) :
- [0790] ATGAAAACCTGAGTCAGGCACAGAGCAAAAACCGAGCAGCCAGCAGTTTAGCTTTACCGGC
AATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGTACGT
GTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATAAAGATATTCTGCAGCGTATTACGGC
CAGCTGTGATTATATCAATAATGCAGTTGAAAAAGGTGAACCGATTTATGGTGTACCAGCGGTT
TTGGTGGTATGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGT
TTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCAGATGTTTCGTGCAGCAATGCTGC
TGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGTATTCGTCTGGAACCTGATTAACGCATGGA
AATCTTTCTGAATGCCGGTGTTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTG
ATCTGGTTCCGCTGAGCTATATTACCGGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGAT
TTTAATGGCAAAGAAATGGACGCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCC
TGCTGCCGAAAGAAGGTCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAA

- [0791] TTGTGTTTATGATACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGG
CACTGAATGGTACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTCAG
CTGTGGGCAGCAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACTGG
ATGGTAAACATGATTATATGGATGGTGAAGTATCCAGGATCGTTATAGCCTGCGTTGTCTGCCG
CAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAATTGAGATTAA
CAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCATGGTGGTAATTTTC
TGGGTCAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTATATCGGTCTGCTGGCAAAACAT
CTGGATGTTTCAGATTGCACTGCTGGCATCACCAGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGT
GGGTAATCGTGAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATT
ATGCCGCTGCTGACCTTTTATGGTAATAGTATTGCAGATCGTTTCCGACCCATGCCGAACAGTT
TAACCAGAATATTAACAGCCAGGGTTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTT
TTCAGAAATTATGTTGCCATTGCCCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAA
AAAACCGGTCATTATGATGCACGTGCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAG
TTCGTCATGTTGTTGGTAAAAAACCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAG
GGTCTGGATGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTCA
GGCCGTTACAGGACATTCTGCCGCCGCTGCAT (SEQ ID NO:11)
- [0792] AvPAL变体42号的多肽序列 (SEQ ID NO:12) :
- [0793] MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKL TINDVVRVARNGTLVSLTNNKDILQRIQASCDYIN
NAVEKGEPIYGVTSFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVRAAMLLRANSHM
RGASGIRLELIKMEIFLNAGVTPYVYEFSGISAGDLVPLSYITGSLIGLDPSFKVDFNGKEMDAPTA
LRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCVDYDTQILTAIAMGVHALDIQALNGTNQSFHPIH
NSKPHPGQLWAADQMISLLAGSQLVRDEL DKGHDYMDGELIQDRYSLRCLP
QYLGPVVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDHLRYIYIGLLAKHLDV
QIALLASPEFSNGLPPLVGNRERKVNMLGKGLQICGNSIMPLLTIFYGNSIADRFPTHAEQFNQINNSQ
GYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKKTGHYDARAQLSPATERLYSAVRHVVGKKP
SSDRPYIWNQDEQGLDEHIARISADIAAGGVIVQAVQDILPLH (SEQ ID NO:12)
- [0794] AvPAL变体43号的多核苷酸序列 (SEQ ID NO:13) :
- [0795] ATGAAAACCTGAGTCAGGCACAGAGCAAAACCGAGCCAGCAGTTTACGCTTACCGGCAATA
GCAGCGCAAATGTGATTATTGGTAATCAGAAACTGACCATCAATGATGTTGTACGTGTTGCCCCGT
AATGGCACCCCTGGTTAGCCTGACCAATAATAAGATATTCTGCAGCGTATTCAGGCCAGCTGTG
ATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTGTACCAGCGGTTTTGGTGGT
ATGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGTTTGGTTTTCT
GAAAACCGGTGCAGGTAAATAAACTGCCGCTGGCAGATGTTCTGTCAGCAATGCTGCTGCGTGCA
AATAGCCACATGCGTGGTGCAAGCGGTATTCTGCTGGAAGTATTAAACGCATGGAAATCTTTCT
GAATGCCGGTGTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTC
CGCTGAGCTATATTACCGGTAGCCTGATTGGCC TGGACCCGAGCTTTAAAGTTGATTTTAATG
GCAAAGAAATGGACGCACCGACCGCACTGC
GTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAAAGAAGGTCTGGCAATGATGAATGGCAC
CAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGATACCCAGATTCTGACCGCAATTGCAA
TGGGTGTTTCATGCACTGGATATTCAGGCACTGAATGGTACAAATCAGAGCTTTCATCCGTTTATC
CATAACAGCAAACCGCATCCGGGTCAGCTGTGGGCAGCAGATCAGATGATTAGCCTGCTGGCCG
GTAGCCAGCTGGTTCGTGATGAACTGGATGGTAAACATGATTATATGGATGGTGAACCTGATCCA
GGATCGTTATAGCCTGCGTTGTCTGCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGA
TTGCCAAACAAATCGAAATTGAGATTAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAAT
CAGGCAAGCTATCATGGTGGTAATTTTCTGGGTGAGTA
TGTTGGTATGGGTATGGATCATCTGCGCTATTATATCGGTCTGCTGGCAAAACATCTGGATGTTT
AGATTGCACTGCTGGCATCACCAGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGTGGGTAATCGT
GAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCT
GACCTTTTATGGTAATAGTATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATA
TTAACAGCCAGGGTTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTTTTCAGAAATTAT
GTTGCCATTGCCCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCA
TTATGATGCACGTGCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTCATGTTG
TTGGTAAAAAACCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGGATGA
ACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTTCAGGCCGTTTCAGGACA
TTCTGCCGAACCTGCAT (SEQ ID NO:13)
- [0797] AvPAL变体43号的多肽序列 (SEQ ID NO:14) :

- [0798] MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVVRVARNGTLVSLTNNDILQRIQASCDYIN
NAVESGEPIYGVTSGFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVRAAMLLRANSHM
RGASGIRLELIKREIFLNAGVTPYVYEFSGISASGDLVPLSYITGSLIGLDPFSFKVDFNGKEMDAPTA
LRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYDQTILTAIAMGVHALDIQALNGTNQSFHPIH
NSKPHPGQLWAADQMISLLAGSQLVRDELGKHDMYMDGELIQDRYSLRCLPQYLGPIVDGISQIAKQ
IEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDHLRYIIGLLAKHLDVQIALLASPEFSNGLPP
SLVGNRERKVNMGKGLQICGNSIMPLLTIFYGNSIADRFPTHAEQFNQNSQGYTSATLARRSVDIF
QNYVAIALMFGVQAVDLRTYKKTGHYDARAQLSPATERLYSAVRHVVGKKPSSDRPYIWNNDNEQG
LDEHIARISADIAAGGVIVQAVQDILPNLH (SEQ ID NO:14)
- [0799] AvPAL变体1002号的多核苷酸序列 (SEQ ID NO:15) :
ATGAAAACCTGAGTCAGGCACAGAGCAAAACAGCAGCCAGCAGTTTAGCTTTACCGGCAATA
GCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCGCGTGTTGCCCGT
AATGGCACCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTATTACAGGCCAGCTGTGA
TTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTGTTACCAGCGGTTTTGGTGGA
TGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGTTTGGTTTCTG
AAAACCGGTGCAGGTAATAAACTGCCGTGGCAGATGTTTCGTGCAGCAATGCTGCTGCGTGCAA
ATAGCCACATGCATGGTGCAAGCGGTATTCGTCTGGAACCTGATTAACGCGCGGAAATCTTTCTG
AATGCCGGTGTTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCC
GCTGAGCTATATTACCGGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCA
AAGAAATGGACGCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAA
AGAAGGTCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTAT
GATACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTACAGGCACTGAATGG
TACAAATCAGAGCTTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTGGGCAG
CAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACTGGATGGTAAACA
TGATTATCGTGATGGTGAATGATCCAGGATCGTTATAGCCTGCGTTGTCTGCCGAGTATCTGG
GTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAATTGAGATTAAACAGCGTTACC
GATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCATGGTGGAATTTTCTGGGTGAGTA
TGTTGGTATGGGTATGGATCATCTGCGCTATTATATCGGTGGCCTGGCAAAACATCTGGATGTTT
AGATTGCACTGCTGGCATCACCGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGTGGGTAATCGT
GAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCT
GACCTTTTATGGTAATAGTATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATA
TTAACAGCCAGGGTTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTGGCCAGAATTAT
GTTGCCATTGCCCTGATGTTTGGTGTTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCA
TTATGATGCACGTGCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTGATGTTG
TTGGTCAGAAACCGAGCTCAGATCGTCCGTATATTGGAATGATAATGAACAGGGTCTGGATGA
ACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTTATTGTTACAGGCCGTTACAGGACA
TTCTGCCGAACCTGCAT (SEQ ID NO:15)
- [0801] AvPAL变体1002号的多肽序列 (SEQ ID NO:16) :
MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGIQAS
CDYINNAVESGEPIYGVTSGFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVRAA
MLLRANSHMHGASGIRLELIKRAEIFLNAGVTPYVYEFSGISASGDLVPLSYITGSLIGLDPFS
KVDNFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYDQTILTAIAMGV
HALDIQALNGTNQSFHPIHNSKPHPGQLWAADQMISLLAGSQLVRDELGKHDMYMDGELIQ
DRYSLRCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDH
LRYIIGLLAKHLDVQIALLASPEFSNGLPPSLVGNRERKVNMGKGLQICGNSIMPLLTIFYGN
SIADRFPTHAEQFNQNSQGYTSATLARRSVDIGQNYVAIALMFGVQAVDLRTYKKTGHYD
ARAQLSPATERLYSAVRHVVGQKPSSDRPYIWNNDNEQGLDE
HIARISADIAAGGVIVQAVQDILPNLH (SEQ ID NO:16)
- [0802] DRYSLRCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDH
LRYIIGLLAKHLDVQIALLASPEFSNGLPPSLVGNRERKVNMGKGLQICGNSIMPLLTIFYGN
SIADRFPTHAEQFNQNSQGYTSATLARRSVDIGQNYVAIALMFGVQAVDLRTYKKTGHYD
ARAQLSPATERLYSAVRHVVGQKPSSDRPYIWNNDNEQGLDE
HIARISADIAAGGVIVQAVQDILPNLH (SEQ ID NO:16)
- [0803] AvPAL变体1008号的多核苷酸序列 (SEQ ID NO:17) :

- ATGAAAACCTGAGTCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGGCAATA
GCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCGCGTGTTGCCCCT
AATGGCACCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTATTACAGGCCAGCTGTGA
TTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTGTTACCAGCGGTTTTGGTGGTA
TGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGTTTGGTTTCTG
AAAACCGGTGCAGGTAATAAACTGCCGCTGGCAGATGTTTCGTGCAGCAATGCTGCTGCGTGCAA
ATAGCCACATGCATGGTGCAAGCGGTATTCGTCTGGAACCTGATTAACGCGCGGAAATCTTTCTG
AATGCCGGTGTTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCC
GCTGAGCTATATTACCGGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCA
AAGAAATGGACGCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAA
AGAAGGTCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTAT
GATACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTACAGGCACTGAATGG
TACAAATCAGAGCTTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTGGGCAG
[0804] CAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACCTGGATGGTAAACA
TGATTATCGTGATGGTGAACCTGATCCAGGATCGTTATAGCCTGCGTTGTCTGCCGCAGTATCTGG
GTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAATTGAGATTAACAGCGTTACC
GATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCATGGTGGTAATTTTCTGGGTGAGTA
TGTGGTATGGGTATGGATCATCTGCGCTATTATATCGGTGGCCTGGCAAAACATCTGGATACCC
AGATTGCACTGCTGGCATCACCGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGTGGGTAATCGT
GAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCT
GACCTTTTATGGTAATAGTATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATA
TTAACAGCCAGGGTTATACCAGCGCAACCCCTGGCACGTCGTAGCGTTGATATTGGCCAGAATTAT
GTTGCCATTGCCCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCA
TTATGATGCACGTGCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTGATGTTG
TTGGTCAGAAACCGAGCTCAGATCGTCCGTATATTGGAATGATAATGAACAGGGTCTGGATGA
ACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTTATTGTTACAGCCGTTTCAGGACA
TTCTGCCGAACCTGCAT (SEQ ID NO:17)
- [0805] AvPAL变体1008号的多肽序列 (SEQ ID NO:18) :
- [0806] MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGI
QASCDYINNAVESGEPIYGVTSFGFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVR
AAMLLRANSHMHGASGIRLELIKRAEIFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDP
SFKVDFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYVDTQILTAIAM
GVHALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLAGSQLVRDELKGKHDYRDGE
LIQDRYSLRCLPQYLGPVVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGM
DHLRYIYIGGLAKHLDLQIALLASPEFSNGLPPSLVGNRERKVNMGKGLQICGNSIMPLLTIFY
[0807] GNSIADRFPHTAEQFNQNSQGYTSATLARRSVDIGQNYVAIALMFGVQAVDLRTRYKKTGH
YDARAQLSPATERLYSAVRHVVGQKPSSDRPYIWNENEQGLDE
HIARISADIAAGGVIVQAVQDILPNLH (SEQ ID NO:18)
- [0808] AvPAL变体1009号的多核苷酸序列 (SEQ ID NO:19) :

ATGAAAACCTGAGTCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGGCAATA
GCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCGCGTGTTGCCCCGT
AATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTATTACAGGCCAGCTGTGA
TTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTGTTACCAGCGGTTTTGGTGGA
TGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGTTTGGTTTCTG
AAAACCGGTGCAGGTAATAAACTGCCGCTGGCAGATGTTTCGTGCAGCAATGCTGCTGCGTGCAA
ATAGCCACATGCATGGTGCAAGCGGTATTCGTCTGGAACCTGATTAACGCGCGGAAATCTTTCTG
AATGCCGGTGTTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCC
GCTGAGCTATATTACCGGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAAATGGCA
AAGAAATGGACGCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAA
AGAAGGTCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTTAT
GATACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTACAGGCACTGAATGG
TACAAATCAGAGCTTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTGGGCAG
[0809] CAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTTCGTGATGAACCTGGATGGTAAACA
TGATTATCGTGATGGTGAACCTGATCCAGGATCGTTATAGCCTGCGTTGTCTGCCGAGTATCTGG
GTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAATTGAGATTAAACAGCGTTACC
GATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCATGGTGGAATTTTCTGGGTGAGTA
TGTTGGTATGGGTATGGATCATCTGCGCTATGAAATCGGTCTGCTGGCAAAACATCTGGATGTTT
AGATTGCACTGCTGGCATCACCGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGTGGGTAATCGT
GAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCT
GACCTTTTATGGTAATAGTATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATA
TTAACAGCCAGGGTTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTGGCCAGAATTAT
GTTGCCATTGCCCTGATGTTTGGTGTTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCA
TTATGATGCACGTGCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTGATGTTG
TTGGTCAGAAACCGAGCTCAGATCGTCCGTATATTGGAATGATAATGAACAGGGTCTGGATGA
ACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTTATTGTTTCAGGCCGTTTCAGGACA
TTCTGCCGAACCTGCAT (SEQ ID NO:19)

[0810] AvPAL变体1009号的多肽序列 (SEQ ID NO:20) :

MKTLQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGI
QASCDYINNAVESGEPIYGVTSFGFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVR
AAMLLRANSHMHGASGIRLELIKRAEIFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDP
[0811] SFKVDNFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANVCYDQTILTAIAM
GVHALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLAGSQLVRDELDDGKHLDYRDGE
LIQDRYSLRCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGM
DHLRYEIGLLAKHLDVQIALLASPEFSNGLPPLVGNRERKVNMLKGLQICGNSIMPLLTFFY
GNSIADRFPTHAEQFNQNSQGYTSATLARRSVDIGQNYVAIALMFGVQAVDLRTYKKTGH
YDARAQLSPATERLYSAVRHVVGQKPSDRPYIWNENEQGLDEHIARISADIAAGGVIVQAV
QDILPNLH (SEQ ID NO:20)

[0812] AvPAL变体1010号的多核苷酸序列 (SEQ ID NO:21) :

ATGAAAACCCTGAGTCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGGCAATA
GCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCGCGTGTGGCCGT
AATGGCACCCTGGTTAGCCTGACCAATAATAACCGATATTCTGCAGGGTATTACAGGCCAGCTGTGA
TTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTGTTACCAGCGGTTTTGGTGGTA
TGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGTTTGGTTTCTG
AAAACCGGTGCAGGTAATAAACTGCCGCTGGCAGATGTTTCGTGCAGCAATGCTGCTGCGTGCAA
ATAGCCACATGCATGGTGCAAGCGGTATTCGTCTGGAACCTGATTAACGCGCGGAAATCTTTCTG
AATGCCGGTGTTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCC
GCTGAGCTATATTACCGGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCA
AAGAAATGGACGCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCGTGCTGCCGAA
AGAAGGTCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTTGTGTTTAT
GATACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTACAGGCACTGAATGG
TACAAATCAGAGCTTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTGAGCTGTGGGCAG
[0813] CAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACTGGATGGTAAACA
TGATTATCGTGATGGTGAACGATCCAGGATCGTTATAGCCTGCGTTGTCTGCCGCAGTATCTGG
GTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAATTGAGATTAAACAGCGTTACC
GATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCATGGTGGAATTTTCTGGGTGAGTA
TGTTGGTATGGGTATGGATCATCTGCGCTATGAAATCGGTCTGCTGGCAAAACATCTGGATACCC
AGATTGCACTGCTGGCATCACCGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGTGGGTAATCGT
GAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCT
GACCTTTTATGGTAATAGTATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATA
TTAACAGCCAGGGTTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTGGCCAGAATTAT
GTTGCCATTGCCCTGATGTTTGGTGTTTCAGGCAGTTGATCTGCGTACCTACAAAAAAACCGGTCA
TTATGATGCACGTGCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTTCATGTTG
TTGGTCAGAAACCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGGATGA
ACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTTATTGTTTCAGGCCGTTTCAGGACA
TTCTGCCGAACCTGCAT (SEQ ID NO: 21)

[0814] AvPAL变体1010号的多肽序列 (SEQ ID NO:22) :

MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGI
QASCDYINNAVESGEPIYGVTSFGFGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVR
AAMLLRANSHMHGASGIRLELIKRAEIFLNAGVTPYVYEFSGSIGAGDLVPLSYITGSLIGLDP
SFKVDFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCYVDTQILTAIAM
GVHALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLAGSQLVRDELDDGKHDIRDGE
[0815] LIQDRYSLRCLPQYLGPVVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGM
DHLRYEIGLLAKHLDLQIALLASPEFSNGLPPSLVGNRERKVNMLKGLQICGNSIMPLLTIFY
GNSIADRFPTHAEQFNQININSQGYTSATLARRSVDIGQNYVAIALMFGVQAVDLRTYKKTGH
YDARAQLSPATERLYSAVRHVVGQKPSSDRPYIWNQGLDE
HIARISADIAAGGVIVQAVQDILPNLH (SEQ ID NO:22)

[0816] AvPAL变体1084号的多核苷酸序列 (SEQ ID NO:23) :

ATGAAAACCCTGAGTCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCCATACCGGCAAT
AGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGTACGTGTTGCCCCG
TAATGGCACCAGCGGTTAGCCTGACCAATAATAAAGATATTCTGCAGCGTATTACAGGCCAGCTGT
[0817] GATTATATCAATAATGCAGTTGAAAAAGGTGAACCGATTTATGGTGTTACCAGCGGTTTTGGTGG
TATGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAACTGCAGACCAATCTGGTTTGGTTTC
TGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCAGATGTTTCGTGCAGCAATGCTGCTGCGTGC
AAATAGCCACATGCGTGGTGCAAGCGGTATTCGTCTGGAACCTGATTAACGCATGGAATCTTT

- [0818] CTGAATGCCGGTGTTACCCCGTATGTTTATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGT
TCCGCTGAGCTATATTACCGGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATG
GCAAAGAAATGGACGCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCAGCC
GAAAGAAGGTCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTT
TATGATACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGGCACTGAA
TGGTACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTCAGCTGTGGG
CAGCAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACTGGATGGTAA
ACATGATTATATGGATGGTGAACCTGATCCAGGATCGTTATAGCCTGCGTTGTCTGCCGCAGTATC
TGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAATTGAGATTAACAGCGTT
ACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCATGGTGGTAATTTTCTGGGTCA
GTATGTTGGTATGGGTATGGATCATCTGCGCTATTATATCGGTCTGCTGGCAAAACATCTGGATG
TTCAGATTGCACTGCTGGCATCACCGGAATTTAGCAATGGTCTGCCTCCGAGTCTGGTGGGTAA
CGTGAACGTAAAGTTAATATGGGTCTGAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCT
GCTGACCTTTTATGGTAATAGTATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGA
ATATTAACAGCCAGGGTTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTTTTCAGAAT
TATGTTGCCATTGCCCTGATGTTTGGTGTTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGG
TCATTATGATGCACGTGCCAGCTGTCACCGGAACCGAACGTCTGTATAGCGCAGTTTCGTCATG
TTGTTGGTAAAAAACCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGGA
TGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTTCAGGCCGTTTCAGG
ACATTCTGCCGCCGCTGCAT
- [0819] AvPAL变体1084号的多肽序列 (SEQ ID NO:24) :
MKTLSQAQSKTSSQFQSHGTGNSSANVIIGNQKLITINDVVRVARNGTAVSLTNNKDILQRI
QASCDYINNAVEKGEPIYGVTSFGGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADV
RAAMLLRANSHMRGASGIRLELIKRMIFLNAGVTPYVYEFSGISASGDLVPLSYITGSLIGLD
PSFKVDFNGKEMDAPTALRQLNLSPLTLQPKGLAMMNGTSVMTGIAANCVYDTQILTAIA
MGVHALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLAGSQLVRDELDTGKHDMYMD
[0820] GELIQDRYSLRCLPQYLGPVVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGM
GMDHLRYIYIGLLAKHLDVQIALLASPEFSNGLPPSLVGNRERKVNMGKGLQICGNSIMPLL
TFYGNISIDRFPTHAEQFNQNSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTYKK
TGHYDARAQLSPATERLYSAVRHVVGKKPSSDRPYIWNDEQGLDE
HIARISADIAAGGVIVQAVQDILPLH (SEQ ID NO:24)
- [0821] AvPAL变体967号的多核苷酸序列 (SEQ ID NO:25) :
ATGAAAACCCCTGAGTCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTACCGG
CAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTGCGC
GTGTTGCCCGTAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGGGTA
TTCAGGCCAGCTGTGATTATATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATGGTG
TTACCAGCGGTTTTGGTGGTATGGCAAATGTTGTAATTAGCCGTGAACAGGCAAGCGAAC
TGCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGGCAG
ATGTTTCGTGCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGTATTC
[0822] GTCTGGAACCTGATTAAACGCGCGGAAATCTTTCTGAATGCCGGTGTTACCCCGTATGTTT
ATGAATTTGGTAGCATTGGTGCCAGCGGTGATCTGGTTCCGCTGAGCTATATTACCGGTA
GCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGACGCAC
CGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAAAGAAGGTCTG
GCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGATACC
CAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTCAGGCACTGAATGGT
ACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTCAGCTGTGG
GCAGCAGATCAGATGATTAGCCTGCTGGCCGGTAGCCAGCTGGTTCGTGATGAACTGGA

- [0823] TGGTAAACATGATTATCGTGATGGTGAACCTGATCCAGGATCGTTATAGCCTGCGTTGTCT
GCCGCAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCGAAAT
TGAGATTAAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCTATCA
TGGTGGTAATTTTCTGGGTCACTATGTTGGTATGGGTATGGATCATCTGCGCTATTATATC
GGTGGCCTGGCAAAACATCTGGATGTTCAAGATTGCACTGCTGGCATCACCGGAATTTAGC
AATGGTCTGCCTCCGAGTCTGGTGGGTAATCGTGAACGTAAAGTTAATATGGGTCTGAAA
GGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAGTATT
GCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATATTAACAGCCAGGGTTAT
ACCAGCGCAACCTTGGCACGTCGTAGCGTTGATATTGGCCAGAATTATGTTGCCATTGCC
CTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCATTATGAT
GCACGTGCCCAGCTGTCACCGGCAACCGAACGTCTGTATAGCGCAGTTCGTCATGTTGTT
GGTCAGAAACCGAGCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTGGA
TGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTGTATTGTTTCAGGCCGT
TCAGGACATTCTGCCGAACCTGCAT (SEQ ID NO:25)
- [0824] 变体967号的多肽序列 (SEQ ID NO:26) :
MKTLSQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVSLTNNTDILQGIQAS
CDYINNAVESGEPIYGVTSGFGGMANVVISREQASELQTNLVWFLKTGAGNKLPLADVRAA
MLLRANSHMRGASGIRLELIKRAEIFLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDPSFK
VDFNGKEMDAPTALRQLNLSPLTLLPKEGLAMMNGTSVMTGIAANCVYDTQILTAIAMGVH
ALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLAGSQLVRDELKGKHDYRDGELIQD
RYSRLCLPQYLGPIVDGISQIAKQIEIEINSVTDNPLIDVDNQASYHGGNFLGQYVGMGMDHL
RYYIGGLAKHLDVQIALLASPEFSNGLPPSLVGNRERKVNMGKGLQICGNSIMPLLTIFYGNS
IADRFPTHAEQFNQNSQGYTSATLARRSVDIGQNYVAIALMFGVQAVDLRTRYKKTGHYDA
RAQLSPATERLYSAVRHVVGQKPSSDRPYIWNENEQGLDE
HIARISADIAAGGVIVQAVQDILPNLH (SEQ ID NO:26)
- [0826] 表达载体pCK100900i:
TGGCCACCATCACCATCACCATTAGGGAAGAGCAGATGGGCAAGCTTGACCTGTGAAGT
GAAAAATGGCGCACATTGTGCGACATTTTTTTTTTGAATTCTACGTAAAAAGCAGCCGATA
CATCGGCTGCTTTTTTTTTTGNNGAGGTTCCAACCTTGTGGTATAATGAAATAAGATCACT
CCGGAGCGTATTTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAGGAACTAAAATGGA
GAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTT
TGAGGCATTTCACTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTTCAGCTGGATATTAC
GGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTATTCACAT
TCTTGCCCGCCTGATGAATGCTCATCCGGAGTTCCGTATGGCAATGAAAGACGGTGAGCT
GGTGATATGGGATAGTGTTACCCCTTGTTACACCGTTTTCCATGAGCAAACCTGAAACGTT
TTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTGCA
AGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATAT
GTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAAGTTTGAATTTAAACGTGGCCAAT
ATGGACAACCTTCTTCGCCCCGTTTTTACCATGGGCAAATATTATACGCAAGGCGACAAG
GTGCTGATGCCGCTGGCGATTTCAGGTTTCATCATGCCGTCTGTGATGGCTTCCATGTCCGC
AGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAACCTGCA
GGAGCTCAAACAGCAGCCTGTATTTCAGGCTGCTTTTTTCGTTTTGGTCTGCGCGTAATCTC
- [0827]

[0828]

TTGCTCTGAAAACGAAAAACCGCCTTGAGGGCGGTTTTTCGAAGGTTCTCTGAGCTAC
CAACTCTTTGAACCGAGGTAAGTGGCTTGGAGGAGCGCAGTCACCAAACTTGTCCTTTC
AGTTTAGCCTTAACCGGCGCATGACTTCAAGACTAACTCCTCTAAATCAATTACCAGTGG
CTGCTGCCAGTGGTGCTTTTGCATGTCTTCCGGGTTGGACTCAAGACGATAGTTACCGG
ATAAGGCGCAGCGGTGCGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGA
ACTGCCTACCCGGAAGTGAAGTGTGAGGCGTGAATGAGACAAACGCGGCCATAACAGCG
GAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCGCCAGG
GGGAAACGCCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCACTGATTTGAGCGTCA
GATTTTCGTGATGCTTGTGAGGGGGGCGGAGCCTATGGAAAAACGGCTTTGCCGCGGCCCT
CTCACTTCCCTGTAAAGTATCTTCTGCGCATCTTCCAGGAAATCTCCGCCCCGTTCTGTAAG
CCATTTCCGCTCGCCGCGAGTCGAACGACCGAGCGTAGCGAGTCAGTGAGCGAGGAAGCG
GAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCAGCCTTTTTCTCCTGCCAC
ATGAAGCACTTCACTGACACCTCATCAGTGAACCACCGCTGGTAGCGGTGGTTTTTTTA
GGCCTATGGCCTTTTTTTTTNTGNNAACCTTTCGCGGTATGGNATNANAGCGCCCGGA
AGAGAGTCAATTAAGAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTGCGAGA
GTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTC
TGCGAAAACGCGGGAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACC
GCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTC
TGGCCCTGCACGCGCCGTCGCAAATTGTGCGGCGGATTAAATCTCGCGCCGATCAACTGG
GTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCG
GTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGAC
CAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCT
CTGACCAGACCCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGGCG
TGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTT
CTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTC
AGCCGATAGCGGAACGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATG
CAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGC
GCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGACATCTCGGT
AGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCA
AACAGGATTTTCGCTGCTGGGGCAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGG
GCCAGGCGGTGAAGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAAGAAAAACCACC
CTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTG
GCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGGTACCCGATAAAAGCGGCTT
CCTGACAGGAGGCCGTTTTGTTTCTCGAGTTAATTAAGGCAGTGAGCGCAACGCAATTAA
TGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATG
TTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTAC
GGATTCACTGGCCGTCGTTTTACAATCTAGAGGCCAGCCTGGCCATAAGGAGATATACAT
ATGGGCCATCATCATCATCATCATCATCACAGCAGCGGCCATATCGAAGGTGCT
CATATGAAAACCTGAGCCAGGCACAGAGCAAAACCAGCAGCCAGCAGTTTAGCTTTAC
CGGCAATAGCAGCGCAAATGTGATTATTGGTAATCAGAACTGACCATCAATGATGTTG
CACGTGTTGCCCATAATGGCACCCCTGGTTAGCCTGACCAATAATACCGATATTCTGCAGG
GTATTACAGGCCAGCTGTGATTATCAATAATGCAGTTGAAAGCGGTGAACCGATTTATG
GTGTTACCAGCGGTTTTGGTGGTATGGCAAATGTTGCAATTAGCCGTGAACAGGCAAGCG
AACTGCAGACCAATCTGGTTTGGTTTCTGAAAACCGGTGCAGGTAATAAACTGCCGCTGG
CAGATGTTCTGTGCAGCAATGCTGCTGCGTGCAAATAGCCACATGCGTGGTGCAAGCGGT
ATTCGTCTGGAAGTGAATTAACGCATGGAAATCTTCTGAATGCCGCTGTTACCCGCTAT
GTTTATGAATTTAGTGCATTTGGTGCCAGCGGTGATCTGGTTCCGCTGAGCTATATTACC
GGTAGCCTGATTGGCCTGGACCCGAGCTTTAAAGTTGATTTTAATGGCAAAGAAATGGAC
GCACCGACCGCACTGCGTCAGCTGAATCTGAGTCCGCTGACCCTGCTGCCGAAAGAAGG
TCTGGCAATGATGAATGGCACCAGCGTTATGACCGGTATTGCAGCAAATTGTGTTTATGA
TACCCAGATTCTGACCGCAATTGCAATGGGTGTTTCATGCACTGGATATTACAGGCACTGAA
TGGTACAAATCAGAGCTTTCATCCGTTTATCCATAACAGCAAACCGCATCCGGGTCAGCT
GTGGGCAGCAGATCAGATGATTAGCCTGCTGGCCAATAGCCAGCTGGTTCGTGATGAAC
TGGATGGTAAACATGATTATCGTGATCATGAACTGATCCAGGATCGTTATAGCCTGCGTT
GTCTGCCGCGAGTATCTGGGTCCGATTGTTGATGGTATTAGCCAGATTGCCAAACAAATCG

[0829] AAATTGAGATTAACAGCGTTACCGATAACCCGCTGATTGATGTTGATAATCAGGCAAGCT
ATCATGGTGGTAATTTTCTGGGTCAGTATGTTGGTATGGGTATGGATCATCTGCGCTATTA
TATCGGTCTGCTGGCAAAACATCTGGATGTTTCAAGTTCGCTGGCATCACCGGAATT
TAGCAATGGTCTGCCTCCGAGTCTGCTGGGTAATCGTGAACGTAAAGTTAATATGGGTCT
GAAAGGTCTGCAGATTTGCGGTAATAGCATTATGCCGCTGCTGACCTTTTATGGTAATAG
TATTGCAGATCGTTTTCCGACCCATGCCGAACAGTTTAACCAGAATATTAACAGCCAGGG
TTATACCAGCGCAACCCTGGCACGTCGTAGCGTTGATATTTTTCAGAATTATGTTGCCATT
GCCCTGATGTTTGGTGTTCAGGCAGTTGATCTGCGTACCTACAAAAAACCGGTCATTAT
GATGCACGTGCCTGTCTGTACCCGGCAACCGAACGTCTGTATAGCGCAGTTCGTCATGTT
GTTGGTCAGAAACCGACCTCAGATCGTCCGTATATTTGGAATGATAATGAACAGGGTCTG
GATGAACATATTGCACGTATTAGTGCAGATATTGCAGCCGGTGGTATTGTTTTCAGGCC
GTTTACAGGACATTCTGCCGTGTCTGCATTAAGGCCAAAC (SEQ ID NO:27)

[0830] 实施例1

[0831] PAL基因获得和表达载体的构建

[0832] 获得多变鱼腥藻苯丙氨酸解氨酶 (AvPAL) 质粒DNA, 并将编码AvPAL的合成基因密码子优化用于在大肠杆菌中表达, 并克隆进大肠杆菌表达载体pET16b中以提供pET16b-AvPAL (SEQ ID NO:1)。AvPAL开放阅读框 (SEQ ID NO:2) 使用寡核苷酸: PAL-pCK-F和PAL-pCK-R通过PCR来扩增, 并被亚克隆进表达载体pCK100900i (SEQ ID NO:27) 中。

引物	序列 5'至 3'	SEQ ID NO:
[0833] PAL-pCK-F	CTAGAGGCCAGCCTGGCCATAAGGAGATATACAT	SEQ ID NO:28
	ATGAAAACCTGAGCCAGGCAC	
PAL-pCK-R	GATGGTGATGGTGGCCAGTTTGGCCTTAATGCAG ACACGGCAGAATG	SEQ ID NO:29

[0834] 将该质粒构建体转化进源自W3110的大肠杆菌菌株中。使用本领域技术人员通常已知的定向进化技术从该质粒构建体产生基因变体文库 (参见例如, 美国专利号8,383,346和W02010/144103)。

[0835] 实施例2

[0836] 高通量 (HTP) 生长和测定

[0837] PAL和PAL变体的高通量 (HTP) 生长

[0838] 将转化的大肠杆菌细胞通过平铺在包含1%葡萄糖和30μg/ml氯霉素的LB琼脂平板上来选择。在37℃下孵育过夜后, 将菌落放入NUNC™ (Thermo-Scientific) 的96孔浅平底板的孔中, 该96孔浅平板填充有180μl/孔的补充有1%葡萄糖和30μg/ml氯霉素的LB。将培养物在摇床中过夜生长18-20小时 (200rpm, 30℃和85%相对湿度; Kuhner)。将过夜生长的样品 (20μL) 转移到填充有380μL的补充有30μg/ml氯霉素的Terrific肉汤的Costar 96孔深板中。将板在摇床中孵育135分钟 (250rpm, 30℃, 和85%相对湿度; Kuhner)。然后将细胞用40μL无菌水中的10mM IPTG诱导, 并在摇床中过夜孵育20-24小时 (250rpm, 30℃, 和85%相对湿度; Kuhner)。将两个重复培养物合并, 使细胞沉淀 (4000rpm x 20min), 将上清液弃去, 并在分析之前将细胞冷冻在-80℃。

[0839] HTP沉淀物的裂解

[0840] 首先, 将500μL裂解缓冲液 (20mM Tris pH 7.5, 1mM MgSO₄, 1mg/ml溶菌酶和0.5mg/ml多粘菌素B硫酸盐) 加入到细胞沉淀物。将混合物在室温下搅拌1.5h, 并沉淀 (4000rpm x 5min), 然后在本文所描述的多种HTP测定中使用澄清的裂解物。这些裂解物通过SDS-PAGE的分析显示, 过表达的蛋白以~60kDa的表观MW存在, 与PAL的期望的MW一致。

[0841] 澄清的裂解物的分析

[0842] PAL变体活性通过测量肉桂酸的形成来确定,如通过290nm处吸光度随时间的变化确定的。对于该测定,将100μL 200mM Tris/50mM苯丙氨酸pH 7.5,或200mM磷酸钠/50mM苯丙氨酸pH 7.0,80μL水,和20μL澄清的裂解物加入到聚-丙烯酸酯96孔板(Costar#3635, Corning)的孔中。将反应物短暂混合,并且活性使用 **SpectraMax[®] Plus³⁸⁴** 或 **SpectraMax[®] 190** (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。

[0843] 用蛋白酶预处理的澄清的裂解物的HTP-分析

[0844] PAL变体用胰凝乳蛋白酶和胰蛋白酶来刺激,以模拟下肠道(lower intestine)环境。首先,将30μL蛋白酶混合物(0.01-100mg/ml胰凝乳蛋白酶(C4129Sigma Aldrich)、0.01-100mg/ml胰蛋白酶(T7409Sigma Aldrich)、1mM CaCl₂和1mM HCl)、0-30μL的在500mM磷酸钠pH 7.0中的20mM牛磺胆酸钠、和90-120μL澄清的裂解物加入到96孔圆底板(Costar#3798, Corning)的孔。将板密封,并在37℃、400rpm、1" throw下孵育1h,然后分析。对于该测定,将100μL的200mM Tris/50mM苯丙氨酸pH 7.5或200mM磷酸钠/50mM苯丙氨酸pH 7.0和100μL蛋白酶处理的裂解物加入到聚-丙烯酸酯96孔板(Costar#3635, Corning)的孔。将该反应物短暂混合,并且活性使用 **SpectraMax[®] Plus³⁸⁴** 或 **SpectraMax[®] 190** (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。结果在下表中提供。

[0845] 用酸处理的澄清的裂解物的HTP-分析

[0846] 在该测定中, PAL变体在酸性条件下被刺激,以模拟胃环境。首先,将20μL的1M柠檬酸钠(pH 3.7-4.5)和30μL水或50μL 400mM柠檬酸钠pH 3.7-4.8,和50μL澄清的裂解物加入到96孔圆底板(Costar#3798, Corning)的孔。将板密封,并在37℃、400rpm、1" throw下孵育1h,然后分析。对于该测定,将100μL 200mM Tris、50mM苯丙氨酸pH 7.5和80μL 1M Tris pH 7.5,或200mM磷酸钠/50mM苯丙氨酸pH 7.0和80μL 1.0M磷酸钠pH 7.0,和20μL酸处理的裂解物加入到聚-丙烯酸酯96孔板(Costar#3635, Corning)的孔。将该反应物短暂混合,并且活性使用 **SpectraMax[®] Plus³⁸⁴** 或 **SpectraMax[®] 190** (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。结果在下表中提供。

[0847] 用胃蛋白酶预处理的澄清的裂解物的HTP分析

[0848] 在另一测定中, PAL变体用酸性条件和胃蛋白酶来刺激,以在模拟胃环境的条件下进一步测试变体。首先,将50μL在400mM柠檬酸钠pH 1.5-4中的0.01-100mg/ml胃蛋白酶,和50μL澄清的裂解物加入到96孔圆底板(Costar#3798, Corning)的孔。将板密封,并在37℃、400rpm、1" throw下孵育1-12h,然后分析。对于该测定,将100μL 200mM Tris/50mM苯丙氨酸pH 7.5和80μL 1M Tris pH 7.5,或200mM磷酸钠/50mM苯丙氨酸pH 7.0和20μL酸处理的裂解物加入到聚-丙烯酸酯96孔板(Costar#3635, Corning)的孔。将该反应物短暂混合,并且活性使用 **SpectraMax[®] Plus³⁸⁴** 或 **SpectraMax[®] 190** (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。

[0849]

表2-1. 当未刺激(U-C)、蛋白酶-刺激(P-C)和 酸-刺激(A-C)时的PAL变体的相对活性。				
变体号	U-C	P-C	A-C	相对于SEQ ID NO:4的 氨基酸差异
1	+	+	+	A39V/A91V/Y158H/S180A/N290G/A394V/S399N/ N474Q/K522Y/T524S
2	+	+	++	A39V/A91V/Y158H/A256G/A394V/P404A/N474Q
3	+	++	++	A39V/A91V/S180A/A394V/K522F/T524S
4	+	+	++	A39V/A91V/Y158H/T243I/A256G/S399N/P404A/L407 V/N474Q
5	+	+	++	E308Q
6	+	++	+	N400K
7	+	+	+	L364Q
8	+	+	++	A256G/N290G/P404A/L407V/N474Q/K522F
9	+	+	++	Y158H/S180A/A394V/T463N/N474Q/T524S

[0850]

表2-1. 当未刺激(U-C)、蛋白酶-刺激(P-C)和 酸-刺激(A-C)时的PAL变体的相对活性。				
变体号	U-C	P-C	A-C	相对于SEQ ID NO:4的 氨基酸差异
10	+	+	++	A39V/A91V/Y158H/S180A/K195E/A256G/S399N/ L407V/Y459F/T463N/K522N/T524S
11	+	+	++	A39V/A91V/Y158H/S180A/K195E/A256G/ N290G/S399N/Y459F/T463N
12	+	++	++	A39V/A91V/Y158H/S180A/K195E/T243I/A394V
13	+	++	++	A39V/A91V/Y158H/K195E/T243I/A256G/A394V/ S399N/N474Q/K522Y/T524S
14	+	++	++	A39V/A91V/Y158H/S180A/K195E/I245L/S399N/ L407V/Y459F/T463N/N474Q/T524S
15	+	++	+	H307M
16	+	+	++	A39V/A91V/Y158H/S180A/T243L/A394V/S399N/ T463N/K522F
17	+	+	++	A91V/N474Q
18	++	+	++	S180A/K195E
19	++	+	++	A91V/Y158H/T243I/A256G/N290G/S399N/L407V/ Y459F/ T463N/N474Q/K522N/T524S
20	++	++	++	A39V/Y158H/S180A/S399N
21	++	++	++	A91V/N270K
22	++	+++	++	H307G
23	++	++	+	N453G
24	++	++	++	H307Q
25	++	++	+	L257W
26	++	+	+++	A91V/Y158H/I245L/A256G/S399N/Y459F/T463N/ K522Y/T524S
27	++	++	+	F443H
28	++	++	++	A91V/Y158H/S180A
29	++	++	+	A91V/ Y158H/ K195E/ S399N
30	++	++	+++	A39V/ A91V/ A256G/ N290G/ A394V/ S399N/ P404A/ L407V/K522Y/ T524S
31	++	++	+	L257A
32	++	+	+++	Y158H/K195E/T243L/A256G/A394V/S399N/ L407V/N474Q/ K522F/T524S
33	++	++		L349M
34	+++	+++	+	I326F
35	+++	+++	+++	T460G/ P528L

[0851] 1. 相对活性被计算为变体的活性/SEQ ID NO:4 (由SEQ ID NO:3编码) 的活性。

[0852] 2. 变体22号具有SEQ ID NO:7的多核苷酸序列和SEQ ID NO:8的多肽序列, 且变体

30号具有SEQ ID NO:5的多核苷酸序列和SEQ ID NO:6的多肽序列。

[0853] 3.+ = 0.1至1.0的相对于野生型AvPAL的相对活性；

[0854] ++ = >1.0至2.0的相对于野生型AvPAL的相对活性；和

[0855] +++ = >2.0的相对于野生型AvPAL的相对活性。

[0856]

表2-2. 未刺激(U-C)、蛋白酶-刺激(P-C)和 酸-刺激(A-C)的PAL变体的相对活性。				
变体号	U-C	P-C	A-C	相对于野生型AvPA L(SEQ ID NO:4)的 氨基酸差异
36	++	+++	++	A39V/A91V/N290G/H307G/L407V/ T524S
37	++	+++	++	A39V/A91V/N290G/H307G/L407V
38	+	++	+	A39V/A91V/A256G/N290G/H307G/S399N/P404A/ L407V/K522Y/T524S
39	++	+	++	A39V/A91V/A256G/N290G/S399N/P404A/L407V/K522 Y
40	+	++	+	A39V/A256G/N290G/H307G/P404A/L407V
41	+	++	+	A39V/A91V/A256G/N290G/H307G/P404A/L407V/T524S

[0857] 1. 相对活性被计算为变体的活性/变体30号的活性。

[0858] 2. 变体36号具有SEQ ID NO:9的多核苷酸序列和SEQ ID NO:10的多肽序列。

[0859] 3.+ = >1.0至3.0的相对于变体30号的相对活性；

[0860] ++ = >3.0至10的相对于变体30号的相对活性；和

[0861] +++ = >10至35的相对于变体30号的相对活性。

[0862]

表2-3. 未刺激(U-C)、蛋白酶-刺激(P-C)和 酸-刺激(A-C)的PAL变体的相对活性。				
变体号	U-C	A-C	P-C	突变(相对于变体36号)
42	+	+	+	T54K/G59R/S73K/R305M/C503Q/Q521K/C565P
43	+	+++	++	T54K/G59R/R305M/C503Q/Q521K/C565N
44	++	+++	++	G59R/C503Q
45	++	+++	++	K32P/G59R/S73K/Q240W/C503Q/C565N
46	+	+	+	K32P/G59R/S73K/Q240W/C565P
47	+	+++	++	K32P/T54K/S73K/R305M/C503Q/Q521K/C565N
48	+++	+++	++	Y304H/D353A
49	++	+++	+++	S73K/D353A
50	+	+	++	A112C/S546R
51	++	+	++	S73K/Q240W/Y304H
52	+	+	++	R134Q/Q240W/Y304H/D353A/E509L

[0863] 1. 相对活性被计算为变体的活性/变体36号的活性。

[0864] 2. 变体42号具有SEQ ID NO:11的多核苷酸序列和SEQ ID NO:12的多肽序列，变体43号具有SEQ ID NO:13的多核苷酸序列和SEQ ID NO:14的多肽序列。

[0865] 3.+ = >0.5至1.5的相对于变体36号的相对活性；

[0866] ++ = >1.5至3的相对于变体36号的相对活性；和

[0867] +++ = >3至10的相对于变体36号的相对活性。

表2-4.
未刺激(U-C)、蛋白酶-刺激(P-C)和
酸-刺激(A-C)的PAL变体的相对活性。

变体号	U-C	P-C	A-C	突变 (相对于变体30号 或野生型AvPAL)
53	+	+	+	D303R
54	+	+	++	E308A
55	+	+	+	E308D
56	++	++	+	G256A
57	+	++	+	H307A
58	+	+++	+	H307D
59	+	++	+	H307E
60	+	++	+	H307F
61	+	+++	+	H307G
62	+	+	+	H307I
63	+	+	+	H307L
64	++	+++	++	H307M
65	+	++	+	H307N
66	++	++	++	H307R
67	+	+	+	H307Y
68	+	+++	+	R305L
69	+	+++	+	R305M
70	+	+++	+	R305Q
71	++	++	+	V91A/G256A
72	+	++	+	Y304H
73	+	+	+	Y304W
74	+	+	+	C503K
75	+	+	+	C503Q
76	+	+	+	C565A
77	+	+	+	C565G
78	+	+	+	C565I
79	+	+	+	C565K
80	+	+	+	C565L
81	+	++	+	C565N
82	+	+	+	C565P
83	+	+	+	C565T

[0868]

[0869] 1. 相对活性被计算为变体的活性/变体30号(变体53-73)或SEQ ID NO:4(由SEQ ID NO:3编码)(变体74-83)的活性。

[0870] 2. +=>0.5至1.5的相对于变体30号的相对活性; ++=>1.5至3的相对于变体30号的相对活性; 和+++=>3至10的相对于变体30号的相对活性。

表2-5.
未刺激(U-C)、蛋白酶-刺激(P-C)和
酸-刺激(A-C)的PAL变体的相对活性。

变体号	U-C	P-C	A-C	相对于变体42的突变
42	+	+	+	
1011	+	+	+	R59G/R134Q/Q240R/K521Q/P564L
1012	+	+	++	R59G/P564M
1013	+	+	+	R59G/Q240W/E509L/K521Q/P564M
1014	++	++	+	Q240W/Y304H/D353N/E509L/K521S/P564L
1015	+	++	+	R59G/R134Q/Q240R/K521Q/P565N
1016	+	+++	+	Y304H/D353A/E509L
1017	+	++	+	R59G/R134Q/Y304H/D353N/K521Q
1018	+	+++	+	R59G/R134Q/Q240R/K521S
1019	+	+	+	R134Q/Y304H/D353A/K521S/P565N
1020	+	+	+	R59G/R134Q/Q240R/Y304H/D353A/K521Q/P564M
1021	++	+++	++	Q240W/D353A/E509L/K521Q
1022	+	++	+	R134Q/Q240W/E509L/K521S
1023	++	++	++	R134Q/Q240R/D353A/K521Q/P564G
1024	+	+	+	R134Q/K521Q
1025	+	+++	+	K521Q
1026	+	+++	+	R134Q/D353A/K521S/P564M
1027	+	+++	+	R59G/R134Q/Q240W/D353A
1028	++	++	+	R59G/K521S
1029	+	++	+	R59G/D353A
1030	+	+	+	Q240R/D353A/K521S
1031	+	+	+	R59G/D353A/E509L/P564L
1032	+	+	+	R59G/D353A/K521Q
1033	+	+	+	R59G/R134Q/D353A/E509L/K521Q
1034	+	+	+	Q240R/Y304H/E509L
1035	+	+	+	R59G/Q240R/D353A/E509L
1036	+	+	+	R59G/R134Q/E509L
1037	+	+	+	Q240W/Y304H/E509L/K521S
1038	+	+	+	Q240W/D353A/E509L/K521S/P564L/P565N
1039	+	+	+	R59G/Q240R/D353A/P564M/P565N
1040	+	+	+	R134Q/D353A/K521S
1041	+	+	+	R134Q/Q240W/Y304H/D353A/P564L
1042	+	+	+	Q240R/E509L/P565N

[0871]

表2-5.
未刺激(U-C)、蛋白酶-刺激(P-C)和
酸-刺激(A-C)的PAL变体的相对活性。

变体号	U-C	P-C	A-C	相对于变体42的突变
1043	+	+	+	Q240R/K521S/P564M/P565N
1044	+	+	+	R134Q/D353A/P564L
1045	+	+	+	R59G/R134Q
1046	+	+	+	R59G/Q240R
1047	+	+	+	Y304H/E509L
1048	+	+	+	R134Q/Q240R/D353A/E509L/K521Q/P564L
1049	+	+	+	Q240W/E509L/P564L/P565N
1050	+	+	+	R59G/E509L/P564L
1051	+	+	+	R134Q/D353A/E509L/P564L
1052	+	+	+	R59G/D353A/P565N
1053	+	+	+	Q240R/E509L
1054	+	+	+	Q240R
1055	+	+	+	D353A/K521Q
1056	+	+	+	Q240W/D353A/E509L
1057	+	+	+	F18H/L47A/L214Q/E540D
1058	+	+	+	F18H/L47A/L214Q/E308K/F450Y/S546R
1059	+	+	+	F18H/L47A/F450Y/P528L/S546R
1060	+	+	+	L214Q/E308Q/T460G
1061	+	+	+	F18H/F450Y/E540D
1062	+	+	+	P528L/S546R
1063	+	+	+	F18H/L214Q/F450Y
1064	+	+	+	E308Q/F450Y/R467G
1065	+	+	+	F18H/E540D
1066	+	+	+	L47A/L214Q/S546R
1067	+	+	+	F18H/L214Q/E308Q
1068	+	+	+	L47A/L214Q/E308Q
1069	+	+	+	F18H/L47E/L214Q/R467G/E540D/S546R
1070	+	+	+	L47A/F450Y
1071	+	+	+	L47A/L214Q
1072	+	+	+	F18H/L47A/L214Q/S546R
1073	+	+	+	F18H/L214Q
1074	+	+	+	L214Q/F450Y/P528L
1075	+	+	+	F18H/L214Q/E308Q/F450Y/R467G

[0872]

表2-5.
未刺激(U-C)、蛋白酶-刺激(P-C)和
酸-刺激(A-C)的PAL变体的相对活性。

[0873]

变体号	U-C	P-C	A-C	相对于变体42的突变
1076	+	+	+	L47A/L214Q/E540D
1077	+	+	+	L214Q/E540D
1078	+	+	+	F18H/L47A/E308Q/S546R
1079	+	+	+	F18H
1080	+	+	+	F18H/L47A/F450Y/S546R
1081	+	+	+	F18H/L47A/L214Q/F450Y
1082	+	+	+	F18H/L47A
1083	+	+	+	L47A
1084	+	+	+	F18H/L47A/L214Q
1085	+	+	+	L214Q/S546R
1086	+	+	+	F18H/L214Q/R467G/S546R
1087	+	+	+	F18H/L47A/L214Q/E308Q
1088	+	+	+	F18H/L214Q/T460G
1089	+	+	+	F18H/L47A/F450Y
1090	+	+	+	F18H/L214Q/E540D

[0874] 1. 相对活性被计算为变体的活性/变体42号的活性。

[0875] 2. -<0.5的相对于变体42号的相对活性; +=>0.5至1.5的相对于变体42号的相对活性; ++=>1.5至3的相对于变体42号的相对活性。

[0876] 实施例3

[0877] 确定PAL变体的蛋白聚集的测定

[0878] PAL变体聚集的倾向根据制造商的说明书使用 **ProteoStat[®]** 蛋白聚集测定试剂盒 (Enzo) 来确定。简要地, 将0-100μM的纯化的PAL与 **ProteoStat[®]** 检测试剂(1:2000)混合, 并通过流式细胞术分析。与 **ProteoStat[®]** 聚集标准品一致地评价样品荧光, 与如本领域已知的(参见例如, Bershtein等, Mol. Cell., 133-144[2013])。

[0879] 实施例4

[0880] 来自摇瓶(SF)培养物的冻干的裂解物

[0881] 将如以上所描述的的生长选择的HTP培养物平铺在具有1%葡萄糖和30μg/ml氯霉素的LB琼脂平板上, 并在37℃下生长过夜。将来自每个培养物的单菌落转移至50ml具有1%葡萄糖和30μg/ml氯霉素的LB。培养物在30℃、250rpm下生长18h, 并以大约1:10的稀释在250ml具有30μg/ml氯霉素的Terrific肉汤中亚培养, 至最终OD₆₀₀为0.2。将培养物在30℃、250rpm下孵育135分钟, 至OD₆₀₀为0.6, 并用1mM IPTG诱导。诱导的培养物在30℃、250rpm下孵育20h。在该孵育期之后, 将培养物以4000rpm x 10min离心。弃去上清液, 将沉淀物重悬于30ml的50mM pH 7.5的磷酸钠中。使细胞沉淀(4000rpm x 10min), 重悬于12ml 50mM pH 7.5的磷酸钠中, 并且使用One Shot Cell Disruption系统(Constant Systems)以17000psi裂解。使裂解物沉淀(10,000rpm x 30min), 并将上清液冷冻并干燥以产生含酶的

粉末。

[0882] 来自摇瓶培养物的PAL的纯化

[0883] 将PAL变体42号在摇瓶培养物中生长至饱和,如以上所述。将饱和的培养物通过离心(4000rpm×20min)沉淀,并在纯化之前将细胞沉淀物储存于-80℃。将细胞沉淀物在室温下解冻,并以5mL缓冲液/g细胞重悬于具有130mM NaCl pH 8的25mM Tris中。样品浆液使用具有110psi的压力设定的微流化器(microfluidizer)来裂解。将所得裂解物通过以10,000rpm离心1小时,随后通过0.2μm PES滤器(Millipore)过滤来澄清。

[0884] 过滤后,在10mM Phe的存在或不存在下,将所得裂解物在70℃-85℃下加热1.5-2小时。将裂解物从热源移开,并在4℃下以10,000rpm离心1小时来澄清。然后将包含可溶性PAL的上清液通过0.2μm PES滤器过滤,然后装载到层析柱中。

[0885] 将经热处理的包含80-100mg总蛋白的过滤的裂解物使用具有1.2M硫酸铵pH 8的25mM Tris稀释两倍。将样品装载到用具有0.6M硫酸铵pH 8的25mM Tris预平衡的HiPrep 16/10苯基FF(hi sub)柱(GE Healthcare)上。在样品装载之后,将柱用三倍柱体积的相同缓冲液来洗涤,随后用一个柱体积的在pH 8的25mM Tris中的0.6M-0M硫酸铵的线性梯度洗涤。将紧密结合的PAL使用三倍柱体积的具有25mM Tris,pH 8的等度洗脱从柱上洗脱。将包含活性和纯PAL的级分汇集。

[0886] 将来自苯基柱的纯化的PAL缓冲液交换入pH 8.5的0.5M Tris中,并浓缩。浓缩的PAL通过SDS-PAGE来分析,并发现以~60kDa条带存在。纯化的PAL样品使用0.45μm PES滤器来过滤,并储存在-80℃下直到准备使用。

[0887] 实施例5

[0888] 纯化的PAL和PAL变体的表征

[0889] 在本实施例中,描述了进行的旨在表征野生型和变体PAL的测定。

[0890] 对酸性pH的耐受性:

[0891] 将包含PAL变体的冻干的粉末以2g/L溶解于pH 7.0的20mM磷酸钠中。然后,将50μL酶溶液与50μL 400mM柠檬酸(pH 4.0-5.2)或100mM磷酸钠混合,并将反应物在37℃以400rpm(1" throw)孵育1h。然后,将20μL溶液与80μL 1M磷酸钠pH 7.0和100μL 200mM Tris/50mM苯丙氨酸pH 7.5短暂混合。在酸性条件下的酶活性使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices)光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。结果在图2中示出。如图2中所示,与野生型PAL相比,变体30号和36号在pH约4至4.8下孵育之后保持更多的活性。

[0892] 测定K_M:

[0893] 为了评价变体PAL中的突变是否已改变了PAL变体对苯丙氨酸的亲合力,测定了野生型酶和变体36的米氏常数。首先,将100μL的100mM Tris pH 8.0中的15μg/ml PAL和100μL的在100mM Tris,pH 8.0中的0-32mM苯丙氨酸加入到聚-丙烯酸酯96孔平板(Costar#3625,Corning)的孔。将该反应物短暂混合,并且初始速率使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices)光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。每个测试的PAL变体的K_M通过如本领域已知的

将数据拟合为Lineweaver-Burke图来确定。结果在图3中示出。如所示的,野生型酶的 K_M 为74 μ M,且变体36的 K_M 为60 μ M。

[0894] 氨基酸特异性:

[0895] 一些苯丙氨酸解氨酶除了苯丙氨酸以外显示针对酪氨酸和/或组氨酸的活性。为了评价存在于PAL变体中的突变是否已改变了PAL变体对苯丙氨酸的特异性,评价了野生型酶和变体36对这三种氨基酸的活性。首先,将100 μ L在10mM磷酸钠pH 7.0中5g/L的含PAL的冻干粉末和100 μ L在200mM磷酸钠pH 7.5中的50mM苯丙氨酸或组氨酸或2.5mM酪氨酸加入到聚-丙烯酸酯96孔板 (Costar#3635, Corning) 的孔。将该溶液短暂混合,并且初始速率使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。结果在图4中示出。如所示的,未观察到WT酶或变体36号对组氨酸或酪氨酸的可检测的活性,表明这些酶是苯丙氨酸特异性的。

[0896] 对猪和牛蛋白酶的耐受性:

[0897] 将如实施例4所描述制备的PAL变体样品以2g/L溶解于100mM磷酸钠pH 7.0中。将猪胰蛋白酶和牛胰凝乳蛋白酶(均为100mg)溶解于2ml 100mM磷酸钠pH 7.0中,并在100mM磷酸钠中连续11次稀释2倍。然后,将80 μ L PAL变体酶溶液与20 μ L胰蛋白酶和胰凝乳蛋白酶溶液混合。将反应混合物在37 $^{\circ}$ C下以400rpm (1" throw) 孵育1h。然后,将20 μ L反应物与80 μ L水和100 μ L 100mM磷酸钠、50mM苯丙氨酸pH 7.0混合。将每种溶液短暂混合,并且活性使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。结果在图2中示出。如该图中所示,所有测试的变体与野生型PAL相比显示出改进的蛋白酶耐受性,变体36号针对蛋白水解最稳定。

[0898] 对人蛋白酶的耐受性:

[0899] 如以上所述,针对猪胰蛋白酶和牛胰凝乳蛋白酶筛选了一些进化的PAL变体以评价它们对通过胃肠道中存在的酶的蛋白水解的耐受性。还使用人酶测试了一些进化的PAL变体,以确认它们耐受猪或牛酶的人同源物。在这些测定中,将WT PAL和变体36号的冻干的粉末(2.4g/L,在100mM磷酸钠pH 7.0中)与人类胰凝乳蛋白酶(Athens Research) 0-80BTEE单位/ml或人胰蛋白酶(ProSpec) (0-10,000BAEE单位/ml) 在37 $^{\circ}$ C下孵育2h。然后,将100 μ L混合物加入到聚-丙烯酸酯96孔板 (Costar#3635, Corning) 的孔,随后加入100 μ L 50mM苯丙氨酸、200mM磷酸钠pH 7.0。将溶液短暂混合,并且初始速率使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移(每12-20s一次,经过5-20min)追踪吸光度来确定。结果在图5中示出。如图5中所示,变体36号比野生型PAL酶更稳定。

[0900] 对胰腺粗提物的耐受性:

[0901] 还测试了进化的PAL变体以确定其对胰腺酶的耐受性。WT PAL、变体36号、变体42号和变体43号的冻干的粉末(如实施例4所描述的制备;12g/L,在50mM磷酸钾pH 6.8中)与猪胰酶制剂(pancreatin) (4x Sigma-Aldrich, St. Louis, MO) 以1:1混合,并在37 $^{\circ}$ C下振荡

(400rpm, 1" throw) 孵育多达23h。在指定的时间点, 将反应物的10 μ L等分试样加入到在聚-丙烯酸酯96孔板 (Costar#3635, Corning) 的孔中的190 μ L 50mM苯丙氨酸、190mM磷酸钠pH 7.0中。将反应物短暂混合, 并且初始速率使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移 (每12-20s一次, 经过5-20min) 追踪吸光度来确定。结果在图6中示出。如图6中所示, 在这些测定条件下, 与野生型PAL酶相比, 变体36号, 变体42号和变体43号都显示显著的稳定性。

[0902] 肠道去垢剂的影响:

[0903] 还测试了进化的PAL变体以确定它们在肠道胆汁酸和脂肪酸的存在或不存在下对蛋白水解的敏感性, 以评价这些酸是否影响它们的稳定性。将包含变体36号的冻干的粉末 (如实施例4中描述的) 以50 μ g/ml溶解于0-16mM牛磺胆酸钠、100mM磷酸钠pH 7.0中。将猪胰蛋白酶和牛胰凝乳蛋白酶 (各80mg) 溶解于2ml的100mM磷酸钠pH 7.0中, 并在100mM磷酸钠中连续11次稀释2倍。对于测定, 将50 μ L PAL溶液与50 μ L蛋白酶溶液混合。将混合物在37 $^{\circ}$ C下以400rpm (1" throw) 孵育1h。然后, 将50 μ L混合物与150 μ L 200mM磷酸钠、50mM苯丙氨酸pH 7.0混合。将每种反应物短暂混合, 并且活性使用SpectraMax[®] Plus³⁸⁴或SpectraMax[®] 190 (Molecular Devices) 光吸收酶标仪通过在290nm处随时间推移 (每12-20s一次, 经过5-20min) 追踪吸光度来确定。结果在图7中示出。如该图中所示的, 添加的牛磺胆酸钠增加了变体36号对蛋白水解的敏感性。

[0904] 实施例6

[0905] 变体PAL的肠道稳定性

[0906] 为了评价当其经过动物肠道时PAL变体的稳定性和活性, 用纯化的酶变体强饲小鼠。将10-12周鼠龄和称重20-26g的健康C57B1/6小鼠维持在代谢笼中, 并且禁食15h。随意提供水。禁食过夜后, 使用21-gauge强饲针用0.3ml的0.5M Tris-HCl pH 8.5和在0.5M Tris-HCl pH 8.5中的8mg/ml WT PAL (如实施例4中描述的制备) 或在0.5M Tris-HCl pH 8.5中的8mg/ml变体42号 (如实施例4中所描述的制备) 的混合物强饲动物。在强饲后0.5、2或6h, 将动物断头, 使用绿顶毛细血管血液收集管 (Ram Scientific) 收集血浆, 并收集胃、十二指肠 (距离胃~1-8cm)、空肠 (距离胃~10-18cm)、回肠 (盲肠以上~8cm) 和结肠 (盲肠以下~5cm) 的内容物。记录这些内容物的重量, 并在分析前将这些内容物储存在-80 $^{\circ}$ C。

[0907] 将胃或肠内容物用100mM磷酸钠pH 7.0稀释4X, 短暂混合, 并以14,000rpm x 2min离心。将上清液转移到350 μ L 0.45 μ M AcroPrep[™]高级96-孔过滤板 (Pall Corp), 并通过真空过滤去除颗粒。如前述实施例中所描述的评价澄清的滤液的酶活性并通过SDS-PAGE评价完整的PAL蛋白的存在。结果表明, 与野生型PAL酶和阴性对照相比, 进化的PAL变体的酶活性在十二指肠和空肠中显示是更高的。

[0908] 实施例7

[0909] 血浆苯丙氨酸的水平

[0910] 评价从实施例6中所描述的小鼠收集的血浆样品, 以确定所测试的小鼠的血液中的苯丙氨酸的量。将小鼠血浆 (50 μ L) 与250 μ L包含0.6mM d1-苯丙氨酸 (环D₅) (即, 包含与芳香族环碳原子键合的氘而非氢的同位素标记的版本的苯丙氨酸; Cambridge同位素实验室) 的乙腈合并。将样品在RT下混合5min, 在4 $^{\circ}$ C下以3200x g离心10min, 并将上清液转移

至板用于样品分析。对于分析,将10 μ L每个样品注入进3200 QTRAP[®] LC/MS/MS系统 (AB Sciex) 经过DISCOVERY[®]C₁₈柱 (150x 2.1mm, 5 μ m珠子) (Supelco, 现在Sigma-Aldrich), 用水中0.1%的甲酸 (A) 和乙腈 (B) 洗脱。跨5min梯度 (t=0, 97%A; 3min, 50%A; 3.5min, 5%A; 4min, 97%A; 5min, 95%A) 洗脱样品, 寻找内源性苯丙氨酸的166至120的转变, 和同位素标记的标准品的171至125的转变。结果表明, 与野生型PAL酶和阴性对照相比, 来自给予了进化的PAL变体 (即, 变体42号) 的小鼠样品中的血浆苯丙氨酸水平在30分钟时间-点时是较低的。

[0911] 实施例8

[0912] 变体PAL的治疗功能

[0913] 为了评价PAL变体是否降低体内的血清Phe水平, 使用了PKU的小鼠模型。在这些实验中, 将PAL蛋白强饲入受影响的动物。首先, 通过从小鼠的饮食中去除苯丙氨酸三天, 随后注射已知量的包含Phe的溶液来在小鼠中建立一致的基线Phe水平。将具有C57B1/6背景的3-至-6个月鼠龄的纯合PAH enu-2小鼠 (参见, McDonald等, Proc. Natl. Acad. Sci. USA 87: 1965-1967 [1990]) 转移至无苯丙氨酸的饮食 (TD. 97152, Harlan), 具有在其饮用水中提供的0.03g/L Phe持续72h。在开始处理之前在时间=0h时, 将小鼠注射0.15mg/g体重的Phe (来自在水中的10g/L的Phe溶液)。注射后55分钟, 通过尾部-静脉穿刺收集大约20 μ L血液, 并点制在滤纸上。随后, 在注射后时间1h、3h和5h时, 将小鼠用0.3ml 50-100g/L WT AvPAL、WT AvPAL加抑肽酶、BSA或变体42号强饲。在注射后6h、7h和9h时, 将另外的血液斑点收集在滤纸上。将血液斑点干燥并在使用本领域已知的方法进行Phe和Tyr水平的LC-MS/MS分析之前储存在-20 $^{\circ}$ C (参见, Chase等, Clin. Chem., 39: 66-71 [1993])。

[0914] 结果在图8中示出。如在该图中所示的, 用无活性蛋白 (BSA) 强饲, 导致增加的血清Phe水平。相比之下, 用蛋白水解敏感的WT AvPAL的处理导致恒定的Phe-水平, 而与蛋白酶抑制剂抑肽酶结合的相同蛋白导致Phe-水平的显著下降。该结果还表明, 在蛋白酶抑制剂不存在下施用工程化-PAL变体42号导致降低的血清Phe水平。

[0915] 实施例9

[0916] PAL的去免疫化

[0917] 在此实施例中, 描述了进行的旨在鉴定将T细胞表位从PAL中去除的多样性的实验。

[0918] 去免疫化多样性的鉴定:

[0919] 为了鉴定当突变时可去除T细胞表位的氨基酸, 使用计算机方法鉴定预测引发T细胞应答的PAL序列。并行地, 还进行了针对允许的、无害突变的实验探究, 特别是针对在未刺激的测定 (例如, 在实施例2中描述的测定) 中保持蛋白活性的氨基酸的实验探究。然后分析了活性变体以确定突变对预测的免疫原性的影响。

[0920] 在变体AvPAL中计算机鉴定推定的T细胞表位:

[0921] 使用如本领域已知的免疫表位数据库 (IEDB; 免疫表位数据库和分析资源网站 (Immune Epitope Database and Analysis Resource website)) 工具和专有的统计分析工具 (参见例如, iedb.org和Vita等, Nucl. Acids Res., 38 (Database issue): D854-62 [2010]. Epub 2009 Nov 11) 鉴定了在AvPAL变体36号中推定的T细胞表位。将该AvPAL变体

解析成所有可能的15-mer分析框架,与每个框架的最后14个氨基酸重叠。使用在IEDB网站上推荐的方法,通过针对与集体覆盖近95%的人类群体的八种常见的II类HLA-DR等位基因(DRB1*0101、DRB1*0301、DRB1*0401、DRB1*0701、DRB1*0801、DRB1*1101、DRB1*1301、和DRB1*1501)(参见例如, Southwood等, J. Immunol., 160:3363-3373[1998])的预测的结合对15-mer分析框架的9-mer核心区域进行评分,评价15-mer分析框架的免疫原性潜力。使用如本领域已知的统计分析工具鉴定了包含在变体AvPAL内的潜在的T-细胞表位簇(即,包含变体AvPAL内的具有异常高的免疫原性潜力的子区域)。将鉴定的T细胞表位簇针对已知表位的IEDB数据库和GenBank蛋白数据库进行筛选。这些筛选鉴定了在变体AvPAL变体36号中的10(十)个推定的T细胞表位。以下这些表位被称为TCE-I、II、III、IV、V、VI、VII、VIII、IX和X。

[0922] 去免疫化文库的设计:

[0923] 首先,创建了包含在如以上描述鉴定的10个推定的T细胞表位区域中的从用于创建变体AvPAL的定向进化轮鉴定出的所有中性和有益突变的组合文库。分析了这些突变对与八种常见的II类HLA-DR等位基因的预测结合的影响。预测多个突变去除或减少TCE-I、II、VI、VII。将这些突变结合到组合文库中。然后使用饱和诱变设计文库,以在剩余的六个T-细胞表位(即,TCE-III、IV、V、VIII、IX和X)内诱变每一个氨基酸。最后,创建了包含从靶向TCE-I、III、IV、VIII和X的多轮进化鉴定的有益多样性连同报道的影响PAL变体的聚集状态的C503和C565两个氨基酸的组合文库。使来自该文库的最佳击中经历靶向TCE-III和VIII的另外的饱和诱变和在几个位置处的另外的靶向的诱变。

[0924] 去免疫化文库的构建和筛选:

[0925] 如以上描述的设计的组合和饱和诱变文库通过本领域已知的方法来构建,并在如实施例2中所描述的未刺激的测定中测试活性。鉴定和测序活性变体。在以下表9-1至9-7中提供了关于AvPAL变体36号和许多AvPAL变体的活性和突变。

[0926] 去免疫化多样性的鉴定:

[0927] 通过评价活性变体与以上描述的八种常见的II类HLA-DR等位基因的结合分析了活性变体以确定它们的免疫原性水平。总免疫原性评分和免疫原性击中计数在表9-1至9-7中显示。总免疫原性评分反映了变体的整体预测的免疫原性(即,较高评分指示较高水平的预测的免疫原性)。免疫原性“击中计数”指示具有异常高的免疫原性潜力(即,较高的击中计数指示较高的免疫原性潜力)的15-mer分析框架的数目。具有较低的总的预测的免疫原性评分和/或免疫原性击中计数小于参考变体的免疫原性击中计数的突变被认为是“去免疫化突变”。将被鉴定为最佳的去免疫化突变重组以产生许多有活性的并预测比起始参考变体AvPAL具有显著更低的免疫原性的变体。在以下表中,FIOP结果来自未刺激的测定;对于总免疫原性评分(TIS)和免疫原性击中计数(IHC),结果针对整个PAL蛋白(表9-1、9-8和9-9)或针对所指定的表位(表9-2至9-7)显示。

[0928]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和免疫原性击中计数(IHC) (“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
变体36		+	680	52
100	I27E/L214E	+	634	51
101	V105C/R134Q/Q205T/P266H/L278D	+	645	48
102	I27E/L214E/C503Q/A547D	+	629	51
103	V105C/Q205T/P210C/L214E/C503Q/A547D	+	619	49
104	I27E/A112C/R134Q/Q205T/I285E/C503Q	+	622	45
105	V39A/P266H	+	665	45
106	I27E/V39A/V105C/R134Q/P210C	+	607	41
107	I27E/V39A/V105C/P210C/L214E/P266H/L278D	+	590	42
108	V39A/V105C/L214E/P266H/A547D/C565N	+	613	42
109	I27E/V39A/A112C/L214E	+	616	42

[0929]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和免疫原性击中计数(IHC) (“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
110	I27E/V39A/R134Q/A153G/L214E/P266H/I285E/C503Q/A551D/C565N	+	572	37
111	I27E/R134Q/L278D/I285E/A551D/C565N	+	626	48
112	I27E/V39A/V105C/A112C/Q205T/P266H/I285E/C503Q/A551D	+	594	39
113	I27E/V105C/L214E/P266H/C503Q	+	614	49
114	I27E/V39A/R134Q	+	637	44
115	C503Q/A547D	+	675	52
116	V105C/C503Q	+	660	50
117	I27E/R134Q/Q205T/P266H/L278D/A547D	+	641	50
118	I27E/V105C/R134Q/P210C/P266H/L278D/I285E/C503Q/A551D/C565N	+	596	45
119	I27E/V39A/V105C/R134Q/Q205T/L278D/I285E/C503Q/A547D/A551D/C565N	+	585	38
120	V105C/R134Q/L214E/C503Q/A547D/A551D	+	619	48
121	I27E/V39A/V105C/L214E/L278D/L309P/C503Q/A547D/A551D	+	587	42
122	V105C/L278D/C503Q/A551D	+	655	50
123	I27E/V105C/R134Q/P210C/L214E/P266H/L278D/A551D/C565N	++	591	48
124	I27E/R134Q/L214E/C503Q/A547D	++	620	50
125	R134Q/P210C/L214E/L278D/C503Q/A547D/C565N	+	630	50
126	I27E/V39A/V105C/Q205T/P210C/L214E/L278D/A547D	+	585	42
127	V39A/Q205T/L278D/A547D/A551D	+	654	44
128	V105C/R134Q/L214E/P266H/I285E/C503Q/A551D/C565N	+	598	45
129	I27E/V105C/L214E/A547D/A551D/C565N	+	609	49
130	I27E/V39A/V105C/Q205T/L278D/C503Q/A547D	+	615	42
131	I27E/V39A/V105C/R134Q/L214E/I285E/C503Q/A547D/A551D	+	564	38
132	C503Q/A547D/A551D/C565N	+	675	52
133	V105C/R134Q/L214E/C503Q/A547D	+	619	48
134	I27E/V105C/Q205T/C503Q/A547D/A551D/C565N	+	630	49
135	I27E/V39A	+	646	45
136	I27E/V39A/V105C/L214E/P266H/C503Q/A547D/C565N	+	594	42
137	A112C/R134Q/L214E/C503Q/A547D/A551D/C565N	+	636	48
138	I27E/A112C/R134Q/A153R/L214E/P266H/C503Q	+	619	47
139	C503Q/A551D	+	675	52
140	I27E/V39A/V105C/R134Q/L278D/C503Q/C565N	+	617	42
141	R134Q/Q205T/L214E/I285E/C503Q/A551D/C565N	+	613	47
142	V105C/L214E/I285E/A547D/C565N	+	607	46
143	I27E/V39A/R43L/L214E/A547D	+	625	45
144	I27E/V39A/P210C/L214E/I285E/C503Q/A551D	+	584	41
145	I27E/R134Q/L214E/C503Q/A547D/A551D	+	620	50
146	V39A/V105C/A153R/P266H/A547D/A551D	+	637	42
147	V39A/C503Q	+	665	45

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)
(“-“=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
148	I27E/V39A/V105C/P210C/I285E/C503Q/A547D/A551D/C565N	+	590	39
149	R134Q/L214E/L278D/C503Q/A551D	+	639	50
150	I27E/V39A/R134Q/C503Q/A547D	+	632	44
151	A153R	+	677	51
152	I27E/V39A/V105C/A112C/R134Q/L214E/L278D/C503Q/A547D/C565N	+	585	41
153	C503Q/A547D/C565N	+	675	52
154	I27E/V39A/L214E/P266H/L278D/C503Q/A547D/A551D/C565N	+	614	44
155	L278D/A547D	+	675	52
156	V39A/G45S/L278D/C503Q/A551D	+	667	45
157	V39A/A153R/C503Q/A547D	+	657	44
158	R134Q/P210C/L214E/C503Q/A547D/A551D	+	630	50
159	I27E/A547D/C565N	+	656	52
160	V39A/R134Q/P210C/L214E/A547D/C565N	+	615	43
161	I27E/V39A/P210C/P266H/I285E/C503Q/A547D	+	610	41
162	I27E/V39A/R134Q/L278D/C503Q/A547D	+	632	44
163	I27E/V105C/R134Q/Q205T/P210C/C503Q	+	622	48
164	I27E/V105C/R134Q/L214E/L278D/C503Q/A547D	+	600	48
165	I27E/V105C/C503Q/A547D/C565N	+	636	50
166	I27E/L214E/C503Q	+	634	51
167	V105C/L214E/L278D/C503Q/A547D/A551D	+	628	49
168	I27E/V105C/Q205T/L214E/P266H	+	609	49
169	V39A/A112C/R134Q/L214E/C503Q/A547D/A551D	+	621	41
170	I27E/R134Q/L214E/C503Q/C565N	+	625	50
171	I27E/V39A/R134Q/A153R/Q205T/L214E/P266H/C503Q	+	602	42
172	I27E/V39A/L214E/C503Q/A551D	+	614	44
173	V39A/V105C/Q205T/A551D	+	634	42
174	I27E/V39A/Q205T/C503Q/A547D/C565N	+	635	44
175	A547D	+	675	52
176	I27E/V39A/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N	+	576	41
177	V39A/P275R/L278D/C503Q/A551D	+	667	45
178	I27E/V39A/V105C/Q205T/P210C/L278D/A547D	+	611	42
179	V105C/A153R/Q205T/L214E/P266H/C503Q/A547D	+	620	48
180	V105C/A112C/R134Q/Q205T/L214E/Y492H/C503Q/A547D	+	614	48
181	I27E/P210C/L278D/C503Q	+	651	51
182	I27E/P210C/C503Q	+	651	51
183	I27E/V39A/R134Q/A153R/L214E/C503Q/A547D	+	602	42
184	I27E/P266H/A547D/A551D	+	656	52
185	V105C/L214E/I285E/C503Q/A547D/A551D/C565N	+	607	46
186	I27E/V105C/P266H/I285E/C503Q/A547D/C565N	+	615	47
187	Q205T/L278D/I285E/A547D/A551D	+	648	48

[0930]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)
("—"=FIOP<0.7;"+"=FIOP 0.7<FIOP<1.4;"++"=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
188	V39A/V105C/Q205T/C503Q	+	639	42
189	I27E/V39A/Q205T/P266H/I285E/A547D/A551D/C565N	+	614	41
190	V105C/L214E/I285E	+	612	46
191	V105C/R134Q/C503Q/A547D/C565N	+	646	49
192	I27E/V39A/V105C/R134Q/C503Q/A551D	+	612	42
193	I27E/R134Q/Q205T/I285E/C503Q/A551D	+	620	47
194	I27E/V39A/A112C/Q205T/L278D/I285E	+	616	39
195	I27E/V39A/A112C/Q205T/L214E/P266H/C503Q/A551D/C565N	+	606	42
196	I27E/V39A/L278D/A547D	+	641	45
197	I27E/V39A/V105C/R134Q/Q205T/L214E/A551D/C565N	+	580	41
198	I27E/G45D/Q205T/P266H/C565N	+	656	51
199	I27E/V39A/A112C/L214E/L278D/C503Q/A547D/A551D	+	611	42
200	I27E/Q205T	+	655	51
201	I27E/A112C/R134Q/Q205T/L278D/C503Q	+	643	48
202	C565N	+	680	52
203	I27E/V39A/V105C/R134Q/Q205T/P210C/L278D/C503Q/A547D	+	602	41
204	V105C/L214E/P266H/L278D/A547D	+	628	49
205	V105C/R134Q/A153R/Q205T/L214E/C503Q	+	616	47
206	V105C/R134Q/C503Q	+	651	49
207	I27E/V39A/Q205T/L278D/C503Q/C565N	+	640	44
208	I27E/V39A/V105C/S131N/R134Q/Q205T/L214E/C503Q/A547D/C565N	+	579	41
209	Q205T/L214E/I285E/C503Q/A551D	+	622	48
210	I27E/L214E/L278D/C503Q	+	634	51
211	V105C/R134Q/L214E/L278D/C565N	+	624	48
212	I27E/V39A/L214E	+	619	44
213	L214E/C503Q/A547D	+	648	51
214	L214E/P266H	+	653	51
215	I27E/V39A/C503Q	+	646	45
216	P210C/L214E	+	644	51
217	V105C	+	660	50
218	A112C/R134Q/A153R/L214E/L278D/I285E/C503Q/A547D/A551D/C565N	+	612	44
219	I27E/Q205T/L278D/A551D	+	650	51
220	I27E/L214E/A551D	+	629	51
221	R134Q/L214E/I285E/C503Q	++	623	47
222	I27E/V39A/V105C/C503Q/A551D	+	621	43
223	A112C/L278D/C503Q/A547D	+	672	50
224	I27E/R134Q/A153R/I285E/C503Q/A547D	++	623	47
225	I27E/V39A/V105C/A153R/I285E	+	602	39
226	I27E/V39A/R134Q/L278D/I285E/C503Q/A547D/A551D	++	611	41

[0931]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)
(“-“=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
227	R134Q/C503Q	+	671	51
228	V39A/C503Q/A551D/C565N	+	660	45
229	I27E/V39A/P266H	+	646	45
230	V105C/I285E	+	639	47
231	I27E/V39A/V105C/A153R/L214E	+	596	41
232	A112C/L214E/L278D	+	650	49
233	I27E/R134Q/P210C	+	642	50
234	I27E/A153R/L214E/L278D/A551D	+	626	50
235	V105C/I285E/A547D	+	634	47
236	I27E/Q205T/L214E/L278D/I285E/C503Q/C565N	+	608	48
237	I27E/V39A/P210C/T212S	+	636	44
238	I27E/P266H/L278D/C503Q	+	661	52
239	I27E/V105C/Q205T/L214E/P266H/A551D/C565N	+	604	49
240	I27E/A112C/P210C/L214E/C503Q/A547D	+	617	49
241	I27E/V39A/R134Q/P210C/C503Q/A551D	+	622	43
242	I27E/V39A/V105C/I285E/A547D	+	600	40
243	I27E/V39A/Q205T/L278D/C503Q/A551D/C565N	+	635	44
244	I27E/R134Q/I285E/C503Q	+	631	48
245	I27E/V39A/P210C/L214E/L278D/C503Q/A551D	+	605	44
246	V39A/A112C/R134Q/Q205T/L214E/L278D	+	621	41
247	I27E/V39A/L214E/L278D/C503Q	+	619	44
248	I27E/V105C/L214E	+	614	49
249	I27E/V39A/V105C/R134Q/P266H/C503Q/A547D/A551D	+	612	42
250	I27E/V39A/P210C/P266H/C503Q/A551D	+	631	44
251	I27E/V39A/V105C/A153R/C503Q/A547D/C565N	+	618	42
252	V39A/V105C/R134Q/L214E/C503Q/A547D/A551D	+	604	41
253	I27E/L214E/L278D	+	634	51
254	I27E/V39A/V105C/R134Q/A153R/P210C/L278D/I285E/C503Q/A547D/A551D	+	578	37
255	V105C/R134Q/Q205T/L214E/A547D	+	614	48
256	I27E/V105C/R134Q/A153R/C503Q	+	629	48
257	V105C/R134Q/C503Q/A547D	+	646	49
258	I27E/V39A/L278D/I285E/C503Q/A547D	+	620	42
259	I27E/V39A/R134Q/A153R/L278D	+	634	43
260	I27E/V39A/A112C/R134Q/L214E/C503Q/A547D	+	602	41
261	I27E/V39A/R134Q/A153R/L278D/A547D/A551D	+	629	43
262	I27E/V39A/L278D/C503Q/C565N	+	646	45
263	Q205T/L214E/C503Q/A547D/C565N	+	643	51
264	R134Q/L214E	+	644	50
265	I27E/V39A/P266H/L278D	+	646	45
266	R134Q/Q205T/C503Q	+	665	50

[0932]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)
(“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
267	V39A/V105C/R134Q/P210C/L214E/A551D	+	595	41
268	V105C/A153R/Q205T/P266H/I285E/A547D/C565N	+	625	45
269	I27E/V39A/R134Q/L214E/P266H/A551D	+	605	43
270	L214E/P266H/C503Q/A547D/A551D/C565N	+	648	51
271	V39A/V105C/P210C/A547D	+	630	42
272	I27E/V105C	+	641	50
273	I27E/V39A/R134Q/Q205T/L214E/P266H/I285E/C503Q/A551D/C565N	+	579	40
274	I27E/V39A/R134Q/Q205T/P266H/C503Q/A551D/C565N	+	626	43
275	I27E/V105C/R134Q/P210C/L214E/C503Q/A551D/C565N	+	591	48
276	I27E/L214E/C503Q/A547D/C565N	+	629	51
277	I27E/V39A/R134Q/A153R/P210C/L214E/L278D/I285E/A547D/C565N	+	572	39
278	I27E/V39A/A112C/I285E	+	622	40
279	I27E/A112C/R134Q/L278D/I285E/C503Q/A551D/C565N	+	623	46
280	I27E/L214E/P266H/L278D/I285E/A551D	+	608	48
281	I27E/V39A/V105C/A153R/A551D/C565N	+	618	42
282	V39A/V105C/R134Q/A153R/Q205T/A551D	+	622	40
283	I27E/V39A/V105C/R134Q/Q205T/L214E	+	585	41
284	I27E/V105C/Q205T	+	635	49
285	I27E/Q205T/P266H/L278D/I285E/A551D/C565N	+	629	48
286	I27E/R134Q/L214E/C503Q	+	625	50
287	V105C/Q205T/L214E/C503Q/A547D/A551D/C565N	+	623	49
288	V39A/V105C/L214E/I285E/C503Q/A551D/C565N	+	592	39
289	I27E/V39A/V105C/L278D/C503Q/A547D/C565N	+	621	43
290	I27E/V39A/V105C/R134Q/I285E/A547D/A551D	+	591	39
291	I27E/V39A/A112C/R134Q/L214E/P266H/A551D	+	602	41
292	I27E/L278D/C503Q/A551D	+	656	52
293	I27E/V39A/R134Q/A153R/A547D	+	629	43
294	P266H/C565N	+	680	52
295	I27E/V105C/I285E/C503Q/A547D/A551D/C565N	+	615	47
296	I27E/V39A/L278D	+	646	45
297	P210C/L214E/P266H	+	644	51
298	I27E/V39A/R134Q/P210C/L214E/C503Q	+	601	43
299	I27E/V39A/R43L/V105C/A153R/L214E/P266H/L278D/C503Q	+	607	42
300	I27E/V105C/Q205T/L214E/L278D/I285E/C503Q/A547D/A551D/C565N	+	583	46
301	I27E/V39A/V105C/C503Q/A547D/A551D/C565N	+	621	43
302	I27E/V105C/L214E/I285E/A551D/C565N	+	588	46
303	I27E/P210C/A551D	+	646	51
304	I27E/V39A/V105C/Q205T/L214E/L278D/C503Q/A547D	+	589	42

[0933]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)
("—"=FIOP<0.7;"+"=FIOP 0.7<FIOP<1.4;"++"=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
305	I27E/V105C/L278D/A547D	+	636	50
306	C503Q	+	680	52
307	I27E/V39A/V105C/R134Q/L214E/L278D/A547D/A551D	+	585	41
308	I27E/R134Q	+	652	51
309	V39A/R134Q	+	656	44
310	I27E/V39A/V105C/A112C/Q205T/P210C/P266H/C503Q/A547D	+	611	42
311	V39A/A112C/A153R/Q205T/L278D/C503Q/A547D	+	648	41
312	I27E/V105C/P266H	+	641	50
313	I27E/V39A/V105C/R134Q/P210C/L214E	+	581	41
314	I27E/V39A/V105C/R134Q/L278D/A551D	+	612	42
315	I27E/V39A/V105C/A112C/R134Q/A153R/Q205T/L214E/P266H/L 278D/C503Q/A551D	+	577	40
316	I27E/V105C/R134Q/A153R/I285E/A547D	+	603	45
317	I27E/V39A/A112C/A547D	+	638	43
318	I27E/V39A/R134Q/A153R/L214E/P266H/L278D/C503Q/A547D/C 565N	+	602	42
319	I27E/V105C/Q205T/P266H/C503Q	+	635	49
320	I27E/V39A/V105C/R134Q/L278D/I285E/C503Q	+	596	39
321	C503Q/A551D/C565N	+	675	52
322	I27E/V39A/A112C/R134Q/Q205T/P210C/L214E/A551D/C565N	+	593	41
323	R134Q/C503Q/A547D/A551D	+	666	51
324	I27E/A112C/Q205T/P266H/L278D/I285E/C503Q	+	631	46
325	P266H/L278D/C503Q	+	680	52
326	I27E/V39A/Q205T/P266H/A551D	+	635	44
327	Q205T/L214E/I285E/C503Q/C565N	+	627	48
328	I27E/V105C/R134Q/Q205T/L214E/P266H/L278D/C503Q/C565N	+	600	48
329	I27E/V39A/V105C/R134Q/P266H/C503Q	+	617	42
330	I27E/V39A/V105C/R134Q/P210C/L214E/I285E/A547D	+	555	38
331	I27E/V39A/V105C/Q205T/P210C/L278D/C503Q	+	616	42
332	I27E/V39A/R134Q/A547D/C565N	+	632	44
333	V105C/A547D/A551D	+	655	50
334	I27E/V105C/R134Q/A153R/P210C/L214E/C503Q/A547D	+	588	47
335	I27E/P210C/L214E/C503Q/A547D	++	620	51
336	I27E/V105C/C503Q/A547D/A551D/C565N	+	636	50
337	I27E/V105C/R134Q/L214E/L278D/C503Q/A547D/A551D/C565N	+	600	48
338	I27E/V39A/V105C/A153R/Q205T/L278D/C503Q/A547D/A551D	+	612	41
339	I27E/V39A/V105C/A112C/L214E/I285E/C503Q/A547D	+	573	39
340	L214E/A547D	+	648	51
341	V105C/Q205T/L214E/L278D	+	628	49
342	I27E/V39A/Q205T/L214E/C503Q/C565N	+	614	44
343	I27E/P210C	+	651	51

[0934]

[0935]

表9-1.靶向TCE-I、II、VI和VII的变体的FIOP、总免疫原性(TIS)和免疫原性击中计数(IHC) (“-“=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
344	V105C/Q205T/C503Q/A551D	+	649	49
345	I27E/V39A/A112C/R134Q/P266H/I285E	+	613	39
346	V39A/P210C/L214E/L278D/I285E/C503Q/A551D	+	603	41
347	I27E/Q205T/L214E/C503Q/A547D/C565N	+	624	51
348	I27E/V39A/L214E/L278D/C503Q/A547D/C565N	+	614	44
349	I27E/P210C/C503Q/C565N	+	651	51
350	I27E/L278D/C503Q/A551D/C565N	+	656	52
351	R134Q/L214E/L278D/C503Q	+	644	50
352	I27E/A153R/L214E/L278D/I285E/A551D/C565N	+	605	47
353	I27E/V105C/L214E/A551D/C565N	+	609	49
354	V39A/R134Q/Q205T/L214E/C503Q/C565N	+	624	43

[0936]

表9-2.靶向TCE-III的变体的FIOP、总免疫原性(TIS)和免疫原性击中计数(IHC) (“-“=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
变体 36		+	71	7
400	V80I/R134C/P564Q	+	60	5
401	V121C	+	44	6
402	R134L	+	78	7
403	A124G	+	58	7
404	R134W	+	63	5
405	L126M	+	56	5
406	N130C/M370I	+	62	6
407	N130Q	+	65	7
408	A123G	+	67	8
409	A129G	+	64	9
410	G135A/A394E	+	76	7
411	A129L	+	82	9
412	M133R	+	52	3
413	G135S	+	78	9
414	L126T	+	62	8
415	L127A	+	46	6
416	R134I	+	82	7
417	R134N/G307C	++	60	5
418	L126I	++	80	12
419	G135C	+	67	6
420	M125L	+	72	8

表9-3.靶向TCE-IV的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)
("-"=FIOP<0.7;"+"=FIOP 0.7<FIOP<1.4;"++"=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
变体 36		+	74	11
500	P157H	+	69	8
501	R140M	+	72	11
502	P157F	+	80	12
503	A153C	+	56	7
504	E142V	+	84	14
505	K145G/P157T	+	62	8
506	R140D	+	63	8
507	E142D/G371D	+	79	13
508	M147A	+	48	1
509	T156K/G483C	+	83	12
510	P157D	+	64	8
511	A62S/M147V	+	62	8
512	R146L	+	81	14
513	Y158E	+	72	11
514	G154Y/L174M/Q321K/S456I/G483C	+	75	10
515	S82I/G135C/P157F/W279L	+	77	11
516	I144N	+	61	7
517	T110I/I139R	+	65	8
518	L47M/I144L	+	67	8
519	L47M/M147G/A383E	+	48	4
520	G20S/I144L	+	67	8
521	R146W/D191Y	+	62	5
522	P157Y	+	73	9
523	L47M/P157C	+	68	8
524	E142P	+	68	9
525	F150K	+	59	6
526	L141T	+	55	7
527	V159H	+	67	8
528	I144L	+	67	8
529	A119E/T156H/A289D	+	78	11
530	Q58K/P157D/G369C	+	64	8
531	L47M/R146E	+	58	6
532	E142H	+	79	13
533	R140N/A199E	+	67	8
534	I144V	+	72	13
535	I149L	+	69	9
536	R146H	+	77	12
537	I139M	+	74	11
538	A153S/H250N	+	73	10
539	L319M	+	74	11
540	R140E/A334S/A551D	+	63	8
541	F150L	+	79	12
542	L143M	+	72	11
543	A153G	+	62	8

[0937]

[0938]

表9-3.靶向TCE-IV的变体的FIOP、总免疫原性(TIS)和 免疫原性击中计数(IHC) (“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
544	I139V	+	74	11
545	Q58H/L143V	+	72	11
546	G154R	+	81	13
547	K145Q	+	72	11
548	L143F	+	74	11
549	R140G	+	67	8
550	V159C	+	72	11
551	Q389K	+	74	11
552	L141P	+	58	5
553	M247I	+	74	11
554	F150M	+	77	12
555	L141Q	+	59	7
556	L151M	+	68	9
557	V159L	+	74	11
558	R94C/I149E	+	53	6
559	V159M	+	74	11
560	L118M/L141H	+	58	7
561	K145N	+	71	11
562	I149R	+	72	11
563	K145R	+	84	15
564	L141K	+	59	7
565	R43S	+	74	11

[0939]

表9-4.靶向TCE-V的变体的FIOP、总免疫原性(TIS)和 免疫原性击中计数(IHC)。 (“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
变体 36		+	36	2
600	Y176R	+	51	4
601	S180C	+	27	0
602	S180T	+	27	0
603	V172L	+	30	1
604	I177V	+	24	0
605	V172I	+	29	1
606	I177M	+	30	1
607	Y176I	+	54	4
608	Y176M	+	33	2
609	Y176V	+	41	3
610	L174M	+	21	0
611	P117T/Y176Q	+	37	2

[0940]

表9-4.靶向TCE-V的变体的FIOP、总免疫原性(TIS)和 免疫原性击中计数(IHC)。 (“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
612	S175G	+	26	0
613	T178L/A477S	+	43	5
614	V172C	+	26	0
615	Y176E	+	20	0

表9-5.靶向TCE-VIII的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)。
(“-“=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
变体 36		+	83	10
700	A232S	+	83	10
701	A112S/M370A/A507E	+	71	10
702	Q240K/H374R	+	83	10
703	S461G	+	83	10
704	H374D	+	83	10
705	M372A	+	74	10
706	L349M	+	83	10
707	Y378S	+	74	7
708	G371H	+	91	11
709	Y377N	+	82	10
710	I379N	+	80	8
711	D191Y/H385N	+	81	10
712	I379L	+	83	10
713	R43S/H374R	+	83	10
714	Q355K/H374S	+	83	10
715	P275T/H374R	+	83	10
716	H374Q/P396Q	+	99	14
717	H385N	+	81	10
718	Y378L	+	78	9
719	I379C	+	70	7
720	M370G	+	64	10
721	M372V	+	76	10
722	K384R	+	87	10
723	A383V	+	88	13
724	M147I/H374S	+	83	10
725	Y378F/P404Q	+	85	10
726	H374S	+	83	10
727	Y378E	+	55	2
728	H374R/G417C	+	83	10
729	L418M	+	83	10
730	S525L	+	83	10
731	Y378D	+	60	4

[0941]

[0942]

表9-5.靶向TCE-VIII的变体的FIOP、总免疫原性(TIS)和 免疫原性击中计数(IHC)。 (“-”=FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++”=FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
732	A383S	+	88	10
733	D387S	+	96	11
734	L382H	+	78	10
735	L382C	+	76	8
736	G371Q	+	84	10
737	H374L	+	85	11
738	Y378C	+	58	5
739	H374A	+	83	10
740	L375M	+	77	10
741	H385C	+	74	8
742	A334S/H374V	+	85	11
743	H374R	+	83	10
744	H385M/P403H	+	83	11
745	Y378I	+	83	10
746	Y377I	+	86	10
747	H385G	+	74	7
748	H385S/P403H	+	79	9
749	H374N	+	83	10
750	S187R/L381V	+	78	7
751	L382S	+	77	9
752	Y377C	+	75	7
753	L381V	+	78	7
754	G371S	+	88	10
755	A256S/L381N	+	70	6
756	I379M	+	80	10
757	R43S/H374K	+	83	10
758	I379H	+	76	7
759	M370S	+	64	10
760	P275Q/M370S	+	64	10
761	G425V	+	83	10
762	A447S	+	83	10
763	L382M	+	80	10
764	G371N	+	84	10
765	L381M/Q560K	+	79	10
767	L382I	+	77	8
768	H374G	+	83	10
769	M370I	+	76	10
770	Q332K/Y377M	+	86	10
771	Y378N	+	64	5
772	L375I	+	83	10
773	H374T	+	83	10
774	L381G	+	60	4

表9-6.靶向TCE-IX的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)。
 (“-“ =FIOP<0.7; “+”=FIOP 0.7<FIOP<1.4; “++” =FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOP C	TIS	IHC
变体 36		+	45	4
800	I439V	+	49	4
801	I439C	+	39	1
802	S438F	+	50	4
803	T433S	+	44	0
804	T433A	+	35	0
805	Y435Q/H446N	+	43	2
806	L431V	+	28	0
807	N437Q	+	40	0
808	L213M/S438L	+	46	4
809	L432C	+	39	2
810	L432V	+	47	4
811	T433L	+	50	4
812	L431P	+	18	0
813	I439F	+	47	4
814	S286R/Y435T	+	53	4
815	S98I	+	45	4
816	S438R	+	48	4
817	S331I	+	45	4
818	S438M	+	44	4
819	Q240K/T433Y	+	41	0
820	T433I	+	46	3
821	T433N	+	38	0
822	L431S	+	27	0
823	N437G	+	38	0
824	L431E	+	26	0
825	L431C	+	25	0
826	G436T	+	45	0
827	G436D	+	42	1
828	T433V	+	59	7
829	N437T/L538M	+	52	4
830	A289S/L431E	+	26	0
831	T433Q	+	37	0
832	I439L	+	46	4
833	F434C	+	34	0
834	T433P	+	35	0
835	A24S/F434M	+	69	8
836	L431G	+	18	0
837	S438C	+	45	1
838	T433W	+	33	0
839	Y435L	+	57	5
840	A62S/T433N	+	38	0
841	N437E	+	46	3
842	A477S	+	45	4
843	G436M	+	87	11
844	T433R	+	70	10
845	S438T	+	46	4

[0943]

表9-7.靶向TCE-X的变体的FIOP、总免疫原性(TIS)和
免疫原性击中计数(IHC)。
 (“-“ =FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++” =FIOP>1.4)

变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
变体 36		+	47	7
900	I471V	+	45	7
901	F472G	+	16	0
902	G483S	+	60	9
903	G483A/S524I	+	64	10
904	Y475C	+	20	0
905	I478N	+	44	6
906	V476I	+	30	2
907	N474D/R490H	+	29	3
908	P275Q	+	47	7
909	Q473K	+	55	7
910	G483H	+	58	8
911	Q473H/A507S	+	52	7
912	N474H	+	49	7
913	Y475F	+	58	7
914	I478S	+	46	6
915	Q473M	+	69	9
916	V476C	+	23	0
917	L104M/V476L	+	36	3
918	A119E/G365A	+	47	7
919	N474W	+	51	7
920	I471K	+	29	2
921	Q292H/A479G	+	35	2
922	G276V	+	47	7
923	I471N	+	26	1
924	A479S	+	53	7
925	A558S	+	47	7
926	Q473S	+	56	8
927	Q473R	+	62	8
928	Y475Q	+	33	5
929	I471G	+	32	3
930	A479G	+	35	2
931	A70S/N474E	+	32	3
932	I471F	+	48	7
933	Q58R/Y475H	+	32	5
934	F482L	+	56	9
935	V476L	+	36	3
936	A24E	+	47	7
937	N474R	+	55	7
938	G483C	+	49	6
939	Q473H	+	52	7
940	I471M	+	46	7
941	I471R	+	31	2
942	P404T/A477V	+	47	6
943	G483R/G537C	+	78	13

[0944]

[0945]

表9-7.靶向TCE-X的变体的FIOP、总免疫原性(TIS)和 免疫原性击中计数(IHC)。 (“-“ =FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++” =FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOPC	TIS	IHC
944	F482C	+	45	6
945	Q355H/I478C	+	31	1
946	L206M	+	47	7
947	N474A	+	48	7
948	Y475L	+	60	7
949	I471W	+	34	4

[0946]

表9-8.靶向TCE-I、III、IV、VIII和X的变体的FIOP、总免疫原性(TIS)和 免疫原性击中计数(IHC) (“-“ =FIOP<0.7;“+”=FIOP 0.7<FIOP<1.4;“++” =FIOP>1.4)				
变体号	活性突变 (与变体36号相比)	FIOP	TIS	IHC
变体 36		+		
950	F472G/C503Q/C565N	+	651	45
951	L381G/F472GC503Q/C565N	+	628	39
952	M133R/C503Q	+	663	48
953	M133R/L381G/C565N	+	640	42
954	V39A/C503Q/C565N	+	667	45
955	V39A/F472G/C503Q/C565N	+	636	38
956	V39A/G248C/L381G/F472G/C503Q/C565N	+	613	32
957	V39A/K115E/M133R/C565N	++	646	39
958	V39A/L381G	+	644	39
959	V39A/M133R/C503Q	+	648	41
960	V39A/M133R/C503Q/C565N	+	648	41
961	V39A/M133R/F472G/C503Q/C565N	+	617	34
962	V39A/M133R/F472G/C565N	+	617	34
963	V39A/M147A/C565N	+	641	35
964	V39A/M147A/F472G/C503Q/C565N	+	610	28
965	V39A/M147A/F472G/C565N	+	610	28
966	V39A/M147A/L381G/C503Q/C565N	+	618	29
967	V39A/M147A/L381G/F472G/C503Q/C565N	+	587	22
968	V39A/M147A/Y378E/C503Q/C565N	+	613	27
969	V39A/M147A/Y378E/C565N	+	613	27
970	V39A/Y378E/C503Q/C565N	+	639	37
971	V39A/Y378E/C565N	+	639	37
972	Y378D/C503Q	+	659	46
973	Y378E/F472G/C503Q/C565N	+	623	37

表9-9.靶向TCE-III和VIII的变体的FIOP、总免疫原性(TIS)和免疫原性击中计数(IHC)

(“-“ = FIOP < 0.7; “+” = FIOP 0.7 < FIOP < 1.4; “++” = FIOP > 1.4)

变体号	活性突变 (与变体967号相比)	FIOP	TIS	IHC
变体 967		+		
974	A129I	-	598	26
975	A129V	-	595	24
976	A136K	++	595	22
977	A24E/G381L	+	604	28
978	A289S	+	590	23
979	A383C	+	577	18
980	A383M	++	602	28
981	G381A	++	598	24
982	G381C	-	580	22
983	G381F	+	601	25
984	G381I	++	610	27
985	G381L	++	610	28
986	G381M	+	606	28
987	G381N	+	597	24
988	G381Q	+	590	23
989	G381S	+	601	24
990	G381T	+	590	23
991	H132L	-	603	25
992	H132S	-	592	23
993	H374G	+	587	22
994	H374M	+	589	23
995	H374Q	-	587	22
996	L127V	-	576	22
997	L431M	+	591	22
998	L563M	-	587	22
999	M372L	-	580	22
1000	R134C	-	576	20
1001	R134F	+	588	21
1002	R134H	+	579	21
1003	R134K	-	585	21
1004	S131C	-	579	21
1005	S131T	+	580	21
1006	V388C	+	577	21

[0947]

[0948]

表9-9.靶向TCE-III和VIII的变体的FIOP、总免疫原性(TIS)和免疫原性击中计数(IHC)				
("-" = FIOP < 0.7; "+" = FIOP 0.7 < FIOP < 1.4; "++" = FIOP > 1.4)				
变体号	活性突变 (与变体967号相比)	FIOP	TIS	IHC
1007	V388T	+	578	21
1008	R134H/V388T	+	570	20
1009	R134H/Y378E/G381L	+	574	19
1010	R134H/Y378E/G381L/V388T	+	566	19

[0949] 去免疫化的PAL变体的体外测试:

[0950] 在树突T细胞测定中测试去免疫化的PAL变体,以经验性地测试其引发T细胞应答的能力。使用标准技术从人供体分离外周血单核细胞(PBMC)。将这些细胞用作在确定成分培养基中培养的单核细胞的来源,以产生未成熟的树突细胞。将这些未成熟的树突细胞(DC)装载去免疫化的PAL变体,并然后通过确定成分培养基中进一步培养诱导成更成熟的表型,以提供抗原致敏的DC。将从与DC相同的供体样品中获得的CD8+T细胞耗尽的供体PBMC用CFSE来标记,然后与抗原致敏的DC培养7天,其后一式八份被测试。每个DC-T细胞培养物包括一组未经处理的对照组(即,阴性对照)。该测定还包括参考抗原对照(即,阳性对照),包括两种有效的全蛋白抗原。利用从具有多样性主要组织相容性复合物II类等位基因的50个人供体分离的细胞的测定提供了PAL变体引发T细胞应答的能力的统计学相关的评价。

[0951] 尽管已经参考具体的实施方案描述了本发明,可做出多种改变并且可替换等同物,以适应特定的情况、材料、物质的组成、方法、方法一个步骤或多个步骤,从而实现本发明的益处,而不偏离所要求保护的范围。

[0952] 本公开内容中引用的每个和每一个出版物和专利文件通过引用被并入本文用于在美国的所有目的,如同每个这样的出版物或文件被明确和分别指出以通过引用并入本文一样。出版物和专利文件的引用不被认为是任何这样的文件是相关的现有技术的指示,也不构成对其内容或日期的承认。

- [0001] 序列表
- [0002] <110> 科德克希思公司
- [0003] 吉伽特·W·哈思曼
- [0004] 尼古拉斯·J·阿加德
- [0005] 本杰明·米杰茨
- [0006] 乔纳森·弗罗姆
- [0007] 张希云
- [0008] <120> 苯丙氨酸解氨酶多肽
- [0009] <130> CX7-131W02
- [0010] <150> US 61/813,586
- [0011] <151> 2013-04-18
- [0012] <150> US 61/897,932
- [0013] <151> 2013-10-31
- [0014] <160> 33
- [0015] <170> PatentIn版本3.5
- [0016] <210> 1
- [0017] <211> 7407
- [0018] <212> DNA
- [0019] <213> 人工序列
- [0020] <220>
- [0021] <223> 合成的多核苷酸
- [0022] <400> 1
- [0023] tctcatgttt gacagcttat catcgataag ctttaatgcg gtagttttatc acagttaaata 60
- [0024] tgctaacgca gtcaggcacc gtgtatgaaa tctaacaatg cgctcatcgt catcctcggc 120
- [0025] accgtcaccc tggatgctgt aggcataaggc ttggttatgc cggtagtacc gggcctcttg 180
- [0026] cgggatatcc ggatatagtt cctcctttca gcaaaaaacc cctcaagacc cgtttagagg 240
- [0027] ccccaagggg ttatgctagt tattgctcag cggtaggcagc agccaactca gcttcctttc 300
- [0028] gggccttggt agcagccgga tccttaatgc agacacggca gaatgtcctg aacggcctga 360
- [0029] acaataacac caccgctgc aatatctgca ctaatacgtg caatatgttc atccagaccc 420
- [0030] tgttcattat cattccaaat atacggacga tctgaggtcg gtttctgacc aacaacatga 480
- [0031] cgaactgcgc tatacagacg ttcggttgcc ggtgacagac aggcacgtgc atcataatga 540
- [0032] ccggtttttt tgtaggtacg cagatcaact gcctgaacac caaacatcag ggcaatggca 600
- [0033] acataattct gaaaaatatc aacgctacga cgtgccaggg ttgcgctggt ataaccctgg 660
- [0034] ctgttaatat tctggttaaa ctgttcggca tgggtcggaa aacgatctgc aatactatta 720
- [0035] ccataaaagg tcagcagcgg cataatgcta ttaccgcaa tctgcagacc ttccagaccc 780
- [0036] atattaactt tacgttcacg attaccacg agactcggag gcagaccatt gctaaattcc 840
- [0037] ggtgatgcca gcagtcaat ctgaacatcc agatgttttg ccagcagacc gatataatag 900
- [0038] cgcagatgat ccatacccat accaacatac tgaccagaa aattaccacc atgatatgct 960
- [0039] gcctgattat caacatcaat cagcgggtta tcggtaacgc tggttaatctc aatttcgatt 1020
- [0040] tgtttgcaa tctggctaata accatcaaca atcggacca gatactgcgg cagacaacgc 1080
- [0041] aggctataac gatcctggat cagttcatga tcacgataat catgtttacc atccagttca 1140

[0042]	tcacgaacca gctggctatt ggccagcagg ctaatcatct gatctgctgc ccacagctga	1200
[0043]	cccggatgcg gtttgcgtgt atggataaac ggatgaaagc tctgatttgt accattcagt	1260
[0044]	gcctgaatat ccagtgcatt aacacccatt gcaattgcgg tcagaatctg ggtatcataa	1320
[0045]	acacaatttg ctgcaatacc ggtcataacg ctgggtccat tcatcattgc cagaccttct	1380
[0046]	ttcggcagca gggtcagcgg actcagattc agctgacgca gtgcggtcgg tgcgtccatt	1440
[0047]	tctttgccat taaaatcaac tttaaagctc ggggtccaggc caatcaggct accggtaata	1500
[0048]	tagctcagcg gaaccagatc accgctggca ccaatgctac caaattcata aacatacggg	1560
[0049]	gtaacaccgg cattcagaaa gatttccatg cgtttaatca gttccagacg aataccgctt	1620
[0050]	gcaccacgca tgtggctatt tgcacgcagc agcattgctg cacgaacatc tgccagcggc	1680
[0051]	agtttattac ctgcaccggg tttcagaaac caaaccagat tggctctgcag ttcgcttgcc	1740
[0052]	tgttcacggc taattgcaac atttgccata ccacaaaac cgctggtaac accataaatc	1800
[0053]	ggttcaccgc tttcaactgc attattgata taatcacagc tggcctgaat accctgcaga	1860
[0054]	atatcggtat tattggtcag gctaaccagg gtgccattac gggcaacacg tgcaacatca	1920
[0055]	ttgatggtea gtttctgatt accaataatc acatttgcgc tgctattgcc ggtaaagcta	1980
[0056]	aactgctggc tgctggtttt gctctgtgcc tggctcaggg ttttcatatg acgaccttcg	2040
[0057]	atatggccgc tgctgtgatg atgatgatga tgatgatgat gatggcccat ggtatatctc	2100
[0058]	cttcttaaag ttaaacaaaa ttatttctag aggggaattg ttatccgctc acaattcccc	2160
[0059]	tatagtgagt cgtattaatt tcgcgggacg gagatctcga tcctctacgc cggacgcac	2220
[0060]	gtggccggca tcaccggcgc cacaggctgcg gttgctggcg cctatatcgc cgacatcacc	2280
[0061]	gatggggaag atcgggctcg ccacttcggg ctcatgagcg cttgtttcgg cgtgggtatg	2340
[0062]	gtggcaggcc ccgtggccgg gggactgttg ggcgccatct ccttgcatgc accattcctt	2400
[0063]	gcggcggcgg tgctcaacgg cctcaacctc ctactgggct gcttcctaata gcaggagtcg	2460
[0064]	cataagggag agcgtcgaga tcccggacac catcgaatgg cgcaaacct ttcgcggtat	2520
[0065]	ggcatgatag cgcccgaag agagtcaatt cagggtggtg aatgtgaaac cagtaacgtt	2580
[0066]	atacgatgic gcagagtatg ccggtgtctc ttatcagacc gtttcccgcg tggatgaacca	2640
[0067]	ggccagccac gtttctgcga aaacgcggga aaaagtggaa gcggcgatgg cggagctgaa	2700
[0068]	ttacattccc aaccgcgtgg cacaacaact ggccggcaca cagtcgttgc tgattggcgt	2760
[0069]	tgccacctcc agtctggccc tgcacgcgcc gtcgcaaat gtcgcggcga ttaaattctc	2820
[0070]	cgccgatcaa ctgggtgcca gcgtgggtgt gtcgatggtg gaacgaagcg gcgtcgaagc	2880
[0071]	ctgtaaagcg gcggtgcaca atcttctcgc gcaacgcgtc agtgggctga tcattaacta	2940
[0072]	tccgctggat gaccaggatg ccattgctgt ggaagctgcc tgcactaatg ttccggcgtt	3000
[0073]	atttcttgat gtctctgacc agacacccat caacagtatt attttctccc atgaagacgg	3060
[0074]	tacgcgactg ggcgtggagc atctggctgc attgggtcac cagcaaatcg cgctgttagc	3120
[0075]	gggcccatta agttctgtct cggcgcgtct gcgtctggct ggctggcata aatatctcac	3180
[0076]	tcgcaatcaa attcagccga tagcggaaac ggaaggcgac tggagtgcc tgtccggttt	3240
[0077]	tcaacaacc atgcaaatgc tgaatgagg catcgttccc actgcgatgc tggttgcaa	3300
[0078]	cgatcagatg gcgtgggcg caatgcgcgc cattaccgag tccgggctgc gcgttggtgc	3360
[0079]	ggatatctcg gtagtgggat acgacgatac cgaagacagc tcatgttata tcccgcggtt	3420
[0080]	aaccaccatc aaacaggatt ttcgctgct ggggcaaac agcgtggacc gcttgctgca	3480
[0081]	actctctcag ggccaggcgg tgaagggcaa tcagctgttg cccgtctcac tggatgaaa	3540
[0082]	aaaaaccacc ctggcgcca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt	3600
[0083]	aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgcgcg aacgcaatta	3660

[0084]	atgtaagtta gctcactcat taggcaccgg gatctcgacc gatgcccttg agagccttca	3720
[0085]	acccagtcag ctctttccgg tgggcgcggg gcatgactat cgtcgccgca cttatgactg	3780
[0086]	tcttctttat catgcaactc gtaggacagg tgccggcagc gctctgggtc attttcggcg	3840
[0087]	aggaccgctt tcgctggagc gcgacgatga tcggcctgtc gcttgcggtta ttcggaatct	3900
[0088]	tgcacgccct cgctcaagcc ttcgtcactg gtccccccac caaacgtttc ggcgagaagc	3960
[0089]	aggccattat cgccggcatg gcggccgacg cgctgggcta cgtcttgctg gcgttcgca	4020
[0090]	cgcgaggtg gatggccttc cccattatga ttctttctgc ttccggcggc atcgggatgc	4080
[0091]	ccgcttgca ggccatgtg tccaggcagg tagatgacga ccatcaggga cagcttcaag	4140
[0092]	gatcgctgc ggctcttacc agcctaactt cgatcactgg accgctgac gtcacggcga	4200
[0093]	tttatgccgc ctccggcagc acatggaacg ggttggtcatg gattgtaggc gccgccctat	4260
[0094]	acctgtctg cctccccgcg ttgcgtcgcg gtgcatggag ccggggccacc tcgacctgaa	4320
[0095]	tggaagccgg cggcacctcg ctaacggatt caccactcca agaattggag ccaatcaatt	4380
[0096]	cttgcggaaga actgtgaatg cgcaaaccaa cccttggcag aacatatcca tcgcgtccgc	4440
[0097]	catctccagc agccgcacgc ggcgcactc gggcagcgtt gggtcctggc cacgggtgcg	4500
[0098]	catgatcgtg ctctgtcgt tgaggaccg gctaggctgg cggggttgcc ttactggta	4560
[0099]	gcagaatgaa tcaccgatac gcgagcgaac gtgaagcgac tgctgctgca aaacgtctgc	4620
[0100]	gacctgagca acaacatgaa tggctctcgg ttcccggtt tcgtaaagtc tggaaacgcg	4680
[0101]	gaagtcagcg ccctgcacca ttatgttccg gatctgcac gcaggatgct gctggctacc	4740
[0102]	ctgtggaaca cctacatctg tattaacgaa gcgctggcat tgaccctgag tgatttttct	4800
[0103]	ctggccccgc cgcattccata ccgccagttg ttaccctca caacgttcca gtaaccgggc	4860
[0104]	atgttcatca tcagtaacct gtatcgtgag catctctct cgtttcatcg gtatcattac	4920
[0105]	cccatgaac agaaatcccc cttacacgga ggcacagtg accaaacagg aaaaaaccgc	4980
[0106]	ccttaacatg gcccgcttta tcagaagcca gacattaacg cttctggaga aactcaacga	5040
[0107]	gctggacgcg gatgaacagg cagacatctg tgaatcgctt cacgaccag ctgatgagct	5100
[0108]	ttaccgcagc tgctcgcgc gtttcgggtga tgacggtgaa aacctctgac acatgcagct	5160
[0109]	cccgagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag ccgctcagg	5220
[0110]	cgcgtcagcg ggtgttggcg ggtgtcgggg cgcagccatg acccagtcac gtagcgatag	5280
[0111]	cggagtgtat actggcttaa ctatgcgga tcagagcaga ttgtactgag agtgcacat	5340
[0112]	atatcggtg tgaaataacc cacagatgcg taaggagaaa ataccgcatc aggcgctctt	5400
[0113]	ccgttctctc gctcactgac tcgtcgcgt cggtcgttcg gctgcggcga gcggtatcag	5460
[0114]	ctactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca	5520
[0115]	tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt	5580
[0116]	tccataggt ccgccccct gacgagcatc aaaaaatcg acgtcaagt cagaggtggc	5640
[0117]	gaaaccgac aggaactataa agataccagg cgtttcccc tggaagctcc ctctgctgct	5700
[0118]	ctctgttcc gaccctgcg cttaccgat acctgtccgc ctttctccct tcgggaagcg	5760
[0119]	tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca	5820
[0120]	agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggttaact	5880
[0121]	atcgtcttga gtccaaccg gtaagacacg acttatcgcc actggcagca gccactggta	5940
[0122]	acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tgggtggccta	6000
[0123]	actacggcta cactagaagg acagtattg gtatctgcgc tctgctgaag ccagttacct	6060
[0124]	tcggaaaaag agttggtagc tcttgatccg gaaacaaac caccgctggt agcggtggtt	6120
[0125]	ttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga	6180

[0126]	tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttgggtca	6240
[0127]	tgagattatc aaaaaggatc ttcacctaga tcctttttaa ttaaaaatga agttttaaat	6300
[0128]	caatctaaag tatatatgag taaacttggc ctgacagtta ccaatgctta atcagtgagg	6360
[0129]	cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt	6420
[0130]	agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag	6480
[0131]	accacagctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc	6540
[0132]	gcagaagtgg tcttgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag	6600
[0133]	ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctgcaggca	6660
[0134]	tcgtgggtgc acgctcgtcg tttggatagg cttcattcag ctccggttcc caacgatcaa	6720
[0135]	ggcgagttac atgatcccc atgttgtgca aaaaagcggc tagctccttc ggtcctccga	6780
[0136]	tcgttgctcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcaactgcata	6840
[0137]	attctcttac tgtcatgcca tccgtaagat gcttttctgt gactgggtgag tactcaacca	6900
[0138]	agtcattctg agaatagttg atgcggcgac cgagttgctc ttgcccggcg tcaacacggg	6960
[0139]	ataataccgc gccacatagc agaactttaa aagtgtcat cattggaaaa cgttcttcgg	7020
[0140]	ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttgatgtaa cccactcgtg	7080
[0141]	cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag	7140
[0142]	gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac	7200
[0143]	tcttctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca	7260
[0144]	tatttgaatg tatttagaaa aataaacaata taggggttcc gcgcacattt ccccgaaaag	7320
[0145]	tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta	7380
[0146]	tcacgaggcc ctttcgtctt caagaat	7407
[0147]	<210> 2	
[0148]	<211> 1701	
[0149]	<212> DNA	
[0150]	<213> 人工序列	
[0151]	<220>	
[0152]	<223> 合成的多核苷酸	
[0153]	<400> 2	
[0154]	atgaaaaccc tgagccagge acagagcaaa accagcagcc agcagtttag ctttaccggc	60
[0155]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgcacgt	120
[0156]	gttgcccgta atggcacctt ggtagcctg accaataata ccgatattct gcagggtatt	180
[0157]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatgggtgt	240
[0158]	accagcgggt ttgggtgtat ggcaaatgtt gcaattagcc gtgaacaggc aagcgaactg	300
[0159]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360
[0160]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggatattcgt	420
[0161]	ctggaactga ttaaacgcat ggaaatcttt ctgaatgccg gtgttaccct gtatgtttat	480
[0162]	gaatttggtg gcattgggtg cagcgggtgat ctggttccgc tgagctatat taccggtagc	540
[0163]	ctgattggcc tggaccggag ctttaaagtt gatttttaatg gcaaagaaat ggacgcaccg	600
[0164]	accgcaactg gtcagctgaa tctgagtcgg ctgaccctgc tgccgaaaga aggtctggca	660
[0165]	atgatgaatg gcaccagcgt tatgaccggc attgcagcaa attgtgttta tgataccag	720
[0166]	attctgaccg caattgcaat ggggtgttcat gcaactggata ttcaggcact gaattgtaca	780
[0167]	aatcagagct ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840

[0168]	gcagatcaga tgattagcct gctggccaat agccagctgg ttcgtgatga actggatggt	900
[0169]	aaacatgatt atcgtgatca tgaactgatac caggatcggt atagcctgcg ttgtctgccg	960
[0170]	cagtatctgg gtccgattgt tgatgggtatt agccagattg ccaaacaat cgaaattgag	1020
[0171]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[0172]	ggtaattttc tgggtcagta tggttggtatg ggtatggatc atctgcgcta ttatatcggt	1140
[0173]	ctgctggcaa aacatctgga tggtcagatt gactgctgg catcacgga atttagcaat	1200
[0174]	ggtctgcctc cgagtctgct gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260
[0175]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[0176]	gatcggtttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[0177]	agcgcaaccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg	1440
[0178]	atgtttggtg ttacggcagt tgatctgcgt acctacaaaa aaaccgggtca ttatgatgca	1500
[0179]	cgtgcctgtc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560
[0180]	cagaaaccga cctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[0181]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[0182]	gacattctgc cgtgtctgca t	1701
[0183]	<210>	3
[0184]	<211>	1701
[0185]	<212>	DNA
[0186]	<213>	多变鱼腥藻 (<i>Anabaena variabilis</i>)
[0187]	<400>	3
[0188]	atgaaaaccc tgagccaggc acagagcaaa accagcagcc agcagtttag ctttaccggc	60
[0189]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgcacgt	120
[0190]	gttgcccgta atggcacctt ggtagcctg accaataata ccgatattct gcagggtatt	180
[0191]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt	240
[0192]	accagcgggt ttggtggtat ggcaaatgtt gcaattagcc gtgaacaggc aagcgaactg	300
[0193]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360
[0194]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggtattcgt	420
[0195]	ctggaactga ttaaacgcat ggaaatctt ctgaatgccg gtgttacctt gtatgtttat	480
[0196]	gaatttggtg gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc	540
[0197]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600
[0198]	accgactgac gtcagctgaa tctgagtcg ctgacctgc tgccgaaaga aggtctggca	660
[0199]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720
[0200]	attctgaccg caattgcaat ggggtgttcat gactggata ttacggcact gaatggtaca	780
[0201]	aatcagagct ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840
[0202]	gcagatcaga tgattagcct gctggccaat agccagctgg ttcgtgatga actggatggt	900
[0203]	aaacatgatt atcgtgatca tgaactgatac caggatcggt atagcctgcg ttgtctgccg	960
[0204]	cagtatctgg gtccgattgt tgatgggtatt agccagattg ccaaacaat cgaaattgag	1020
[0205]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[0206]	ggtaattttc tgggtcagta tggttggtatg ggtatggatc atctgcgcta ttatatcggt	1140
[0207]	ctgctggcaa aacatctgga tggtcagatt gactgctgg catcacgga atttagcaat	1200
[0208]	ggtctgcctc cgagtctgct gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260
[0209]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320

[0210]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[0211]	agcgcaaccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg	1440
[0212]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500
[0213]	cgtgcctgtc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtgtg	1560
[0214]	cagaaaccga cctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[0215]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[0216]	gacattctgc cgtgtctgca t	1701
[0217]	<210>	4
[0218]	<211>	567
[0219]	<212>	PRT
[0220]	<213>	多变鱼腥藻
[0221]	<400>	4
[0222]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[0223]	1 5 10 15	
[0224]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[0225]	20 25 30	
[0226]	Leu Thr Ile Asn Asp Val Ala Arg Val Ala Arg Asn Gly Thr Leu Val	
[0227]	35 40 45	
[0228]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys	
[0229]	50 55 60	
[0230]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val	
[0231]	65 70 75 80	
[0232]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Ala Ile Ser Arg Glu Gln	
[0233]	85 90 95	
[0234]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala	
[0235]	100 105 110	
[0236]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg	
[0237]	115 120 125	
[0238]	Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile	
[0239]	130 135 140	
[0240]	Lys Arg Met Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr	
[0241]	145 150 155 160	
[0242]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr	
[0243]	165 170 175	
[0244]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe	
[0245]	180 185 190	
[0246]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu	
[0247]	195 200 205	
[0248]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly	
[0249]	210 215 220	
[0250]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln	
[0251]	225 230 235 240	

[0252]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[0253]		245	250 255
[0254]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[0255]		260	265 270
[0256]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[0257]		275	280 285
[0258]	Ala Asn Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[0259]		290	295 300
[0260]	Arg Asp His Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[0261]		305	310 315 320
[0262]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[0263]		325	330 335
[0264]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[0265]		340	345 350
[0266]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[0267]		355	360 365
[0268]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[0269]		370	375 380
[0270]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[0271]		385	390 395 400
[0272]	Gly Leu Pro Pro Ser Leu Leu Gly Asn Arg Glu Arg Lys Val Asn Met		
[0273]		405	410 415
[0274]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[0275]		420	425 430
[0276]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[0277]		435	440 445
[0278]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[0279]		450	455 460
[0280]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu		
[0281]		465	470 475 480
[0282]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[0283]		485	490 495
[0284]	His Tyr Asp Ala Arg Ala Cys Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[0285]		500	505 510
[0286]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Thr Ser Asp Arg Pro		
[0287]		515	520 525
[0288]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[0289]		530	535 540
[0290]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[0291]		545	550 555 560
[0292]	Asp Ile Leu Pro Cys Leu His		
[0293]		565	

[0294]	<210> 5
[0295]	<211> 1701
[0296]	<212> DNA
[0297]	<213> 人工序列
[0298]	<220>
[0299]	<223> 合成的多核苷酸
[0300]	<400> 5
[0301]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc 60
[0302]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgtagct 120
[0303]	gttgcccgta atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt 180
[0304]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt 240
[0305]	accagcggtt ttggtggtat ggcaaatgtt gtaattagcc gtgaacagcg aagcgaactg 300
[0306]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat 360
[0307]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggatattcgt 420
[0308]	ctggaactga ttaaaccgat ggaaatcttt ctgaatgccg gtgttaccct gtatgtttat 480
[0309]	gaatttggtg gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc 540
[0310]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg 600
[0311]	accgcactgc gtcagctgaa tctgagtcgc ctgaccctgc tgccgaaaga aggtctggca 660
[0312]	atgatgaatg gcaccagcgt tatgaccggg attgcagcaa attgtgttta tgataccag 720
[0313]	attctgaccg caattgcaat ggggtgttcat gactggata ttcagggaact gaattgtaca 780
[0314]	aatcagagct ttcattcgtt tatccataac agcaaaccgc atccgggtca gctgtgggca 840
[0315]	gcagatcaga tgattagcct gctggccggg agccagctgg ttcgtgatga actggatggt 900
[0316]	aaacatgatt atcgtgatca tgaactgatc caggatcgtt atagcctgcg ttgtctgccg 960
[0317]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaatttag 1020
[0318]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt 1080
[0319]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt 1140
[0320]	ctgctggcaa aacatctgga tggtcagatt gactgctgg tatcaccgga atttaacaat 1200
[0321]	ggtctgcctg cgagtcgtgt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt 1260
[0322]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca 1320
[0323]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc 1380
[0324]	agcgcaaccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg 1440
[0325]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca 1500
[0326]	cgtgcctgtc tgtcaccgcg aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt 1560
[0327]	cagtatccga gtcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa 1620
[0328]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag 1680
[0329]	gacattctgc cgtgtctgca t 1701
[0330]	<210> 6
[0331]	<211> 567
[0332]	<212> PRT
[0333]	<213> 人工序列
[0334]	<220>
[0335]	<223> 合成的多肽

[0336]	<400> 6																
[0337]	Met	Lys	Thr	Leu	Ser	Gln	Ala	Gln	Ser	Lys	Thr	Ser	Ser	Gln	Gln	Phe	
[0338]	1				5					10					15		
[0339]	Ser	Phe	Thr	Gly	Asn	Ser	Ser	Ala	Asn	Val	Ile	Ile	Gly	Asn	Gln	Lys	
[0340]				20					25					30			
[0341]	Leu	Thr	Ile	Asn	Asp	Val	Val	Arg	Val	Ala	Arg	Asn	Gly	Thr	Leu	Val	
[0342]			35					40					45				
[0343]	Ser	Leu	Thr	Asn	Asn	Thr	Asp	Ile	Leu	Gln	Gly	Ile	Gln	Ala	Ser	Cys	
[0344]		50					55					60					
[0345]	Asp	Tyr	Ile	Asn	Asn	Ala	Val	Glu	Ser	Gly	Glu	Pro	Ile	Tyr	Gly	Val	
[0346]	65					70				75						80	
[0347]	Thr	Ser	Gly	Phe	Gly	Gly	Met	Ala	Asn	Val	Val	Ile	Ser	Arg	Glu	Gln	
[0348]					85					90					95		
[0349]	Ala	Ser	Glu	Leu	Gln	Thr	Asn	Leu	Val	Trp	Phe	Leu	Lys	Thr	Gly	Ala	
[0350]				100					105					110			
[0351]	Gly	Asn	Lys	Leu	Pro	Leu	Ala	Asp	Val	Arg	Ala	Ala	Met	Leu	Leu	Arg	
[0352]			115					120					125				
[0353]	Ala	Asn	Ser	His	Met	Arg	Gly	Ala	Ser	Gly	Ile	Arg	Leu	Glu	Leu	Ile	
[0354]		130					135					140					
[0355]	Lys	Arg	Met	Glu	Ile	Phe	Leu	Asn	Ala	Gly	Val	Thr	Pro	Tyr	Val	Tyr	
[0356]	145					150				155						160	
[0357]	Glu	Phe	Gly	Ser	Ile	Gly	Ala	Ser	Gly	Asp	Leu	Val	Pro	Leu	Ser	Tyr	
[0358]					165					170					175		
[0359]	Ile	Thr	Gly	Ser	Leu	Ile	Gly	Leu	Asp	Pro	Ser	Phe	Lys	Val	Asp	Phe	
[0360]				180					185					190			
[0361]	Asn	Gly	Lys	Glu	Met	Asp	Ala	Pro	Thr	Ala	Leu	Arg	Gln	Leu	Asn	Leu	
[0362]			195					200					205				
[0363]	Ser	Pro	Leu	Thr	Leu	Leu	Pro	Lys	Glu	Gly	Leu	Ala	Met	Met	Asn	Gly	
[0364]		210					215					220					
[0365]	Thr	Ser	Val	Met	Thr	Gly	Ile	Ala	Ala	Asn	Cys	Val	Tyr	Asp	Thr	Gln	
[0366]	225					230				235						240	
[0367]	Ile	Leu	Thr	Ala	Ile	Ala	Met	Gly	Val	His	Ala	Leu	Asp	Ile	Gln	Gly	
[0368]					245					250					255		
[0369]	Leu	Asn	Gly	Thr	Asn	Gln	Ser	Phe	His	Pro	Phe	Ile	His	Asn	Ser	Lys	
[0370]				260					265					270			
[0371]	Pro	His	Pro	Gly	Gln	Leu	Trp	Ala	Ala	Asp	Gln	Met	Ile	Ser	Leu	Leu	
[0372]			275					280					285				
[0373]	Ala	Gly	Ser	Gln	Leu	Val	Arg	Asp	Glu	Leu	Asp	Gly	Lys	His	Asp	Tyr	
[0374]		290					295					300					
[0375]	Arg	Asp	His	Glu	Leu	Ile	Gln	Asp	Arg	Tyr	Ser	Leu	Arg	Cys	Leu	Pro	
[0376]	305					310				315						320	
[0377]	Gln	Tyr	Leu	Gly	Pro	Ile	Val	Asp	Gly	Ile	Ser	Gln	Ile	Ala	Lys	Gln	

[0378]	325	330	335
[0379]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[0380]	340	345	350
[0381]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[0382]	355	360	365
[0383]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[0384]	370	375	380
[0385]	His Leu Asp Val Gln Ile Ala Leu Leu Val Ser Pro Glu Phe Asn Asn		
[0386]	385	390	395
[0387]	Gly Leu Pro Ala Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[0388]	405	410	415
[0389]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[0390]	420	425	430
[0391]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[0392]	435	440	445
[0393]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[0394]	450	455	460
[0395]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu		
[0396]	465	470	475
[0397]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[0398]	485	490	495
[0399]	His Tyr Asp Ala Arg Ala Cys Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[0400]	500	505	510
[0401]	Ser Ala Val Arg His Val Val Gly Gln Tyr Pro Ser Ser Asp Arg Pro		
[0402]	515	520	525
[0403]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[0404]	530	535	540
[0405]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[0406]	545	550	555
[0407]	Asp Ile Leu Pro Cys Leu His		
[0408]	565		
[0409]	<210> 7		
[0410]	<211> 1701		
[0411]	<212> DNA		
[0412]	<213> 人工序列		
[0413]	<220>		
[0414]	<223> 合成的多核苷酸		
[0415]	<400> 7		
[0416]	atgaaaaccc tgagccaggc acagagcaaa accagcagcc agcagtttag ctttaccggc 60		
[0417]	aatagcagcg caaatgtgat tattggtaat cagaaactga ccatcaatga tggtgcacgt 120		
[0418]	gttggccgta atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt 180		
[0419]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt 240		

[0420]	accagcggtt ttggtggtat ggcaaatggt gcaattagcc gtgaacaggc aagcgaactg	300
[0421]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360
[0422]	gttcgtgcag caatgctgct gcgtgcaa atgcccacatgc gtggtgcaag cggttattcgt	420
[0423]	ctggaactga ttaaaccgat ggaaatcttt ctgaatgccg gtgttaccct gtatgtttat	480
[0424]	gaatttggtg gcatgtgtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc	540
[0425]	ctgattggcc tggacccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600
[0426]	accgcactgc gtcagctgaa tctgagtcgc ctgacctgc tgccgaaaga aggtctggca	660
[0427]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720
[0428]	attctgaccg caattgcaat ggggtgttcat gactggata ttcaggcact gaatggtaca	780
[0429]	aatcagagct ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840
[0430]	gcagatcaga tgattagcct gctggccaat agccagctgg ttcgtgatga actggatggt	900
[0431]	aaacatgatt atcgtgatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960
[0432]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag	1020
[0433]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[0434]	ggtaattttc tgggtcagta tggttggtat ggtatggatc atctgcgcta ttatatcggt	1140
[0435]	ctgctggcaa aacatctgga tggtcagatt gactgctgg catcaccgga atttagcaat	1200
[0436]	ggtctgcctc cgagtctgct gggtaatcgt gaacgtaaag ttaatatggg tctgaaagg	1260
[0437]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[0438]	gatcgttttc cgacctatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[0439]	agcgaacccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg	1440
[0440]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500
[0441]	cgtgcctgtc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560
[0442]	cagaaaccga cctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[0443]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[0444]	gacattctgc cgtgtctgca t	1701
[0445]	<210>	8
[0446]	<211>	567
[0447]	<212>	PRT
[0448]	<213>	人工序列
[0449]	<220>	
[0450]	<223>	合成的多肽
[0451]	<400>	8
[0452]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[0453]	1 5 10 15	
[0454]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[0455]	20 25 30	
[0456]	Leu Thr Ile Asn Asp Val Ala Arg Val Ala Arg Asn Gly Thr Leu Val	
[0457]	35 40 45	
[0458]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys	
[0459]	50 55 60	
[0460]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val	
[0461]	65 70 75 80	

[0462]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Ala Ile Ser Arg Glu Gln		
[0463]		85	90 95
[0464]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala		
[0465]		100	105 110
[0466]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg		
[0467]		115	120 125
[0468]	Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile		
[0469]		130	135 140
[0470]	Lys Arg Met Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr		
[0471]		145	150 155 160
[0472]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr		
[0473]		165	170 175
[0474]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe		
[0475]		180	185 190
[0476]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu		
[0477]		195	200 205
[0478]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly		
[0479]		210	215 220
[0480]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln		
[0481]		225	230 235 240
[0482]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[0483]		245	250 255
[0484]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[0485]		260	265 270
[0486]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[0487]		275	280 285
[0488]	Ala Asn Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[0489]		290	295 300
[0490]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[0491]		305	310 315 320
[0492]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[0493]		325	330 335
[0494]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[0495]		340	345 350
[0496]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[0497]		355	360 365
[0498]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[0499]		370	375 380
[0500]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[0501]		385	390 395 400
[0502]	Gly Leu Pro Pro Ser Leu Leu Gly Asn Arg Glu Arg Lys Val Asn Met		
[0503]		405	410 415

[0504]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu
[0505]	420 425 430
[0506]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu
[0507]	435 440 445
[0508]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu
[0509]	450 455 460
[0510]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu
[0511]	465 470 475 480
[0512]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly
[0513]	485 490 495
[0514]	His Tyr Asp Ala Arg Ala Cys Leu Ser Pro Ala Thr Glu Arg Leu Tyr
[0515]	500 505 510
[0516]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Thr Ser Asp Arg Pro
[0517]	515 520 525
[0518]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg
[0519]	530 535 540
[0520]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln
[0521]	545 550 555 560
[0522]	Asp Ile Leu Pro Cys Leu His
[0523]	565
[0524]	<210> 9
[0525]	<211> 1701
[0526]	<212> DNA
[0527]	<213> 人工序列
[0528]	<220>
[0529]	<223> 合成的多核苷酸
[0530]	<400> 9
[0531]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc 60
[0532]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgtagct 120
[0533]	gttgcccgta atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt 180
[0534]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtggt 240
[0535]	accagcggtt ttggtggtat ggcaaatggt gtaattagcc gtgaacaggc aagcgaactg 300
[0536]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat 360
[0537]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggtattcgt 420
[0538]	ctggaactga ttaaacgcat ggaaatcttt ctgaatgccg gtgttaccct gtatgtttat 480
[0539]	gaatttggtg gcatggtgct cagcgggtgat ctggttccgc tgagctatat taccggtagc 540
[0540]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg 600
[0541]	accgcactgc gtcagctgaa tctgagtccg ctgacctgc tgccgaaaga aggtctggca 660
[0542]	atgatgaatg gcaccagct tatgaccggt attgcagcaa attgtgttta tgataccag 720
[0543]	attctgaccg caattgcaat ggggtgttcat gactggata ttcaggcact gaatggtaca 780
[0544]	aatcagagct ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca 840
[0545]	gcagatcaga tgattagcct gctggccggg agccagctgg ttcgtgatga actggatggt 900

[0546]	aaacatgatt atcgtgatgg tgaactgac caggatcggt atagcctgcg ttgtctgccg	960
[0547]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag	1020
[0548]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[0549]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt	1140
[0550]	ctgctggcaa aacatctgga tgttcagatt gcactgctgg catcaccgga atttagcaat	1200
[0551]	ggctcgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260
[0552]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[0553]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[0554]	agcgcaaccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg	1440
[0555]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccgggtca ttatgatgca	1500
[0556]	cgtgcctgtc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttgttggt	1560
[0557]	cagaaaccga gctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[0558]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[0559]	gacattctgc cgtgtctgca t	1701
[0560]	<210>	10
[0561]	<211>	567
[0562]	<212>	PRT
[0563]	<213>	人工序列
[0564]	<220>	
[0565]	<223>	合成的多肽
[0566]	<400>	10
[0567]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[0568]	1 5 10 15	
[0569]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[0570]	20 25 30	
[0571]	Leu Thr Ile Asn Asp Val Val Arg Val Ala Arg Asn Gly Thr Leu Val	
[0572]	35 40 45	
[0573]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys	
[0574]	50 55 60	
[0575]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val	
[0576]	65 70 75 80	
[0577]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln	
[0578]	85 90 95	
[0579]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala	
[0580]	100 105 110	
[0581]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg	
[0582]	115 120 125	
[0583]	Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile	
[0584]	130 135 140	
[0585]	Lys Arg Met Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr	
[0586]	145 150 155 160	
[0587]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr	

[0588]		165		170		175
[0589]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe					
[0590]		180		185		190
[0591]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu					
[0592]		195		200		205
[0593]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly					
[0594]		210		215		220
[0595]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln					
[0596]	225		230		235	240
[0597]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala					
[0598]		245		250		255
[0599]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys					
[0600]		260		265		270
[0601]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu					
[0602]		275		280		285
[0603]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr					
[0604]		290		295		300
[0605]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro					
[0606]	305		310		315	320
[0607]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln					
[0608]		325		330		335
[0609]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val					
[0610]		340		345		350
[0611]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val					
[0612]		355		360		365
[0613]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys					
[0614]		370		375		380
[0615]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn					
[0616]	385		390		395	400
[0617]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met					
[0618]		405		410		415
[0619]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu					
[0620]		420		425		430
[0621]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu					
[0622]		435		440		445
[0623]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu					
[0624]		450		455		460
[0625]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu					
[0626]	465		470		475	480
[0627]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly					
[0628]		485		490		495
[0629]	His Tyr Asp Ala Arg Ala Cys Leu Ser Pro Ala Thr Glu Arg Leu Tyr					

[0630]	500	505	510
[0631]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Ser Ser Asp Arg Pro		
[0632]	515	520	525
[0633]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[0634]	530	535	540
[0635]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[0636]	545	550	555 560
[0637]	Asp Ile Leu Pro Cys Leu His		
[0638]	565		
[0639]	<210> 11		
[0640]	<211> 1701		
[0641]	<212> DNA		
[0642]	<213> 人工序列		
[0643]	<220>		
[0644]	<223> 合成的多核苷酸		
[0645]	<400> 11		
[0646]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc	60	
[0647]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgtagct	120	
[0648]	gttgcccgta atggcaccct ggtagcctg accaataata aagatattct gcagcgtatt	180	
[0649]	caggccagct gtgattatat caataatgca gttgaaaaag gtgaaccgat ttatggtgtt	240	
[0650]	accagcggtt ttggtggtat ggcaaatgtt gtaattagcc gtgaacagcg aagcgaactg	300	
[0651]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggtg ataaactgcc gctggcagat	360	
[0652]	gttcgtgcag caatgtctgt gcgtgcaaat agccacatgc gtggtgcaag cggtagtctg	420	
[0653]	ctggaactga ttaaacgcat ggaaatcttt ctgaatgccg gtgttaccgc gtatgtttat	480	
[0654]	gaatttggtg gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc	540	
[0655]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600	
[0656]	accgcactgc gtcagctgaa tctgagtccg ctgaccctgc tgccgaaaga aggtctggca	660	
[0657]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720	
[0658]	attctgaccg caattgcaat ggggtgttcat gcaactggata ttcaggcaact gaatggtaca	780	
[0659]	aatcagagct ttcattccgt tatccataac agcaaaccgc atccgggtca gctgtgggca	840	
[0660]	gcagatcaga tgattagcct gctggccggt agccagctgg ttcgtgatga actggatggt	900	
[0661]	aaacatgatt atatggatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960	
[0662]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaatttag	1020	
[0663]	attaacagcg ttaccgataa cccgtgatt gatgttgata atcaggcaag ctatcatggt	1080	
[0664]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt	1140	
[0665]	ctgctggcaa aacatctgga tgttcagatt gcaactgctg catcaccgga atttagcaat	1200	
[0666]	ggtctgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260	
[0667]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320	
[0668]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380	
[0669]	agcgcaaccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg	1440	
[0670]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500	
[0671]	cgtgcccagc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560	

[0672] aaaaaaccga gctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa 1620
 [0673] catattgcac gtattagtgc agatattgca gccggtgggtg ttattgttca ggccgttcag 1680
 [0674] gacattctgc cgccgttgc t 1701
 [0675] <210> 12
 [0676] <211> 567
 [0677] <212> PRT
 [0678] <213> 人工序列
 [0679] <220>
 [0680] <223> 合成的多肽
 [0681] <400> 12
 [0682] Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe
 [0683] 1 5 10 15
 [0684] Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys
 [0685] 20 25 30
 [0686] Leu Thr Ile Asn Asp Val Val Arg Val Ala Arg Asn Gly Thr Leu Val
 [0687] 35 40 45
 [0688] Ser Leu Thr Asn Asn Lys Asp Ile Leu Gln Arg Ile Gln Ala Ser Cys
 [0689] 50 55 60
 [0690] Asp Tyr Ile Asn Asn Ala Val Glu Lys Gly Glu Pro Ile Tyr Gly Val
 [0691] 65 70 75 80
 [0692] Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln
 [0693] 85 90 95
 [0694] Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala
 [0695] 100 105 110
 [0696] Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg
 [0697] 115 120 125
 [0698] Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile
 [0699] 130 135 140
 [0700] Lys Arg Met Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr
 [0701] 145 150 155 160
 [0702] Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr
 [0703] 165 170 175
 [0704] Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe
 [0705] 180 185 190
 [0706] Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu
 [0707] 195 200 205
 [0708] Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly
 [0709] 210 215 220
 [0710] Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln
 [0711] 225 230 235 240
 [0712] Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala
 [0713] 245 250 255

[0714]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[0715]	260	265	270
[0716]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[0717]	275	280	285
[0718]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[0719]	290	295	300
[0720]	Met Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[0721]	305	310	315
[0722]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[0723]	325	330	335
[0724]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[0725]	340	345	350
[0726]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[0727]	355	360	365
[0728]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[0729]	370	375	380
[0730]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[0731]	385	390	395
[0732]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[0733]	405	410	415
[0734]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[0735]	420	425	430
[0736]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[0737]	435	440	445
[0738]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[0739]	450	455	460
[0740]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu		
[0741]	465	470	475
[0742]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[0743]	485	490	495
[0744]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[0745]	500	505	510
[0746]	Ser Ala Val Arg His Val Val Gly Lys Lys Pro Ser Ser Asp Arg Pro		
[0747]	515	520	525
[0748]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[0749]	530	535	540
[0750]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[0751]	545	550	555
[0752]	Asp Ile Leu Pro Pro Leu His		
[0753]	565		
[0754]	<210> 13		
[0755]	<211> 1701		

[0756]	<212> DNA
[0757]	<213> 人工序列
[0758]	<220>
[0759]	<223> 合成的多核苷酸
[0760]	<400> 13
[0761]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc 60
[0762]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgtagct 120
[0763]	gttgcccgta atggcacctt ggtagcctg accaataata aagatattct gcagcgtatt 180
[0764]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtggt 240
[0765]	accagcgggt ttggtggtat ggcaaatgtt gtaattagcc gtgaacaggc aagcgaactg 300
[0766]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat 360
[0767]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggatattct 420
[0768]	ctggaactga ttaaagcgtt ggaaatcttt ctgaatgccg gtgttaccct gtatgtttat 480
[0769]	gaatttggtg gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc 540
[0770]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg 600
[0771]	accgcactgc gtcagctgaa tctgagtcgc ctgaccctgc tgccgaaaga aggtctggca 660
[0772]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag 720
[0773]	attctgaccg caattgcaat ggggtgttcat gcaactggata ttcaggcact gaattgtaca 780
[0774]	aatcagagct ttcattccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca 840
[0775]	gcagatcaga tgattagcct gctggccggg agccagctgg ttcgtgatga actggatggt 900
[0776]	aaacatgatt atatggatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg 960
[0777]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag 1020
[0778]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt 1080
[0779]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt 1140
[0780]	ctgctggcaa aacatctgga tgttcagatt gcaactgctg catcaccgga atttagcaat 1200
[0781]	ggtctgcctc cgagctcgtt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt 1260
[0782]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca 1320
[0783]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc 1380
[0784]	agcgaacccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg 1440
[0785]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca 1500
[0786]	cgtgcccgag tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt 1560
[0787]	aaaaaaccca gctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa 1620
[0788]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag 1680
[0789]	gacattctgc cgaacctgca t 1701
[0790]	<210> 14
[0791]	<211> 567
[0792]	<212> PRT
[0793]	<213> 人工序列
[0794]	<220>
[0795]	<223> 合成的多肽
[0796]	<400> 14
[0797]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe

[0798]	1	5	10	15
[0799]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys			
[0800]	20	25	30	
[0801]	Leu Thr Ile Asn Asp Val Val Arg Val Ala Arg Asn Gly Thr Leu Val			
[0802]	35	40	45	
[0803]	Ser Leu Thr Asn Asn Lys Asp Ile Leu Gln Arg Ile Gln Ala Ser Cys			
[0804]	50	55	60	
[0805]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val			
[0806]	65	70	75	80
[0807]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln			
[0808]	85	90	95	
[0809]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala			
[0810]	100	105	110	
[0811]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg			
[0812]	115	120	125	
[0813]	Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile			
[0814]	130	135	140	
[0815]	Lys Arg Met Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr			
[0816]	145	150	155	160
[0817]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr			
[0818]	165	170	175	
[0819]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe			
[0820]	180	185	190	
[0821]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu			
[0822]	195	200	205	
[0823]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly			
[0824]	210	215	220	
[0825]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln			
[0826]	225	230	235	240
[0827]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala			
[0828]	245	250	255	
[0829]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys			
[0830]	260	265	270	
[0831]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu			
[0832]	275	280	285	
[0833]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr			
[0834]	290	295	300	
[0835]	Met Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro			
[0836]	305	310	315	320
[0837]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln			
[0838]	325	330	335	
[0839]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val			

[0840]	340	345	350
[0841]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[0842]	355	360	365
[0843]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[0844]	370	375	380
[0845]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[0846]	385	390	395
[0847]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[0848]	405	410	415
[0849]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[0850]	420	425	430
[0851]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[0852]	435	440	445
[0853]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[0854]	450	455	460
[0855]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu		
[0856]	465	470	475
[0857]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[0858]	485	490	495
[0859]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[0860]	500	505	510
[0861]	Ser Ala Val Arg His Val Val Gly Lys Lys Pro Ser Ser Asp Arg Pro		
[0862]	515	520	525
[0863]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[0864]	530	535	540
[0865]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[0866]	545	550	555
[0867]	Asp Ile Leu Pro Asn Leu His		
[0868]	565		
[0869]	<210> 15		
[0870]	<211> 1701		
[0871]	<212> DNA		
[0872]	<213> 人工序列		
[0873]	<220>		
[0874]	<223> 合成的多核苷酸		
[0875]	<400> 15		
[0876]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc 60		
[0877]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgcgcg 120		
[0878]	gttgcccgta atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt 180		
[0879]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt 240		
[0880]	accagcgggt ttggtggtat ggcaaatgtt gtaattagcc gtgaacaggc aagcgaactg 300		
[0881]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat 360		

[0882]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc atggtgcaag cggtattcgt	420
[0883]	ctggaactga ttaaaccgcgc ggaaatcttt ctgaatgccg gtgttacccc gtatgtttat	480
[0884]	gaatttggtg gcattggtgc cagcggatgat ctggttccgc tgagctatat taccggtagc	540
[0885]	ctgattggcc tggacccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600
[0886]	accgcactgc gtcagctgaa tctgagtcgc ctgaccctgc tgccgaaaga aggtctggca	660
[0887]	atgatgaatg gcaccagcgt tatgaccggg attgcagcaa attgtgttta tgatacccg	720
[0888]	attctgaccg caattgcaat ggggtgttcat gactggata ttcaggcact gaatggtaca	780
[0889]	aatcagagct ttcattccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840
[0890]	gcagatcaga tgattagcct gctggccggg agccagctgg ttcgtgatga actggatggt	900
[0891]	aaacatgatt atcgtgatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960
[0892]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag	1020
[0893]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[0894]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt	1140
[0895]	ggcctggcaa aacatctgga tgttcagatt gcaactgctg catcaccgga atttagcaat	1200
[0896]	ggtctgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260
[0897]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[0898]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[0899]	agcgcaaccc tggcacgtcg tagcgttgat attggccaga attatgttgc cattgccctg	1440
[0900]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccgggtca ttatgatgca	1500
[0901]	cgtgcccgag tgtcaccgga aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560
[0902]	cagaaaccga gctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[0903]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[0904]	gacattctgc cgaacctgca t	1701
[0905]	<210>	16
[0906]	<211>	567
[0907]	<212>	PRT
[0908]	<213>	人工序列
[0909]	<220>	
[0910]	<223>	合成的多肽
[0911]	<400>	16
[0912]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[0913]	1 5 10 15	
[0914]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[0915]	20 25 30	
[0916]	Leu Thr Ile Asn Asp Val Ala Arg Val Ala Arg Asn Gly Thr Leu Val	
[0917]	35 40 45	
[0918]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys	
[0919]	50 55 60	
[0920]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val	
[0921]	65 70 75 80	
[0922]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln	
[0923]	85 90 95	

[0924]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala		
[0925]	100	105	110
[0926]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg		
[0927]	115	120	125
[0928]	Ala Asn Ser His Met His Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile		
[0929]	130	135	140
[0930]	Lys Arg Ala Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr		
[0931]	145	150	155
[0932]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr		
[0933]	165	170	175
[0934]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe		
[0935]	180	185	190
[0936]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu		
[0937]	195	200	205
[0938]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly		
[0939]	210	215	220
[0940]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln		
[0941]	225	230	235
[0942]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[0943]	245	250	255
[0944]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[0945]	260	265	270
[0946]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[0947]	275	280	285
[0948]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[0949]	290	295	300
[0950]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[0951]	305	310	315
[0952]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[0953]	325	330	335
[0954]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[0955]	340	345	350
[0956]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[0957]	355	360	365
[0958]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Gly Leu Ala Lys		
[0959]	370	375	380
[0960]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[0961]	385	390	395
[0962]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[0963]	405	410	415
[0964]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[0965]	420	425	430

[0966]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu	
[0967]	435	440 445
[0968]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu	
[0969]	450	455 460
[0970]	Ala Arg Arg Ser Val Asp Ile Gly Gln Asn Tyr Val Ala Ile Ala Leu	
[0971]	465	470 475 480
[0972]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly	
[0973]	485	490 495
[0974]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr	
[0975]	500	505 510
[0976]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Ser Ser Asp Arg Pro	
[0977]	515	520 525
[0978]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg	
[0979]	530	535 540
[0980]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln	
[0981]	545	550 555 560
[0982]	Asp Ile Leu Pro Asn Leu His	
[0983]	565	
[0984]	<210> 17	
[0985]	<211> 1701	
[0986]	<212> DNA	
[0987]	<213> 人工序列	
[0988]	<220>	
[0989]	<223> 合成的多核苷酸	
[0990]	<400> 17	
[0991]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc	60
[0992]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tggtgcgcgt	120
[0993]	gttgcctgta atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt	180
[0994]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt	240
[0995]	accagcgggt ttggtggtat ggcaaatgtt gtaattagcc gtgaacagcg aagcgaactg	300
[0996]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360
[0997]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc atggtgcaag cggtattcgt	420
[0998]	ctggaactga ttaaacgcgc ggaaatcttt ctgaatgccg gtgttacccc gtatgtttat	480
[0999]	gaatttggtg gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc	540
[1000]	ctgattggcc tggacccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600
[1001]	accgcaactgc gtcagctgaa tctgagtcgc ctgaccctgc tgccgaaaga aggtctggca	660
[1002]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720
[1003]	attctgaccg caattgcaat ggggtgttcat gactggata ttcaggcact gaatggtaca	780
[1004]	aatcagagct ttcacccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840
[1005]	gcagatcaga tgattagcct gctggccggg agccagctgg ttcgtgatga actggatggt	900
[1006]	aaacatgatt atcgtgatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960
[1007]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag	1020

[1008]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[1009]	ggtaattttc tgggtcagta tggttgatg ggtatggatc atctgcgcta ttatatcggt	1140
[1010]	ggcctggcaa aacatctgga taccagatt gactgctgg catcaccgga atttagcaat	1200
[1011]	ggtctgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260
[1012]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[1013]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[1014]	agcgcaaccc tggcacgtcg tagcgttgat attggccaga attatgttgc cattgccctg	1440
[1015]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500
[1016]	cgtgcccagc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560
[1017]	cagaaaccga gtcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[1018]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[1019]	gacattctgc cgaacctgca t	1701
[1020]	<210>	18
[1021]	<211>	567
[1022]	<212>	PRT
[1023]	<213>	人工序列
[1024]	<220>	
[1025]	<223>	合成的多肽
[1026]	<400>	18
[1027]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[1028]	1 5 10 15	
[1029]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[1030]	20 25 30	
[1031]	Leu Thr Ile Asn Asp Val Ala Arg Val Ala Arg Asn Gly Thr Leu Val	
[1032]	35 40 45	
[1033]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys	
[1034]	50 55 60	
[1035]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val	
[1036]	65 70 75 80	
[1037]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln	
[1038]	85 90 95	
[1039]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala	
[1040]	100 105 110	
[1041]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg	
[1042]	115 120 125	
[1043]	Ala Asn Ser His Met His Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile	
[1044]	130 135 140	
[1045]	Lys Arg Ala Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr	
[1046]	145 150 155 160	
[1047]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr	
[1048]	165 170 175	
[1049]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe	

[1050]	180	185	190
[1051]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu		
[1052]	195	200	205
[1053]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly		
[1054]	210	215	220
[1055]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln		
[1056]	225	230	235
[1057]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[1058]	245	250	255
[1059]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[1060]	260	265	270
[1061]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[1062]	275	280	285
[1063]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[1064]	290	295	300
[1065]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[1066]	305	310	315
[1067]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[1068]	325	330	335
[1069]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1070]	340	345	350
[1071]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[1072]	355	360	365
[1073]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Gly Leu Ala Lys		
[1074]	370	375	380
[1075]	His Leu Asp Thr Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[1076]	385	390	395
[1077]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[1078]	405	410	415
[1079]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[1080]	420	425	430
[1081]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[1082]	435	440	445
[1083]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[1084]	450	455	460
[1085]	Ala Arg Arg Ser Val Asp Ile Gly Gln Asn Tyr Val Ala Ile Ala Leu		
[1086]	465	470	475
[1087]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[1088]	485	490	495
[1089]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[1090]	500	505	510
[1091]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Ser Ser Asp Arg Pro		

[1092]	515	520	525
[1093]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[1094]	530	535	540
[1095]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[1096]	545	550	555 560
[1097]	Asp Ile Leu Pro Asn Leu His		
[1098]	565		
[1099]	<210> 19		
[1100]	<211> 1701		
[1101]	<212> DNA		
[1102]	<213> 人工序列		
[1103]	<220>		
[1104]	<223> 合成的多核苷酸		
[1105]	<400> 19		
[1106]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc	60	
[1107]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgcgcg	120	
[1108]	gttgcccgta atggcacctt ggtagcctg accaataata ccgatattct gcagggtatt	180	
[1109]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtggt	240	
[1110]	accagcgggt ttggtggtat ggcaaatggt gtaattagcc gtgaacaggc aagcgaactg	300	
[1111]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360	
[1112]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc atggtgcaag cggtattcgt	420	
[1113]	ctggaactga ttaaacgcgc ggaaatcttt ctgaatgccg gtgttaccct gtatgtttat	480	
[1114]	gaatttggtg gcatgtgtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc	540	
[1115]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600	
[1116]	accgcactgc gtcagctgaa tctgagtccg ctgacctgc tgccgaaaga aggtctggca	660	
[1117]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720	
[1118]	attctgaccg caattgcaat ggggtgttcat gactggata ttcaggcact gaatggtaca	780	
[1119]	aatcagagct ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840	
[1120]	gcagatcaga tgattagcct gctggccggt agccagctgg ttcgtgatga actggatggt	900	
[1121]	aaacatgatt atcgtgatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960	
[1122]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag	1020	
[1123]	attaacagcg ttaccgataa cccgtgatt gatgttgata atcaggcaag ctatcatggt	1080	
[1124]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta tgaaatcgg	1140	
[1125]	ctgctggcaa aacatctgga tgttcagatt gcaactgctg catcaccgga atttagcaat	1200	
[1126]	ggtctgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260	
[1127]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320	
[1128]	gatcgttttc cgacctatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380	
[1129]	agcgcaaccc tggcacgtcg tagcgttgat attggccaga attatgttgc cattgccttg	1440	
[1130]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500	
[1131]	cgtgcccagc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560	
[1132]	cagaaaccga gtcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620	
[1133]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680	

[1134]	gacattctgc cgaacctgca t 1701															
[1135]	<210> 20															
[1136]	<211> 567															
[1137]	<212> PRT															
[1138]	<213> 人工序列															
[1139]	<220>															
[1140]	<223> 合成的多肽															
[1141]	<400> 20															
[1142]	Met	Lys	Thr	Leu	Ser	Gln	Ala	Gln	Ser	Lys	Thr	Ser	Ser	Gln	Gln	Phe
[1143]	1				5					10					15	
[1144]	Ser	Phe	Thr	Gly	Asn	Ser	Ser	Ala	Asn	Val	Ile	Ile	Gly	Asn	Gln	Lys
[1145]				20					25					30		
[1146]	Leu	Thr	Ile	Asn	Asp	Val	Ala	Arg	Val	Ala	Arg	Asn	Gly	Thr	Leu	Val
[1147]			35					40					45			
[1148]	Ser	Leu	Thr	Asn	Asn	Thr	Asp	Ile	Leu	Gln	Gly	Ile	Gln	Ala	Ser	Cys
[1149]		50					55				60					
[1150]	Asp	Tyr	Ile	Asn	Asn	Ala	Val	Glu	Ser	Gly	Glu	Pro	Ile	Tyr	Gly	Val
[1151]	65					70					75					80
[1152]	Thr	Ser	Gly	Phe	Gly	Gly	Met	Ala	Asn	Val	Val	Ile	Ser	Arg	Glu	Gln
[1153]				85						90					95	
[1154]	Ala	Ser	Glu	Leu	Gln	Thr	Asn	Leu	Val	Trp	Phe	Leu	Lys	Thr	Gly	Ala
[1155]				100					105					110		
[1156]	Gly	Asn	Lys	Leu	Pro	Leu	Ala	Asp	Val	Arg	Ala	Ala	Met	Leu	Leu	Arg
[1157]			115					120					125			
[1158]	Ala	Asn	Ser	His	Met	His	Gly	Ala	Ser	Gly	Ile	Arg	Leu	Glu	Leu	Ile
[1159]		130						135					140			
[1160]	Lys	Arg	Ala	Glu	Ile	Phe	Leu	Asn	Ala	Gly	Val	Thr	Pro	Tyr	Val	Tyr
[1161]	145					150					155					160
[1162]	Glu	Phe	Gly	Ser	Ile	Gly	Ala	Ser	Gly	Asp	Leu	Val	Pro	Leu	Ser	Tyr
[1163]				165						170						175
[1164]	Ile	Thr	Gly	Ser	Leu	Ile	Gly	Leu	Asp	Pro	Ser	Phe	Lys	Val	Asp	Phe
[1165]				180					185					190		
[1166]	Asn	Gly	Lys	Glu	Met	Asp	Ala	Pro	Thr	Ala	Leu	Arg	Gln	Leu	Asn	Leu
[1167]			195					200					205			
[1168]	Ser	Pro	Leu	Thr	Leu	Leu	Pro	Lys	Glu	Gly	Leu	Ala	Met	Met	Asn	Gly
[1169]		210						215				220				
[1170]	Thr	Ser	Val	Met	Thr	Gly	Ile	Ala	Ala	Asn	Cys	Val	Tyr	Asp	Thr	Gln
[1171]	225					230					235					240
[1172]	Ile	Leu	Thr	Ala	Ile	Ala	Met	Gly	Val	His	Ala	Leu	Asp	Ile	Gln	Ala
[1173]				245						250						255
[1174]	Leu	Asn	Gly	Thr	Asn	Gln	Ser	Phe	His	Pro	Phe	Ile	His	Asn	Ser	Lys
[1175]				260					265					270		

[1176]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[1177]	275	280	285
[1178]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[1179]	290	295	300
[1180]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[1181]	305	310	315
[1182]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[1183]	325	330	335
[1184]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1185]	340	345	350
[1186]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[1187]	355	360	365
[1188]	Gly Met Gly Met Asp His Leu Arg Tyr Glu Ile Gly Leu Leu Ala Lys		
[1189]	370	375	380
[1190]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[1191]	385	390	395
[1192]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[1193]	405	410	415
[1194]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[1195]	420	425	430
[1196]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[1197]	435	440	445
[1198]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[1199]	450	455	460
[1200]	Ala Arg Arg Ser Val Asp Ile Gly Gln Asn Tyr Val Ala Ile Ala Leu		
[1201]	465	470	475
[1202]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[1203]	485	490	495
[1204]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[1205]	500	505	510
[1206]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Ser Ser Asp Arg Pro		
[1207]	515	520	525
[1208]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[1209]	530	535	540
[1210]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[1211]	545	550	555
[1212]	Asp Ile Leu Pro Asn Leu His		
[1213]	565		
[1214]	<210> 21		
[1215]	<211> 1701		
[1216]	<212> DNA		
[1217]	<213> 人工序列		

[1218]	<220>	
[1219]	<223>	合成的多核苷酸
[1220]	<400>	21
[1221]	atgaaaaccc	tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc 60
[1222]	aatagcagcg	caaatgtgat tattggtaat cagaaactga ccatcaatga tgttgcgcg 120
[1223]	gttgcccgta	atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt 180
[1224]	caggccagct	gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt 240
[1225]	accagcggtt	ttggtggtat ggcaaatgtt gtaattagcc gtgaacagcg aagcgaactg 300
[1226]	cagaccaatc	tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat 360
[1227]	gttcgtgcag	caatgtctgt gcgtgcaaat agccacatgc atggtgcaag cggatttcgt 420
[1228]	ctggaactga	ttaaaccgcg ggaaatcttt ctgaatgccg gtgttaccgc gtatgtttat 480
[1229]	gaatttggtg	gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc 540
[1230]	ctgattggcc	tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg 600
[1231]	accgcactgc	gtcagctgaa tctgagtcgc ctgaccctgc tgccgaaaga aggtctggca 660
[1232]	atgatgaatg	gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag 720
[1233]	attctgaccg	caattgcaat ggggtgttcat gcaactggata ttcaggcact gaatggtaca 780
[1234]	aatcagagct	ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca 840
[1235]	gcagatcaga	tgattagcct gctggccggt agccagctgg ttcgtgatga actggatggt 900
[1236]	aaacatgatt	atcgtgatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg 960
[1237]	cagtatctgg	gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag 1020
[1238]	attaacagcg	ttaccgataa cccgtgatt gatgttgata atcaggcaag ctatcatggt 1080
[1239]	ggtaattttc	tgggtcagta tgttggtatg ggtatggatc atctgcgcta tgaaatcgg 1140
[1240]	ctgctggcaa	aacatctgga taccagatt gcaactgctg catcacggga atttagcaat 1200
[1241]	ggctctgcctc	cgagctctgg gggtaatcgt gaacgtaaag ttaatatggg tctgaaagg 1260
[1242]	ctgcagattt	gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca 1320
[1243]	gatcgttttc	cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc 1380
[1244]	agcgcaaccc	tggcacgtcg tagcgttgat attggccaga attatgttgc cattgccctg 1440
[1245]	atgtttggtg	ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca 1500
[1246]	cgtgcccgag	tgaccacggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt 1560
[1247]	cagaaaccga	gctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa 1620
[1248]	catattgcac	gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag 1680
[1249]	gacattctgc	cgaacctgca t 1701
[1250]	<210>	22
[1251]	<211>	567
[1252]	<212>	PRT
[1253]	<213>	人工序列
[1254]	<220>	
[1255]	<223>	合成的多肽
[1256]	<400>	22
[1257]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[1258]	1	5 10 15
[1259]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	

[1260]	20	25	30
[1261]	Leu Thr Ile Asn Asp Val Ala Arg Val Ala Arg Asn Gly Thr Leu Val		
[1262]	35	40	45
[1263]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys		
[1264]	50	55	60
[1265]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val		
[1266]	65	70	75
[1267]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln		
[1268]	85	90	95
[1269]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala		
[1270]	100	105	110
[1271]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg		
[1272]	115	120	125
[1273]	Ala Asn Ser His Met His Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile		
[1274]	130	135	140
[1275]	Lys Arg Ala Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr		
[1276]	145	150	155
[1277]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr		
[1278]	165	170	175
[1279]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe		
[1280]	180	185	190
[1281]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu		
[1282]	195	200	205
[1283]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly		
[1284]	210	215	220
[1285]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln		
[1286]	225	230	235
[1287]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[1288]	245	250	255
[1289]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[1290]	260	265	270
[1291]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[1292]	275	280	285
[1293]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[1294]	290	295	300
[1295]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[1296]	305	310	315
[1297]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[1298]	325	330	335
[1299]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1300]	340	345	350
[1301]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		

[1302]	355	360	365
[1303]	Gly Met Gly Met Asp His Leu Arg Tyr Glu Ile Gly Leu Leu Ala Lys		
[1304]	370	375	380
[1305]	His Leu Asp Thr Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[1306]	385	390	395 400
[1307]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[1308]	405	410	415
[1309]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[1310]	420	425	430
[1311]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[1312]	435	440	445
[1313]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[1314]	450	455	460
[1315]	Ala Arg Arg Ser Val Asp Ile Gly Gln Asn Tyr Val Ala Ile Ala Leu		
[1316]	465	470	475 480
[1317]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[1318]	485	490	495
[1319]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[1320]	500	505	510
[1321]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Ser Ser Asp Arg Pro		
[1322]	515	520	525
[1323]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		
[1324]	530	535	540
[1325]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[1326]	545	550	555 560
[1327]	Asp Ile Leu Pro Asn Leu His		
[1328]	565		
[1329]	<210> 23		
[1330]	<211> 1701		
[1331]	<212> DNA		
[1332]	<213> 人工序列		
[1333]	<220>		
[1334]	<223> 合成的多核苷酸		
[1335]	<400> 23		
[1336]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ccataccggc	60	
[1337]	aatagcagcg caaatgtgat tattgtaaat cagaaactga ccatcaatga tgttgtacgt	120	
[1338]	gttgcccgtg atggcaccgc ggtagcctg accaataata aagatattct gcagcgtatt	180	
[1339]	cagccagct gtgattatat caataatgca gttgaaaaag gtgaaccgat ttatggtgtt	240	
[1340]	accagcggtt ttggtggtat ggcaaatgtt gtaattagcc gtgaacaggc aagcgaactg	300	
[1341]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360	
[1342]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggatattcgt	420	
[1343]	ctggaactga ttaaacgcat ggaaatcttt ctgaatgccg gtgttacccc gtatgtttat	480	

[1344]	gaatttggtgta gcattggtgc cagcgggtgat ctggttccgc tgagctatat taccggtagc	540
[1345]	ctgattggcc tggacccgag ctttaaagtt gatttttaatg gcaaagaaat ggacgcaccg	600
[1346]	accgcactgc gtcagctgaa tctgagtcgc ctgaccctgc agccgaaaga aggtctggca	660
[1347]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720
[1348]	attctgaccg caattgcaat ggggtgttcat gcaactggata ttcaggcact gaatggtaca	780
[1349]	aatcagagct ttcattccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840
[1350]	gcagatcaga tgattagcct gctggccggt agccagctgg ttcgtgatga actggatggt	900
[1351]	aaacatgatt atatggatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960
[1352]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaaat cgaaattgag	1020
[1353]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[1354]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt	1140
[1355]	ctgctggcaa aacatctgga tgttcagatt gcaactgctg catcaccgga atttagcaat	1200
[1356]	ggtctgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaagg	1260
[1357]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[1358]	gatcgttttc cgacctatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[1359]	agcgcaaccc tggcacgtcg tagcgttgat atttttcaga attatgttgc cattgccctg	1440
[1360]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500
[1361]	cgtgcccagc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560
[1362]	aaaaaaccga gtcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[1363]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[1364]	gacattctgc cgccgctgca t	1701
[1365]	<210>	24
[1366]	<211>	567
[1367]	<212>	PRT
[1368]	<213>	人工序列
[1369]	<220>	
[1370]	<223>	合成的多肽
[1371]	<400>	24
[1372]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[1373]	1 5 10 15	
[1374]	Ser His Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[1375]	20 25 30	
[1376]	Leu Thr Ile Asn Asp Val Val Arg Val Ala Arg Asn Gly Thr Ala Val	
[1377]	35 40 45	
[1378]	Ser Leu Thr Asn Asn Lys Asp Ile Leu Gln Arg Ile Gln Ala Ser Cys	
[1379]	50 55 60	
[1380]	Asp Tyr Ile Asn Asn Ala Val Glu Lys Gly Glu Pro Ile Tyr Gly Val	
[1381]	65 70 75 80	
[1382]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln	
[1383]	85 90 95	
[1384]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala	
[1385]	100 105 110	

[1386]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg		
[1387]	115	120	125
[1388]	Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile		
[1389]	130	135	140
[1390]	Lys Arg Met Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr		
[1391]	145	150	155
[1392]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr		
[1393]	165	170	175
[1394]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe		
[1395]	180	185	190
[1396]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu		
[1397]	195	200	205
[1398]	Ser Pro Leu Thr Leu Gln Pro Lys Glu Gly Leu Ala Met Met Asn Gly		
[1399]	210	215	220
[1400]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln		
[1401]	225	230	235
[1402]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[1403]	245	250	255
[1404]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[1405]	260	265	270
[1406]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[1407]	275	280	285
[1408]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[1409]	290	295	300
[1410]	Met Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[1411]	305	310	315
[1412]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[1413]	325	330	335
[1414]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1415]	340	345	350
[1416]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[1417]	355	360	365
[1418]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[1419]	370	375	380
[1420]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[1421]	385	390	395
[1422]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[1423]	405	410	415
[1424]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[1425]	420	425	430
[1426]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[1427]	435	440	445

[1428]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu	
[1429]	450	455 460
[1430]	Ala Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Val Ala Ile Ala Leu	
[1431]	465	470 475 480
[1432]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly	
[1433]	485	490 495
[1434]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr	
[1435]	500	505 510
[1436]	Ser Ala Val Arg His Val Val Gly Lys Lys Pro Ser Ser Asp Arg Pro	
[1437]	515	520 525
[1438]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg	
[1439]	530	535 540
[1440]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln	
[1441]	545	550 555 560
[1442]	Asp Ile Leu Pro Pro Leu His	
[1443]	565	
[1444]	<210> 25	
[1445]	<211> 1701	
[1446]	<212> DNA	
[1447]	<213> 人工序列	
[1448]	<220>	
[1449]	<223> 合成的多核苷酸	
[1450]	<400> 25	
[1451]	atgaaaaccc tgagtcaggc acagagcaaa accagcagcc agcagtttag ctttaccggc	60
[1452]	aatagcagcg caaatgtgat tatttgtaat cagaaactga ccatcaatga tgttgcgcg	120
[1453]	gttgcccgta atggcaccct ggtagcctg accaataata ccgatattct gcagggtatt	180
[1454]	caggccagct gtgattatat caataatgca gttgaaagcg gtgaaccgat ttatggtgtt	240
[1455]	accagcgggt ttggtggtat ggcaaatgtt gtaattagcc gtgaacaggc aagcgaactg	300
[1456]	cagaccaatc tggtttggtt tctgaaaacc ggtgcaggta ataaactgcc gctggcagat	360
[1457]	gttcgtgcag caatgctgct gcgtgcaaat agccacatgc gtggtgcaag cggatattcgt	420
[1458]	ctggaactga ttaaacgcgc ggaaatcttt ctgaatgccg gtgttaccgc gtatgtttat	480
[1459]	gaatttggtg gcattggtgc cagcggtagt ctggttccgc tgagctatat taccggtagc	540
[1460]	ctgattggcc tggaccgag ctttaaagtt gattttaatg gcaaagaaat ggacgcaccg	600
[1461]	accgcactgc gtcagctgaa tctgagtcgc ctgaccctgc tgccgaaaga aggtctggca	660
[1462]	atgatgaatg gcaccagcgt tatgaccggt attgcagcaa attgtgttta tgataccag	720
[1463]	attctgaccg caattgcaat ggggtgttcat gactggata ttcaggcact gaatggtaca	780
[1464]	aatcagagct ttcatccgtt tatccataac agcaaaccgc atccgggtca gctgtgggca	840
[1465]	gcagatcaga tgattagcct gctggccgggt agccagctgg ttcgtgatga actggtggt	900
[1466]	aaacatgatt atcgtgatgg tgaactgac caggatcgtt atagcctgcg ttgtctgccg	960
[1467]	cagtatctgg gtccgattgt tgatggtatt agccagattg ccaaacaat cgaaattgag	1020
[1468]	attaacagcg ttaccgataa cccgctgatt gatgttgata atcaggcaag ctatcatggt	1080
[1469]	ggtaattttc tgggtcagta tgttggtatg ggtatggatc atctgcgcta ttatatcggt	1140

[1470]	ggcctggcaa aacatctgga tggtcagatt gcactgctgg catcaccgga atttagcaat	1200
[1471]	ggtctgcctc cgagtctggt gggtaatcgt gaacgtaaag ttaatatggg tctgaaaggt	1260
[1472]	ctgcagattt gcggtaatag cattatgccg ctgctgacct tttatggtaa tagtattgca	1320
[1473]	gatcgttttc cgacccatgc cgaacagttt aaccagaata ttaacagcca gggttatacc	1380
[1474]	agcgcaaccc tggcacgtcg tagcgttgat attggccaga attatgttgc cattgccctg	1440
[1475]	atgtttggtg ttcaggcagt tgatctgcgt acctacaaaa aaaccggtca ttatgatgca	1500
[1476]	cgtgcccagc tgtcaccggc aaccgaacgt ctgtatagcg cagttcgtca tgttggtggt	1560
[1477]	cagaaaccga gctcagatcg tccgtatatt tggaatgata atgaacaggg tctggatgaa	1620
[1478]	catattgcac gtattagtgc agatattgca gccggtggtg ttattgttca ggccgttcag	1680
[1479]	gacattctgc cgaacctgca t	1701
[1480]	<210>	26
[1481]	<211>	567
[1482]	<212>	PRT
[1483]	<213>	人工序列
[1484]	<220>	
[1485]	<223>	合成的多肽
[1486]	<400>	26
[1487]	Met Lys Thr Leu Ser Gln Ala Gln Ser Lys Thr Ser Ser Gln Gln Phe	
[1488]	1 5 10 15	
[1489]	Ser Phe Thr Gly Asn Ser Ser Ala Asn Val Ile Ile Gly Asn Gln Lys	
[1490]	20 25 30	
[1491]	Leu Thr Ile Asn Asp Val Ala Arg Val Ala Arg Asn Gly Thr Leu Val	
[1492]	35 40 45	
[1493]	Ser Leu Thr Asn Asn Thr Asp Ile Leu Gln Gly Ile Gln Ala Ser Cys	
[1494]	50 55 60	
[1495]	Asp Tyr Ile Asn Asn Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val	
[1496]	65 70 75 80	
[1497]	Thr Ser Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Arg Glu Gln	
[1498]	85 90 95	
[1499]	Ala Ser Glu Leu Gln Thr Asn Leu Val Trp Phe Leu Lys Thr Gly Ala	
[1500]	100 105 110	
[1501]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg	
[1502]	115 120 125	
[1503]	Ala Asn Ser His Met Arg Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile	
[1504]	130 135 140	
[1505]	Lys Arg Ala Glu Ile Phe Leu Asn Ala Gly Val Thr Pro Tyr Val Tyr	
[1506]	145 150 155 160	
[1507]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr	
[1508]	165 170 175	
[1509]	Ile Thr Gly Ser Leu Ile Gly Leu Asp Pro Ser Phe Lys Val Asp Phe	
[1510]	180 185 190	
[1511]	Asn Gly Lys Glu Met Asp Ala Pro Thr Ala Leu Arg Gln Leu Asn Leu	

[1512]	195	200	205
[1513]	Ser Pro Leu Thr Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly		
[1514]	210	215	220
[1515]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Thr Gln		
[1516]	225	230	235
[1517]	Ile Leu Thr Ala Ile Ala Met Gly Val His Ala Leu Asp Ile Gln Ala		
[1518]	245	250	255
[1519]	Leu Asn Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Asn Ser Lys		
[1520]	260	265	270
[1521]	Pro His Pro Gly Gln Leu Trp Ala Ala Asp Gln Met Ile Ser Leu Leu		
[1522]	275	280	285
[1523]	Ala Gly Ser Gln Leu Val Arg Asp Glu Leu Asp Gly Lys His Asp Tyr		
[1524]	290	295	300
[1525]	Arg Asp Gly Glu Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[1526]	305	310	315
[1527]	Gln Tyr Leu Gly Pro Ile Val Asp Gly Ile Ser Gln Ile Ala Lys Gln		
[1528]	325	330	335
[1529]	Ile Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1530]	340	345	350
[1531]	Asp Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[1532]	355	360	365
[1533]	Gly Met Gly Met Asp His Leu Arg Tyr Tyr Ile Gly Gly Leu Ala Lys		
[1534]	370	375	380
[1535]	His Leu Asp Val Gln Ile Ala Leu Leu Ala Ser Pro Glu Phe Ser Asn		
[1536]	385	390	395
[1537]	Gly Leu Pro Pro Ser Leu Val Gly Asn Arg Glu Arg Lys Val Asn Met		
[1538]	405	410	415
[1539]	Gly Leu Lys Gly Leu Gln Ile Cys Gly Asn Ser Ile Met Pro Leu Leu		
[1540]	420	425	430
[1541]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[1542]	435	440	445
[1543]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Thr Ser Ala Thr Leu		
[1544]	450	455	460
[1545]	Ala Arg Arg Ser Val Asp Ile Gly Gln Asn Tyr Val Ala Ile Ala Leu		
[1546]	465	470	475
[1547]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Lys Thr Gly		
[1548]	485	490	495
[1549]	His Tyr Asp Ala Arg Ala Gln Leu Ser Pro Ala Thr Glu Arg Leu Tyr		
[1550]	500	505	510
[1551]	Ser Ala Val Arg His Val Val Gly Gln Lys Pro Ser Ser Asp Arg Pro		
[1552]	515	520	525
[1553]	Tyr Ile Trp Asn Asp Asn Glu Gln Gly Leu Asp Glu His Ile Ala Arg		

[1554]	530	535	540
[1555]	Ile Ser Ala Asp Ile Ala Ala Gly Gly Val Ile Val Gln Ala Val Gln		
[1556]	545	550	555 560
[1557]	Asp Ile Leu Pro Asn Leu His		
[1558]	565		
[1559]	<210> 27		
[1560]	<211> 4945		
[1561]	<212> DNA		
[1562]	<213> 人工序列		
[1563]	<220>		
[1564]	<223> 合成的多核苷酸		
[1565]	<220>		
[1566]	<221> misc_feature		
[1567]	<222> (140) .. (142)		
[1568]	<223> n是a、c、g或t		
[1569]	<220>		
[1570]	<221> misc_feature		
[1571]	<222> (1758) .. (1758)		
[1572]	<223> n是a、c、g或t		
[1573]	<220>		
[1574]	<221> misc_feature		
[1575]	<222> (1761) .. (1762)		
[1576]	<223> n是a、c、g或t		
[1577]	<220>		
[1578]	<221> misc_feature		
[1579]	<222> (1781) .. (1781)		
[1580]	<223> n是a、c、g或t		
[1581]	<220>		
[1582]	<221> misc_feature		
[1583]	<222> (1784) .. (1784)		
[1584]	<223> n是a、c、g或t		
[1585]	<220>		
[1586]	<221> misc_feature		
[1587]	<222> (1786) .. (1786)		
[1588]	<223> n是a、c、g或t		
[1589]	<400> 27		
[1590]	tgccaccat caccatcacc attaggaag agcagatggg caagcttgac ctgtgaagtg 60		
[1591]	aaaaatggcg cacattgtgc gacatttttt ttgaattct acgtaaaaag cagccgatac 120		
[1592]	atcggtgct ttttttttgn nngaggttcc aacttgttgt ataataaaat aagatcactc 180		
[1593]	cggagcgtat tttttgagtt atcgagattt tcaggagcta aggaggaaact aaaatggaga 240		
[1594]	aaaaaatcac tggatatacc accgttgata tatcccaatg gcatcgtaaa gaacattttg 300		
[1595]	aggcatttca gtcagttgct caatgtacct ataaccagac cgttcagctg gatattacgg 360		

[1596]	cctttttaaa gaccgtaaag aaaaataagc acaagtttta tccggccttt attcacattc	420
[1597]	ttgccccgct gatgaatgct catccggagt tccgtatggc aatgaaagac ggtgagctgg	480
[1598]	tgatatggga tagtggtcac cttgtttaca cegttttcca tgagcaaact gaaacgtttt	540
[1599]	catcgctctg gagtgaatac cagcagcatt tccggcagtt tctacacata tattcgcaag	600
[1600]	atgtggcgtg ttacggtgaa aacctggcct atttccctaa agggttttatt gagaatatgt	660
[1601]	ttttcgtctc agccaatccc tgggtgagtt tcaccagttt tgattttaaac gtggccaata	720
[1602]	tggacaactt cttcgcccc gttttcacca tgggcaaata ttatacgcaa ggcgacaagg	780
[1603]	tgctgatgcc gctggcgatt caggttcacg atgccgtctg tgatggcttc catgtcggca	840
[1604]	gaatgcttaa tgaattacaa cagtactgcg atgagtggca gggcggggcg taactgcagg	900
[1605]	agctcaaaca gcagcctgta ttcaggctgc ttttttcgtt ttggctctgcg cgtaatctct	960
[1606]	tgctctgaaa acgaaaaaac cgccttgagc ggcggttttt cgaaggttct ctgagctacc	1020
[1607]	aactctttga accgaggtaa ctggcttgga ggagcgcagt caccaaaact tgtcctttca	1080
[1608]	gtttagcctt aaccggcgca tgacttcaag actaatcct ctaaatacat taccagtggc	1140
[1609]	tgctgccagt ggtgcttttg catgtctttc cgggttgagc tcaagacgat agttaccgga	1200
[1610]	taaggcgagc cggctcgact gaacgggggg ttcgtgcata cagtccagct tggagcgaac	1260
[1611]	tgcttaccgc gaactgagtg tcaggcgtgg aatgagacaa acgcggccat aacagcggaa	1320
[1612]	tgacaccggt aaaccgaaag gcaggaacag gagagcgac gagggagccg ccagggggaa	1380
[1613]	acgcctggta tctttatagt cctgtcgggt ttcgccacca ctgatttgag cgtcagattt	1440
[1614]	cgtgatgctt gtcagggggg cggagcctat ggaaaaacgg ctttgccgcg gccctctcac	1500
[1615]	ttccctgtta agtatcttcc tggcatcttc caggaaatct ccgcccggt cgtaagccat	1560
[1616]	ttccgctgcg cgcagtcgaa cgaccgagcg tagcgagtca gtgagcgagg aagcgggaata	1620
[1617]	tatcctgtat cacatattct gctgacgcac cggctgcagc ttttttctcc tgccacatga	1680
[1618]	agcacttcac tgacaccctc atcagtgaac caccgctggt agcgggtggt tttttaggcc	1740
[1619]	tatggccttt ttttttntg nnaaaccttt cgcggtatgg natnanagcg cccggaagag	1800
[1620]	agtcaattaa gaggtggtg aatgtgaaac cagtaacgtt atacgatgtc gcagagtatg	1860
[1621]	ccggtgtctc ttatcagacc gtttcccgcg tggatgaacca ggccagccac gtttctgcga	1920
[1622]	aaacgcggga aaaagtggaa gcggcgatgg cggagctgaa ttacattccc aaccgcgtgg	1980
[1623]	cacaacaact ggcgggcaaa cagtcgttgc tgattggcgt tgccacctcc agtctggccc	2040
[1624]	tgcacgcgcc gtcgcaaatt gtcgcggcga ttaaattctc gcgcgatcaa ctgggtgcca	2100
[1625]	gcgtggtggt gtcgatggta gaacgaagcg gcgtcgaagc ctgtaaagcg gcggtgcaca	2160
[1626]	atcttctgcg gcaacgcgtc agtgggctga tcattaacta tccgctggat gaccaggatg	2220
[1627]	ccattgctgt ggaagctgcc tgcactaatg ttccggcgtt atttcttgat gtctctgacc	2280
[1628]	agacacccat caacagtatt attttctccc atgaagacgg tacgcgactg ggcgtggagc	2340
[1629]	atctggtgcg attgggtcac cagcaaactc cgtgttagc gggccatta agttctgtct	2400
[1630]	cggcgcgtct gcgtctggct ggctggcata aatatctcac tcgcaatcaa attcagccga	2460
[1631]	tagcggaacg ggaaggcgac tggagtgcc atgtccggtt tcaacaaacc atgcaaactc	2520
[1632]	tgaatgaggg catcgttccc actgcgatgc tggttgcca cgatcagatg gcgctgggag	2580
[1633]	caatgcgcgc cattaccgag tccgggctgc gcgttggtgc ggacatctcg gtagtgggat	2640
[1634]	acgacgatac cgaagacagc tcatgttata tcccgccgtt aaccaccatc aaacaggatt	2700
[1635]	ttgcctgct ggggcaaacc agcgtggacc gcttgctgca actctctcag ggccaggcgg	2760
[1636]	tgaagggcaa tcagctgttg cccgtctcac tggatgaaaag aaaaaccacc ctggcgccca	2820
[1637]	atacgcaaac cgctctctcc cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg	2880

[1638]	tttcccgact ggaaagcggg cagtgcgagg taccgataa aagcggcttc ctgacaggag	2940
[1639]	gccgttttgt ttctcgagtt aattaaggca gtgagcgcaa cgcaattaat gtgagttagc	3000
[1640]	tcactcatta ggcacccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa	3060
[1641]	ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac ggattcactg	3120
[1642]	gccgtcgttt tacaatctag aggccagcct ggccataagg agatatacat atggggccatc	3180
[1643]	atcatcatca tcatcatcat catcacagca gcggccatat cgaaggctcgt catatgaaaa	3240
[1644]	ccctgagcca ggcacagagc aaaaccagca gccagcagtt tagctttacc ggcaatagca	3300
[1645]	gcgcaaagt gattattggt aatcagaaac tgaccatcaa tgatgttgca cgtgttgccc	3360
[1646]	gtaatggcac cctggtttagc ctgaccaata ataccgatat tctgcagggt attcaggcca	3420
[1647]	gctgtgatta tatcaataat gcagttgaaa gcggtgaacc gatttatggt gttaccagcg	3480
[1648]	gttttggtgg tatggcaaat gttgcaatta gccgtgaaca ggcaagcgaa ctgcagacca	3540
[1649]	atctggtttg gtttctgaaa accggtgcag gtaataaact gccgtggca gatgttcgtg	3600
[1650]	cagcaatgct gctgcgtgca aatagccaca tgcgtggtgc aagcgggtatt cgtctggaac	3660
[1651]	tgattaaacg catggaaatc tttctgaatg ccggtgttac cccgtatgtt tatgaatttg	3720
[1652]	gtagcattgg tgccagcggg gatctggttc cgctgagcta tattaccggg agcctgattg	3780
[1653]	gcctggaccc gagctttaa gttgatttta atggcaaaga aatggacgca ccgaccgcac	3840
[1654]	tgcgtcagct gaatctgagt ccgctgaccc tgctgccgaa agaaggctctg gcaatgatga	3900
[1655]	atggcaccag cgttatgacc ggtattgcag caaattgtgt ttatgatacc cagattctga	3960
[1656]	ccgcaattgc aatgggtgtt catgcactgg atattcaggc actgaatggt acaaatcaga	4020
[1657]	gctttcatcc gtttatccat aacagcaaac cgcacccggg tcagctgtgg gcagcagatc	4080
[1658]	agatgattag cctgctggcc aatagccagc tggttcgtga tgaactggat ggtaaactg	4140
[1659]	attatcgtga tcatgaactg atccaggatc gttatagcct gcgttgtctg ccgcagtatc	4200
[1660]	tgggtccgat tgttgatggt attagccaga ttgccaaaca aatcgaaatt gagattaaca	4260
[1661]	gcgttaccga taaccgcgtg attgatgttg ataatcaggc aagctatcat ggtggttaatt	4320
[1662]	ttctgggtca gtatgttggt atgggtatgg atcatctgcg ctattatatac ggtctgctgg	4380
[1663]	caaaacatct ggatgttcag attgcactgc tggcatcacc ggaatttagc aatgggtctgc	4440
[1664]	ctccgagtct gctgggtaat cgtgaacgta aagttaatat gggctctgaaa ggtctgcaga	4500
[1665]	tttgcggtaa tagcattatg ccgctgctga ctttttatgg taatagtatt gcagatcgtt	4560
[1666]	ttccgaccca tgccgaacag tttaaccaga atattaacag ccagggttat accagcgcaa	4620
[1667]	ccctggcacg tcgtagcgtt gatatttttc agaattatgt tgccattgcc ctgatgtttg	4680
[1668]	gtgttcaggc agttgatctg cgtacctaca aaaaaaccgg tcattatgat gcacgtgcct	4740
[1669]	gtctgtcacc ggcaaccgaa cgtctgtata gcgcagttcg tcatgttggt ggtcagaaac	4800
[1670]	cgacctcaga tcgtccgtat atttggaatg ataatgaaca gggctctggat gaacatattg	4860
[1671]	cacgtattag tgcagatatt gcagccggtg gtgtttattgt tcaggccgtt caggacattc	4920
[1672]	tgccgtgtct gcattaaggc caaac	4945
[1673]	<210>	28
[1674]	<211>	56
[1675]	<212>	DNA
[1676]	<213>	人工序列
[1677]	<220>	
[1678]	<223>	合成的多核苷酸
[1679]	<400>	28

[1680]	ctagaggcca gcctggccat aaggagatat acatatgaaa accctgagcc aggcac 56
[1681]	<210> 29
[1682]	<211> 47
[1683]	<212> DNA
[1684]	<213> 人工序列
[1685]	<220>
[1686]	<223> 合成的多核苷酸
[1687]	<400> 29
[1688]	gatggtgatg gtggccagtt tggccttaat gcagacacgg cagaatg 47
[1689]	<210> 30
[1690]	<211> 569
[1691]	<212> PRT
[1692]	<213> 点状念珠藻 (Nostoc punctiforme)
[1693]	<400> 30
[1694]	Met Asn Ile Thr Ser Leu Gln Gln Asn Ile Thr Arg Ser Trp Gln Ile
[1695]	1 5 10 15
[1696]	Pro Phe Thr Asn Ser Ser Asp Ser Ile Val Thr Val Gly Asp Arg Asn
[1697]	20 25 30
[1698]	Leu Thr Ile Asp Glu Val Val Asn Val Ala Arg His Gly Thr Gln Val
[1699]	35 40 45
[1700]	Arg Leu Thr Asp Asn Ala Asp Val Ile Arg Gly Val Gln Ala Ser Cys
[1701]	50 55 60
[1702]	Asp Tyr Ile Asn Asn Ala Val Glu Thr Ala Gln Pro Ile Tyr Gly Val
[1703]	65 70 75 80
[1704]	Thr Ser Gly Phe Gly Gly Met Ala Asp Val Val Ile Ser Arg Glu Gln
[1705]	85 90 95
[1706]	Ala Ala Glu Leu Gln Thr Asn Leu Ile Trp Phe Leu Lys Ser Gly Ala
[1707]	100 105 110
[1708]	Gly Asn Lys Leu Ser Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg
[1709]	115 120 125
[1710]	Ala Asn Ser His Leu Tyr Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile
[1711]	130 135 140
[1712]	Gln Arg Ile Glu Thr Phe Leu Asn Ala Gly Val Thr Pro His Val Tyr
[1713]	145 150 155 160
[1714]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr
[1715]	165 170 175
[1716]	Ile Thr Gly Ala Leu Ile Gly Leu Asp Pro Ser Phe Thr Val Asp Phe
[1717]	180 185 190
[1718]	Asp Gly Lys Glu Met Asp Ala Val Thr Ala Leu Ser Arg Leu Gly Leu
[1719]	195 200 205
[1720]	Pro Lys Leu Gln Leu Gln Pro Lys Glu Gly Leu Ala Met Met Asn Gly
[1721]	210 215 220

[1722]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Tyr Asp Ala Lys		
[1723]	225	230	235 240
[1724]	Val Leu Leu Ala Leu Thr Met Gly Val His Ala Leu Ala Ile Gln Gly		
[1725]		245	250 255
[1726]	Leu Tyr Gly Thr Asn Gln Ser Phe His Pro Phe Ile His Gln Cys Lys		
[1727]		260	265 270
[1728]	Pro His Pro Gly Gln Leu Trp Thr Ala Asp Gln Met Phe Ser Leu Leu		
[1729]		275	280 285
[1730]	Lys Asp Ser Ser Leu Val Arg Glu Glu Leu Asp Gly Lys His Glu Tyr		
[1731]	290	295	300
[1732]	Arg Gly Lys Asp Leu Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Ala		
[1733]	305	310	315 320
[1734]	Gln Phe Ile Gly Pro Ile Val Asp Gly Val Ser Glu Ile Thr Lys Gln		
[1735]		325	330 335
[1736]	Ile Glu Val Glu Met Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1737]		340	345 350
[1738]	Glu Asn Gln Val Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val		
[1739]		355	360 365
[1740]	Gly Val Thr Met Asp Arg Leu Arg Tyr Tyr Ile Gly Leu Leu Ala Lys		
[1741]	370	375	380
[1742]	His Ile Asp Val Gln Ile Ala Leu Leu Val Ser Pro Glu Phe Ser Asn		
[1743]	385	390	395 400
[1744]	Gly Leu Pro Pro Ser Leu Val Gly Asn Ser Asp Arg Lys Val Asn Met		
[1745]		405	410 415
[1746]	Gly Leu Lys Gly Leu Gln Ile Ser Gly Asn Ser Ile Met Pro Leu Leu		
[1747]		420	425 430
[1748]	Ser Phe Tyr Gly Asn Ser Leu Ala Asp Arg Phe Pro Thr His Ala Glu		
[1749]		435	440 445
[1750]	Gln Phe Asn Gln Asn Ile Asn Ser Gln Gly Tyr Ile Ser Ala Asn Leu		
[1751]	450	455	460
[1752]	Thr Arg Arg Ser Val Asp Ile Phe Gln Asn Tyr Met Ala Ile Ala Leu		
[1753]	465	470	475 480
[1754]	Met Phe Gly Val Gln Ala Val Asp Leu Arg Thr Tyr Lys Met Lys Gly		
[1755]		485	490 495
[1756]	His Tyr Asp Ala Arg Thr Cys Leu Ser Pro Asn Thr Val Gln Leu Tyr		
[1757]		500	505 510
[1758]	Thr Ala Val Cys Glu Val Val Gly Lys Pro Leu Thr Ser Val Arg Pro		
[1759]		515	520 525
[1760]	Tyr Ile Trp Asn Asp Asn Glu Gln Cys Leu Asp Glu His Ile Ala Arg		
[1761]	530	535	540
[1762]	Ile Ser Ala Asp Ile Ala Gly Gly Gly Leu Ile Val Gln Ala Val Glu		
[1763]	545	550	555 560

[1764]	His Ile Phe Ser Ser Leu Lys Ser Thr
[1765]	565
[1766]	<210> 31
[1767]	<211> 571
[1768]	<212> PRT
[1769]	<213> 胶须藻属(Rivularia sp.)PCC 7116
[1770]	<400> 31
[1771]	Met Asn Thr Val Arg Leu Thr Lys Asn Thr Val Ser Gln Lys Ser Phe
[1772]	1 5 10 15
[1773]	Ser Phe Leu Asn Asn Ser Asp Ala Ser Val Ile Val Gly Asp Arg Gln
[1774]	20 25 30
[1775]	Leu Thr Ile Glu Glu Val Val Ser Val Ala Arg Tyr Arg Ala Arg Val
[1776]	35 40 45
[1777]	Lys Leu Thr Glu Asn Leu Glu Lys Leu Ala Asn Val Gln Ala Ser Cys
[1778]	50 55 60
[1779]	Asp Phe Ile Arg Asp Ala Val Glu Ser Gly Glu Pro Ile Tyr Gly Val
[1780]	65 70 75 80
[1781]	Thr Thr Gly Phe Gly Gly Met Ala Asn Val Val Ile Ser Pro Glu Ser
[1782]	85 90 95
[1783]	Ala Thr Leu Leu Gln Asn Asn Leu Met Cys Tyr His Lys Val Gly Ala
[1784]	100 105 110
[1785]	Gly Asn Lys Leu Pro Leu Ala Asp Val Arg Ala Ala Met Leu Leu Arg
[1786]	115 120 125
[1787]	Ala Asn Ser His Val Ala Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile
[1788]	130 135 140
[1789]	Lys Arg Met Leu Ile Phe Leu Asn Ala Gly Val Thr Pro His Val Pro
[1790]	145 150 155 160
[1791]	Glu Phe Gly Ser Ile Gly Ala Ser Gly Asp Leu Thr Pro Leu Ala Tyr
[1792]	165 170 175
[1793]	Ile Ser Gly Ala Leu Ile Gly Leu Asn Ser Ser Tyr Ile Val Asp Phe
[1794]	180 185 190
[1795]	Asp Gly Glu Glu Met Asp Ala Pro Thr Ala Leu Gln Lys Leu Gly Leu
[1796]	195 200 205
[1797]	Glu Pro Leu Gln Leu Leu Pro Lys Glu Gly Leu Ala Met Met Asn Gly
[1798]	210 215 220
[1799]	Thr Ser Val Met Thr Gly Ile Ala Ala Asn Cys Val Gln Asp Thr Arg
[1800]	225 230 235 240
[1801]	Ile Leu Leu Ala Leu Ser Val Ala Thr His Ala Leu Thr Ile Gln Gly
[1802]	245 250 255
[1803]	Leu Glu Gly Thr Asn Gln Ser Phe His Glu Tyr Ile His Lys Leu Lys
[1804]	260 265 270
[1805]	Pro His Ser Gly Gln Ile Trp Ala Ala Ser Gln Met Leu Glu Leu Leu

[1806]	275	280	285
[1807]	Ala Gly Ser Gly Leu Ile Arg Asp Glu Leu Asp Gly Ser His Asp Tyr		
[1808]	290	295	300
[1809]	Arg Gly Lys Asn Pro Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro		
[1810]	305	310	315
[1811]	Gln Tyr Met Gly Pro Ile Val Asp Gly Ile Glu Asp Ile Ala Lys Gln		
[1812]	325	330	335
[1813]	Val Glu Ile Glu Ile Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Val		
[1814]	340	345	350
[1815]	Glu Asn Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Ile		
[1816]	355	360	365
[1817]	Gly Val Gly Met Asp Arg Leu Arg Tyr His Ile Gly Met Leu Ala Lys		
[1818]	370	375	380
[1819]	His Leu Asp Val Gln Ile Ala Tyr Leu Val Ala Pro Glu Phe Asn Asn		
[1820]	385	390	395
[1821]	Gly Leu Ser Pro Ser Leu Val Gly Asn Gln Gln Arg Thr Val Asn Met		
[1822]	405	410	415
[1823]	Gly Leu Lys Gly Leu Gln Ile Thr Gly Asn Ser Ile Met Pro Leu Leu		
[1824]	420	425	430
[1825]	Thr Phe Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu		
[1826]	435	440	445
[1827]	Gln Tyr Asn Gln Asn Ile Asn Ser Gln Gly Phe Ala Ser Ala Asn Leu		
[1828]	450	455	460
[1829]	Ala Arg Thr Ser Val Glu Ile Phe Gln Gln Tyr Ile Ala Leu Ala Leu		
[1830]	465	470	475
[1831]	Met Phe Gly Val Gln Ser Val Asp Leu Arg Thr Tyr Ala Ile Ala Gly		
[1832]	485	490	495
[1833]	His Tyr Asp Ala Arg Ala Thr Leu Ser Pro Ala Thr Gln Asp Leu Tyr		
[1834]	500	505	510
[1835]	Met Ala Val Arg Asn Val Val Gly Arg Pro Pro Ser Lys Glu Arg Ala		
[1836]	515	520	525
[1837]	Tyr Ile Trp Asp Asp Asn Glu Gln Gly Leu Asp Ser His Ile Ser Lys		
[1838]	530	535	540
[1839]	Ile Ala Asp Asp Ile Ala Tyr Gly Gly Gln Ile Val Thr Ala Ile Ser		
[1840]	545	550	555
[1841]	Glu Val Leu Ser Ala Leu Lys Ser Val Asn Asn		
[1842]	565	570	
[1843]	<210> 32		
[1844]	<211> 565		
[1845]	<212> PRT		
[1846]	<213> 颤藻属(<i>Oscillatoria</i> sp.)PCC 6506		
[1847]	<400> 32		

[1848]	Met Ser Thr Asn Leu Thr Glu Pro Ile Ser Gln Lys Leu Leu Gln Trp
[1849]	1 5 10 15
[1850]	Leu Glu Pro Asn Leu Ser Ala Val Ile Val Gly Asn Arg Arg Leu Ser
[1851]	20 25 30
[1852]	Val Glu Glu Val Ala Arg Val Ala Arg Cys Gly Thr Arg Val Arg Leu
[1853]	35 40 45
[1854]	Asn Asp Glu Ser Asp Val Ala Glu Arg Val Gln Ala Ser Cys Asp Tyr
[1855]	50 55 60
[1856]	Ile Thr Asp Ala Val Glu Ser Gly Lys Ala Ile Tyr Gly Val Thr Thr
[1857]	65 70 75 80
[1858]	Gly Phe Gly Gly Met Ala Asn Thr Glu Ile Ala Pro Glu Glu Ala Ala
[1859]	85 90 95
[1860]	Ser Leu Gln Asn Asn Leu Ile Trp Phe Leu Lys Ala Gly Ala Gly Gln
[1861]	100 105 110
[1862]	Lys Leu Pro Thr Ala Cys Val Arg Ala Ala Met Leu Leu Arg Ile Asn
[1863]	115 120 125
[1864]	Ser His Leu Gln Gly Ala Ser Gly Ile Arg Leu Glu Leu Ile Lys Arg
[1865]	130 135 140
[1866]	Met Ile Val Phe Leu Asn Ala Gly Val Thr Pro His Val Cys Glu Leu
[1867]	145 150 155 160
[1868]	Gly Ser Ile Gly Ala Ser Gly Asp Leu Val Pro Leu Ala Gln Ile Thr
[1869]	165 170 175
[1870]	Gly Ala Leu Ile Gly Leu Asp Asp Ser Phe Thr Val Asp Phe Asn Gly
[1871]	180 185 190
[1872]	Arg Glu Met Ser Ala Ile Gln Ala Leu Glu Leu Leu Asp Leu Pro Lys
[1873]	195 200 205
[1874]	Ile Asp Leu Arg Pro Lys Glu Gly Leu Ala Met Val Asn Gly Thr Ser
[1875]	210 215 220
[1876]	Val Met Thr Gly Ile Ala Ala Asn Cys Val Arg Asp Ser Gln Val Met
[1877]	225 230 235 240
[1878]	Leu Ala Leu Ala Met Gly Thr His Ala Leu Met Ile Gln Gly Leu Gly
[1879]	245 250 255
[1880]	Ala Thr Asn Gln Ser Phe His Pro Phe Ile His Lys Leu Lys Pro His
[1881]	260 265 270
[1882]	Phe Gly Gln Val Trp Ala Ala Ser Gln Met Val Glu Leu Leu Ala Gly
[1883]	275 280 285
[1884]	Ser Cys Leu Asn Arg Asp Glu Leu Asp Gly Gln His Asp Ala Arg Gly
[1885]	290 295 300
[1886]	Glu His Pro Ile Gln Asp Arg Tyr Ser Leu Arg Cys Leu Pro Gln Tyr
[1887]	305 310 315 320
[1888]	Ile Gly Pro Ile Val Asp Gly Ile Ala Glu Ile Ala Gln Gln Ile Glu
[1889]	325 330 335

[1890]	Val Glu Val Asn Ser Val Thr Asp Asn Pro Leu Ile Asp Thr Glu Asn
[1891]	340 345 350
[1892]	Gln Ala Ser Tyr His Gly Gly Asn Phe Leu Gly Gln Tyr Val Gly Val
[1893]	355 360 365
[1894]	Gly Met Asp Arg Leu Arg Tyr Leu Leu Gly Leu Leu Ala Lys His Leu
[1895]	370 375 380
[1896]	Asp Val Gln Ile Ala Leu Leu Ala Ala Pro Glu Phe Asn Asn Gly Leu
[1897]	385 390 395 400
[1898]	Ser Pro Ser Leu Val Gly Asn Thr Ser Arg Lys Val Asn Met Gly Leu
[1899]	405 410 415
[1900]	Lys Gly Leu Gln Ile Ala Gly Asn Ser Ile Met Pro Leu Leu Thr Phe
[1901]	420 425 430
[1902]	Tyr Gly Asn Ser Ile Ala Asp Arg Phe Pro Thr His Ala Glu Gln Phe
[1903]	435 440 445
[1904]	Asn Gln Asn Ile Asn Ser Gln Gly Phe Ala Ser Ala Asn Leu Ala Arg
[1905]	450 455 460
[1906]	Arg Ser Ile Glu Leu Phe Gln Gln Tyr Met Ala Ile Ser Leu Met Phe
[1907]	465 470 475 480
[1908]	Ala Val Gln Ala Val Asp Leu Arg Thr His Glu Val Ala Gly His Tyr
[1909]	485 490 495
[1910]	Asp Ala Arg Glu Cys Leu Ser Pro Leu Ser Leu Pro Leu Tyr Glu Ala
[1911]	500 505 510
[1912]	Val Arg Glu Val Val Gly Gln Pro Pro Asn Val Asp Arg Ser Tyr Ile
[1913]	515 520 525
[1914]	Trp Asn Asp Asn Glu Gln Ser Leu Asp Ile His Ile Ala Met Ile Ala
[1915]	530 535 540
[1916]	Ala Asp Ile Ala Gln Glu Gly Arg Ile Val Gln Ala Val Asn Gln Ile
[1917]	545 550 555 560
[1918]	Leu Ser Ser Leu Lys
[1919]	565
[1920]	<210> 33
[1921]	<211> 559
[1922]	<212> PRT
[1923]	<213> 粘球藻属(Gloeocapsa sp.)PCC 7428
[1924]	<400> 33
[1925]	Met Asn Thr Ala Val Gln Thr Gln Ala Ser Thr Asn Leu Asn Ala Asp
[1926]	1 5 10 15
[1927]	Thr Ile Leu Leu Gly Asp Arg Asn Leu Thr Ile Asp Glu Val Val Ser
[1928]	20 25 30
[1929]	Val Ala Arg His Gly Ala Lys Val Asn Ile Ser Thr Ala Asp Asn Val
[1930]	35 40 45
[1931]	Ala Gln Arg Ile Gln Ala Ser Cys Asp Tyr Ile Ala Glu Ala Val Ala

[1932]	50	55	60
[1933]	Thr Gly Arg Pro Ile Tyr Gly Val Thr Ser Gly Phe Gly Gly Met Ala		
[1934]	65	70	75
[1935]	Asn Val Val Ile Ser Arg Glu Tyr Ala Asp Leu Leu Gln His Asn Leu		
[1936]	85	90	95
[1937]	Val Trp Tyr His Lys Val Gly Ala Gly Arg Lys Leu Pro Leu Thr Asp		
[1938]	100	105	110
[1939]	Val Arg Ala Ala Met Leu Leu Arg Val Asn Ser His Leu His Gly Ala		
[1940]	115	120	125
[1941]	Ser Gly Ile Arg Arg Glu Ile Val Gln Arg Met Glu Met Phe Leu Asn		
[1942]	130	135	140
[1943]	Ala Arg Val Thr Pro His Val Pro Glu Tyr Gly Ser Ile Gly Ala Ser		
[1944]	145	150	155
[1945]	Gly Asp Leu Thr Pro Leu Ser Tyr Ile Thr Gly Ala Leu Ile Gly Leu		
[1946]	165	170	175
[1947]	Asp Asp Arg Tyr Lys Val Asp Phe Asp Gly Glu Glu Ile Asp Ala Ile		
[1948]	180	185	190
[1949]	Thr Ala Leu Glu Arg Leu Gly Leu Pro Gln Leu Gln Leu Gln Ala Lys		
[1950]	195	200	205
[1951]	Glu Gly Leu Ala Met Met Asn Gly Thr Ser Val Met Thr Gly Ile Ala		
[1952]	210	215	220
[1953]	Ala Asn Cys Val Tyr Asp Thr Arg Leu Leu Met Ala Leu Thr Met Gly		
[1954]	225	230	235
[1955]	Ala His Ala Leu Ile Leu Gln Gly Leu Asn Gly Thr Asn Gln Ser Phe		
[1956]	245	250	255
[1957]	His Pro Phe Ile His Lys Leu Lys Pro His Pro Gly Gln Lys Trp Ala		
[1958]	260	265	270
[1959]	Ala Ser Thr Met Leu Asp Leu Leu Ala Gly Ser Arg Leu Ile Arg Glu		
[1960]	275	280	285
[1961]	Glu Leu Asp Gly Thr His Glu Tyr Arg Gly Gln Ala Pro Ile Gln Asp		
[1962]	290	295	300
[1963]	Arg Tyr Ser Leu Arg Cys Leu Ala Gln Tyr Met Gly Pro Ile Val Asp		
[1964]	305	310	315
[1965]	Gly Val Ser Gln Val Ala Gln Gln Val Glu Ile Glu Met Asn Ser Ala		
[1966]	325	330	335
[1967]	Thr Asp Asn Pro Leu Ile Asp Val Glu Asn Gln Ala Ser Tyr His Gly		
[1968]	340	345	350
[1969]	Gly Asn Phe Leu Gly Gln Tyr Ile Gly Met Gly Met Asp His Leu Arg		
[1970]	355	360	365
[1971]	Tyr Tyr Ile Gly Met Met Ala Lys His Leu Asp Val Gln Ile Ala Tyr		
[1972]	370	375	380
[1973]	Leu Val Ala Pro Glu Phe Asn Asn Gly Leu Pro Ala Ser Leu Val Gly		

[1974]	385	390	395	400
[1975]	Asn Lys Glu Arg Ile Val Asn Met Gly Leu Lys Gly Leu Gln Ile Thr			
[1976]		405	410	415
[1977]	Gly Asn Ser Ile Met Pro Leu Leu Ser Phe Tyr Gly Asn Ser Ile Ala			
[1978]		420	425	430
[1979]	Asp Arg Tyr Pro Thr His Ala Glu Gln Tyr Asn Gln Asn Ile Asn Ser			
[1980]		435	440	445
[1981]	Gln Gly Phe Ala Ala Ala Asn Leu Thr Arg Asn Ala Val Glu Ile Phe			
[1982]		450	455	460
[1983]	Gln Gln Tyr Met Ala Ile Ala Leu Met Phe Gly Val Gln Ala Val Asp			
[1984]	465	470	475	480
[1985]	Leu Arg Thr Tyr Ala Tyr Ala Gly His Tyr Asp Ala Ser Glu Cys Leu			
[1986]		485	490	495
[1987]	Ser Pro Thr Thr Arg Arg Leu Tyr Gln Ala Val Arg Glu Val Val Gly			
[1988]		500	505	510
[1989]	Gln Pro Ser Ser Ala Thr Arg Pro Tyr Ile Trp Asp Asp Arg Glu Gln			
[1990]		515	520	525
[1991]	Pro Leu Asp Glu His Ile Ala Lys Ile Ala Ala Asp Ile Ala Ala Glu			
[1992]		530	535	540
[1993]	Gly Val Ile Val Ala Ala Val Lys Asp Leu Leu Thr Ser Leu Lys			
[1994]	545	550	555	

YP_324488.1	MKTL	SAQSKTSSQ	QFSTG	SSANVIG	NKLTIND	VARNGTIV	SLTNN	TDIL	QHQASCDY	INNA	VESE	GEPI	YGVTS	SGFG	MANVA	ISRE	QASEL	100																														
YP_001865631.1	MNITS	LQONITR	SWQIP	FTNSS	DSIVT	VCGRN	LITIDE	VNVAR	HGTQ	RLTNAD	VI	RGVQASCDY	INNA	VE	TAQPI	YGVTS	SGFG	MADV	VSRE	QAAEL	90																											
YP_007056096.1	MNTV	RLTKNT	VSQK	SFSL	NNSDAS	VIVG	DRQ	LTIEE	VS	VARV	RA	RKLTEN	LEKLAN	VQASCD	EIRDA	VESE	GEPI	YGVTS	SGFG	MANV	VSPE	80																										
ZP_07108482.1	MST	-NLTEP	-ISQ	LLQW	LEPN	LSAV	IVGN	RLSVEE	VARV	ARCC	TRVRLIN	DESD	VAER	VQASCD	YITDA	VE	SGKAI	YGVTS	SGFG	MANTE	IAPEE	70																										
YP_007127054.1	MNTAV	QTQAS	-----	TN	L	NADT	IL	CD	RNL	TIDE	V	SVAR	HGAK	VNI	STAD	NVAQR	IQASCD	YIAE	AVATGR	PIYGVTS	SGFG	MANV	VSRE	60																								
YP_324488.1	QTNL	VWFL	KTCAG	NKLI	PLAD	VRA	AMLL	RAN	SHM	FGAS	CIR	LEI	IRME	IF	NA	GVTP	VVVE	FGS	IGAS	CDI	VPL	SYIT	GS	IG	SL	IG	LI	GP	SFK	V	D	ENG	K	END	AP	200												
YP_001865631.1	QTNL	IWFL	KSGAG	NKLI	SLAD	VRA	AMLL	RAN	SHLY	GCAS	CIR	LEI	IR	ET	FL	NA	GVTP	PHV	VE	FGS	IGAS	CDI	VPL	SYIT	GC	AL	IG	LI	GP	SFT	V	D	ED	G	K	END	AV	210										
YP_007056096.1	QNNL	MCYH	KVGAG	NKLI	PLAD	VRA	AMLL	RAN	SHV	AGAS	CIR	LEI	IRML	IF	NA	GVTP	PHV	PE	FGS	IGAS	CDI	TP	PLAY	TS	CA	LI	GIN	SSY	IV	D	ED	G	E	END	AP	220												
ZP_07108482.1	QNNL	IWFL	KAGAC	OKIPTAC	VRA	AMLL	RAN	SHL	QGCAS	CIR	LEI	IR	MI	V	FL	NA	GVTP	PHV	CE	LG	SSIGAS	CDI	VPL	AO	IT	GC	AL	IG	LI	GP	SFT	V	D	ED	G	REMSAI	230											
YP_007127054.1	QHNL	VWYH	KVGAG	RKLI	PLTD	VRA	AMLL	RAN	SHL	HGCAS	CIR	REI	VO	RM	EM	FL	NA	RVTP	PHV	PEY	GSIGAS	CDI	TP	LYT	TC	AL	IG	LI	GP	DRY	KV	D	ED	G	E	EIDAI	240											
YP_324488.1	TALR	QNL	SPL	TL	LPKE	GLAM	NGT	SV	MTG	IA	NCV	VD	TQ	IL	TA	AM	GV	HALD	IQ	AL	NG	TN	QS	FH	PS	PH	NS	KPH	PG	Q	W	A	D	Q	MS	L	AN	SQ	IV	R	D	E	L	D	G	250		
YP_001865631.1	TALS	R	GL	PK	LQ	LPKE	GLAM	NGT	SV	MTG	IA	NCV	VD	AK	V	IL	AL	TM	GV	HAL	IQ	CL	Y	TN	QS	FH	PS	PH	NS	KPH	PG	Q	W	A	D	Q	MS	L	AN	SQ	IV	R	D	E	L	D	G	260
YP_007056096.1	TALQ	K	GLE	PI	Q	LPKE	GLAM	NGT	SV	MTG	IA	NCV	VD	TR	I	IL	AL	SV	ATH	LI	Q	GLE	TN	QS	FH	PS	PH	NS	KPH	PG	Q	W	A	D	Q	MS	L	AN	SQ	IV	R	D	E	L	D	G	270	
ZP_07108482.1	QALE	L	LP	K	D	LPKE	GLAM	NGT	SV	MTG	IA	NCV	VD	SQ	V	IL	AL	AM	G	TH	LI	Q	GLE	TN	QS	FH	PS	PH	NS	KPH	PG	Q	W	A	D	Q	MS	L	AN	SQ	IV	R	D	E	L	D	G	280
YP_007127054.1	TALR	GL	PK	LQ	LPKE	GLAM	NGT	SV	MTG	IA	NCV	VD	TR	I	IL	AL	TM	GV	HAL	IQ	CL	Y	TN	QS	FH	PS	PH	NS	KPH	PG	Q	W	A	D	Q	MS	L	AN	SQ	IV	R	D	E	L	D	G	290	

图1

YP_324488.1	KEDYRDHELIO	DRYSIRCLPQYI	GPIVDGISQIAKQIEIEHNSVT	DNPLIDVENQAS	YHGGNFI	CGYV	CGMDHLRYI	YIGLIAKHEDVQI	ALLASPEFSN	310	320	330	340	350	360	370	380	390	400
YP_001865631.1	KHEYRGKDLIO	DRYSIRCLPQYI	GPIVDGSEITKQIEVENNSVT	DNPLIDVENQAS	YHGGNFI	CGYV	CGMDHLRYI	YIGLIAKHEDVQI	ALLVSPFSN	310	320	330	340	350	360	370	380	390	400
YP_007056096.1	SHDYRGKNPIQ	DRYSIRCLPQYI	GPIVDGIEDIAKQVEIEHNSVT	DNPLIDVENQAS	YHGGNFI	CGYI	CGMDRLRYHIGMI	AKHEDVQI	AYLVAPEFNN	310	320	330	340	350	360	370	380	390	400
ZP_07108482.1	QHDARGEHPIQ	DRYSIRCLPQYI	GPIVDGIAEIAQQIEVENNSVT	DNPLIDTENQAS	YHGGNFI	CGYV	CGMDRLRYLGLIAKHEDVQI	ALLA	AAPEFNN	310	320	330	340	350	360	370	380	390	400
YP_007127054.1	THEYRGOAPIQ	DRYSIRCLPQYI	GPIVDGVSQVQAQQVEIEHNSVT	DNPLIDVENQAS	YHGGNFI	CGYI	CGMDHLRYI	YIGMAKHEDVQI	AYLVAPEFNN	310	320	330	340	350	360	370	380	390	400
YP_324488.1	GLPPSILGNRER	KVNMGILKGLQICGNS	IMPLLTFYGNSHADREF	THAEQENQIN	SQGT	SATARRSVD	IFQNVMAIAL	MFGVQAVD	LRTYKKTGHYDA	410	420	430	440	450	460	470	480	490	500
YP_001865631.1	GLPPSLVGN	SDRKVNMGILKGLQISGNS	IMPLLTFYGNSHADREF	THAEQENQIN	SQGYI	SANLTRRSVD	IFQNVMAIAL	MFGVQAVD	LRTYKMKGHYDA	410	420	430	440	450	460	470	480	490	500
YP_007056096.1	GLSPSIVGN	QORTVNMGILKGLQITGNS	IMPLLTFYGNSHADREF	THAEQXNQIN	SQGFASANDPART	SVEIFQVTHALAL	MFGVQSV	VDLRTYA	IAGHYDA	410	420	430	440	450	460	470	480	490	500
ZP_07108482.1	GLSPSIVGNT	SRKVNMGILKGLQIAGNS	IMPLLTFYGNSHADREF	THAEQENQIN	SQGFASANDPART	SIELFQVMAI	SIMFAVQAVD	LRTHEV	AGHYDA	410	420	430	440	450	460	470	480	490	500
YP_007127054.1	GLPASIVGN	KERIIVNMGILKGLQITGNS	IMPLLTFYGNSHADREF	THAEQXNQIN	SQGFANL	TRNAVEIFQVMAIAL	MFGVQAVD	LRTYA	AGHYDA	410	420	430	440	450	460	470	480	490	500
YP_324488.1	RACLSPATERLY	SAVRHVQKPTSDR	PIWINDNEQGLDEHIA	RIASADIACG	VIVQAVQDIL	PCILH----	(SEQ ID NO:4)			510	520	530	540	550	560	570			
YP_001865631.1	RTCLSPNTVQL	YTAVCEVVGKPLTSVR	PIWINDNEQCLDEHIA	RIASADIACG	LIVQAVEHIF	SSIKST--	(SEQ ID NO:29)			510	520	530	540	550	560	570			
YP_007056096.1	RATLSPATQDLY	MAVRNVGRPPSKER	PIWINDNEQGLSHISKI	ADDIANGGQIVTA	ISEVLSALKSVNN		(SEQ ID NO:30)			510	520	530	540	550	560	570			
ZP_07108482.1	RECLSPLSPL	LYEAVREVGP	PNVDRSYIWNDEQSLDI	HIAMIAADIAQ	ECRIVQAVNQIL	SSIK----	(SEQ ID NO:31)			510	520	530	540	550	560	570			
YP_007127054.1	SECLSPTRRL	LYQAVREVGP	SSATRPPIWDDREQ	PLDEHIAKIAADIA	EGVIAAVKDL	ITSIK----	(SEQ ID NO:32)			510	520	530	540	550	560	570			

图1续

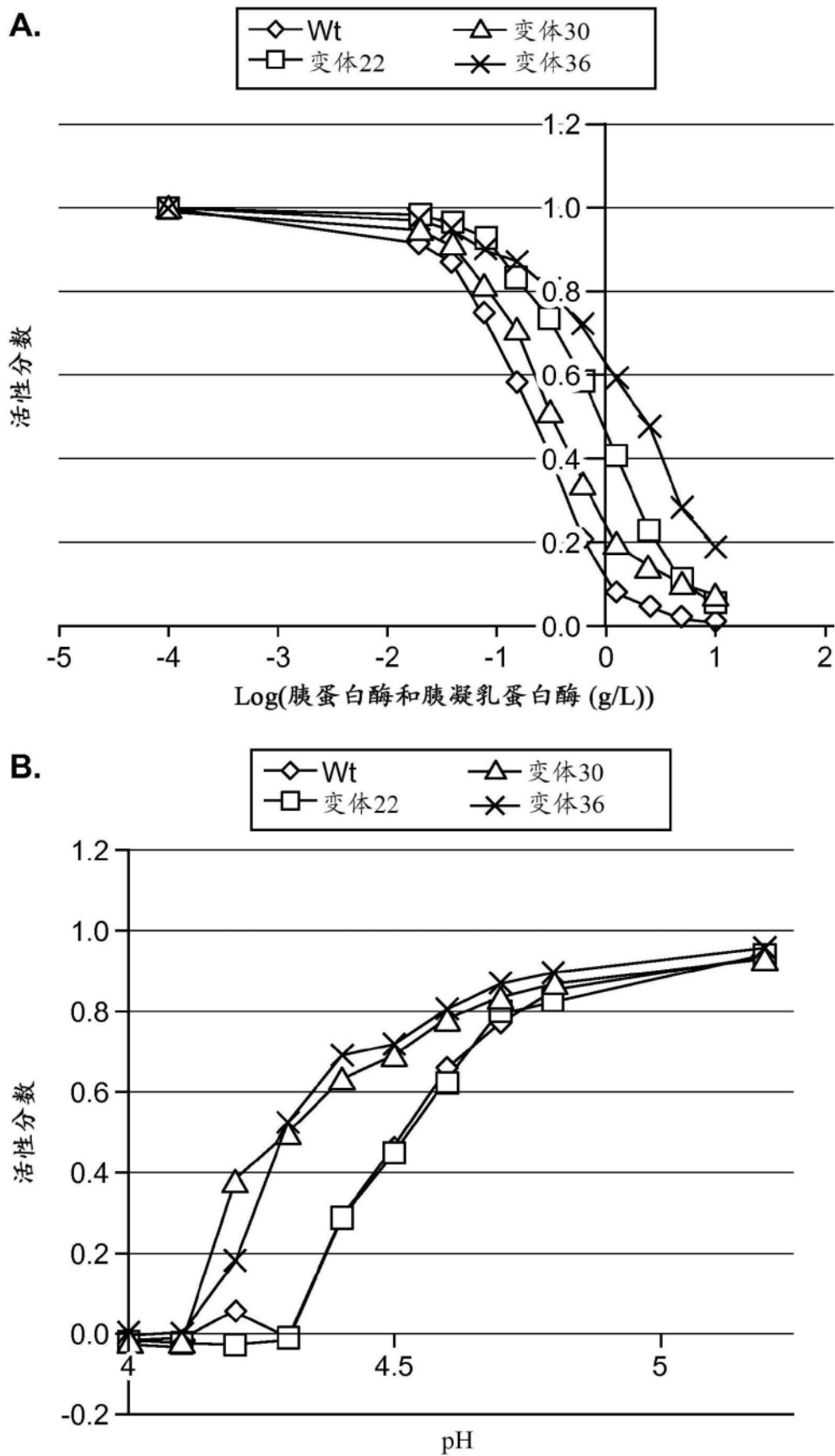


图2

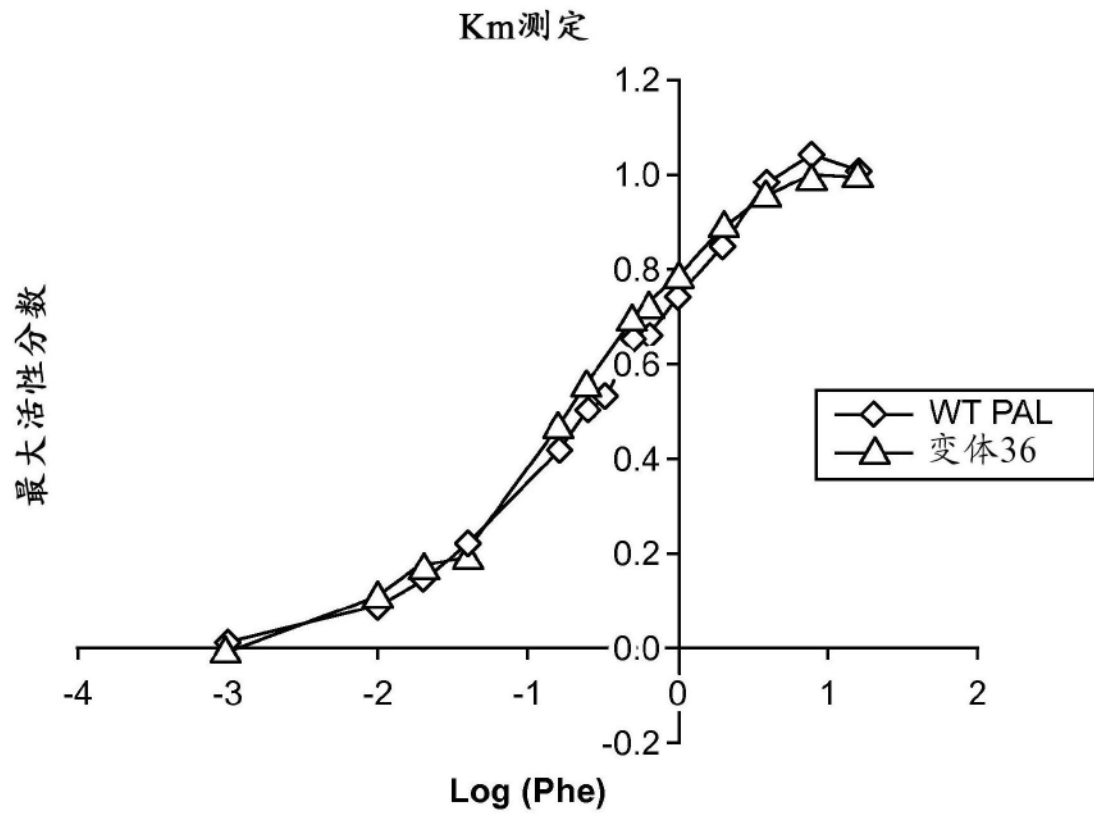


图3

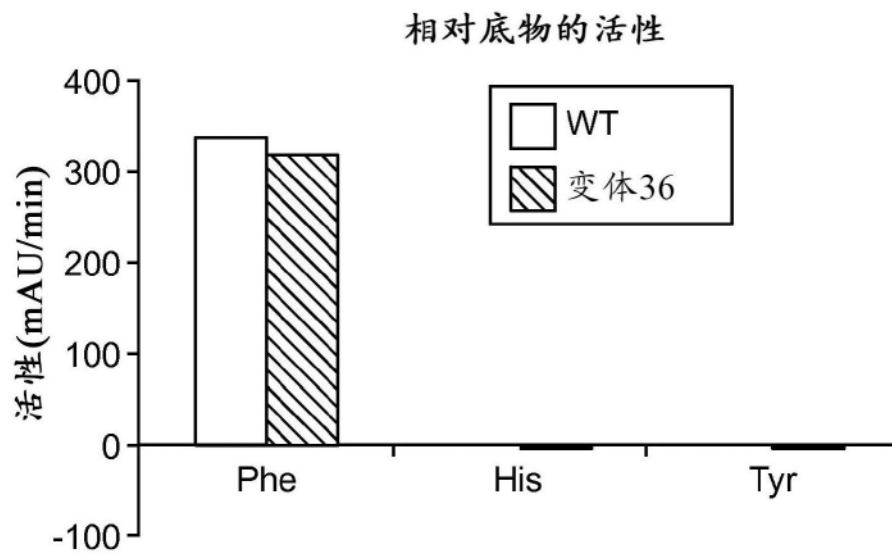


图4

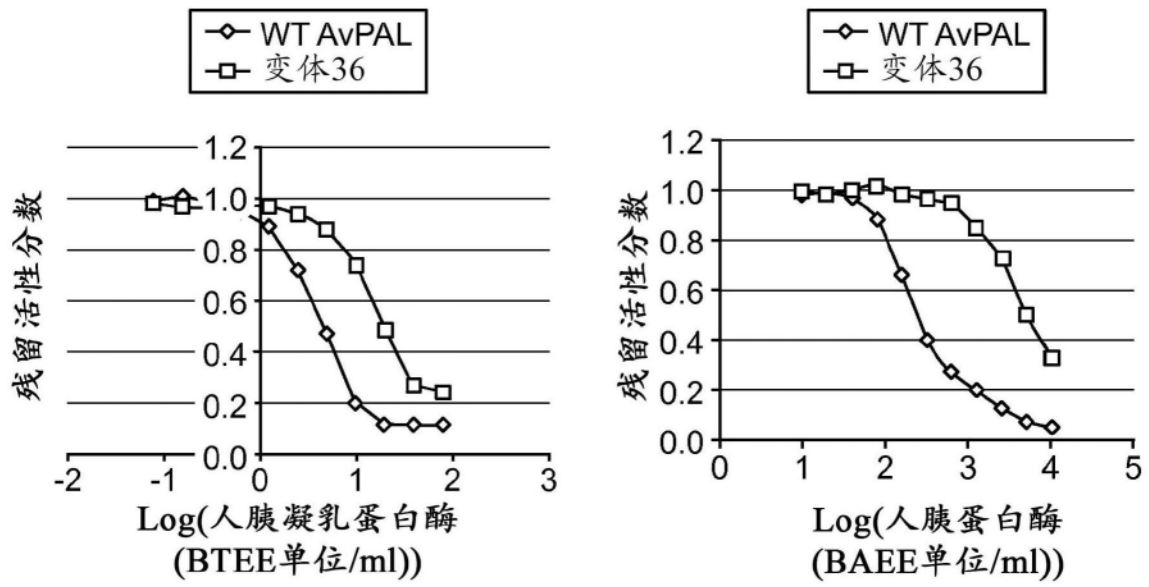


图5

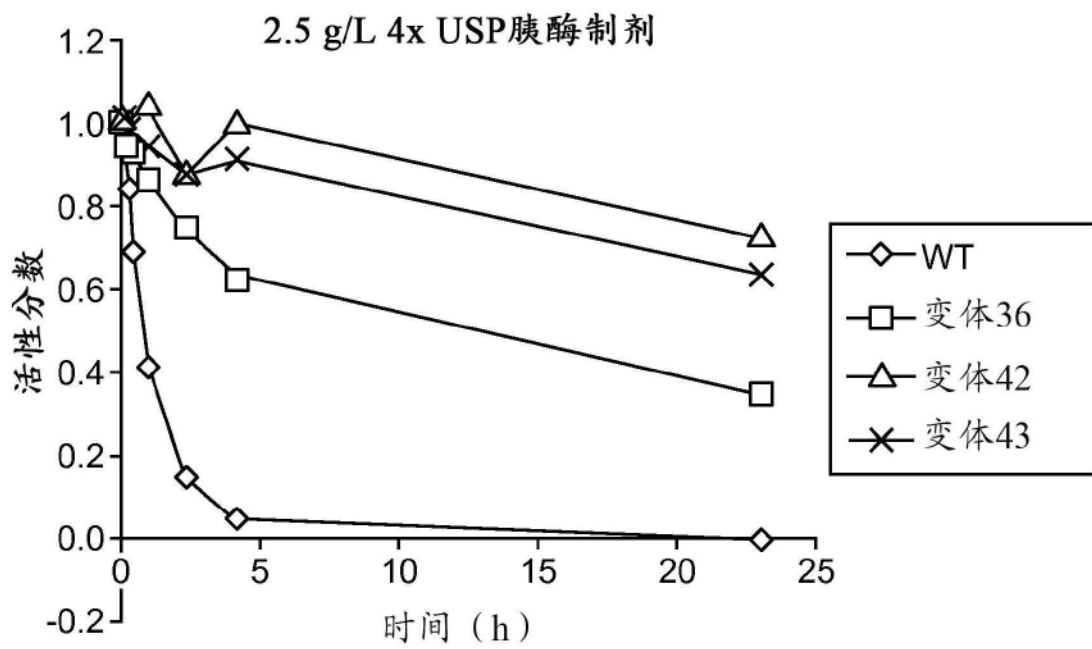


图6

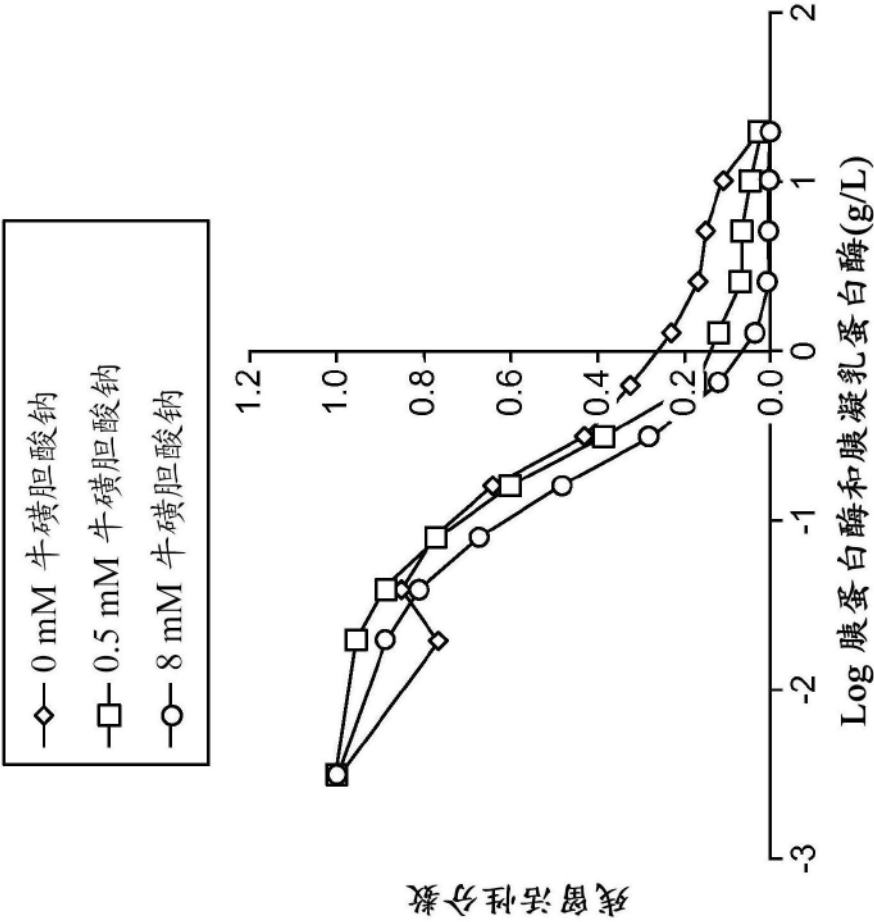


图7

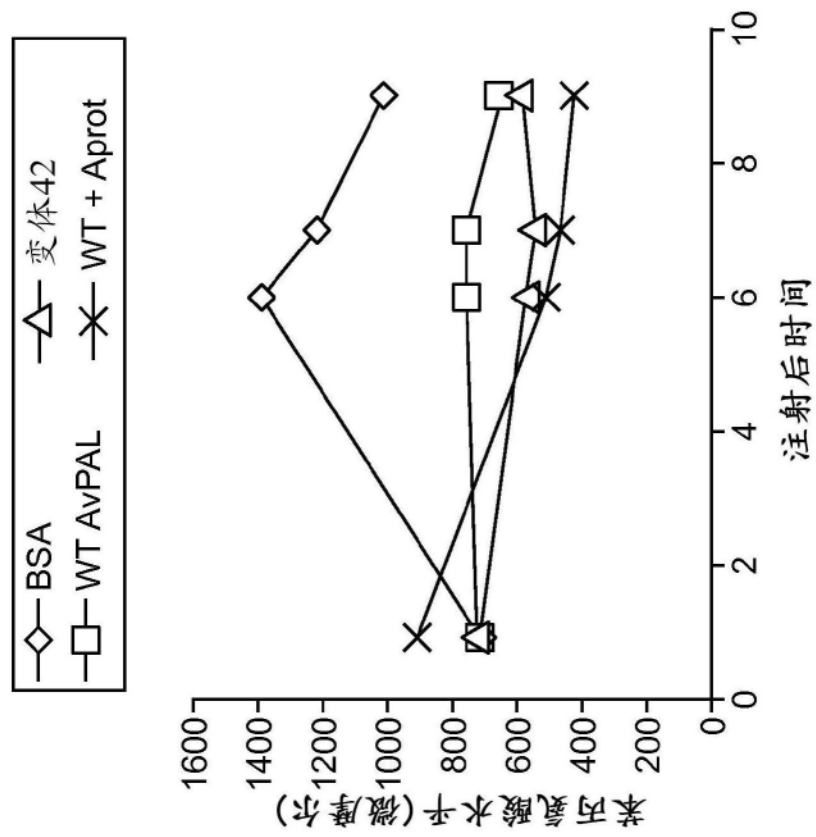


图8