



(12)发明专利

(10)授权公告号 CN 103348003 B

(45)授权公告日 2016.09.07

(21)申请号 201180064714.3

专利权人 格利科斯芬兰公司

(22)申请日 2011.11.24

(72)发明人 J·纳蒂南 A·康埃尔瓦

(65)同一申请的已公布的文献号

J·伊尔蒂南 M·萨洛埃莫

申请公布号 CN 103348003 A

H·维斯卡里 A·于斯科南

(43)申请公布日 2013.10.09

(74)专利代理机构 北京市金杜律师事务所

11256

(30)优先权数据

代理人 陈文平 徐志明

61/417,144 2010.11.24 US

(85)PCT国际申请进入国家阶段日

(51)Int.Cl.

2013.07.10

C12N 9/10(2006.01)

(86)PCT国际申请的申请数据

C12P 21/00(2006.01)

PCT/EP2011/070956 2011.11.24

C12N 5/10(2006.01)

(87)PCT国际申请的公布数据

C12N 1/14(2006.01)

W02012/069593 EN 2012.05.31

C12N 15/80(2006.01)

审查员 徐丹

(73)专利权人 诺瓦提斯国际药品有限公司

权利要求书4页 说明书66页

地址 百慕大汉密尔顿

序列表78页 附图47页

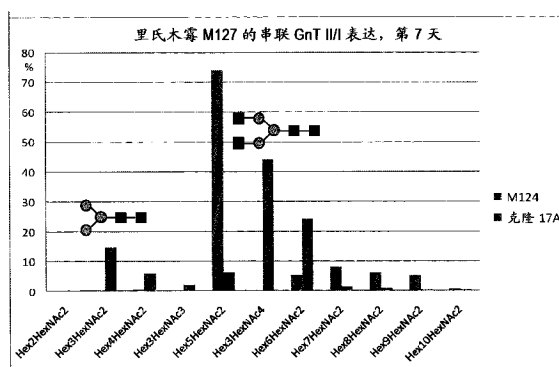
(54)发明名称

具有N-乙酰葡萄糖胺基转移酶活性的融合酶

(57)摘要

本公开涉及具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质。本公开还涉及用于制备复合的N-聚糖的方法,其包括以下步骤:提供含有这种重组蛋白质的宿主细胞和培养该宿主细胞使得该重组蛋白质被表达。

D.



1. 一种具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基上和催化N-乙酰葡萄糖胺转移到末端Man α 6残基,以及其中所述重组蛋白质是由N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域及在N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域之间的间隔区组成的融合蛋白质,且其中所述N-乙酰葡萄糖胺基转移酶I催化结构域完全来自SEQ ID NO:1且包含与SEQ ID NO:1的氨基酸残基105-445 100%相同的序列和其中所述N-乙酰葡萄糖胺基转移酶II催化结构域完全来自SEQ ID NO:20且包含与SEQ ID NO:20的氨基酸残基30-447 100%相同的序列。

2. 如权利要求1的重组蛋白质,其中所述N-乙酰葡萄糖胺基转移酶II催化结构域处于N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端。

3. 如权利要求2的重组蛋白质,其中所述间隔区包含来自非催化结构域的序列。

4. 如权利要求1-3任一项的重组蛋白质,其中所述间隔区包含选自于SEQ ID NO:118、SEQ ID NO:120、SEQ ID NO:122和SEQ ID NO:124的序列。

5. 如权利要求4的重组蛋白质,其中所述间隔区包含选自于SEQ ID NO:118、SEQ ID NO:120和SEQ ID NO:124的序列。

6. 如权利要求4的重组蛋白质,其中所述间隔区包含SEQ ID NO:124或SEQ ID NO:120的序列。

7. 一种具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基上和催化N-乙酰葡萄糖胺转移到末端Man α 6残基,以及其中所述重组蛋白质是由(i)N-乙酰葡萄糖胺基转移酶I催化结构域;(ii)N-乙酰葡萄糖胺基转移酶II催化结构域;(iii)在N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域之间的间隔区和在所述重组蛋白质N-末端的靶向肽组成的融合蛋白质,且其中所述N-乙酰葡萄糖胺基转移酶I催化结构域完全来自SEQ ID NO:1且包含与SEQ ID NO:1的氨基酸残基105-445 100%相同的序列和其中所述N-乙酰葡萄糖胺基转移酶II催化结构域完全来自SEQ ID NO:20且包含与SEQ ID NO:20的氨基酸残基30-447 100%相同的序列。

8. 如权利要求7的重组蛋白质,其中所述靶向肽包含非催化结构域。

9. 如权利要求3的重组蛋白质,其中所述非催化结构域是来自选自于甘露糖苷酶和糖基转移酶的蛋白质。

10. 如权利要求8的重组蛋白质,其中所述非催化结构域是来自选自于甘露糖苷酶和糖基转移酶的蛋白质。

11. 如权利要求10的重组蛋白质,其中所述蛋白质来自选自于枝顶孢属(*Acremonium*)、曲霉属(*Aspergillus*)、短梗霉属(*Aureobasidium*)、隐球菌属(*Cryptococcus*)、金孢子菌属(*Chrysosporium*)、线黑粉酵母属(*Filibasidium*)、镰孢菌属(*Fusarium*)、赤霉属(*Gibberella*)、腐质霉属(*Humicola*)、梨孢菌属(*Magnaporthe*)、毛霉属(*Mucor*)、毁丝霉属(*Myceliophthora*)、漆斑菌属(*Myrothecium*)、新考马脂霉属(*Neocallimastix*)、脉孢霉属(*Neurospora*)、拟青霉属(*Paecilomyces*)、青霉属(*Penicillium*)、瘤胃壶菌属(*Piromyces*)、裂褶菌属(*Schizophyllum*)、踝节菌属(*Talaromyces*)、嗜热子囊菌属(*Thermoascus*)、梭孢壳属(*Thielavia*)、弯颈霉属(*Tolypocladium*)和木霉属(*Trichoderma*)的生物体。

12. 如权利要求7的重组蛋白质,其中所述靶向肽是Kre2靶向肽。
13. 如权利要求8的重组蛋白质,其中所述靶向肽包含连接至非催化结构域的N-末端的跨膜结构域。
14. 如权利要求8的重组蛋白质,其中所述靶向肽包含连接到非催化结构域的N-末端的胞质域。
15. 如权利要求13的重组蛋白质,其中所述靶向肽包含连接到跨膜结构域的N末端的胞质域。
16. 如权利要求7的重组蛋白质,其中所述N-乙酰葡萄糖胺基转移酶II催化结构域位于所述N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端,所述间隔区序列包含人N-乙酰葡萄糖胺基转移酶I非催化结构域且位于所述催化结构域之间,所述靶向肽位于N-乙酰葡萄糖胺基转移酶II催化结构域的N-末端,该靶向肽包含胞质结构域、跨膜结构域和来自人N-乙酰葡萄糖胺基转移酶II的非催化结构域。
17. 如权利要求1的重组蛋白质,其由与SEQ ID NO:95、SEQ ID NO:119、SEQ ID NO:123或SEQ ID NO:125 100%相同的序列组成。
18. 如权利要求7的重组蛋白质,其中所述N-乙酰葡萄糖胺基转移酶II催化结构域位于所述N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端,所述间隔区位于所述催化结构域之间,该间隔区包含选自于SEQ ID NO:118、SEQ ID NO:120、SEQ ID NO:122和SEQ ID NO:124的序列,所述靶向肽位于N-乙酰葡萄糖胺基转移酶II催化结构部的N-末端且该靶向肽包含胞质结构域、跨膜结构域和来自人N-乙酰葡萄糖胺基转移酶II的非催化结构域。
19. 如权利要求18的重组蛋白质,其中所述间隔区来自选自于SEQ ID NO:118、SEQ ID NO:120和SEQ ID NO:124的序列。
20. 如权利要求18的重组蛋白质,其中所述间隔区来自SEQ ID NO:124或SEQ ID NO:120的序列。
21. 如权利要求18的重组蛋白质,其中所述间隔区来自SEQ ID NO:124的序列。
22. 一种分离的多核苷酸,其编码上述权利要求中任一项的重组蛋白质。
23. 一种表达载体,其包含可操作地与启动子连接的如权利要求22所述的分离的多核苷酸。
24. 一种包含如权利要求23所述的表达载体的宿主细胞。
25. 一种制备复合的N-聚糖的方法,其包括:
 - (1)提供宿主细胞,其中该宿主细胞包含编码权利要求1的重组蛋白质的多核苷酸;和
 - (2)培养该宿主细胞以使该重组蛋白质被表达,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和催化N-乙酰葡萄糖转移到末端Man α 6残基,以产生复合的N-聚糖。
26. 如权利要求25的方法,其中所述受体聚糖被连接于异源多肽。
27. 如权利要求25或权利要求26的方法,其中所述复合的N-聚糖是GlcNAc β 2Man β 3(GlcNAc β 2Man β 6)Man β 4GlcNAc β 4GlcNAc。
28. 如权利要求25或权利要求26的方法,其中所述宿主细胞是选自于木霉菌、枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、线黑粉酵母属、镰孢菌属、赤霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂

褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属和弯颈霉属的丝状真菌细胞。

29. 如权利要求25或权利要求26的方法,其中所述宿主细胞还包含编码UDP-GlcNAc转运蛋白的多核苷酸。

30. 如权利要求25或权利要求26的方法,其中所述宿主细胞具有与野生型宿主细胞中的活性水平相比降低的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶的活性水平。

31. 如权利要求30的方法,其中所述宿主细胞具有与野生型宿主细胞中的表达水平相比降低的alg3基因的表达水平。

32. 如权利要求30的方法,其中从所述宿主细胞删除该alg3基因。

33. 如权利要求25或权利要求26的方法,其中所述宿主细胞具有与野生型宿主细胞中的活性水平相比降低的 α -1,6-甘露糖基转移酶的活性水平。

34. 如权利要求33的方法,其中所述宿主细胞具有与野生型宿主细胞中的表达水平相比降低的och1基因的表达水平。

35. 如权利要求25或权利要求26的方法,其中所述宿主细胞还包含编码 α -1,2-甘露糖苷酶的多核苷酸。

36. 如权利要求25或权利要求26的方法,其中所述宿主细胞还包含编码 $\alpha\beta$ -1,4-半乳糖基转移酶的多核苷酸。

37. 如权利要求25或权利要求26的方法,其中所述宿主细胞还包含编码唾液酸转移酶的多核苷酸。

38. 一种制备复合的N-聚糖的方法,其包括:

(1)提供木霉属宿主细胞,其中该宿主细胞具有与野生型宿主细胞中的表达水平相比降低的alg3基因的表达水平并包含权利要求1的重组蛋白质;以及

(2)培养该宿主细胞以产生复合的N-聚糖。

39. 一种制备复合的N-聚糖的方法,包括在缓冲液中将权利要求1的重组蛋白质、受体聚糖和N-乙酰葡萄糖胺供体一起孵育,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到末端Man α 6残基以产生复合的N-聚糖。

40. 一种具有与野生型丝状真菌细胞中的表达水平相比降低的alg3基因的表达水平的丝状真菌细胞,其中所述丝状真菌细胞包含如权利要求1-21中任一项所述的重组蛋白质。

41. 如权利要求40的丝状真菌细胞,其中所述alg3基因包含突变。

42. 如权利要求41的丝状真菌细胞,其中所述alg3基因突变是alg3基因的缺失。

43. 如权利要求40-42中任一项的丝状真菌细胞,其中所述重组蛋白质是由可操作地与启动子连接的多核苷酸编码的。

44. 如权利要求43的丝状真菌细胞,其中所述启动子是诱导型启动子。

45. 如权利要求44的丝状真菌细胞,其中所述诱导型启动子是cbh1启动子。

46. 如权利要求40-42中任一项的丝状真菌细胞,其中所述丝状真菌细胞还包含编码UDP-GlcNAc转运蛋白的多核苷酸。

47. 如权利要求40-42中任一项的丝状真菌细胞,其中所述丝状真菌细胞具有与野生型丝状真菌细胞中的活性水平相比降低的 α -1,6-甘露糖基转移酶的活性水平。

48. 如权利要求47的丝状真菌细胞,其中所述丝状真菌细胞具有与野生型宿丝状真菌

细胞中的表达水平相比降低的och1基因的表达水平。

49.如权利要求40-42中任一项的丝状真菌细胞,其中所述丝状真菌细胞还包含编码 α -1,2-甘露糖苷酶的多核苷酸。

50.如权利要求40-42中任一项的丝状真菌细胞,其中所述丝状真菌细胞还包含编码 β -1,4-半乳糖基转移酶的多核苷酸。

51.如权利要求40-42中任一项的丝状真菌细胞,其中所述丝状真菌细胞还包含编码唾液酸转移酶的多核苷酸。

52.如权利要求40-42中任一项的丝状真菌细胞,其中所述丝状真菌细胞选自于木霉菌、枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、线黑粉酵母属、镰孢菌属、赤霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属和弯颈霉属。

53.如权利要求9或10的重组蛋白质,其中所述糖基转移酶是甘露糖基转移酶。

54.如权利要求53的重组蛋白质,其中所述甘露糖基转移酶选自MNN2、MNN4、MNN6、MNN9、MNN10、MNS1、KRE2、VAN1和OCH1。

具有N-乙酰葡萄糖胺基转移酶活性的融合酶

[0001] 相关申请的交叉参考

[0002] 本申请要求2010年11月24日提交的61/417,144号美国临时申请的利益,其通过引用全部在此引入。

[0003] 提交ASCII文本文件的序列列表

[0004] 下列以ASCII文本文件提交的内容通过引用全部引入本文:序列列表的计算机可读形式(CRF)(文件名称:619672001040SEQLIST.txt,记录日期:2011年11月22日,大小:305KB)。

发明领域

[0005] 本公开涉及可用于制备N-聚糖的组合物和方法。

[0006] 发明背景

[0007] 蛋白质的翻译后修饰通常是正确的蛋白质折叠和功能所必需的。常见的蛋白质修饰是在内质网中将寡糖(聚糖)添加到新生多肽上以形成糖蛋白,该过程被称为糖基化。N-糖基化在产生用于治疗目的的重组蛋白质中是特别重要的。由于标准的原核表达系统缺乏这种修饰所必需的适合的机制,这些治疗性蛋白质的产生必须使用替代的表达系统。酵母和真菌是用于表达蛋白质的引人注目的选择,因为它们在简单的培养基中可以很容易地大规模生长,这使得能够获得低的生产成本。此外,用来操纵酵母和真菌细胞的相对简单的遗传构成以及更复杂的真核细胞,如哺乳动物或昆虫细胞的工具是可得的(De Pourcq等,Appl Microbiol Biotechnol,87(5):1617-31)。

[0008] 真菌细胞和哺乳动物细胞共有导致甘露糖(8)N-乙酰葡萄糖胺(2)(Man8GlcNAc2)形成的糖基化早期阶段中的共同步骤。然而,在该过程的后期阶段中存在着显著的差异。例如,在酵母中,通过甘露糖基转移酶和甘露聚糖聚合酶向Man8GlcNAc2添加额外的甘露糖亚基以得到高甘露糖型的N-聚糖。相反,甘露糖从人Man8GlcNAc2除去以产生Man5GlcNAc2,接着是涉及酶N-乙酰葡萄糖胺基转移酶I(GnTI)、甘露糖苷酶II(Mns II)和N-乙酰葡萄糖胺基转移酶II(GnTII)的三个顺序反应,以将Man5GlcNAc2转化成GlcNAc2Man3GlcNAc2。

[0009] 在哺乳动物和真菌细胞中糖基化过程之间的差异为糖基化哺乳动物蛋白在真菌细胞中的表达提出了挑战,因为具有高甘露糖型N-聚糖的糖蛋白不适合于人类中的治疗应用(De Pourcq等,2010;Wildt和Gerngross,Nature Reviews Microbiology,3:119-128)。因此,已经进行了对酵母和真菌物种中的糖基化途径重新工程化使他们能够表达重组人类蛋白质的研究。酵母或真菌细胞的糖工程(glycoengineering)的一般途径是破坏高甘露糖型N-聚糖的形成中所涉及的内源性基因。可以将这些基因的破坏与内源性甘露糖苷酶和/或来自不同物种的糖基转移酶和糖苷酶的过表达相结合(Chiba等,1998,J Biol Chem 273:26298-304;Kainz等,2008,Appl Environ Microbiol 74:1076-86;Maras等,1997,Euro J Biochem 249:701-07;Maras等,1999,Febs Letters 452:365-70;Hamilton等,2003,Science301:1244-6;De Pourcq等,2010)。然而,在非哺乳动物细胞中糖基化的哺乳动物蛋白质产生仍然需要复杂的和费时的基因工程并且可能在制备所需的糖蛋白时是低

效的。

[0010] 因此,在本领域中仍然需要在非哺乳动物细胞中表达复合的N-聚糖的更简单和更有效的系统。

发明内容

[0011] 本文描述的是包含具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质的组合物。本文还描述了制备复合的N-聚糖的方法和制备Man3GlcNAc2聚糖的方法。

[0012] 因此,一个方面包括具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和催化N-乙酰葡萄糖胺转移到末端Man α 6残基,以及其中该重组蛋白质包含来自至少两种不同的酶的催化结构域。在某些实施方式中,该受体聚糖被连接到选自于氨基酸、肽或多肽的分子上。在某些实施方式中,该分子是异源多肽。在某些可以与前面的实施方式相结合的实施方式中,该受体聚糖是Man3。在某些可以与前面的实施方式相结合的实施方式中,该重组蛋白质是包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质。在某些实施方式中,该N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域来自人类的酶。在某些实施方式中,该N-乙酰葡萄糖胺基转移酶I催化结构域包含与SEQ ID NO:1的氨基酸残基105-445至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。在某些可以与前面的实施方式相结合的实施方式中,该N-乙酰葡萄糖胺基转移酶II催化结构域包含与SEQ ID NO:21的氨基酸残基30-447至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。在某些可以与前面的实施方式相结合的实施方式中,该N-乙酰葡萄糖胺基转移酶I催化结构域是在N-乙酰葡萄糖胺基转移酶II催化结构域的N-末端。在某些可以与前面的实施方式相结合的实施方式中,该N-乙酰葡萄糖胺基转移酶II催化结构域是在N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端。

[0013] 在某些可以与前面的实施方式相结合的实施方式中,该重组蛋白还包含在N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域之间间隔区。在某些实施方式中,该间隔区包含来自非催化结构域(stem domain)的序列。在某些可以与前面的实施方式相结合的实施方式中,该间隔区的长度是至少5、至少10、至少15、至少20、至少30、至少40或至少50个氨基酸。在某些可以与前面的实施方式相结合的实施方式中,该间隔区包含选自于SEQ ID NO:118、SEQ ID NO:120、SEQ ID NO:122和SEQ ID NO:124的序列。在某些实施方式中,该间隔区包含选自于SEQ ID NO:118、SEQ ID NO:120和SEQ ID NO:124的序列。在某些实施方式中,该间隔区包含SEQ ID NO:120或SEQ ID NO:124的序列。在某些实施方式中,该间隔区包含SEQ ID NO:124的序列。

[0014] 在某些可以与前面的实施方式相结合的实施方式中,该重组蛋白质还包含连接到催化结构域的N-末端的靶向肽。在某些实施方式中,该靶向肽包含非催化结构域。在某些实施方式中,该非催化结构域来自N-乙酰葡萄糖胺基转移酶I酶或N-乙酰葡萄糖胺基转移酶II酶。在某些实施方式中,N-乙酰葡萄糖胺基转移酶I酶和N-乙酰葡萄糖胺基转移酶II酶是人类的酶。在某些可以与前面的实施方式相结合的实施方式中,该非催化结构域是来自选自于甘露糖苷酶、甘露糖基转移酶、糖基转移酶、2型高尔基蛋白、MNN2、MNN4、MNN6、MNN9、MNN10、MNS1、

KRE2、VAN1或OCH1的蛋白质。在某些实施方式中,该蛋白质来自选自于枝顶孢属(*Acremonium*)、曲霉属(*Aspergillus*)、短梗霉属(*Aureobasidium*)、隐球菌属(*Cryptococcus*)、金孢子菌属(*Chrysosporium*)、*Chrysosporium lucknowense*、线黑粉酵母属(*Filibasidium*)、镰孢菌属(*Fusarium*)、赤霉属(*Gibberella*)、腐质霉属(*Humicola*)、梨孢菌属(*Magnaporthe*)、毛霉属(*Mucor*)、毁丝霉属(*Myceliophthora*)、漆斑菌属(*Myrothecium*)、新考马脂霉属(*Neocallimastix*)、脉孢霉属(*Neurospora*)、拟青霉属(*Paecilomyces*)、青霉属(*Penicillium*)、瘤胃壶菌属(*Piromyces*)、裂褶菌属(*Schizophyllum*)、踝节菌属(*Talaromyces*)、嗜热子囊菌属(*Thermoascus*)、梭孢壳属(*Thielavia*)、弯颈霉属(*Tolytocladium*)或木霉属(*Trichoderma*)的生物体。在某些可以与前面的实施方式相结合的实施方式中,该靶向肽是Kre2靶向肽。在某些实施方式中,该靶向肽包含跨膜结构域。在某些可以与前面的实施方式相结合的实施方式中,该靶向肽还包含连接到非催化结构域的N-末端的跨膜结构域。在某些可以与前面的实施方式相结合的实施方式中,该跨膜结构域来自N-乙酰葡萄糖胺基转移酶I酶或N-乙酰葡萄糖胺基转移酶II酶。在某些实施方式中,该N-乙酰葡萄糖胺基转移酶I酶和N-乙酰葡萄糖胺基转移酶II酶是人类的酶。在某些可以与前面的实施方式相结合的实施方式中,该跨膜结构域是来自选自于甘露糖苷酶、甘露糖基转移酶、糖基转移酶、2型高尔基蛋白、MNN2、MNN4、MNN6、MNN9、MNN10、MNS1、KRE2、VAN1或OCH1的蛋白质。在某些实施方式中,该蛋白质来自选自于枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、*Chrysosporium lucknowense*、线黑粉酵母属、镰孢菌属、赤霉属、腐质霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属、弯颈霉属或木霉属的生物体。在某些实施方式中,该靶向肽包含胞质结构域。在某些可以与前面的实施方式相结合的实施方式中,该靶向肽还包含连接到非催化结构域的N-末端的胞质结构域。在某些可以与前面的实施方式相结合的实施方式中,该靶向肽还包含连接到跨膜结构域的N末端的胞质结构域。在某些可以与前面的实施方式相结合的实施方式中,该胞质结构域来自N-乙酰葡萄糖胺基转移酶I酶或N-乙酰葡萄糖胺基转移酶II酶。在某些实施方式中,该N-乙酰葡萄糖胺基转移酶I酶和N-乙酰葡萄糖胺基转移酶II酶是人类的酶。在某些可以与前面的实施方式相结合的实施方式中,该胞质结构域是来自选自于甘露糖苷酶、甘露糖基转移酶、糖基转移酶、2型高尔基蛋白、MNN2、MNN4、MNN6、MNN9、MNN10、MNS1、KRE2、VAN1或OCH1的蛋白质。在某些实施方式中,该蛋白质来自选自于枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、*Chrysosporium lucknowense*、线黑粉酵母属、镰孢菌属、赤霉属、腐质霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属、弯颈霉属或木霉属的生物体。

[0015] 另一个方面包括重组蛋白质,其包含其中N-乙酰葡萄糖胺基转移酶II催化结构域位于N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端的人N-乙酰葡萄糖胺基转移酶II催化结构域和人N-乙酰葡萄糖胺基转移酶I催化结构域、含有来自人N-乙酰葡萄糖胺基转移酶I非催化结构域的序列的位于催化结构域之间的间隔区序列以及位于N-乙酰葡萄糖胺基转移酶II催化结构域的N-末端的靶向肽,其中该靶向肽包含胞质结构域、跨膜结构域和来自人N-乙酰葡萄糖胺基转移酶II的非催化结构域。另一个方面包括包含与SEQ ID NO:95至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或

100%相同的序列的重组蛋白质。

[0016] 另一个方面包括重组蛋白质,其包含其中N-乙酰葡萄糖胺基转移酶II催化结构域位于N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端的N-乙酰葡萄糖胺基转移酶II催化结构域和N-乙酰葡萄糖胺基转移酶I催化结构域,其中包含选自于SEQ ID NO:118、SEQ ID NO:120、SEQ ID NO:122和SEQ ID NO:124的序列的位于催化结构域之间的间隔区;以及位于N-乙酰葡萄糖胺基转移酶II催化结构域的N-末端的靶向肽,其中该靶向肽包含胞质结构域、跨膜结构域和来自人N-乙酰葡萄糖胺基转移酶II的非催化结构域。在某些实施方式中,该间隔区包含选自于SEQ ID NO:118、SEQ ID NO:120和SEQ ID NO:124的序列。在某些实施方式中,该间隔区包含SEQ ID NO:120或SEQ ID NO:124的序列。在某些实施方式中,该间隔区包含SEQ ID NO:124的序列。

[0017] 另一个方面包括编码上述实施方式中任一项的重组蛋白质的分离的多核苷酸。另一个方面包括含有可操作地与启动子连接的前述实施方式的分离多核苷酸的表达载体。在某些实施方式中,该启动子是组成型启动子。在某些实施方式中,该启动子是诱导型启动子。在某些实施方式中,启动子是来自选自于gpdA、cbh1、米曲霉(*Aspergillus oryzae*) TAKA淀粉酶、米赫根毛霉(*Phizomucor miehei*)天冬氨酸蛋白酶、黑曲霉(*Aspergillus niger*)中性 α -淀粉酶、黑曲霉酸稳定的 α -淀粉酶、黑曲霉葡糖淀粉酶(glaA)、泡盛曲霉(*Aspergillus awamori*)glaA、米赫根毛霉脂肪酶、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、米曲霉乙酰胺酶、尖孢镰刀菌(*Fusarium oxysporum*)胰蛋白酶样蛋白酶、真菌内切 α -L-阿拉伯糖酶(abnA)、真菌 α -L-阿拉伯呋喃糖苷酶(abfA)、真菌 α -L-阿拉伯呋喃糖苷酶B(abfA)、真菌木聚糖酶(xlnA)、真菌肌醇六磷酸酶、真菌ATP合成酶、真菌亚基9(oliC)、真菌磷酸丙糖异构酶(tpi)、真菌醇脱氢酶(adhA)、真菌 α -淀粉酶(amy)、真菌淀粉葡萄糖苷酶(glaA)、真菌乙酰胺酶(amdS)、真菌甘油醛-3-磷酸脱氢酶(gpd)、酵母醇脱氢酶、酵母醇氧化酶、酵母乳糖酶、酵母3-磷酸甘油酸激酶、酵母磷酸丙糖异构酶、细菌 α -淀粉酶、细菌Spo2或SSO的基因。另一个方面包括含有前述实施方式中任一项的表达载体的宿主细胞。

[0018] 另一个方面包括产生前述实施方式中任一项重组蛋白质的方法,其包括以下步骤:将编码重组蛋白质的分离的多核苷酸引入宿主细胞,和培养该宿主细胞以使得表达该重组蛋白质。在某些实施方式中,该方法还包括从宿主细胞纯化重组蛋白质的步骤。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞是真菌细胞。在某些实施方式中,该真菌细胞选自于酵母或丝状真菌。

[0019] 另一个方面包括制备复合的N-聚糖的方法,其包括以下步骤:提供宿主细胞,其中该宿主细胞包含编码融合蛋白质的多核苷酸,该融合蛋白质包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域,和培养该宿主细胞以使得表达该融合蛋白质,其中该融合蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到末端Man α 6残基以产生复合的N-聚糖。在某些实施方式中,该复合的N-聚糖被连接于选自于氨基酸、肽或多肽的分子。在某些实施方式中,该分子是异源多肽。在某些可以与前面的实施方式相结合的实施方式中,该受体聚糖是Man3。在某些可以与前面的实施方式相结合的实施方式中,该复合的N-聚糖是GlcNAc β 2Man α 3(GlcNAc β 2Man α 6)Man β 4GlcNAc β 4GlcNAc。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞是真核

细胞。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞是真菌细胞。在某些实施方式中,该真菌细胞是选自于酿酒酵母(*S.cerevisiae*)、乳酸克鲁维酵母(*K.lactis*)、毕赤酵母(*P.pastoris*)、汉逊酵母(*H.polymorpha*)、白假丝酵母(*C.albicans*)、裂殖酵母(*Schizosaccharomyces*)或耶氏酵母(*Yarrowia*)的酵母细胞。在某些可以与前面的实施方式相结合的实施方式中,该真菌细胞是选自于木霉菌、枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、*Chrysosporium lucknowense*、线黑粉酵母属、镰孢菌属、赤霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属或弯颈霉属的丝状真菌细胞。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞还包含编码UDP-GlcNAc转运蛋白的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞具有与野生型宿主细胞中的活性水平相比降低的多萜醇基(dolichyl)-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶的活性水平。在某些实施方式中,该宿主细胞具有与野生型宿主细胞中的表达水平相比降低的alg3基因的表达水平。在某些实施方式中,从宿主细胞删除该alg3基因。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞具有与野生型宿主细胞中的活性水平相比降低的 α -1,6-甘露糖基转移酶的活性水平。在某些实施方式中,该宿主细胞具有与野生型宿主细胞中的表达水平降低的och1基因的表达水平。在某些实施方式中,从宿主细胞中删除了och1基因。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞还包含编码 α -1,2-甘露糖苷酶的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞还包含编码 β -1,4-半乳糖基转移酶的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞还包含编码唾液酸转移酶的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞是木霉属的细胞,其具有与野生型的木霉属细胞中的活性水平相比降低的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶的活性水平。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞是酵母或真菌细胞,其具有与野生型酵母细胞中的活性水平相比降低的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶的活性水平和降低的 α -1,6-甘露糖基转移酶的活性水平,并进一步包含编码 α -1,2-甘露糖苷酶的多核苷酸。

[0020] 另一个方面包括制备复合的N-聚糖的方法,其包括以下步骤:提供木霉属的宿主细胞,其中该宿主细胞具有与野生型宿主细胞中的表达水平降低的alg3基因的表达水平且包含编码N-乙酰葡萄糖胺基转移酶I催化结构域的第一多核苷酸和编码N-乙酰葡萄糖胺基转移酶II催化结构域的第二多核苷酸;以及培养该宿主细胞以产生复合的N-聚糖。

[0021] 另一个方面包括制备复合的N-聚糖的方法,其包括以下步骤:在缓冲液中将包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质、受体聚糖和N-乙酰葡萄糖胺供体一起孵育,其中该融合蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到末端Man α 6残基以产生复合的N-聚糖。在某些实施方式中,该受体聚糖被连接于选自于氨基酸、肽或多肽的分子上。在某些实施方式中,该分子是异源多肽。在某些实施方式中,该受体聚糖是Man3。在某些可以与前面的实施方式相结合的实施方式中,该N-乙酰葡萄糖胺供体是UDP-GlcNAc转运蛋白。

[0022] 另一个方面包括丝状真菌细胞,其包含alg3的突变和Man3GlcNAc2,其中该Man3GlcNAc2包括细胞分泌的中性N-聚糖的至少50%、至少60%、至少70%、至少80%、至少

90%或100%(摩尔%)。该中性N-聚糖可连接到选自于氨基酸、肽和多肽的分子上。在某些实施方式中,alg3的突变是alg3的缺失。在某些可以与前面的实施方式相结合的实施方式中,该细胞是里氏木霉(*Trichoderma reesei*)细胞。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞还包含编码N-乙酰葡萄糖胺基转移酶I催化结构域的第一多核苷酸和编码N-乙酰葡萄糖胺基转移酶II催化结构域的第二多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞还包含编码含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质的多核苷酸。

[0023] 另一个方面包括在宿主细胞中产生Man3GlcNAc2聚糖的方法,其包括以下步骤:提供具有与野生型宿主细胞中的活性水平相比降低的甘露糖基转移酶的活性水平的宿主细胞并培养该宿主细胞以产生Man3GlcNAc2聚糖,其中该Man3GlcNAc2聚糖包括宿主细胞分泌的中性N-聚糖的至少50%、至少60%、至少70%、至少80%、至少90%或100%(摩尔%)。该中性N-聚糖可连接到选自于氨基酸、肽和多肽的分子上。在某些实施方式中,该Man3GlcNAc2聚糖被连接到异源多肽上。在某些可以与前面的实施方式相结合的实施方式中,该甘露糖基转移酶是多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞具有与野生型宿主细胞中的表达水平相比降低的alg3基因的表达水平。在某些实施方式中,从宿主细胞中删除alg3基因。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞是木霉属细胞。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞中的 α -1,6-甘露糖基转移酶的活性水平与野生型宿主细胞中的活性水平相比降低。在某些可以与前面的实施方式相结合的实施方式中,该宿主细胞含有编码 α -1,2-甘露糖苷酶的内源性多核苷酸。

[0024] 另一个方面包括丝状真菌细胞,其具有与野生型丝状真菌细胞中的表达水平相比降低的alg3基因的表达水平,其中该丝状真菌细胞包含前述实施方式中任一项的重组蛋白质。在某些实施方式中,该alg3基因包含突变。优选地,该重组蛋白质具有N-乙酰葡萄糖胺基转移酶活性,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和催化N-乙酰葡萄糖胺转移到末端Man α 6残基,和其中该重组蛋白质是包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质。在某些实施方式中,alg3基因的突变是alg3基因的缺失。在某些可以与前面的实施方式相结合的实施方式中,该融合蛋白质是由可操作地与启动子连接的多核苷酸编码的。在某些实施方式中,该启动子是诱导型启动子。在某些实施方式中,该诱导型启动子是cbh1启动子。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞还包含编码UDP-GlcNAc转运蛋白的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞具有与野生型的丝状真菌细胞中的活性水平相比降低的 α -1,6-甘露糖基转移酶的活性水平。在某些实施方式中,该丝状真菌细胞具有与野生型宿丝状真菌细胞中的表达水平相比降低的och1基因的表达水平。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞还包含编码 α -1,2-甘露糖苷酶的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞还包含编码 β -1,4-半乳糖基转移酶的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞还包含编码唾液酸转移酶的多核苷酸。在某些可以与前面的实施方式相结合的实施方式中,该丝状真菌细胞选自于木霉菌、枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、*Chrysosporium lucknowense*、线黑粉酵母属、镰孢

菌属、赤霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属和弯颈霉属。

附图说明

[0025] 图1显示了里氏木霉菌株M44、M81、M84、M109、MH0、M131、M132、M133、M134和M124上一般糖基化(average glycosylation)的中性N-聚糖质谱图。

[0026] 图2显示单磷酸化的Man7Gn2的碎片分析。仅显示了单磷酸化的Man7Gn2的一种示例性结构。

[0027] 图3显示了里氏木霉菌株M44、M81、M84、M109、M110、M131、M132、M133、M134和M124的酸性聚糖质谱图。

[0028] 图4显示了在发酵罐中培养131.4小时(分批补料)的里氏木霉菌株M44的中性(a)和酸性(b)N-聚糖图谱。

[0029] 图5显示了里氏木霉培养基中中性(a)和酸性(b)N-聚糖的质谱图。

[0030] 图6显示了里氏木霉分泌的蛋白质的膜印迹。

[0031] 图7显示了在发酵罐中培养的里氏木霉M44的分析蛋白质带的实例。蛋白质的糖基化与里氏木霉的一般糖基化没有显著不同。谱带聚焦于较少的基线信号,和与其他信号相比,频谱的主要信号是非定量的。

[0032] 图8显示了来自亲本菌株和来自Alg3敲除菌株的DNA采用alg3探针获得的Southern印迹。

[0033] 图9A显示具有预测的限制性产物大小的pTTv38构建体片段的限制性内切酶图谱。图9B显示了用EcoRI+PvuI(E+P)或KpnI+NheI(K+N)消化的来自亲本菌株和Alg3敲除菌株的基因组DNA的Southern印迹。对照DNA是用NotI消化的pTTv38质粒DNA。该印迹是用AmdS探针探测。

[0034] 图10显示了中性N-聚糖的MALDI分析。部分A显示亲本菌株M124。部分B显示Alg3敲除的4A。方块代表N-乙酰葡萄糖胺,且圆圈代表甘露糖,除了标记为葡萄糖的一个。

[0035] 图11显示了来自4A Alg3敲除菌株的Man3Gn2的碎片分析。

[0036] 图12显示了来自Alg3敲除菌株4A(部分A)和亲本菌株M124(部分B)的Hex5Gn2的碎片分析。用方框标记的信号仅作为来自Alg3敲除菌株的异构体存在。

[0037] 图13显示了 α -甘露糖苷酶消化后的来自Alg3敲除菌株4A的中性N-聚糖。

[0038] 图14显示了通过液相色谱分离来自Alg3敲除菌株的两种主要聚糖。

[0039] 图15显示了Hex3HexNAc2(A部分)和Hex6HexNAc2(B部分)级分的质子NMR谱。在40°C下,使用配备有冷冻探针的Varian Unity INOVA 600MHz谱仪收集质谱。

[0040] 图16显示了亲本菌株M124(部分A)和Alg3敲除菌株4A(B)的酸性级分。用星号标记具有两个磷酸根单元的N-聚糖。

[0041] 图17显示了来自在烧瓶中培养5天的里氏木霉Alg3敲除菌株4A的上清液的中性N-聚糖。

[0042] 图18显示了来自在发酵罐中培养10天的里氏木霉Alg3敲除菌株4A的上清液的中性N-聚糖。

[0043] 图19显示了GnTI反应混合物的MALDI图谱。GnTI已经将54%的受体转化成具有一

个额外的HexNAc的产物。

[0044] 图20显示了GnTII表达的Western印迹分析。在12%SDS-PAGE凝胶上跑样和在硝酸纤维素膜上进行印迹。使用小鼠 α -HIS单克隆抗体在膜上检测到组氨酸标记的GnTII。在左侧所示的数字是分子量标志物蛋白质的大小(kDa)。

[0045] 图21显示了GnTII反应混合物的MALDI谱。83%的受体(m/z 913.340)转化成产物(m/z 1136.433)。

[0046] 图22显示了对于GnTI/GnTII融合蛋白质观察到的GnTI活性。

[0047] 图23显示了通过靶向于alg3基因座得到的GnTI/GnTII里氏木霉转化体中存在的N-聚糖。

[0048] 图24显示了来自GnTII/GnTI融合蛋白质的酶活性测试的纯化反应混合物的MALDI谱。

[0049] 图25显示了 β 1-2,3,4,6-N-乙酰氨基葡萄糖苷酶反应混合物的图谱。

[0050] 图26显示了 β 1-4GalT反应混合物的MALDI谱。

[0051] 图27显示了观察到的来自里氏木霉M127pTTv110转化体(alg3基因座的gnt II/I)在第3天(A)、第5天(B)和第7天(C和D)的上清液蛋白质的N-聚糖。克隆17A在第7天产生了最多的G0。(E)来自在摇瓶中培养7天的里氏木霉菌株M127GnTII/I转化体克隆17A的上清液蛋白质的中性N-聚糖质谱。标记星号的信号来源于培养基。

[0052] 图28显示了来自里氏木霉M202GnT II/I转化体克隆(A)9A-1和(B)31A-1(两者都用大豆胰蛋白酶抑制剂培养)的利妥昔单抗的中性N-聚糖,及(C)从在大豆胰蛋白酶抑制剂存在下在摇瓶中培养5天的里氏木霉M202GnT II/I转化体克隆9A-1纯化的利妥昔单抗的中性N-聚糖的质谱。

[0053] 图29显示了间隔区修饰的GnTII/GnTI融合反应混合物的MALDI谱。部分(A)显示了具有3xG4S间隔区修饰的GnTII/GnTI的反应混合物。36%的受体已被转化为具有两个额外的HexNAc的产物。部分(B)显示了具有2xG4S间隔区修饰的GnTII/GnTI的反应混合物。38%的受体已被转化为具有两个额外的HexNAc的产物。在任一图谱中未检测到GnTI产物Hex3HexNAc2(计算m/z 933.318)的[M+Na]⁺信号的计算m/z值,因为所有的GnTI产物直接转化成Hex3HexNAc3(计算m/z 1136.318)。

[0054] 图30显示了GnTII/I间隔区变异体的细胞沉淀(A)和上清液(B)的Western印迹。泳道1:GnTII阳性对照,2:GY3模拟株,3:GY7-2野生型GnTII/I,4:GY32-5 3xG4S间隔区,5:GY32-9 3xG4S间隔区,6:GY33-7 2xG4S间隔区,7:GY33-8 2xG4S间隔区,8:GY49-3CBHI间隔区和9:GY50-10 EGIV间隔区。

[0055] 图31显示了在(A)第3天的表达阶段和(B)第4天的表达阶段后来自上清液的和在蛋白酶抑制剂存在下表达的野生型GnTII/I和间隔区变异体的GnT活性。x轴标示了样品身份(wt=野生型;_1、_2=间隔区变异体的平行克隆),且y轴标示了形成的产物的百分比(GnTI和GnTII反应产物加在一起)。

[0056] 图32显示了上清液、细胞和溶解产物中的GnTII/I融合蛋白质(具有野生型间隔区)的GnT活性。GnTI和GnTII产物被加在一起。

[0057] 图33显示了(A)上清液、(B)细胞和(C)溶解产物中GnTII/I野生型和间隔区变异体的GnT活性。

[0058] 图34显示了亲本菌株M124和GnTI转化体在第5天的中性N-聚糖的示例性图谱。用箭头标记Gn增加的信号(m/z 1460)(具有cbh1启动子的pTTv11,具有gpdA启动子的pTTv13)。

[0059] 图35显示了第3天和第5天四个阳性GnTI转化体中Man5和Gn1Man5的量。定量相对于内部校准物(Hex2HexNAc4,2pmol)进行。

[0060] 图36显示了具有内部校准物(NeuAcHex4HexNAc2,0.5pmol)的亲本M124菌株和GnTI转化体的磷酸化N-聚糖的示例性图谱。GnTI产物用箭头标记。

[0061] 图37显示了来自第5天的不同GnTII菌株/克隆的中性N-聚糖的图表。部分(A)显示pTTv140克隆。部分(B)显示pTTv142克隆。部分(C)显示pTTv143克隆。部分(D)显示pTTv141克隆。

[0062] 图38显示了来自第3天、第5天和第7天的不同GnTII菌株/克隆和亲本菌株M198的中性N-聚糖的实例。部分(A)显示克隆1-117A。部分(B)显示克隆3-11A。部分(C)显示克隆30A。部分(D)显示亲本菌株M198。

[0063] 图39显示了里氏木霉菌株M198和GnTII克隆3-17A的经分离的蛋白质的膜。用箭头标记50kDa的蛋白质。

[0064] 图40显示了亲本菌株M198(A)和GnTII克隆3-17A(B)的总的分泌蛋白质相对单个分泌蛋白质的柱形图。

[0065] 图41显示了从第3天到第7天发酵罐培养的GnTII菌株M329和第5天的摇瓶培养菌株M329的柱形图。

[0066] 图42显示了里氏木霉ALG3和ALG3同源物的多个氨基酸序列的比对。

[0067] 详细说明

[0068] 本发明涉及具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质,其中该重组蛋白质催化N-乙酰葡萄糖胺(GlcNAc)转移到受体聚糖的末端Man α 3残基和催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 6残基上,以及其中该重组蛋白质包含来自至少两种不同的酶的催化结构域。

[0069] 在一些实施方式中,本发明的重组蛋白质包括两个催化结构域,其中一个催化结构域具有N-乙酰葡萄糖胺基转移酶I(GnTI)活性(例如,与末端Man α 3残基反应),和另一个催化结构域具有N-乙酰葡萄糖胺基转移酶II(GnTII)活性(例如,与末端Man α 6残基反应)。

[0070] 在一些实施方式中,本发明的重组蛋白质催化基本上顺序发生的反应。例如,本发明的重组蛋白质可以首先催化GlcNAc转移到末端Man α 3残基上,然后催化GlcNAc转移到受体聚糖的末端Man α 6残基上。在一个实施方式中,该基本上顺序的反应比按相反顺序的两个反应更有效至少10倍、至少20倍、至少30倍、至少40倍、至少50倍、至少60倍、至少70倍、至少80倍、至少90倍或至少100倍。在某些实施方式中,顺序反应意思是,如果GlcNAc还没有被转移到末端Man α 3残基上,则基本上或完全没有GlcNAc可以被转移到末端Man α 6残基上。在具体的实施方式中,受体聚糖包含GlcNAc β 2Man α 3分支。

[0071] 在一些实施方式中,重组蛋白质与Man α 3和Man α 6残基(任选地在分支受体聚糖中)特异性地反应,但是基本上不或绝对不与其他Man α 结构(例如Man α 单糖偶联物)、与Man α 苯甲基和/或Man α Ser/T-肽反应。非实质反应性优选地是低于与末端Man α 3和Man α 6残基的反应的0.1mM受体聚糖浓度的 V_{max} 的10%、低于8%、低于6%、低于4%、低于2%、低于1%或低

于0.1%。在具体实施方式中,重组蛋白质与受体聚糖的末端Man α 3残基(优选为GnTI反应)和末端Man α 6残基(优选为GnTII反应)具有基本上相似的反应性。优选地,在相同的条件下,任一种催化活性与另一种催化活性相比没有超过10倍、5倍、3倍或2倍的反应效率的差异。

[0072] 在具体实施方式中,GlcNAc转移至末端Man α 3和Man α 6上产生至少10%、至少25%、至少50%、至少70%、至少90%或至少95%的Man3聚糖转化为带有两个末端GlcNAc的聚糖的转化率。反应的效率可以通过如本文公开的实施例中所述的体外或体内分析来测量。可以基本上如实施例中所述,或作为0.1mM的受体浓度和饱和供体浓度下的最大反应速率V_{max}来测量GlcNAc转移反应的效率。在具体实施方式中,用连接到氨基酸、肽或多肽上的Man3受体聚糖来测量反应的效率。

[0073] 本公开还涉及到一种制备复合的N-聚糖的方法,包括以下步骤:提供宿主细胞,其中该宿主细胞包含编码其中包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质的多核苷酸,以及培养该宿主细胞以使得该融合蛋白质被表达,其中该融合蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 6残基和以产生复合的N-聚糖。

[0074] 本发明还涉及一种具有与野生型丝状真菌细胞中的表达水平相比降低的alg3基因表达水平的丝状真菌细胞,其中该丝状真菌细胞包含本发明的重组蛋白质。

[0075] 定义

[0076] 如本文所使用的“重组蛋白质”是指由重组核酸产生的任何蛋白质。如本文所使用的“重组核酸”是指其中至少以下一项为真的核酸多聚物:(a)核酸的序列对于给定的宿主细胞是外来的(即,不是天然存在于其中的);(b)序列可以在给定的宿主细胞中天然地存在,但以非自然的(例如,大于预期的)量存在或以多于或少于天然表达水平的水平表达;或(c)核酸的序列包括两个或更多的子序列,其不是以在自然界发现的相同相互关系而存在。例如,对于情形(c),重组核酸序列将具有两个或更多个来自不相关的基因的序列,其进行排列以形成新的功能性核酸。在另一个实例中,重组核酸序列将包含天然地不是彼此相邻的启动子序列和基因编码序列。

[0077] 如本文所使用的“N-乙酰葡萄糖胺基转移酶活性”是指将N-乙酰葡萄糖胺基残基(GlcNAc)转移到受体聚糖的酶的活性。通常,具有这一活性的酶是N-乙酰葡萄糖胺基转移酶(GlcNAc转移酶)。在某些实施方式中,GlcNAc转移酶是真核的。在某些实施方式中,GlcNAc转移酶是形成从GlcNAc-残基的1-位至末端甘露糖残基的 β -连接的哺乳动物酶。在某些实施方式中,GlcNAc转移酶是将 β 2-连接的GlcNAc-残基转移至聚糖的2-位末端甘露糖残基(特别是至N-连接的聚糖)的 β 2-N-乙酰葡萄糖胺基转移酶。在某些实施方式中, β 2-GlcNAc转移酶是具有GnTI活性和GnTII活性的酶。GnTI活性将GlcNAc残基转移至Man α 3分支。该Man α 3分支可以是N-连接的聚糖核心结构上的Man α 3(R-Man α 6)Man β 分支,如Man3GlcNAc2或Man3或Man5GlcNAc2或Man5。GnTI酶可以是哺乳动物酶、植物酶或低等真核生物酶。GnTII活性将GlcNAc残基转移到N-连接的聚糖核心结构的Man α 6-分支上如Man α 6(GlcNAc β 2Man α 3)Man β -分支。这样的Man α 6-分支的例子是GlcNAc1Man3GlcNAc2。

[0078] 如本文所使用的“N-乙酰葡萄糖胺”是指N-乙酰葡萄糖胺残基(GlcNAc)。GlcNAc可以是聚糖结构的部分。胺基团在2位上,具有D-构型,且具有作为残基的吡喃糖结构。或者它可以被命名为2-乙酰氨基-2-脱氧-D-吡喃葡萄糖(D-GlcpNAc)。GlcNAc也可以是游离的还原性

单糖(即不是聚糖的部分)。

[0079] 如本文所使用的“Man”是指甘露糖残基。“末端Man α 3”或“末端Man α 6”是指未被另一个单糖残基或多个残基取代为非还原性末端的末端残基的甘露糖。

[0080] 如本文所使用的“聚糖”是指可连接到载体,如氨基酸、肽、多肽、脂质或还原性末端偶联物上的寡糖链。在某些实施方式中,本发明涉及通过至天冬酰胺残基(Asn)侧链的酰胺氮的N-连接与多肽N-糖基化位点(例如-Asn-Xxx-Ser/Thr-)偶联的N-连接聚糖,其中Xxx是除脯氨酸外的任何氨基酸残基。本发明还可以涉及作为多萜醇基-磷酸-寡糖(Dol-P-P-OS)前体脂质结构的部分的聚糖,其是真核细胞的内质网中N-连接聚糖的前体。该前体寡糖由其还原性末端连接至多萜醇脂质上的两个磷酸残基。例如, α 3甘露糖基转移酶Alg3修饰N-聚糖的Dol-P-P-寡糖前体。一般情况下,本文所描述的聚糖结构是末端聚糖结构,其中非还原性残基不被其他的一个或多个单糖残基修饰。

[0081] 如本文所使用的“糖蛋白”是指连接至聚糖的肽或多肽。聚糖可以在共翻译或翻译后的修饰中连接于肽或多肽上。

[0082] 如本文所使用的“糖脂”是指连接到聚糖上的脂质并包括甘油糖脂、鞘糖脂和糖基磷脂酰肌醇。

[0083] 如整个本公开中所使用的,糖脂和碳水化合物的命名基本上是根据由IUPAC-IUB生物化学命名委员会的推荐(例如Carbohydrate Res.1998,312,167;Carbohydrate Res.1997,297,1;Eur.J.Biochem.1998,257,29)。假定Gal(半乳糖)、Glc(葡萄糖)、GlcNAc(N-乙酰葡萄糖胺)、GalNAc(N-乙酰半乳糖胺)、Man(甘露糖)和Neu5Ac是D-构型的,Fuc是L-构型的,且所有单糖单元为吡喃糖形式(D-Galp、D-Glcp、D-GlcpNAc、D-GalpNAc、D-Manp、L-Fucp、D-Neup5Ac)。胺基团是如对于天然半乳糖和GalNAc或GlcNAc的2-位上的葡糖胺限定的。糖苷键部分地以较短的命名表示,部分地以较长的命名表示,唾液酸SA/Neu5X-残基 α 3和 α 6的意思分别与 α 2-3和 α 2-6相同,并且对于己糖的单糖残基, α 1-3、 α 1-6、 β 1-2、 β 1-3、 β 1-4和 β 1-6可以分别缩写为 α 3、 α 6、 β 2、 β 3、 β 4和 β 6。乳糖胺是指II型N-乙酰基乳糖胺、Gal β 4GlcNAc和/或I型N-乙酰基乳糖胺, Gal β 3GlcNAc,且唾液酸(SA)是指N-乙酰基神经氨酸(Neu5Ac)、N-羟乙酰神经氨酸(Neu5Gc)或任何其它天然的唾液酸,包括Neu5X的衍生物。唾液酸是指NeuNX或Neu5X,其中优选X是Ac或GC。有时Neu5Ac/Gc/X可被称为NeuNAc/NeuNGc/NeuNX。

[0084] 本发明的重组蛋白质

[0085] 本发明涉及具有N-乙酰葡萄糖胺基转移酶活性的重组蛋白质,其中该重组蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和催化N-乙酰葡萄糖胺转移到末端Man α 6残基上。本发明的重组蛋白质可以包括但不限于,具有N-乙酰葡萄糖胺基转移酶活性的全长蛋白质、具有N-乙酰葡萄糖胺基转移酶活性的蛋白质片段、具有N-乙酰葡萄糖胺基转移酶活性的催化结构域和具有N-乙酰葡萄糖胺基转移酶活性的融合蛋白质。本发明的单一重组蛋白质具有催化N-乙酰葡萄糖胺的两种转移的能力。N-乙酰葡萄糖胺转移到末端Man α 3残基可以在N-乙酰葡萄糖胺转移到末端Man α 6残基之前或之后发生。或者,该转移可以同时发生。

[0086] 受体聚糖可以连接到如氨基酸、肽或多肽的分子上。在某些实施方式中,所述氨基酸为天冬酰胺残基。天冬酰胺残基可以由侧链酰胺形成氨基糖苷键(生物的哺乳动物多肽N-聚糖连接结构),并可以是肽链(如二肽、寡肽或多肽)的部分。聚糖可以是还原性末端的

衍生物如还原性GlcNAc或Man的N-、O-或C-连接的(优选的糖苷的)衍生物,例如具有某些聚糖连接结构(选自于氨基酸、烷基、杂烷基、酰基、烷氧基、芳基、芳基烷基或杂芳基烷基)的间隔体或末端有机残基。间隔体还可以连接到多价载体或固相上。在某些实施方式中,含烷基的结构包括:甲基、乙基、丙基和C4-C26烷基、脂质如甘油脂质、磷脂、多萜醇磷脂和神经酰胺及衍生物。还原性末端还可以通过还原性胺化为仲胺连接或衍生结构而衍生化。某些载体包括生物多聚-或寡聚物,如(多)肽、多糖如葡聚糖、纤维素、直链淀粉或葡糖胺聚糖和其它有机聚合物或低聚物,如塑料,包括聚乙烯、聚丙烯、聚酰胺(例如,尼龙或聚苯乙烯)、聚丙烯酰胺和聚乳酸、树状聚合物如PAMAM、星型聚合物或海星树状聚合物或聚赖氨酸及聚烷基二醇如聚乙二醇(PEG)。固相可以包括微量滴定孔、二氧化硅颗粒、玻璃、金属(包括钢、金和银),聚合物珠粒如聚苯乙烯或树脂珠、聚乳酸珠、聚糖珠或含有磁珠的有机间隔物。

[0087] 在某些实施方式中,受体聚糖被连接到异源多肽。如本文所使用的“肽”和“多肽”是包括多个连续聚合的氨基酸残基的氨基酸序列。通常情况下,对于本发明来说,肽是包括最多50个氨基酸残基的那些分子,以及多肽包括超过50个氨基酸残基。肽或多肽可以包括修饰的氨基酸残基、不被密码子所编码的天然存在的氨基酸残基以及非天然存在的氨基酸残基。如本文所使用的“蛋白质”可以指任何大小的肽或多肽。术语“异源多肽”是指不是在给定宿主细胞中天然发现的,或对于给定的宿主细胞不是内源性的多肽。在某些实施方式中,异源多肽是治疗性蛋白质。例如,治疗性蛋白质可包括单克隆抗体、促红细胞生成素、干扰素、生长激素、酶或凝血因子。例如,受体聚糖可以连接到治疗性蛋白质如利妥昔单抗上。

[0088] 受体聚糖

[0089] 在某些实施方式中,受体聚糖的结构具有下式: $[R_1]_y\text{Man}\alpha 3([R_2]_z\text{Man}\alpha 6)\text{Man}\{\beta 4\text{GlcNAc}(\text{Fuc}\alpha x)_n[\beta 4\text{GlcNAc}]_m\}_q$,其中q、y、z、n和m是0或1;x是任选的海藻糖残基的连接位3或6; R_1 是GlcNAc,优选GlcNAc $\beta 2$;和 R_2 是分支结构的 $\text{Man}\alpha 3(\text{Man}\alpha 6)$,条件是当z是1时,则y为0,和当z为0时,则y为0或1。()限定在正则的N-聚糖核心结构中的分支,其存在或不存在。[]和{}限定线性序列中存在或不存在的聚糖结构的部分。当z是0和y是0时,那么该结构是 Man_3 聚糖,和当z是0和y是1时,该结构是 GlcNAcMan_3 聚糖。当y为0和z为1时,该聚糖是 Man_5 聚糖。受体聚糖可以 β -糖苷连接到Asn残基上,优选从还原性末端GlcNAc。在一个实施方式中,受体聚糖是多肽连接的N-聚糖,其中m和q是1,且受体结构包含 $[R_1]_y\text{Man}\alpha 3([R_2]_z\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}(\text{Fuc}\alpha x)_n\beta 4\text{GlcNAc}$ 的衍生物。任选的衍生物包括单糖残基如GlcNAc或木糖的取代。

[0090] 受体聚糖可以是 Man_3 、 GlcNAcMan_3 或 Man_5 。在某些实施方式中,受体聚糖是 Man_3 或 GlcNAcMan_3 。 Man_3 是三甘露糖基的聚糖,其包含至少一个 $\text{Man}\alpha 3$ 或 $\text{Man}\alpha 6$ 残基,和优选是支链寡糖,如 $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}$ 。其他的某些 Man_3 寡糖是 $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta$ 、 $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}$ 和多肽连接的 $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}\beta 4\text{GlcNAc}$ 。此外,根据宿主细胞的不同,聚糖可以在Man和/或GlcNAc残基中包含Fuc、Xyl或GlcNAc,如 $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}\beta 4(\text{Fuc}\alpha x)_n\text{GlcNAc}$,其中x是3或6和n是0或1,也由指明了末端甘露糖结构和还原性末端成的单糖组成式如 $\text{Man}_3\text{GlcNAc}_2$ (n为0)和 $\text{Man}_3\text{GlcNAc}_2\text{Fuc}$ (n是1)描述。在某些实施方式中,特别是那些具有多肽连接结构的, Man_3 结构是 $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}\beta 4(\text{Fuc}\alpha 6)_n\text{GlcNAc}$ 。在某些实施方式中,多肽连接的 GlcNAcMan_3 的结构是 $\text{GlcNAc}\beta 2\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}\beta 4(\text{Fuc}\alpha 6)_n\text{GlcNAc}$,也由单糖组成式 $\text{GlcNAcMan}_3\text{GlcNAc}_2$ (n为0)和 $\text{GlcNAcMan}_3\text{GlcNAc}_2\text{Fuc}$ (n

为1)描述。在某些实施方式中,多肽连接的Man5结构是 $\text{Man}\alpha 3\{\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\alpha 6\}\text{Man}\beta 4\text{GlcNAc}\beta 4(\text{Fuc}\alpha 6)_n\text{GlcNAc}$,其中{}和()表示分支和n为0或1,也由单糖组成式 Man5GlcNAc2 (n是0)和 Man5GlcNAc2Fuc (n是1)描述。

[0091] 因此,某些Man3聚糖具有根据下式的结构, $\text{Man}\alpha 3(\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}(\text{Fuc}\alpha x)_n\beta 4\text{GlcNAc}$,其中n是0或1,表示分子部分的存在或不存在,其中x是3或6,和其中()限定结构中的分支。在其中受体聚糖是Man3的本发明实施方式中,该重组蛋白质催化N-乙酰葡萄糖胺转移至Man3的末端 $\text{Man}\alpha 3$ 和 $\text{Man}\alpha 6$,从而产生 GlcNAc2Man3 , $\text{GlcNAc}\beta 2\text{Man}\alpha 3(\text{GlcNAc}\beta 2\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}\beta 4(\text{Fuc}\alpha x)_n\text{GlcNAc}$,其中n是0或1,也通过单糖组成式 $\text{GlcNAc2Man3GlcNAc2}$ (n为0)和 $\text{GlcNAc2Man3GlcNAc2Fuc}$ (n是1)来描述。

[0092] 在其中受体聚糖是Man5的本发明实施方式中,重组蛋白质催化N-乙酰葡萄糖胺转移到Man5的末端 $\text{Man}\alpha 3$ 。已经将2个甘露糖从 GlcNAcMan5 (例如,通过甘露糖苷酶II)移除而形成 GlcNAcMan3 后,该重组蛋白质催化N-乙酰葡萄糖胺转移到末端 $\text{Man}\alpha 6$,因此形成 GlcNAc2Man3 (其具有结构 $\text{GlcNAc}\beta 2\text{Man}\alpha 3(\text{GlcNAc}\beta 2\text{Man}\alpha 6)\text{Man}\beta 4\text{GlcNAc}\beta 4(\text{Fuc}\alpha x)_n\text{GlcNAc}$,其中n是0或1,如果连接到抗体上也被称为G0)。

[0093] 含有N-乙酰葡萄糖胺基转移酶催化结构域的融合蛋白质

[0094] 在某些实施方式中,本发明的重组蛋白质含有N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质。术语“融合蛋白质”是指含有与异源氨基酸连接的蛋白质或多肽的任何蛋白质或多肽。

[0095] N-乙酰葡萄糖胺基转移酶I(GlcNAc-TI ;GnTI;EC 2.4.1.101)催化反应 $\text{UDP-N-乙酰基-D-葡萄糖胺}+3-(\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R}\rightleftharpoons\text{UDP}+3-(2-(\text{N-乙酰基}-\beta\text{-D-葡萄糖胺基})-\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R}$,其中R代表聚糖受体中N-连接的寡糖的其余部分。N-乙酰葡萄糖胺基转移酶I催化结构域是能够催化该反应的N-乙酰葡萄糖胺基转移酶I酶的任何部分。SEQ ID NO:1-19中列出了来自各种生物体的N-乙酰葡萄糖胺基转移酶I酶的氨基酸序列。另外的GnTI酶在CAZy数据库中的糖基转移酶家族13(cazy.org/GT13_all)中列出。酶学特征性的种类包括拟南芥(*A.thaliana*)AAR78757.1(US6 653459)、秀丽隐杆线虫(*C.elegans*)AAD03023.1(Chen S.等J.Biol.Chem1999;274(1):288-97)、黑腹果蝇(*D.melanogaster*)AAF57454.1(Sarkar和Schachter Biol Chem.2001 Feb;382(2):209-17)、中国仓鼠(*C.griseus*)AAC52872.1(Puthalakath H.等J.Biol.Chem 1996 271(44):27818-22)、智人(*H.sapiens*)AAA52563.1(Kumar R.等Proc Natl Acad Sci U S A.1990 Dec;87(24):9948-52)、金黄地鼠(*M.auratus*)AAD04130.1(Opat As等Biochem J.1998 Dec 15;336(Pt 3):593-8),(包括失活突变体的例子)、兔,穴兔(*O.cuniculus*)AAA31493.1(Sarkar M等.Proc Natl Acad Sci U S A.1991 Jan 1;88(1):234-8)。可以在cazy.org/GT13_characterized中找到特征性活性酶的其他例子。在Unligil UM等EMBO J.2000 Oct 16;19(20):5269-80中通过X-射线晶体学描述了兔GnTI的催化结构域的三维结构。GnTI的蛋白质数据库(PDB)结构是1F08、1F09、1FOA、2AM3、2AM4、2AM5和2APC。在某些实施方式中,N-乙酰葡萄糖胺基转移酶I催化结构域是来自人N-乙酰葡萄糖胺基转移酶I酶(SEQ ID NO:1)或其变异体。在某些实施方式中,N-乙酰葡萄糖胺基转移酶I催化结构域包含与SEQ ID NO:1的氨基酸残基84-445至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。在一些实施方式中,可以使用较

短的序列作为催化结构域(如人类酶的氨基酸残基105-445或兔子的酶的氨基酸残基107-447;Sarkar等(1998)Glycoconjugate J 15:193-197)。可以用作GnTI催化结构域的另外的序列包括来自人类酶的大致氨基酸30至445的氨基酸残基或者在氨基酸残基30至105之间开始并延续至人类酶的大致氨基酸445的任何C-末端非催化结构域,或者另一个GnTI的相应同源序列或其催化活性的变异体或突变体。催化结构域可以包括酶的N-末端部分例如非催化结构域、跨膜结构域或胞质结构域的全部或部分。

[0096] 如本文所使用的“胞质的”是用来指与细胞的细胞质相互作用的蛋白质的部分。

[0097] N-乙酰葡萄糖胺基转移酶II(GlcNAc-TII;GnTII;EC 2.4.1.143)催化反应 $\text{UDP-N-乙酰基-D-葡萄糖胺} + 6-(\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R} \rightleftharpoons \text{UDP} + 6-(2-(\text{N-乙酰基-}\beta\text{-D-葡萄糖胺基})-\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R}$,其中R代表聚糖中N-连接的寡糖的其余部分。N-乙酰葡萄糖胺基转移酶II催化结构域是能够催化该反应的N-乙酰葡萄糖胺基转移酶II酶的任何部分。SEQ ID NO:20-33中列出了来自各种生物体的N-乙酰葡萄糖胺基转移酶II酶的氨基酸序列。在某些实施方式中,N-乙酰葡萄糖胺基转移酶II催化结构域是来自人类N-乙酰葡萄糖胺基转移酶II酶(SEQ ID NO:20)。另外的GnTII种类在CAZy数据库中糖基转移酶家族16(cazy.org/GT16_all)中列出。酶学特征性种类包括,秀丽隐杆线虫、黑腹果蝇、智人、褐家鼠(Rattus norvegicus)、野猪(Sus scrofa)(cazy.org/GT16_characterized)的GnTII。在某些实施方式中,N-乙酰葡萄糖胺基转移酶II催化结构域包含与SEQ ID NO:21的氨基酸残基大致30至大致447至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。催化结构域可以包括酶的N-末端部分例如非催化结构域、跨膜结构域或胞质结构域的全部或部分。

[0098] 在某些实施方式中,N-乙酰葡萄糖胺基转移酶I催化结构域处于N-乙酰葡萄糖胺基转移酶II催化结构域的N-末端。在其它的实施方式中,N-乙酰葡萄糖胺基转移酶II催化结构域处于N-乙酰葡萄糖胺基转移酶I催化结构域的N-末端。术语“N-末端”是指与一组参照的氨基酸残基相比更接近以具有游离氨基(-NH₂)的氨基酸结尾的多肽末端的一组氨基酸残基的定位。

[0099] 间隔区

[0100] 在本发明的某些实施方式中,该重组蛋白质包含在N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域之间间隔区。术语“间隔区”是指分隔N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的任何序列的任何数目的连续氨基酸,使得该间隔区对催化结构域的酶学功能没有影响。通常,该间隔区为至少5、至少10、至少15、至少20、至少30、至少40或至少50个氨基酸长度。在某些实施方式中,间隔区包含来自非催化结构域的序列。“非催化结构域”是指其位置邻近天然酶如糖基转移酶或糖基水解酶的跨膜结构域和任选地将酶靶向于或协助酶保留在ER/高尔基体中的蛋白质结构域或其片段。非催化结构域一般从疏水性跨膜结构域之后的第一个氨基酸开始,且结束于催化结构域。示例性的非催化结构域包括但不限于,人GnTI的非催化结构域,人GnTII的从约30至约83或约30至约105的氨基酸残基,或里氏木霉KRE2的从约26至约106或从约26至约83的氨基酸残基。在其中间隔区包含来自非催化结构域的序列的某些实施方式中,间隔区包括人GnTI序列的氨基酸30-83(SEQ ID NO:34)。在其它实施方式中,间隔区可以包括SEQ ID NO:35-38中所列出的任何序列。

[0101] 合适的间隔区的另外的例子包括但不限于,柔性间隔区3XG4S(SEQ ID NO:118)、柔性间隔区2XG4S(SEQ ID NO:120)、里氏木霉CBHI的间隔区(SEQ ID NO:122)以及里氏木霉EGIV纤维素酶的间隔区(SEQ ID NO:124)。

[0102] 在某些实施方式中,间隔区的长度与GnTI非催化结构域的长度大致相同。在某些实施方式中,该长度为约74个氨基酸残基,加或减约37个氨基酸。例如,间隔区的长度是约30个氨基酸至约110个氨基酸,或从约35个氨基酸至约100个氨基酸,或如本文所描述的实施例中所示例的,加上或减去2、3、4或5个氨基酸。在一个实施方式中,间隔区长度与截短的GnTI非催化结构域相对应,例如,从氨基酸25至氨基酸104,或在氨基酸30到氨基酸101之间开始,到GnTI非催化结构域的末端。在某些实施方式中,间隔区可以包括人GnTI非催化结构域的部分,其可以从位于氨基酸70至氨基酸87(根据SEQ ID NO:34中的编号)之间或氨基酸76和氨基酸104之间的氨基酸开始,或者从氨基酸30、35、40、45、50、60、70、73、74、75、76、80、83、84、85、86、87、100、101、102、103或104开始,至人GnTI非催化结构域的末端。在其它实施方式中,间隔区可以包括异源的间隔肽,其可以包括真菌的间隔肽和/或重复的寡聚间隔肽。

[0103] 通常,间隔区是没有特定构象的细长的肽,并包含允许高灵活性(例如,甘氨酸和丙氨酸)、亲水性(例如,丝氨酸和苏氨酸)的氨基酸残基及任选的阻止构象的脯氨酸。间隔区可以被糖基化。在某些实施方式中,间隔区是被O-糖基化,包括真菌O-甘露糖基化。在某些实施方式中,间隔区是内源性真菌、丝状真菌或木霉属的间隔肽,如天然地分隔蛋白结构域的间隔区。间隔区可以衍生自真菌,如丝状真菌(例如,里氏木霉)的分泌的或纤维素水解酶、其片段,或间隔区的多聚体和/或它的片段或其突变的同源物或等同物。天然的真菌间隔区可以含有二聚的或低聚的脯氨酸和/或甘氨酸和/或丝氨酸和/或苏氨酸,和/或选自于丝氨酸、苏氨酸、甘氨酸、脯氨酸或丙氨酸的多个氨基酸残基或它们的任意组合。在某些实施方式中,间隔区是重复的寡聚物,其含有具有1-10或1-5个氨基酸残基(选自于丝氨酸、苏氨酸、甘氨酸、脯氨酸或丙氨酸且任选地为选自于带负电荷的残基谷氨酸或天冬氨酸或带正电荷的残基赖氨酸或精氨酸的带电氨基酸残基)的单体。在某些实施方式中,带电的残基是带负电荷的。在某些实施方式中,该单体包含二聚或寡聚的氨基酸残基,和/或选自于丝氨酸、苏氨酸、甘氨酸、脯氨酸和丙氨酸的多个单一氨基酸残基。在某些实施方式中,寡聚物包含甘氨酸的二聚物或寡聚物的单体以及选自于丝氨酸、苏氨酸、甘氨酸、脯氨酸或丙氨酸的单一残基。在某些实施方式中该单一残基是丝氨酸或苏氨酸。在某些实施方式中,该残基是丝氨酸。在某些实施方式中,重复间隔区的序列是 $\{(Yyy)_nXxx\}_m$,其中n是2-10,m是2-10,且Xxx和Yyy选自于丝氨酸、苏氨酸、甘氨酸、脯氨酸和丙氨酸,条件是Xxx和Yyy不是相同的氨基酸残基。在某些实施方式中,重复间隔区是 $\{(Gly)_nXxx\}_m$,其中n是2-10,m是2-10,和Xxx选自于丝氨酸、苏氨酸、甘氨酸、脯氨酸和丙氨酸。在某些实施方式中,Xxx是丝氨酸或苏氨酸。在某些实施方式中,Xxx是丝氨酸。

[0104] 靶向肽

[0105] 在某些实施方式中,本发明的重组蛋白质包括连接到催化结构域的靶向肽。本文所用的术语“连接的”意思是对于多肽而言的氨基酸残基的两个聚合物或对于多核苷酸而言的核苷酸的两个聚合物或者彼此直接相邻地偶联,或者处于同一多肽或多核苷酸中但被介于其间的氨基酸残基或核苷酸分隔。本文所用的“靶向肽”是指重组蛋白质的任何数目

的连续氨基酸残基,其能够使重组蛋白质定位在宿主细胞内的内质网(ER)或高尔基器(Golgi)中。靶向肽可以是催化结构域的N-末端或C-末端。在某些实施方式中,该靶向肽是催化结构域的N-末端。在某些实施方式中,靶向肽提供与ER或高尔基体成分例如与甘露糖苷酶 II 酶的结合。在其它实施方式中,靶向肽提供与内质网或高尔基体膜的直接结合。

[0106] 靶向肽的部件可以来自通常驻留在ER或高尔基器中的任何酶。这样的酶包括甘露糖苷酶、甘露糖基转移酶、糖基转移酶、2型高尔基体蛋白质及MNN2、MNN4、MNN6、MNN9、MNN10、MNS1、KRE2、VAN1和OCH1酶。这样的酶可以来自酵母或真菌物种,例如枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、*Chrysosporium lucknowense*、线黑粉酵母属、镰孢菌属、赤霉属、腐质霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属、弯颈霉属和木霉属。这类酶的序列可以在GenBank序列数据库中找到。

[0107] 在某些实施方式中,靶向肽来自与重组蛋白质的催化结构域之一相同的酶和生物体。例如,如果该重组蛋白质包括人GnTII催化结构域,重组蛋白质的靶向肽是来自人GnTII酶。在其它实施方式中,靶向肽可以来自与重组蛋白质的催化结构域不同的酶和/或生物体。

[0108] 可以在靶向本发明的重组蛋白质中使用的用于将蛋白质靶向于内质网或高尔基体的各种靶向肽的实例包括:与半乳糖基转移酶融合的Kre2/Mnt1 N-末端肽(Schwientek, JBC 1996,3398)、用于甘露糖苷酶定位至酵母细胞的ER以产生Man5的UDEL(Chiba, JBC 1998,26298-304;Callewaert,FEBS Lett 2001,173-178)、与GnTI催化结构域融合的OCH1靶向肽(Yoshida等,Glycobiology 1999,53-8)、与 α 2-甘露糖苷酶融合的Mns1的酵母N-末端肽(Martinet等,Biotech Lett 1998,1171)、与GnTI或 β 4GalT的催化结构域连接的Kre2的N-末端部分(Vervecken,Appl. Environ Microb 2004,2639-46),各种方法综述于Wildt和Gerngross(Nature Rev Biotech 2005,119)、构巢曲霉中的全长GnTI(Kalsner等,Glycocon.J 1995,360-370)、米曲霉中的全长GnTI(Kasajima等,Biosci Biotech Biochem 2006,2662-8)、曲霉属中与秀丽隐杆线虫GnTI融合的酵母Sec12定位结构的部分(Kainz等2008)、曲霉属中与人GnTI融合的酵母Mnn9的N末端部分(Kainz等2008)、与人GnTI融合的曲霉属Mnn10的N末端部分(Kainz等,Appl. Environ Microb 2008,1076-86)和里氏木霉中的全长人GnTI(Maras等,FEBS Lett 1999,365-70)。

[0109] 在某些实施方式中,靶向肽是具有SEQ ID NO:115或SEQ ID NO:116的氨基酸序列的Kre2/Mnt1(即,Kre2)靶向肽。

[0110] 可用于靶向肽的序列的其它例子包括在下面的表1中列出的序列。

[0111] 表1:靶向肽。下划线是推定的跨膜结构域。在KRE2中,下划线和双下划线是能够实现高尔基体定位的非催化结构域。Other01和Other02是推定的甘露糖基化相关蛋白质。

[0112]

同源子	胞质的	跨膜的	内腔的
KRE2 estExt_fgenes1_ pm.C_30039	MASTNARYVR SEQ ID NO: 39	<u>YLLIAFFTILVFYE</u> <u>VSN</u> SEQ ID NO: 40	<u>SKYEGVDLNKGFTAPDSTKTPKPPATGDAKDFPLALTPNDP</u> <u>GFNDLVGIAPGPRMNATEVTLARNSDVWDIARSIRQ</u> VEDRFNRRYNY DWVFLNDKPPFDNTFKKVTTSLVSGKTHYGEIAPEHWSFPDWIDQDKA KKVREDMAERKIYGDSSVSYRHMCRFESGFFFRQPLMMNYEYYWRV EPSIELYCDIHYDPFRLMVEQGKKYSFVISLYEYPATIALTWEKTKFM KNHPEHIAPDNSMRFLSDDGGETYNNCHFWSNFEIGSLEWLRSKQYI DFFESLDKDGFFFYERWGDAPVHSIAAGLMLNRSEIHFFNDIAYWHV PFTHCPTGEKTRLDLKHCDPKENFDWKGYSCSRFFEMNGMDKPE GWENQQD SEQ ID NO: 41
KRE2 可选 1 e_gw1.28.231.1	MAIARPVR SEQ ID NO: 42	<u>ALGGLAAILWCF</u> <u>FLY</u> SEQ ID NO: 43	QLLRPSSSYNSPGDRYINFERDPNLDPTGEPEGILVRTSDRYAPDAK DTRASATLLALVRNEEVDDMVASMDLERTWNSKFNPWTFNDK PFSEEFKKKTSAVTNATCNELIPKEHWDAPSWIDPAIFEESAAYLKK NGVQYANMMSYHOMCRWNSGMFYKHPALKDVRYYWRVEPKVHFF CDVDYDVFYRMQDNNKTYGFTINLYDDPHTLPTLWPQTAKFLADHPN YLHEHSAIKWVIDDARRPQHNREAGGFSTCHFWSNFEVADMEFWRS KVYEDYFEHLDRAGGFFYERWGDAPVHSIALGLFEDSSKIHWRDIG YQHIPPFFNCPNSPKCKGCVTGRLTDGEPFLHREDCRPNWFKYAGMG SEQ ID NO: 44
OCH1 e_gw1.16.371.1	MLNPRRR SEQ ID NO: 45	<u>ALIAAAFILTVFF</u> <u>LI</u> SEQ ID NO: 46	SRSHNSEASTSEPKDAEAEALSAANAQRAAPPPPPQKPMIDMSG MSTYDKLAYAYEYDIESKFPAYIWQTRKTPSEGDFEFREQEASWSI EHPGFIHEVITDSVADTLLQLLYGSIPEVLEAYHALPLPVLKADLFYRIL YARGGYISDIDTYAIRSALEWIPPQIPKETVGLVIGIEADPRPDWADW YSRRIQFCQWTIQSKPGHPVLRDIISRTNQTLMMKSGKLSAFQGNR VVDLTGPAVWTDTIMDYFNDERYFDMENSKGRIDYRNFTGMETSKRV GDVVVLPITSFSPGVGQMGAQDYDDPMAFVKHDFEGTWKPESEHRI GEIVQELGEGQGEAPKEQ SEQ ID NO: 47
OCH1 可选 1 fgenes1_pm.C_s caffold_13000080	MGMGQCQWSFF RNKVPTQMRRC SEQ ID NO: 48	<u>LPLYITYVCVFLV</u> <u>IV</u> SEQ ID NO: 49	NFDWILAI PNPA SVLRREP KAPPLPGSTFPQKIWQTWKVDPLNFERD LVTARTWTTINPGMRYEYVTDANEMAYIEDRYGPNGDRPDIVEFYK MINLPIIKADLLRYMIMYAEAGGIYADIDVETMKPFHFRIPDRYDEKIDIII GVEIDQPDFKDHPILGKKSMSFCQWTFVARPQQPVMMRLIENIMKWF KTVARDQGVPLGEVQLDFDQVISGTGPSAFTKAMLEEMNRKTGPKV TWDAFHNLDKSLVGGVLVLTVEAFCAQGGHSDSGNHNARNALVKH HFHASNWPSRHPRYKHPAYGQVEDCNWVPECVRKWDEDTSNWDK YSENEQKKILODIENARLERERQQCALALP SEQ ID NO: 50

[0113]

MNN9 e_gw1.5.262.1	MARPMGSVRLKK ANPST SEQ ID NO: 51	<u>LILGAVLCIFIIIFL</u> <u>V</u> SEQ ID NO: 52	SPSSPASASRLSIVSAQHHLSPPTSPYQSPRSGAVQGPPPVTRYN LNKVTVTSDPVRNQEHLILTPMARFYQEYWDNLLRLNYPHEL ITLGFILPKTKEGNQATSMQLQKQIKTQNYGPEKDRFKSHILR QDFDPAVVSQDESERHKLANKARREVMKARNLLFTTLGP STSWVLWLDADITETAPTLIQDLASHDKPIIVANCFQKYYPDES KKMAERPYPDFNSWQDSETALKMAEQMGPDILLEGYAEAMAT YRTLLAYMSTPGGSKDLVVPLDGVGGTALLVKADVIRDGAM FPFFAFYHILIESEGFAMAKRLGWQPYGLPNYKVYHYNE SEQ ID NO: 53
MNN9 可选 1 estExt_GeneWise Plus.C_230146	MLLPKGGLDWRS ARAQIPTRAL WNAVTRTR SEQ ID NO: 54	<u>FILLVGITGLILL</u> <u>W</u> SEQ ID NO: 55	RGVSTSASEMQSFCWGPAPPMEMSPNEHNRWNGHLQTPVI FNHHAPVEVNSSTIEHVDLNPINSTKQAVTKEERILITPLKDA APYLSKYFELLAELTYPHRLIDLAFLVSDSTDDTLAVLASELDR IQKRPDQIPFHSATVIEKDFGFKLSQNVEERHSFEAQGPRRKA MGRARNYLLYTALKPEHSWVYWRDVIDVDSPTGILEDFIAHD RDILVPNIWFHRYRDGVIEGRFDYNSWVESDKGRKLANSLDK DVLVAEGYKQYDTGRTYMAKMGDWRENKDVELELDGIGGV NILVKADVHRSGINFPCYAFENQAETEGFAKMAKRAGYEVVG LPNYVVWHIDTEEKGGNA SEQ ID NO: 56
MNN9 可选 2 estExt_GeneWise Plus.C_400029	MMPRHSSGFNS GYPRADTFEI SPHRFQPRATLPP HRKRKRRTAIR SEQ ID NO: 57	<u>VGIADVVLVLVL</u> <u>WFG</u> SEQ ID NO: 58	QPRSVASLISLGLSGYDDLKLETVRYYDLNSVQGTARGWERE ERILLCVPLRDAEQHLPFFSHLKNFTYPHNLIDLAFLVSDSKD HTLESLTEHLEAIQADDPKQPYGEISIEKDFGQKVNQDVESR HGFAAQASRRKLMAQARNWLLSAALRPYIISWVYWRDVIDVET APFTILEDLMRHINKDVIVPNVWRPLPDWLGGEQPYDLNSWQE SETALALADTLDEDAVIVEGYAEYATWRPHLAYLRDPYGDPD MEMEIDGVGGVSILAKAKVFRAGVHFAFSEKHAETEGFGK MAKRMHFSVVGGLPHYTIWHLIEPSVDDIKHMEEMERERIARE KEEERKKKEAQIKEEFGDANSQWEQDKQMQMDIKLQDRGG DKEAAAAGVNQCAAAKAAGAMEGQKN SEQ ID NO: 59
MNN10 fgenesh5_pg.C_s caffold_5000342	MSLSRSPSPVPG GGWSSPGLNINS GRSSPSNAAGSS VSWEAKMRKQ GNGYPSFSTQNG GFFTRHMRRRI SSSLPRFAAGPG NTYAEREKYG GHSPHAGGGRRLR AFLARIGRRLKW R SEQ ID NO: 60	<u>ILLPLIICTIVAY</u> <u>Y</u> SEQ ID NO: 61	GTHEAPGFVHWWRRISMGGGGEKFFVILGANVGGGVMEWK AREWAIERDSVRNKRKYATRWGYDLEIVDMKTKRYAHEWR ESWEKVDFIRAAMRKYPKAEWFWWLDLNTYVMEPSYSLQRH LFNHLDRHVYRDINVFNPLNITHPTTEEYLDAAERSPVGDGNIN SVNLMLTQDCSGFNLGSSFFIRSAWTEQLLDIWWDPVLYEQK HMEWEHKEQDALEQLYRTQPWIRQHTGFLPQRLINSFPPAAC ADESGLNNTRIHYNKEDRDFVVMAGCEWGRDCWGEYHY REFSYWLNRRNPWELFKEIVAVIWKLTGQRVKL SEQ ID NO: 62
MNN10 可选 1 estExt_GeneWise Plus.CJ 50339	MHFAYPSRKSSN PPFPRSTRLP LRRSRIKT SEQ ID NO: 63	<u>IGIVLEFLVLATLW</u> <u>EE</u> SEQ ID NO: 64	SNPRVPRPDPERVPSGRPPVVLVTVIDPTQYPNAYLKTIKENRE QYAAKHGVEAFIVKAYDYDTQCAPQSWSKLMAMRHALTKEFP ECRFVWYLDQDAYIMDMSKSLEEQLLNQKLESLMIKNYPV PPDSIIKTFSHLRPDEVDLIVSQDSSGLVAGSVVVRNSQWSKFL LETWMDPLYRSYNFQKAERHIALEHIVQWHPTILSKLALVPQR TLGPYTRTDQGDAYQDGFVVMFTGCTKSQECSCTVSASY QKWSSSL SEQ ID NO: 65

[0114]

MNS1 fgenes1_pm.C_s caffold_3000175	MIRDPFGIHSKNA FKATALRAARDIK EAATQAGANALE MSFSLPKHVPDF GDPSRALEDRAW AALLPMYKDKPY AYAPSMRLRPWW R RRK SEQ ID NO: 66	<u>VLGMIAAAVMFV</u> <u>L YVTGFF</u> SEQ ID NO: 67	SSGQTEEAKKKASGSFAFSWLGLSQERGGVDWDERRSVVEAF EVWDAYERYAWGKDEFHPISKNGRNMAPKGLGWIIHDSLDTM MLMNQTTRLQHAREWISTSLTWDQDQDVNTFETTIRMLGGL LSAHYLSTFEPELAPLTEDDEGAPGEDLYLEKAKDLADRLLSA FESESGIPYASVNIGEYKGPSHSDNGASSTAEATTLQLEFKYLA KL TGEKNFWDKVEKVMFVDDNQPEDGLVPIYIYATTGEFRG QNIRLGSRGDSYIEYLIKQYLQTNKQEPYIEMWDEALAGVR KHLVYTEPSEFTIAERPDLGLEHPMSPKMDHLVCFMPGTIAL AATGGLTEAEARKLSTWNKKKDDMQLARELMHTCWGMYK YMKTLGAPEIMYFNIPNPPPESSAPHQAPAAFDDEPHAERKRD FVVHSDNVHNLQRPETVESLFYMWRTGDVKYREWGWD MFK SFVNYTAVEDQGGFTSLLDANSIPPTPKDNMESFWLAETLKYM YLLFSPNDVLP LHKIVLNTEAHFPFRFDMGPLFSTGWKRKPRD GSAKKKATTAATTTDAE SEQ ID NO: 68
MNS1 可选 1 estExt_fgenes1_ pm.C_80182	MARRRYR SEQ ID NO: 69	<u>LFMICAAYILFLL</u> <u>YR</u> SEQ ID NO: 70	VSQNTWDDSAIIYATLRHPPASNPPAAGGESPLKPAKPEHEH EHENGYAPESKPKPQSEPKPEKPAPEHAAGGQKSQGKPSYED DEETGKNPPKSAVIPSOTRLPPDNKVHWRPVKEHFPVPSSEVIS LPTGKPLKVPVRVQHEFGVESPEAKSRRVARQERVGKEIERAW SGYKFFAWMHDELSPVSAKHDPFCGWAATLVDSLDTLWIA GLKEQFDEAARAVEQIDFTTTPRNIPVFETTIRYLGGLGAFD VSGGHDGGYPMLLTAKAVELAEILMGIFDTPNRMPIYYQWQP EYASQPHRAGSVGIAELGTLMEFTRLAQLTSQYKYDAYDRI TDALIELQKQGT SIPGLFPENLDASGCNHTATLRSSLEAAQK QMDEDLSENKPNYRPGKNSKADPQTVEKQPAKKQNEPVEKA KQVPTQQTAKRGKPPFGANGFTANWDCVPQGLVVGYYGFQQ YHMGGGQDSAYEYFPKEYLLLGLESKYQKLYDVAEAIN WLLYRPMTDGDWDILFPAKVSTAGNPSQDLVATFEVTHILTCFI GGMYGLGGKIFGREKDLTAKRLTDGCVWAYQSTVSGIMPE GSQVLACPTLEKCDFNETLWWEKLDPAKDWDRDKQVADDDKDK ATVGEALKETANSHDAAGGSKAVHKRAAVPLPKPGADDDVGS ELPQSLKDKIGFKNGEQKKPTGSSVGIQRDPDAPVDSVLEAHR LPPQEPEEQQVILPDKPQTHEEFVKQRIAEIMGFAPGVVHIQSR QYILRPEAIESVWYMYRITGDPWMEKGWKMFEATIRATRTEI ANSAIDVNSEEPGLKDEMESFWLAETLKYYYLLFSEPSVISLD EWVLNTEAHFPKRPGGSVIGHSI SEQ ID NO: 71
MNS1 可选 2 estExt_GeneWise Plus.C_120298	MLNQLQGRVPRR Y SEQ ID NO: 72	<u>IALVAFAFFVAFL</u> <u>LW</u> SEQ ID NO: 73	SGYDFVPRTATVGRFKYVPSSYDWSKAKVYYPVKDMKTL PQG TPVTFPRLQLRNQSEAQDDTTKARKQAVKDAFVKSWEAYKTY AWTKDQLQPLSLSGKETFSGWSAQLVDALDTLWIMDLKDDFF LAVKEVAVIDWSKTKDNKVINLFEVTIRYLGGLIAAYDLSQEP VLRAKAIELGDTLYATFDTPNRLPSHWLDYSKAKKGTQRADD SMGGAAGGTLCEMFTRLSQITGDPKYDATERIKQFFYRFQNE TTLPGMWVPMMNYREETMVESRYSMGGSADSLYEYLVKMPA LLGGLDPQYPEMAIRALDTARDNLLFRPMTTEKGDNILALGNA LVDHGNVQRITEMQHLCFAGGMYAMAGKLFKRDDYVDLGS RISSGCVWAYDSFSPGIMPESADMAACAKLDGCPYDEVKAPV DPDGRPHGFIHVKSRYLLRPEAIESVFFYMWRTGDQVWRD TAWRMWENIVREAETEHAFIIVEDVTRTASKLTNNYLLQTFW LAETLKIFYLIFDDESAIDLKWFVNTTEAHFPKRPV SEQ ID NO: 74
MNS1 可选 3 estExt_GeneWise	MLVVGRPRLVRN S	<u>ILTLAILSIWHLG</u> <u>LL</u>	SRTPTSASALVSASVSASSEWSRLERLMNRGAPLTPYPDSNSSF DWSAIFRYPPHNTTHLPPRHKQPPLPRIQHRFGPESPAAAKE RIKRLKAVKQVFLRAWQAYKGYAWKQDALLPISGGGREQFS

[0115]

Plus.CJ 60228	SEQ ID NO: 75	SEQ ID NO: 76	GWAATLVDA LDTLWIMGLREEFDEAVA AAEIDFGSSTSSRVN IFETNRYLGGLLAAYDLSGREVLLKKAVELGDLIYAGFNTEN GMPVDFLNFYSAKS GEGLVVESSVVSASPGTLSLELAHLSQVT GDDKYYS AVSQVMDFYQCGQNKTRLPGVWPIDVNMRAKDVV SGSRFTLGGCADSLYEYLPKMHQLLGGGEPKYETMSRTFLQA ADRHVFRPMLPGA EEDVLMPGNVNVEDDSGEAVLDPETEHL ACFVGGMFGLAGRLFSRPDDVETGVRLTNGCVYAYRAFPTGM MPERLDLAPCRDRSSRCPWDEEHWLEERAKRPEWEPHLPRGF TSAKDPYLLRPEAIESVFYSYRITGRQEFQTAAWDMFTAVEK GTRTQFANA AVLDTRAADELPQEDYMESFWLAETLKIFYLML FTTPDIISLDDYVLNTEAHPFKLVG SEQ ID NO: 77
MNS1 可选 4 e_gw1.13.279.1		<u>MVMLVAIALAWL</u> <u>GCSLL</u> SEQ ID NO: 78	RPVDAMRADYLAQLRQETVDMFYHGYSNYMEHAFPEDEL RPI SCTPLTRDRDNPGRISLNDALGNYSLTLDLSLT LAILAGGPQN GPYTGPQALSDFDQGVAEFRHYYGDRSGPAGIRARGFDLDS KVQVFETVIRGVGGLLSAHLFAIGELPITGYVPRPEGVAGDDP LELAPIWPNGFRYDQQLRLALDL SERLLPAFYTPGTIPYPRV NLRSGIPFYVNSPLHQN LGAEVEEQSGRPEITETCSAGAGSLVL EFTVLSRLTG DARFEQA AKRAFWEVWHRRSEIGLIGNGIDAER GLWIGPHAGIGAGMDSFFEYALKSHILLSGLGMPNASTSRRQS TTSWLDPNSLHPLPPEMIHSDAFLQAWHQAHASVKRYLYTD RSHFPYYSNNHRATGQPYAMWIDSLGAFYPGLLALAGEVEEAI EANLVYTALWTRY SALPERWSYREGNVEAGIGWWPGRPFES TYHIYRATRDPWYLVHVGEMVLRDIRRRCYAE CGWAGLQDVQ TGEKQDRMESFFLGETAKYMYLLFDPDHLNKLDAAYVFTE GHPLIIPKSKRGSGSHNRQDRARKAKKSRDVAVYTYDESFTN SCPAPRPPSEHHLIGSATAARPDLFSVSRTDLYRTPNVIIGPLE KVEMRDKKKGRVVRYRATSNHTIFPWTLP PAMPLPENGTCAP PERIISLIEFPANDITSGITSRFGNHL SWQTHLGPTVNILEGLR QLEQVSDPATGEDKWRITHIGNTQLGRHETVFFHAEHVRHLK DEVFSCRRRRDAVEIELLVDPKPSDTNNNTLASSDDVVVDAK AEEQDGM LADDDGDTLNAETLSSNSLFQSLLR AVSSVFEPVYT AIPESDPSAGTAKVYSFDAYTSTGPGAYPMPSISDTPIPGNPFYN FRNPASNFPWSTVFLAGQACEGPLPASAPREHQVIVMLRGCCS FSRKL DNIPSFSPHDRA LQLVVVLDEPPPPPPPPANDRRDVTR PLLDTEQTPKGMKRLHGIPMVLVRAARGDYELFGHAIGVGM RRKYRVESQGLVVENAVVL SEQ ID NO: 79
VAN1 estExt_GeneWise Plus.C_400029	MMPRIHSSGFSN GYPRADTFEISPH RFQPRATLPPHR K RKRTAIR SEQ ID NO: 80	<u>VGIAVVVLVLVL</u> <u>WFG</u> SEQ ID NO: 81	QPRSVASLISL GILSGYDDLKLET VRYDYLSNVQGTARGWERE ERILLCVPLRDAEQHLPMFFSHLKNFTYPHNLIDLAF LVSDSKD HTLESLTEHLEAIQADPDPKQPYGEISIEKDFGQKVNQDVESR HGFAAQASRRKLMAQARNWLLSAALRPYHSWVYWRD VDET APFTILEDLMRHNKDVIVPNVWRPLPDWLGGEQPYDLNSWQE SETALALADTLDEDAVIVEGYAEYATWRPHLAYLRDPYGD PD MEMEIDGVGGVSILAKAKVFRAGVHFP AFSEKHAETEGFGK MAKRMHFSVVG LPHYTIWHL YEPSVDDIKHMEEMERERIARE KEEEERKKKEAQIKEEFGDANSQWEQDKQMQDLKLQDRGG DKEAAAAGVNQGA AAKAAGAMEGQKN SEQ ID NO: 82
VAN1 可选 1 estExt_GeneWise Plus.C_230146	MLLPKGGLDWRS ARAQIPPTR ALWNAVTRTR SEQ ID NO: 83	<u>FILLVGITGLILL</u> <u>W</u> SEQ ID NO: 84	GVSTSASEMQSFYCWGPAKPPMEMSPNEHNRWNGHLQTPVIF NHHAPVEVNSSTIEHVDLNPINSTKQAVTKEERILITPLKDAA PYLSKYFELLAELTYPHRLIDLAF LVSDSTD TLA VLASELDRI QKRDPQIPFHSATVIEKDFGFKLSQNVEERHSFEAQGPRRKAM GRARNYLLYTALKPEHSWVYWRDVIDVSPTGILEDFIAHDRD

[0116]

			ILVPMNIWFHRYRDGVDIEGRFDYNSWVESDKGRKLANSLDKDV VLAEGYKQYDTGRITYMAKMGDWRENKDVELELDGIGGVNIL VKADVHRSGINFPCYAFENQAETEGFAKMAK RAGYEVY G LP NYVVW H 1 DTE EKG G NA SEQ ID NO: 85
VAN1 可选 2 e_gw1.5.262.1	MARPMGSRVRLKK ANPST SEQ ID NO: 86	<u>LILGAVLCIFHILF</u> Y SEQ ID NO: 87	SPSSPASASRLSIVSAQHHLSPPTSPYQSPRSGAVQGPPPVTRYN LNKVTVTSDPVRNQEHILITPMARFYQEYWDNLLRLNYPHEL ITLGFILPKTKEGNQATSM LQKQIKTQNYGPEKDRFKSHILR QDFDPAVVSQDESERHKLANKARREYMAKARNSLFLTTLGP STSWVLWLDADITETAPT LIQDLASHDKPIIVANCFQKYDPES KKMAERPYPDFNSWQDSETALKMAEQMGPDILLEGYAEMAT YRTLLAYMSTPGGSKDLVVPLDGVGGTALLVKADVHRDGAM FPPFAFYHLIESEGFMAKRLGWQPYGLPNYKVYHYNE SEQ ID NO: 88
Other01 estExt_GeneWise Plus.CJ 50339	MHFAYPSRKSSN PPFRPRSTRLP G LRRSRIKT SEQ ID NO: 89	<u>IGIVLFLVLATLW</u> FF SEQ ID NO: 90	SNPRVPRPDPERVPSPGRPPVVLVTVIDPTQYPNAYLKTIKENRE QYAAKHGYEAFIVKAYDYDTQGAPQSWSKLMAMRHALT KFP ECRFVWYLDQDAYIMDMSKSLEEQLLNQKLESLMIKNYPVV PPDSIIKTFSHLRPDEVDLIVSQDSSGLVAGSVVVRNSQWSKFL LETWMDPLYRSYNFQKAERHALEHIVQWHPTILSKLALVPQR TLGPYTRTDQGDAYQDGDVVMFTGCTKSGEQSCETVSASY QKWSSSL SEQ ID NO: 91
Other02 fgenes h5_pg.C_s caffold_5000342	MSLSRSPSPVPG GGWSSPGLNINS GRSSPSNAAGSS VSWESAKMRKQ GNGYPSFSTQNQ GFFTRHMRRISSS LPRAAGPGNTY AEREKYERGGHS PHAGGCRRLRAFL ARIGRRLKWR SEQ ID NO: 92	<u>ILLPLIHCTIVAY</u> YG SEQ ID NO: 93	THEAPGFVHWWRRISMGGGGGEKFVILGANVGGGVMEWKGA REWAIERDSVRNKRKYATRWGYDLEIVDMKTKKRYAHEWRE SWEKVD FIRAAMRKYPKAEFWWLDLNTYVMEPSYSLQRHL FNILDRHVYRDINVFNLNTHIPPTEEYLD AEARSVPGDGNINS VNLMLTQDCSGFNLGFFIRRS AWTEQLLDIWWDPVLYEQKH MEWEHKEQDALEQLYRTQPWIRQHTGFLPQRLINSFPFAACA DESGLNNTRIHYNEKDRDFVNNMAGCEWGRDCWGE MYHYR EFSYWLNRNPWELFKEEIVAVIWKLTGQRVKL SEQ ID NO: 94

[0117] 可以通过在宿主细胞中按糖基化途径表达蛋白质(其中蛋白质之一包含作为唯一的靶向肽的非特征性序列)和测量基于聚糖生物合成的胞质定位产生的聚糖(例如,如 Schwientek JBC 1996 3398中),或通过表达与靶向肽融合的荧光报道蛋白质和利用免疫荧光或分级高尔基体的胞质膜和测量蛋白质的定位来分析蛋白质在高尔基体中的定位而测试非特征性序列用作靶向肽。

[0118] 靶向肽可以包括非催化结构域。在某些实施方式中,非催化结构域来自N-乙酰葡萄糖胺基转移酶I酶或N-乙酰葡萄糖胺基转移酶II酶。特别是在某些实施方式中,非催化结构域来自人N-乙酰葡萄糖胺基转移酶I酶或人N-乙酰葡萄糖胺基转移酶II酶。与来自人N-乙酰葡萄糖胺基转移酶I酶的非催化结构域对应的序列是SEQ ID NO:34。与来自人N-乙酰葡萄糖胺基转移酶II酶的非催化结构域对应的序列是SEQ ID NO:20的残基30-85。

[0119] 靶向肽可以包括跨膜结构域。“跨膜结构域”是指以三维结构在膜中热力学稳定的氨基酸残基的任何序列。在其中靶向肽还包括非催化结构域的实施方式中,跨膜结构域处于非催化结构域的N-末端。在某些实施方式中,跨膜结构域来自N-乙酰葡萄糖胺基转移酶I酶或N-乙酰葡萄糖胺基转移酶II酶。特别地在某些实施方式中,跨膜结构域来自人N-乙酰葡萄糖胺基转移酶I酶或人N-乙酰葡萄糖胺基转移酶II酶。与来自人N-乙酰葡萄糖胺基转移酶I酶的跨

膜结构域对应的序列是SEQ ID NO:1的残基7-29。与来自人N-乙酰葡萄糖胺基转移酶 II 酶的跨膜结构域对应的序列是SEQ ID NO:20的残基10-29。

[0120] 靶向肽可以包括胞质结构域。术语“胞质结构域”指在细胞质环境中作为三维结构体热力学稳定的氨基酸序列。在其中靶向肽还包括非催化结构域的实施方式中,胞质结构域处于非催化结构域的N-末端。在其中靶向肽还包括跨膜结构域的实施方式中,胞质结构域处于跨膜结构域的N-末端。在某些实施方式中,胞质结构域来自N-乙酰葡萄糖胺基转移酶I酶或N-乙酰葡萄糖胺基转移酶 II 酶。特别地在某些实施方式中,胞质结构域来自人N-乙酰葡萄糖胺基转移酶I酶或人N-乙酰葡萄糖胺基转移酶 II 酶。与来自人N-乙酰葡萄糖胺基转移酶I酶的胞质结构域对应的序列是SEQ ID NO:1的残基1-6。与来自人N-乙酰葡萄糖胺基转移酶 II 酶的胞质结构域对应的序列是SEQ ID NO:20的残基1-9。

[0121] 在某些实施方式中,该重组蛋白质包含人GnTI催化结构域N末端的人GnTII催化结构域,在催化结构域之间具有包含人GnTI非催化结构域序列的间隔区序列。在本实施方式中,该重组蛋白质还包括GnTII催化结构域N-末端的靶向肽,具有来自人GnTII的胞质结构域、跨膜结构域和非催化结构域。在本实施方式中的重组蛋白质的序列与SEQ ID NO:95至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同,且编码此实施方式的重组蛋白质的可能cDNA的序列是SEQ ID NO:96。

[0122] 在其它实施方式中,该重组蛋白质包含具有间隔区序列的人GnTI催化结构域N-末端的人GnTII催化结构域。间隔区序列可以包括但不限于,与SEQ ID NO:118、120、122或124至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。在本实施方式中,重组蛋白质还包括GnTII催化结构域N末端的靶向肽,具有来自人GnTII的胞质结构域、跨膜结构域和非催化结构域。因此,在某些实施方式中,该重组蛋白质的序列与选自于SEQ ID NO:119、121、123和125的序列至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同。在某些实施方式中,编码SEQ ID NO:119的重组蛋白质的可能的cDNA序列是SEQ ID NO:141。在其它实施方式中,编码SEQ ID NO:121的重组蛋白质的可能的cDNA序列是SEQ ID NO:139。在再其它的实施方式中,编码SEQ ID NO:123的重组蛋白质的可能的cDNA序列是SEQ ID NO:143。在进一步的实施方式中,编码SEQ ID NO:125的重组蛋白质的可能的cDNA序列是SEQ ID NO:145。

[0123] 本发明重组蛋白质的制备

[0124] 本发明的另一个方面包括编码本发明的重组蛋白质的分离的多核苷酸。如本文所使用的术语“多核苷酸”、“核酸序列”、“核酸的序列”及其变型应当是多聚脱氧核糖核苷酸(含有2-脱氧-D-核糖)、多聚核糖核苷酸(含有D-核糖)、作为嘌呤或嘧啶碱基的N-糖苷的任何其它类型的多核苷酸以及包含非核苷酸骨架的其它聚合物(条件是该聚合物在构造中包含如在DNA和RNA中发现的允许碱基配对和碱基堆积的核碱基)的通称。因此,这些术语包括已知类型的核酸序列修饰,例如,用类似物代替一个或多个天然存在的核苷酸;核苷酸间修饰,例如,如具有不带电的连接(例如,甲基膦酸酯、磷酸三酯、氨基膦酸酯、氨基甲酸酯等)、具有带负电的连接(例如,硫代膦酸酯,二硫代膦酸酯等)以及具有带正电的连接(例如,氨基烷基磷酰胺酯、氨基烷基磷酸三酯)的那些;含有侧基部分的那些,诸如,例如蛋白质(包

括核酸酶、毒素、抗体、信号肽、聚-L-赖氨酸等);带有嵌入剂(例如,吡啶,补骨脂素等)的那些;和含有螯合剂(例如金属、放射性金属、硼、氧化性金属等)的那些。如本文所用的核苷酸和多核苷酸的符号是由IUPAC-IUB生物化学命名委员会所推荐的(Biochem.9:4022,1970)。

[0125] 分离的多核苷酸的序列通过本领域普通技术人员已知的任何合适的方法制备,包括例如,直接化学合成或克隆。对于直接化学合成,核酸聚合物的形成通常涉及3'-封闭和5'-封闭的核苷酸单体顺序添加到延长的核苷酸链的末端5'-羟基基团上,其中各次添加通过延长链的末端5'-羟基基团对添加的单体的3'-位的亲核攻击实现,添加的单体通常是磷衍生物,如磷酸三酯、亚磷酰胺等等。这样的方法是本领域普通技术人员已知的,并在相关的文本和文献中描述[例如,在Matteucci等,(1980)Tetrahedron Lett 21:719-722;4,500,707、5,436,327和5,700,637号美国专利中]。此外,可以从天然来源通过使用适当的限制性内切酶分裂DNA,用凝胶电泳分离片段,以及之后通过本领域普通技术人员已知的技术如利用聚合酶链反应(PCR,例如,4,683,195号美国专利)从凝胶中回收所需的核酸序列来分离所需的序列。

[0126] 本发明的各个多核苷酸可以整合到表达载体中。“表达载体”或“载体”是指转导、转化或感染宿主细胞,从而使细胞表达该细胞天然的核酸和/或蛋白质以外的其它核酸和/或蛋白质,或以该细胞非天然的方式表达核酸和/或蛋白质的化合物和/或组合物。“表达载体”包含宿主细胞要表达的核酸(通常为RNA或DNA)的序列。任选地,所述表达载体还包含帮助核酸成功进入到宿主细胞中的材料,如病毒、脂质体、蛋白质包膜等等。设想本发明中使用的表达载体包括核酸序列可以与任何特定的或所需的操作元件一起插入其中的那些。此外,表达载体必须是可以被转移到宿主细胞中并在其中复制的表达载体。某些表达载体是质粒,特别是具有已经有记录的和包含核酸序列的转录必须的或所需的操作元件的限制性酶切位点的那些。这样的质粒以及其他的表达载体是本领域普通技术人员公知的。

[0127] 单个多核苷酸的整合可以通过已知的方法完成,包括,例如,使用限制性内切酶(如BamHI、EcoRI、HhaI、XhoI、XmaI等等)来切割表达载体中,例如质粒中的特定位点。限制性内切酶产生可以与多核苷酸退火的单链末端,其中该多核苷酸具有或经合成以具有其序列与切割的表达载体的末端互补的末端。利用适当的酶,如DNA连接酶进行退火。正如在本技术领域普通技术人员将理解的,通常用相同的限制性内切酶切割表达载体和所需的多核苷酸,从而保证表达载体的末端与多核苷酸的末端是彼此互补的。此外,可以使用DNA接头以利于核酸序列连接到表达载体中。

[0128] 也可以通过利用本领域普通技术人员已知的方法(例如,4,683,195号美国专利)组合一系列单个多核苷酸。

[0129] 例如,可以首先在独立的PCR中产生各个所需的多核苷酸。此后,设计使得PCR产物的末端含有互补序列的特异性引物。当PCR产物混合、变性和复性时,在其3'末端具有匹配的序列的链重叠并可以作为彼此的引物。通过DNA聚合酶延伸该重叠产生其中原始序列“拼接”在一起的分子。以这种方式,一系列的单个多核苷酸可以“拼接”在一起,且随后同时转导入宿主细胞中。因此,实现多个多核苷酸中每一个的表达。

[0130] 然后将单个的多核苷酸或“拼接的”多核苷酸整合到表达载体中。本发明对于多核苷酸整合到表达载体中的过程并没有限制。本领域的那些普通技术人员熟悉多核苷酸整合到表达载体中的必要步骤。典型的表达载体包含跟随一个或多个调节区的所需多核苷酸,

以及核糖体结合位点,例如3-9个核苷酸长度的位于大肠杆菌中起始密码子上游3-11个核苷酸的核苷酸序列。参阅Shine和Dalgarno(1975)Nature254(5495):34-38以及Steitz(1979)Biological Regulation and Development(ed.Goldberger,R.F),1:349-399(Plenum,New York)。

[0131] 在本文中所使用的术语“可操作地连接的”是指其中将控制序列相对于DNA序列或多核苷酸的编码序列置于适当的位置以使得控制序列指导多肽的表达的构型。

[0132] 调控区域包括,例如,包含启动子和操纵子的那些区域。启动子可操作地连接到所需多核苷酸或者编码多肽的多核苷酸部分,从而通过RNA聚合酶启动多核苷酸或编码多肽的多核苷酸部分的转录。操纵子是邻近启动子的核酸序列,其包含阻遏蛋白可以结合的蛋白质结合域。在不存在阻遏蛋白的情况下,通过启动子启动转录。当存在时,特异于操纵子的蛋白质结合域的阻遏蛋白与操纵子结合,从而抑制转录。以这种方式,根据所用的特定的调节区和相应的阻遏蛋白的存在或不存在实现对转录的控制。例子包括乳糖启动子(与乳糖接触时Lad阻遏蛋白的构象发生改变,从而阻止Lad阻遏蛋白结合到操纵子)和色氨酸启动子(当与色氨酸复合时,TrpR阻遏蛋白具有与操纵子结合的构象,在没有色氨酸时,TrpR阻遏蛋白具有不与操纵子结合的构象)。另一个例子是tac启动子(见de Boer等(1983)Proc Natl Acad Sci USA 80(1):21-25)。正如本领域普通技术人员将理解的,这些和其他的调节区可以被用在本发明中,且本发明在这一方面并不受限。

[0133] 用于与编码本发明的重组蛋白质的分离的多核苷酸连接的某些启动子的实例包括来自以下基因的启动子:gpdA、cbh1、米曲霉TAKA淀粉酶、米黑根毛霉天冬氨酸蛋白酶、黑曲霉中性 α -淀粉酶、黑曲霉酸稳定 α -淀粉酶、黑曲霉葡糖淀粉酶(glaA)、泡盛曲霉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,构巢曲霉乙酰胺酶、米曲霉乙酰胺酶、尖孢镰刀菌胰蛋白样蛋白酶、真菌内切 α -L-阿拉伯糖酶(abnA)、真菌 α -L-阿拉伯呋喃糖苷酶A(abfA)、真菌 α -L-阿拉伯呋喃糖苷酶B(abfB)、真菌木聚糖酶(xlnA)、真菌肌醇六磷酸酶、真菌ATP合成酶、真菌亚基9(oliC)、真菌磷酸丙糖异构酶(tpi)、真菌醇脱氢酶(adhA)、真菌 α -淀粉酶(amy)、真菌淀粉葡萄糖苷酶(glaA)、真菌乙酰胺酶(amdS)、真菌甘油醛-3-磷酸脱氢酶(gpd)、酵母醇脱氢酶、酵母乳糖酶、酵母3-磷酸甘油酸激酶、酵母磷酸丙糖异构酶、细菌 α -淀粉酶、细菌Spo2和SSO。在某些实施方式中,编码本发明的重组蛋白质的分离的多核苷酸可操作地连接组成型启动子。在其它实施方式中,编码本发明的重组蛋白质的分离的多核苷酸可操作地连接诱导型启动子。在某些优选的实施方式中,诱导型启动子来自cbh1基因。

[0134] 虽然任何合适的表达载体可用于整合所需的序列,容易得到的表达载体包括但不限于:质粒,如pSC101、pBR322、pBBR1MCS-3、pUR、pEX、pMR100、pCR4、pBAD24、pUC19;噬菌体,如M13噬菌体和 λ 噬菌体。当然,这样的表达载体可能只适用于特定的宿主细胞。然而,本领域的普通技术人员可以通过常规实验很容易地确定任何特定的表达载体是否适合于任何给定的宿主细胞。例如,表达载体可以被引入到宿主细胞中,然后监测其威力和载体中包含的序列的表达。此外,可以参照描述表达载体和其对任何特定宿主细胞的适用性的有关文章和文献。

[0135] 本发明的另一个方面包括含有表达载体的宿主细胞,该表达载体包含编码本发明的重组蛋白质的分离的多核苷酸。本文所用的“宿主细胞”指可以通过插入重组的DNA或RNA

被转化的活的生物细胞。这种重组DNA或RNA可以是在表达载体中。因此,如本文所述的宿主细胞可以是原核生物体(例如,真细菌界的生物体)或真核细胞。正如本领域的普通技术人员将理解的,原核细胞缺乏膜界定的核,而真核细胞具有膜界定的核。在某些实施方式中,用于制备本发明的重组蛋白质的宿主细胞是真菌细胞,如酵母或丝状真菌。在其它实施方式中,宿主细胞是哺乳动物细胞。这样的细胞可以是人的或非人类的。

[0136] 本发明的另一个方面包括本发明的重组蛋白质的制备方法。该方法包括以下步骤:向宿主细胞中引入编码重组蛋白质的分离的多核苷酸,和培养该宿主细胞使其表达重组蛋白质。该方法还可以包括从宿主细胞纯化重组蛋白质的步骤。

[0137] 本发明的重组蛋白质的制备方法可包括向宿主细胞中引入或转移含有本发明的重组多核苷酸的表达载体。这种用于向宿主细胞中转移表达载体的方法是本领域普通技术人员公知的。例如,用表达载体转化大肠杆菌的一种方法涉及氯化钙处理,其中表达载体通过钙沉淀引入。也可以按照类似的程序使用其它的盐,例如,磷酸钙。此外,电穿孔(即,应用电流以增加细胞对核酸序列的渗透性)可用于转染宿主细胞。此外,核酸序列的显微注射提供转染宿主细胞的能力。也可以采用其他方式,例如脂质复合物、脂质体和树状聚合物。本领域的那些普通技术人员可以使用这些或其他的方法用期望的序列转染宿主细胞。

[0138] 载体可以是自主复制的载体,即,作为染色体外实体存在的载体,其复制独立于染色体的复制,例如质粒、染色体外元件、微染色体或人工染色体。该载体可以包含用于确保自我复制的任何工具。或者,载体可以是当其引入宿主中时被整合到基因组中并与它已经被整合到其中的染色体一起复制的载体。此外,也可使用单一载体或质粒或一起包含待引入到宿主的基因组中的总DNA的两个或多个载体或质粒,或转座子。

[0139] 载体可以包含一种或多种选择标记,其使得容易地选择转化的宿主。选择标记是基因,其产物提供,例如,生物杀灭剂或病毒抗性、重金属抗性、营养缺陷型的原养型等等。细菌细胞的选择可以基于已通过基因,如amp、gpt、neo和hyg基因赋予的抗菌素抗性。

[0140] 酵母宿主的适当标记是,例如,ADE2、HIS3、LEU2、LYS2、MET3、TRP1和URA3。用于丝状真菌宿主的选择标记包括但不限于,amdS(乙酰胺酶)、argB(鸟氨酸氨甲酰基转移酶)、bar(草胺磷乙酰基转移酶)、hph(潮霉素磷酸转移酶)、niaD(硝酸盐还原酶)、pyrG(乳清苷5'-磷酸脱羧酶)、sC(硫酸盐腺苷转移酶)和trpC(邻氨基苯甲酸合酶)以及它们的等同物。那些用于曲霉属中的是构巢曲霉或米曲霉的amdS和pyrG基因以及吸水链霉菌(*Streptomyces hygroscopicus*)的bar基因。那些用于木霉中的是bar、pyr4和amdS。

[0141] 载体可以包含允许载体整合到宿主的基因组中或载体在细胞中独立于基因组自主复制的元件。

[0142] 对于整合入宿主基因组中,载体可以依赖于该基因的序列或用于通过同源或非同源重组使载体整合到基因组中的载体的任何其它元件。或者载体可以包含用于引导通过同源重组整合到宿主基因组中的另外的核苷酸序列。该另外的核苷酸序列使载体能够在染色体中的精确位置整合到宿主基因组中。为了增加在精确的位置整合的可能性,该整合元件可以包含足够数目的核酸,如100至10,000个碱基对,优选为400至10,000个碱基对,最优选为800至10,000个碱基对,其与相应的目标序列高度同源以提高同源重组的概率。整合元件可以是与宿主基因组中的靶序列同源的任何序列。此外,整合元件可以是非编码的或编码的核苷酸序列。另一方面,载体可以通过非同源重组被整合到宿主基因组中。

[0143] 对于自主复制,该载体可以进一步包含复制起点,其使载体能够在所述宿主中自主复制。复制起点可以是在细胞中发挥功能的介导自主复制的任何质粒复制子。在本文中定义的术语“复制起点”或“质粒复制子”是使质粒或载体能够体内复制的序列。用于酵母宿主中的复制起点的实例是2 μ 复制起点,ARS1,ARS4,ARS1和CEN3的组合,及ARS4和CEN6的组合。用于丝状真菌细胞中的复制起点的实例是AMA1和ANSI(Gems等,1991;Cullen等,1987;WO 00/24883)。根据WO 00/24883中公开的方法,可以完成AMA1基因的分离以及包含该基因的质粒或载体的构建。

[0144] 对于其他宿主,可以,例如,对于克鲁维酵母(*Kluyveromyces*)在Jeremiah D.Read等,Applied and Environmental Microbiology, Aug.2007, p.5088-5096中,对于发酵单胞菌(*Zymomonas*) in Osvaldo Delgado等, FEMS Microbiology Letters 132, 1995, 23-26中,对于树干毕赤酵母(*Pichia stipitis*) in US 7,501,275中,和对于梭状芽孢杆菌(*Clostridium*) in WO 2008/040387中找到转化方法。

[0145] 可以将多于一个拷贝的基因插入到宿主中来增加基因产物的产生。可以通过将至少一个附加的基因拷贝整合入宿主基因组中或通过在其中含有可通过在适当的选择剂存在下培养细胞进行选择的扩增拷贝的选择标记基因的细胞的情况下包含带有核苷酸序列的可扩增选择标记基因并因此包含附加的基因拷贝,来获得基因拷贝数的增加。

[0146] 用于接合上述元件来构建本发明的重组表达载体所使用的程序是本领域技术人员公知的(参见,例如,Sambrook等人,1989,同上文)。

[0147] 用至少一种表达载体转化宿主细胞。当只使用单一表达载体(不添加中间体)时,该载体将包含所有必要的核酸序列。

[0148] 一旦已经用表达载体转化了宿主细胞,就使该宿主细胞生长。本发明的方法可包括培养所述宿主细胞,以使得在细胞中表达重组核酸。对于微生物宿主,这个过程导致在合适的培养基中培养细胞。通常情况下,细胞在35℃下在适当的培养基中生长。在本发明中某些生长培养基包括,例如,常见的市售的培养基,例如Luria-Bertani(LB)肉汤、Sabouraud Dextrose(SD)肉汤或酵母培养基(YM)肉汤。其它详细说明书的或合成的生长培养基也可被使用,并且微生物学或发酵领域中的技术人员知道用于特定的宿主细胞生长的合适的培养基。适宜生长的温度范围和其他条件在本领域中是已知的(见,例如,Bailey和Ollis 1986)。

[0149] 从宿主细胞纯化本发明的重组蛋白质的方法在本领域中是众所周知的(见E.L.V.Harris和S.Angel,Eds.(1989)Protein Purification Methods:A Practical Approach, IRL Press, Oxford, England)。这样的方法包括但不限于,制备型圆盘凝胶电泳、等电聚焦、高效液相色谱(HPLC)、反相HPLC、凝胶过滤、离子交换和分配色谱法以及逆流分配,以及它们的组合。在某些实施方式中,重组蛋白质带有额外的序列标记以便于纯化。这些标记包括表位标签和蛋白质标签。表位标签的非限制性实例包括:c-myc、血凝素(HA)、多聚组氨酸(6x-HIS)、GLU-GLU和DYKDDDDK(FLAG)(SEQ ID NO:117)表位标签。可以通过多种已建立的方法将表位标签添加到肽上。表位标签的DNA序列可以作为寡核苷酸或通过PCR扩增所用的引物插入到重组蛋白质编码序列中。作为替代方案,可以将多肽编码序列克隆到产生具有表位标签的融合体的特定载体中,例如,pRSET载体(Invitrogen Corp., San Diego, Calif)。蛋白标签的非限制性实例包括谷胱甘肽-S-转移酶(GST)、绿色荧光蛋白

(GFP)和麦芽糖结合蛋白(MBP)。通过一些公知的方法将蛋白质标签连接到肽或多肽。在一种方法中,多肽或肽的编码序列可以被克隆到产生多肽或肽与目标蛋白质标签之间的融合体的载体中。合适的载体包括,但不限于,示例性的质粒、pGEX(Amersham Pharmacia Biotech, Inc., Piscataway, N.J.)、pEGFP(CLONTECH Laboratories, Inc., Palo Alto, Calif)和pMAL.TM.(New England BioLabs, Inc., Beverly, Mass.)。表达后,可通过在适当的固相基质上的色谱法从宿主细胞的粗分解产物中纯化表位或蛋白质标记的多肽或肽。在某些情况下,可能优选地在纯化后移除表位或蛋白质标签(即,通过蛋白酶切割)。

[0150] 复合聚糖的制备方法

[0151] 本发明的另一个方面包括复合的N-聚糖的制备方法,其包括以下步骤:提供宿主细胞,其中该宿主细胞包含编码融合蛋白质的多核苷酸,该融合蛋白质包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域,和培养该宿主细胞以表达融合蛋白质,其中融合蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到末端Man α 6残基以产生复合的N-聚糖。在某些实施方式中,这一方面包括在木霉菌细胞中产生人样N-聚糖的方法。

[0152] 如本文所使用的术语“复合的N-聚糖”是指含有末端GlcNAc₂Man₃结构的N-聚糖。

[0153] 复合的N-聚糖包括具有下式的任何聚糖:[GlcNAc β 2]_zMan α 3([GlcNAc β 2]_wMan α 6)Man{ β 4GlcNAc β (Fuc α x)_n[4GlcNAc]_m}_p,其中n、m和p是0或1,表示分子部分的存在或不存在,条件是当m为0时,则n是0(海藻糖是连接到GlcNAc的分支),其中x是3或6,其中()限定结构中的分支,其中[]限定聚糖结构的部分,其以线性序列存在或不存在,和其中z和w是0或1。优选地,w和z是1。在某些实施方式中,在复合的N-聚糖包括GlcNAc β 2Man α 3(GlcNAc β 2Man α 6)Man β 4GlcNAc β 4GlcNAc、GlcNAc β 2Man α 3(Man α 6)Man β 4GlcNAc β 4GlcNAc、GlcNAc β 2Man α 3(GlcNAc β 2Man α 6)Man β 4GlcNAc β 4(Fuc α 6)GlcNAc、GlcNAc β 2Man α 3(Man α 6)Man β 4GlcNAc β 4(Fuc α 6)GlcNAc和Man α 3(Man α 6)Man β 4GlcNAc β 4GlcNAc。在某些实施方式中,复合的N-聚糖是真菌非岩藻糖基化GlcNAcMan₃、GlcNAc₂Man₃和/或Man₃。

[0154] 在某些实施方式中,复合的N-聚糖的制备方法会产生不同聚糖的混合物。该复合的N-聚糖可以占到这样的聚糖混合物的至少1%、至少3%、至少5%、至少10%、至少15%、至少20%、至少25%、至少50%或至少75%或更多。

[0155] 可将受体聚糖和因此复合的N-聚糖连接到分子如氨基酸、肽或多肽上。在某些实施方式中,氨基酸衍生物是天冬酰胺残基。该天冬酰胺残基可以为侧链酰胺的氨基糖苷键(生物的哺乳动物多肽N-聚糖连接结构),并可以是如二肽、寡肽或多肽的肽链的部分。聚糖可以是还原末端衍生物如还原性GlcNAc或Man的N-、O-或C-连接的,优选糖苷的,衍生物,如具有选自于氨基酸、烷基、杂烷基、酰基、烷氧基、芳基、芳基烷基和杂芳基烷基的某一聚糖连接结构的间隔体或末端有机残基。间隔体可再连接至多价载体或固相上。在某些实施方式中,含烷基的结构包括甲基、乙基、丙基和C4-C26的烷基、脂质如甘油酯、磷脂、多萜醇-磷脂和神经酰胺及衍生物。还原性末端也可以通过还原胺化为仲胺连接或衍生结构来衍生化。特定载体包括生物多聚或者低聚体,例如(多)肽、聚(糖)如葡聚糖、纤维素、直链淀粉或葡萄糖胺聚糖,以及其他的有机聚合物或寡聚物如塑料,包括聚乙烯、聚丙烯、聚酰胺(如尼龙或聚苯乙烯)、聚丙烯酰胺和聚乳酸,树状聚合物如PAMAM、星形或海星形树状聚合物,或者聚赖氨酸,和聚烷基二醇类如聚乙二醇(PEG)。固相可以包括微滴定孔、二氧化硅颗粒、玻

璃、金属(包括钢、金和银)、聚合物珠如聚苯乙烯或树脂珠、聚乳酸珠、聚糖珠或含有磁珠的有机间隔体。

[0156] 在某些实施方式中,受体聚糖连接到异源多肽。在某些实施方式中,异源多肽是治疗性蛋白质。治疗性蛋白质可包括单克隆抗体、促红细胞生成素、干扰素、生长激素、酶或凝血因子,并可以用于人类或动物的治疗。例如,受体聚糖可以连接到治疗性蛋白,如利妥昔单抗。

[0157] 该受体聚糖可以是在题为“本发明的重组蛋白质”一节中所述的任何受体聚糖。

[0158] 在某些实施方式中,受体聚糖可以是Man5。在这样的实施方式中,使用随机整合或通过靶向整合到已知不影响Man5糖基化的已知位点用GnTII/GnTI融合酶转化表达Man5的里氏木霉菌株。选择产生GlcNAcMan5的菌株。用能够切割Man5结构以生成GlcNAcMan3的甘露糖苷酶II型甘露糖苷酶的催化结构域进一步转化选择的菌株。在某些实施方式中,甘露糖苷酶II型酶属于糖苷水解酶家族38(cazy.org/GH38_all.html)。特征性酶包括列在cazy.org/GH38_characterized.html中的酶。特别有用的酶是裂解糖蛋白的高尔基体型酶,如 α -甘露糖苷酶II(Man2A1;ManA2)亚家族。这样的酶的实例包括人类酶AAC50302、黑腹果蝇酶(Van den Elsen J.M.等(2001)EMBO J.20:3008-3017)、具有根据PDB索引1HTY的三维结构的那些和指为具有PDB中催化结构域的其他酶。对于细胞质表达,甘露糖苷酶的催化结构域通常与N-末端靶向肽融合,或与动物或植物的甘露糖苷酶II酶的内源性动物或植物高尔基体靶向结构一起表达。用甘露糖苷酶II型甘露糖苷酶的催化结构域转化后,选择有效地产生GlcNAc2Man3的菌株。

[0159] 宿主细胞

[0160] 复合的N-聚糖的制备方法包括提供宿主细胞的第一步。只要在用核酸序列转化后保持存活力,任何原核或真核宿主细胞都可以在本发明中使用。优选地,所述宿主细胞不受必要的核酸序列转导、随后的重组蛋白质表达或者产生的中间体的不利影响。合适的真核细胞包括,但不限于,真菌、植物、昆虫或哺乳动物细胞。

[0161] 在某些实施方式中,宿主是真菌菌株。本文所用的“真菌”包括子囊菌门、担子菌门、壶菌门和接合菌门(如Hawksworth等,In,Ainsworth and Bisby's Dictionary of The Fungi,8th edition,1995,CAB International,University Press,Cambridge,UK所定义的)以及卵菌门(如Hawksworth等,1995,同上,171页中所提及的)和所有有丝分裂孢子真菌(Hawksworth等人,1995,同上)。

[0162] 在具体的实施方式中,真菌宿主是酵母菌株。本文所用的“酵母”包括产子囊酵母(内孢霉目)、basidiosporogenous酵母和属于半知菌的酵母(芽孢纲(Blastomycetes))。由于酵母的分类可能在未来发生改变,为了本发明的目的,酵母应定义为如酵母的生物学和活性(Biology and Activities of Yeast)(Skinner,F.A.,Passmore,S.M.,和Davenport,R.R.,eds,Soc.App.Bacteriol.Symposium Series No.9,1980)中所描述的。

[0163] 在某些实施方式中,酵母宿主是念球菌属、汉逊酵母属、克鲁维酵母属、毕赤酵母属、酵母菌属、裂殖酵母属或耶罗威亚酵母菌株。

[0164] 在某些实施方式中,酵母宿主为酿酒酵母、乳酸克鲁维酵母、毕赤酵母、白色念珠菌、多形汉逊酵母、裂殖酵母属或耶罗威亚酵母。

[0165] 在另一具体实施方式中,真菌宿主细胞是丝状真菌菌株。“丝状真菌”包括真菌亚

门和卵菌亚门(Hawksworth等,1995年,如前,所定义的)的所有丝状形式。丝状真菌一般特征是由几丁质、纤维素、葡聚糖、脱乙酰壳多糖、甘露聚糖和其它复合多糖组成的菌丝体壁。营养生长是通过菌丝伸长且碳分解代谢是专性好氧的。与此相反,酵母例如酿酒酵母的营养生长是通过单细胞菌体出芽且碳分解代谢可以是发酵的。

[0166] 丝状真菌宿主细胞可以是,例如,枝顶孢属、曲霉属、镰孢菌属、腐质霉属、毛霉属、毁丝霉属、脉孢霉属、青霉属、柱霉属(*Scytalidium*)、梭孢壳属、弯颈霉属或木霉属菌株。

[0167] 在某些实施方式中,丝状真菌宿主细胞是木霉菌、枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、*Chrysosporium lucknowense*、线黑粉酵母属、镰孢菌属、赤霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、踝节菌属、嗜热子囊菌属、梭孢壳属和弯颈霉属菌株。

[0168] 在某些实施方式中,宿主细胞是哺乳动物细胞。这样的细胞可以是人类的或非人类的。

[0169] 在其他某些实施方式中,宿主细胞是原核的,并且在某些实施方式中,该原核生物是大肠杆菌、枯草芽孢杆菌(*Bacillus subtilis*)、运动发酵单胞菌(*Zymomonas mobilis*)、梭状芽孢杆菌、发酵植物多糖梭菌(*Clostridium phytofermentans*)、热纤梭菌(*Clostridium thermocellum*)、拜氏梭菌(*Clostridium beijerinckii*)、丙酮丁醇梭菌(*Clostridium acetobutylicum*)、热醋穆尔氏菌(*Moorella thermoacetica*)、*Thermoanaerobacterium saccharolyticum*或产酸克雷伯氏菌(*Klebsiella oxytoca*)。在其它实施方式中,原核宿主细胞是*Carboxydocella* sp.、谷氨酸棒状杆菌(*Corynebacterium glutamicum*)、肠杆菌科(*Enterobacteriaceae*)、菊欧文氏菌(*Erwinia chrysanthemi*)、乳杆菌属(*Lactobacillus* sp.)、乳酸片球菌(*Pediococcus acidilactici*)、荚膜红假单胞菌(*Rhodopseudomonas capsulata*)、乳酸链球菌(*Streptococcus lactis*)、弗尼斯弧菌(*Vibrio furnissii*)、弗尼斯弧菌M1、*Caldicellulosiruptor saccharolyticus*或野油菜黄单胞菌(*Xanthomonas campestris*)。在其他实施方式中,宿主细胞是蓝细菌。细菌宿主细胞的其他例子包括,但不限于,属于埃希氏菌属(*Escherichia*)、肠杆菌属、固氮菌(*Azotobacter*)、欧文氏菌属(*Erwinia*)、芽孢杆菌属(*Bacillus*)、假单胞菌属(*Pseudomonas*)、克雷伯菌属(*Klebsiella*)、变形杆菌属(*Proteus*)、沙门氏菌属(*Salmonella*)、沙雷氏菌属(*Serratia*)、志贺氏菌(*Shigella*)、根瘤菌(*Rhizobia*)、透明颤菌(*Vitreoscilla*)、聚球菌属(*Synechococcus*)、集胞藻(*Synechocystis*)和副球菌属(*Paracoccus*)的分类学分类的这些物种。

[0170] 在本发明的制备复合N-聚糖的方法中,该方法包括培养宿主细胞使得该融合蛋白质表达的步骤。对于微生物宿主,这个过程需要在合适的培养基中培养所述细胞。通常,细胞在35℃下在适当的培养基中生长。在本发明中的某些生长培养基包括,例如,常见的市售的培养基,例如Luria-Bertani(LB)肉汤,沙氏葡萄糖(SD)肉汤或酵母培养基(YM)肉汤。其它详细说明书的或合成的生长培养基也可被使用,并且微生物学或发酵领域中的技术人员知道用于特定的宿主细胞生长的合适的培养基。适宜生长的温度范围和其他条件在本领域中是已知的(见,例如,Bailey和Ollis 1986)。在某些实施方式中,细胞培养物的pH值是在3.5和7.5之间、在4.0和7.0之间、在4.5和6.5之间,在5和5.5之间或在5.5。

[0171] 制备复合的N-聚糖的方法中使用的宿主细胞包含编码在如题为“本发明的重组蛋

白质”一节中所述的本发明的任何重组蛋白质的多核苷酸。在某些实施方式中,该宿主细胞包含编码融合蛋白质的多核苷酸,该融合蛋白质包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域,其中融合蛋白质催化N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 6残基以产生复合的N-聚糖。

[0172] 在某些实施方式中,该宿主细胞包含编码UDP-GlcNAc转运蛋白的多核苷酸。编码UDP-GlcNAc转运蛋白的多核苷酸可以是对于宿主细胞内源性的(例如,天然存在的),或者它可以是对于宿主细胞异源的。

[0173] 在某些实施方式中,该宿主细胞包含编码 α -1,2-甘露糖苷酶的多核苷酸。编码 α -1,2-甘露糖苷酶的多核苷酸在宿主细胞中可以是内源性的,或者它可以是对于宿主细胞异源的。这些多核苷酸特别可用于表达从高尔基体转移到内质网的高甘露糖聚糖而没有有效的外切- α -2-甘露糖苷酶切割的宿主细胞。 α -1,2-甘露糖苷酶可以是属于糖苷水解酶家族47(cazy.org/GH47_all.html)的甘露糖苷酶I型酶。在某些实施方式中, α -1,2-甘露糖苷酶是在cazy.org/GH47_characterized.html中列出的酶。特别地, α -1,2-甘露糖苷酶可以是切割糖蛋白的ER型酶,如ER α -甘露糖苷酶I EC 3.2.1.113酶亚家族中的酶。这样的酶的例子包括人 α -2-甘露糖苷酶IB(AAC26169)、哺乳动物ER甘露糖苷酶的组合或丝状真菌酶如 α -1,2-甘露糖苷酶(MDS1)(里氏木霉AAF34579;Maras M等J Biotech.77,2000,255)。对于胞质表达,甘露糖苷酶的催化结构域通常与靶向肽融合,如HDEL、KDEL或者ER或早期高尔基蛋白的部分,或与动物或植物甘露糖苷酶I酶的内源性ER靶向结构一起表达。

[0174] 在某些实施方式中,该宿主细胞包含编码半乳糖基转移酶的多核苷酸。半乳糖基转移酶将 β -连接的半乳糖残基转移到末端N-乙酰葡萄糖胺残基上。在某些实施方式中,半乳糖基转移酶是 β -4-半乳糖基转移酶。一般情况下, β -4-半乳糖基转移酶属于CAZy糖基转移酶家族7(cazy.org/GT7_all.html),且包括 β -N-乙酰葡萄糖胺基-糖肽 β -1,4-半乳糖基转移酶(EC 2.4.1.38),其也被称为N-乙酰乳糖胺合酶(EC2.4.1.90)。可用的亚家族包括 β 4-GalT1、 β 4-GalT-II、-III、-IV、-V和-VI,如哺乳动物或人 β 4-GalT1或 β 4-GalT-II、-III、-IV、-V和-VI或它们的任意组合。 β 4-GalT1、 β 4-GalT-II或 β 4GalT1III对于N-聚糖(如GlcNAcMan3、GlcNAc2Man3或GlcNAcMan5)上末端GlcNAc β 2结构的半乳糖基化是特别有用的(Guo S.等Glycobiology 2001,11:813-20)。催化区域的三维结构是已知的(例如,(2006) J.Mol.Biol.357:1619-1633),并且该结构已经在PDB数据库中用代码2FYD表示。CAZy数据库包括某些酶的实例。特征性的酶也在CAZy数据库中被列在cazy.org/GT7_characterized.html中。可用的 β 4GalT酶的实例包括 β 4GalT1,例如牛(*Bos taurus*)酶AAA30534.1(Shaper N.L.等Proc.Natl.Acad.Sci.U.S.A.83(6),1573-1577(1986))、人类的酶(Guo S.等Glycobiology 2001,11:813-20)和小家鼠(*Mus musculus*)酶AAA37297(Shaper,N.L.等1998 J.Biol.Chem.263(21),10420-10428); β 4GalT II酶如人 β 4GalT II BAA75819.1、中国仓鼠(*Cricetulus griseus*)AAM77195、小家鼠酶BAA34385和日本青鳉鱼(*Oryzias latipes*)BAH36754;及 β 4-GalT1III酶,如人 β 4-GalT1IIIBAA75820.1、中国仓鼠(*Cricetulus griseus*)AAM77196和小家鼠酶AAF22221。

[0175] 半乳糖基转移酶可以在宿主细胞的细胞质中表达。可以使用异源靶向肽,如Schwientek,J.Biol.Chem 1996 3398中描述的Kre2肽。可用于半乳糖基转移酶表达的启动

子包括组成型启动子如gpd,内源性糖基化酶和糖基转移酶(如在高尔基体或ER中合成N-聚糖的甘露糖基转移酶)的启动子,和高产率的内源性蛋白质的诱导型启动子,如cbh1启动子。

[0176] 在其中宿主细胞包含编码半乳糖基转移酶的多核苷酸的本发明某些实施方式中,宿主细胞还包含编码UDP-Gal和/或UDP-Gal转运蛋白的多核苷酸。在其中宿主细胞包含编码半乳糖基转移酶的多核苷酸的本发明某些实施方式中,在培养宿主细胞时,乳糖可以代替葡萄糖用作碳源。培养基可以是在pH值4.5和7.0之间,或在5.0和6.5之间。在其中宿主细胞含有编码半乳糖基转移酶的多核苷酸和编码UDP-Gal和/或UDP-Gal转运蛋白的多核苷酸的本发明的某些实施方式中,二价阳离子如 Mn^{2+} 、 Ca^{2+} 或 Mg^{2+} 可被加入到细胞培养基中。

[0177] 在某些实施方式中,宿主细胞中含有编码唾液酸转移酶的多核苷酸。唾液酸转移酶将 $\alpha 3$ -或 $\alpha 6$ -连接的唾液酸,如Neu5Ac,转移到半乳糖基化的复合聚糖的末端Gal。合适的唾液酸转移酶的实例可以在糖基化蛋白家族29(cazy.org/GT29.html)中找到。可用的 $\alpha 3$ 或 $\alpha 6$ -唾液酸转移酶包括 β -半乳糖苷 α -2,6-唾液酸转移酶(EC 2.4.99.1)与某些亚家族ST6Gal-I,和N-乙酰氨基乳糖苷 α -2,3-唾液酸转移酶(EC2.4.99.6),具有与 β -半乳糖苷 α -2,3-唾液酸转移酶(EC 2.4.99.4)的可能的交叉反应性。 $\alpha 3$ -唾液酸转移酶的可用亚型包括ST3Gal-III和ST3Gal-IV。它们的酶学特征性的种类如在糖基化酶的CAZy数据库中(cazy.org/GT29_characterized.html)中表征的进行列举。编码 $\alpha 3$ -或 $\alpha 6$ -连接的唾液酸转移酶的多核苷酸可能对于宿主细胞是内源性的,或者它可以对于宿主细胞是异源的。宿主细胞中的唾液酸化作用可能需要合成供体CMP-唾液酸如CMP-Neu5Ac的酶的表达,特别是在真菌、植物、线虫/寄生虫或昆虫细胞中。

[0178] 宿主细胞可以具有提高的或降低的各种内源酶的活性水平。活性水平降低可以通过用抑制剂、抗体等抑制内源性酶的活性来提供。在某些实施方式中,宿主细胞被遗传修饰来提高或降低各种内源性酶的活性。“遗传修饰的”是指用于建立以提高的水平、降低的水平或以突变的形式表达多肽的原核或真核宿主细胞的任何重组DNA或RNA方法。换句话说,宿主细胞已经用重组多核苷酸分子转染、转化或转导,并由此被改变以使得该细胞改变所需蛋白质的表达。

[0179] 导致基因或基因产物导致基因表达、基因的功能或基因产物(即,由该基因编码的蛋白质)的功能降低的遗传修饰可被称为基因的灭活(完全或部分)、缺失、干扰、堵断、沉默或下调或表达的减弱。例如,导致该基因所编码的蛋白质的功能降低的基因中的遗传修饰可以是基因完全缺失(即,该基因不存在,并因此该蛋白质不存在)、会导致蛋白质不完整或没有翻译(例如,蛋白质不表达)的基因突变或者减少或消除蛋白质天然功能(例如,表达具有降低的或没有酶活性或作用的蛋白质)的基因突变的结果。更具体地,涉及降低本文所讨论的蛋白质的作用一般是指所述宿主细胞中的任何遗传修饰,其导致蛋白质的表达和/或功能性(生物活性)的降低,并包括蛋白质的活性降低(例如,降低的催化作用)、增加蛋白质的抑制或降解以及蛋白质表达的减少或消除。例如,通过阻断或减少蛋白质的产生、降低蛋白质的作用或抑制蛋白质的作用可以降低本发明蛋白质的作用或活性。这些修饰中一些修饰的组合也是可能的。阻断或减少蛋白质的产生可以包括将编码该蛋白质的基因置于需要在生长培养基中存在诱导化合物的启动子的控制下。通过建立使得从培养基耗尽诱导剂的条件,编码蛋白质的基因的表达(和因此蛋白质的合成)可以被关闭。阻断或降低蛋白质的

作用也可以包括使用切除技术的方法,其与4,743,546号美国专利中描述的相似。使用这种方法,编码目的蛋白质的基因被克隆到特定的基因序列之间,这使得基因从基因组中特异性地受控切除。可以通过例如培养物的培养温度的改变,如在4,743,546号美国专利中,或者通过其他一些物理或营养信号来引起切除。

[0180] 一般来说,根据本发明,突变体或修饰蛋白质的给定特性(例如,酶活性)的增加或减少是参照来自于相同的生物体(来自相同的来源或亲本序列)的野生型(即,正常的,非修饰的)蛋白质的相同特性做出的,其是在相同或同等的条件下测定或建立的。同样地,遗传修饰的宿主细胞的特性的提高或降低(例如,蛋白质的表达和/或生物活性,或产物的产生)是参照同一物种且优选相同菌株的野生型宿主细胞在相同或同等的条件下的相同特性作出的。这些条件包括在蛋白质的活性(例如,表达或生物活性)或宿主细胞的其他特性进行测量的培养条件(例如,培养基成分、温度、pH值等),以及使用的检测分析类型、被评估的宿主细胞等等。如上面所讨论的,等同条件是相似但并不一定是相同的(例如,可以耐受条件的一些保守性变化)的条件(例如,培养条件),且与在相同条件下进行的操作相比,基本上不改变对细胞生长或者酶表达或生物活性的影响。

[0181] 优选地,与在野生型宿主细胞中的野生型蛋白质的活性相比,具有提高或降低给定蛋白质(例如,酶)的活性的遗传修饰的遗传修饰宿主细胞分别具有蛋白质活性或作用(例如,表达,产生和/或生物活性)至少约5%、更优选至少约10%、更优选至少约15%、更优选至少约20%、更优选至少约25%、更优选至少约30%、更优选至少约35%、更优选至少约40%、更优选至少约45%、更优选至少约50%、更优选至少约55%、更优选至少约60%、更优选至少约65%、更优选至少约70%、更优选至少约75%、更优选至少约80%、更优选至少约85%、更优选至少约90%和更优选至少约95%,或者5%和100%之间的全部整数的任何百分比(例如6%、7%、8%等)的提高或降低。当将分离的修饰核酸分子或蛋白质直接与分离的野生型核酸分子或蛋白质进行比较时(例如,如果是在体外与体内进行比较),相同的差异是特定的。

[0182] 在本发明的另一个方面,与在野生型宿主细胞中的野生型蛋白质的活性相比,具有提高或降低给定蛋白质(例如,酶)的活性的遗传修饰的遗传修饰宿主细胞分别具有蛋白质活性或作用(例如,表达,产生和/或生物活性)至少约2倍、和更优选至少约5倍、和更优选至少约10倍、和更优选为约20倍、和更优选至少约30倍、和更优选至少约40倍、和更优选至少约50倍、和更优选至少约75倍、和更优选至少约100倍、和更优选至少约125倍和更优选至少约150倍,或者从至少约2倍开始的全部整数的任何增量(例如,3倍、4倍、5倍、6倍等)的提高或降低。

[0183] 在某些实施方式中,与野生型宿主细胞中的活性水平相比,该宿主细胞具有降低的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶活性水平。多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶(EC 2.4.1.130)将 α -D-甘露糖残基从多萜醇-磷酸D-甘露糖转移到膜脂连接的寡糖。通常情况下,多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶是由alg3基因所编码的。在某些实施方式中,与野生型宿主细胞中的表达水平相比,宿主细胞具有降低的alg3基因表达水平。在某些实施方式中,从宿主细胞中删除了alg3基因。

[0184] 在某些实施方式中,与野生型宿主细胞中的活性水平相比,该宿主细胞具有降低

的 α -1,6-甘露糖基转移酶的活性水平。 α -1,6-甘露糖基转移酶(EC 2.4.1.232)将 α -D-甘露糖残基从GDP-甘露糖转移到蛋白质结合寡糖中,从而在高尔基器中形成产生 α -(1->6)-D-甘露糖基-D-甘露糖的连接。通常情况下, α -1,6-甘露糖基转移酶由och1基因编码。在某些实施方式中,与野生型宿主细胞中的表达水平相比,宿主细胞具有降低的och1基因表达水平。在某些实施方式中,从宿主细胞中删除och1基因。

[0185] 在某些实施方式中,该宿主细胞具有降低的蛋白酶活性水平。在某些实施方式中,从宿主细胞中删除编码各种蛋白酶的基因。这些基因包括,例如,基因编码蛋白酶如pep1(曲霉属中的pepA)和纤维素分解酶,如纤维二糖水解酶1(cbh1)。

[0186] 在某些实施方式中,该宿主细胞可以具有降低的参与非同源末端连接(NHEJ)的蛋白质活性水平以提高同源重组的效率。在某些实施方式中,从宿主细胞中删除编码这些蛋白质的基因。该基因和它们的同源物包括,但不限于,在例如,Ninomiya等2004,Ishibashi等2006,Villalba等2008,和Mizutani等2008中所描述的Ku70、Ku80、Lig4、Rad50、Xrs2、Sir4、Lif1或Neil。

[0187] 在制备复合N-聚糖的方法的某些实施方式中,宿主细胞是木霉属的细胞,其与野生型的木霉属细胞中的活性水平相比,具有降低的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶的活性水平。

[0188] 在制备复合N-聚糖的方法的其它某些实施方式中,宿主细胞是酵母细胞,其与野生型的酵母细胞中的活性水平相比,具有降低的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶的活性水平和降低的 α -1,6-甘露糖基转移酶的活性水平,并且还包含编码 α -1,2-甘露糖苷酶的多核苷酸。

[0189] 制备复合的N-聚糖的体外方法

[0190] 在另一个方面,本发明提供了复合的N-聚糖的制备方法,包括将包含N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域的融合蛋白质、受体聚糖及N-乙酰葡萄糖胺供体一起在缓冲液中孵育的步骤,其中所述融合蛋白质催化N-乙酰葡萄糖胺转移至受体聚糖的末端Man α 3残基和N-乙酰葡萄糖胺转移到受体聚糖的末端Man α 6残基以产生复合的N-聚糖。在某些实施方式中,受体聚糖被连接到氨基酸、肽或多肽。在某些实施方式中,受体聚糖被连接到异源多肽。在某些实施方式中,受体聚糖是Man3。在某些实施方式中,N-乙酰葡萄糖胺供体是UDP-GlcNAc转运蛋白。通常情况下,缓冲液包含浓度为1 μ M至100mM、100 μ M至50mM或0.1mM至25mM的二价阳离子,如Mn²⁺、Ca²⁺或Mg²⁺。通常使用过量摩尔的N-乙酰葡萄糖胺供体,如相对于受体聚糖上的反应性受体位点过量1.1-100倍。受体聚糖的浓度通常介于1 μ M至100mM、100 μ M至50mM或1至25mM之间。当受体聚糖被连接到多肽上时,因为其较高的分子量,浓度范围通常是在低端。可以根据其在缓冲液中的溶解度调节反应物组分的浓度。可以调节酶活性(单位)的量以允许在合理的反应时间内实现有效的反应。合理的反应时间通常是从数分钟至数天。在某些实施方式中,反应时间为从约0.5个小时至一天或从1至6个小时。

[0191] 有用的缓冲液包括适用于融合蛋白质的缓冲液,如TRIS、HEPES、MOPS,pH值范围为约5至8.5、5.5至8.0或6.0和7.5。TRIS、HEPES或MOPS缓冲液的浓度通常在5至150mM之间,10-100mM之间或10-60mM之间,其经调节以维持该pH。可以通过加入盐如10-200mM的NaCl和/或稳定的酶但非可糖基化的蛋白质(例如,纯的非糖基化的聚糖或含有白蛋白的非受体

聚糖)优化该反应。在一些实施方式中,体外反应经调节以在细胞培养基中进行。可以使用磷酸盐缓冲液以降低反应速度。

[0192] 制备Man₃GlcNAc₂聚糖的细胞与方法

[0193] 在另一个方面,本发明提供了含有alg3突变和Man₃GlcNAc₂的丝状真菌细胞,其中Man₃GlcNAc₂占由细胞分泌的中性N-聚糖的至少50%、至少60%、至少70%、至少80%、至少90%或100%(摩尔%)。中性N-聚糖可连接到氨基酸、肽或多肽。可通过本领域中已知的任何手段突变alg3基因,如点突变或删除整个alg3基因。优选地,通过alg3的突变降低或消除alg3蛋白质的功能。丝状真菌细胞可以是枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、Chrysosporium lucknowense、线黑粉酵母属、镰孢菌属、赤霉属、腐质霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、柱霉属、踝节菌属、嗜热子囊菌属、梭孢壳属、弯颈霉属或木霉属的细胞。在某些实施方式中,丝状真菌细胞是里氏木霉细胞。在某些实施方式中,丝状真菌细胞进一步包含一个或多个编码任何本发明的重组蛋白质的多核苷酸。例如,丝状真菌细胞可以进一步含有编码N-乙酰葡萄糖胺基转移酶I催化结构域的第一多核苷酸和编码N-乙酰葡萄糖胺基转移酶II催化结构域的第二多核苷酸。或者,丝状真菌细胞可以进一步含有编码融合蛋白质的多核苷酸,其中该融合蛋白质包括N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域。

[0194] 在再另一个方面中,本发明提供了在宿主细胞中制备Man₃GlcNAc₂聚糖的方法,包括以下步骤:提供与野生型宿主细胞中的活性水平相比具有降低的甘露糖基转移酶活性水平的宿主细胞,和培养该宿主细胞以产生Man₃GlcNAc₂聚糖,其中Man₃GlcNAc₂聚糖构成由宿主细胞分泌的中性N-聚糖的至少50%、至少60%、至少70%、至少80%、至少90%或100%(摩尔%)。

[0195] 该Man₃GlcNAc₂聚糖可以连接到分子如氨基酸、肽或多肽上。在某些实施方式中,所述氨基酸为天冬酰胺残基。天冬酰胺残基可以与侧链酰胺形成氨基糖苷连接(生物的哺乳动物蛋白质N-聚糖连接结构),并可以是如二肽、寡肽或多肽的肽链的部分。聚糖可以是还原性末端衍生物如还原性GlcNAc或Man的N-、O-或C-连接的,优选的糖苷的,衍生物,例如具有某些选自于氨基酸、烷基、杂烷基、酰基、烷氧基、芳基、芳基烷基或杂芳基烷基的聚糖连接结构的间隔体或末端有机残基。间隔体还可以连接到多价载体或固相上。在某些实施方式中,含烷基的结构包括:甲基、乙基、丙基和C₄-C₂₆的烷基、脂质如甘油脂质、磷脂、长醇-磷脂和神经酰胺及衍生物。还原性末端还可以通过还原性胺化衍生为仲胺连接或衍生结构而衍生化。某些载体包括生物多聚-或低聚物,如(多)肽、多(糖)如葡聚糖、纤维素、直链淀粉或糖胺聚糖和其它有机聚合物或低聚物如塑料,包括聚乙烯、聚丙烯、聚酰胺(例如,尼龙或聚苯乙烯)、聚丙烯酰胺和聚乳酸、树状聚合物如PAMAM,星型聚合物或海星树状聚合物、或聚赖氨酸和聚烷基二醇如聚乙二醇(PEG)。固相可以包括微量滴定孔、二氧化硅颗粒、玻璃、包括钢、金和银的金属、如聚苯乙烯或树脂珠的聚合物珠粒、聚乳酸珠、多糖珠或含有磁珠的有机间隔体。

[0196] 在某些实施方式中,Man₃GlcNAc₂聚糖连接到异源多肽上。在某些实施方式中,异源多肽是治疗性蛋白质。治疗性蛋白质可包括单克隆抗体、促红细胞生成素、干扰素、生长激素、酶或凝血因子,并且可以用于治疗人或动物。例如,Man₃GlcNAc₂聚糖可以连接到治疗性

蛋白如利妥昔单抗上。通常情况下,将Man₃GlcNAc₂聚糖进一步修饰以成为复合的聚糖。这样的修改可以在宿主细胞中体内发生或通过体外的方法发生。

[0197] 在某些实施方式中,甘露糖基转移酶是多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶。通常情况下,多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶是由alg3基因所编码的。在某些实施方式中,与野生型宿主细胞中的表达水平相比,宿主细胞具有降低的alg3基因的表达水平。在某些实施方式中,从宿主细胞中删除了alg3基因。SEQ ID NO:97和98分别提供了里氏木霉中的alg3基因的核酸序列和氨基酸序列。

[0198] 在某些实施方式中,与野生型宿主细胞中的活性水平相比,该宿主细胞中的 α -1,6-甘露糖基转移酶的活性水平没有降低。通常情况下, α -1,6-甘露糖基转移酶是由och1基因编码的。在某些实施方式中,该宿主细胞含有编码 α -1,2-甘露糖苷酶的内源性多核苷酸。

[0199] 在某些实施方式中,宿主细胞是木霉属的细胞,并在某些实施方式中,宿主细胞是里氏木霉细胞。

[0200] 本发明的丝状真菌细胞

[0201] 在进一步的方面中,本发明提供了丝状真菌细胞,其与野生型丝状真菌细胞中的alg3基因的表达水平相比,具有降低的本发明alg3基因的表达水平,其中该丝状真菌细胞还包含如题为“本发明的重组蛋白质”一节中所描述的任何本发明的重组蛋白质。例如,在某些实施方式中,丝状真菌细胞还含有编码融合蛋白质的多核苷酸,其中该融合蛋白质包括N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域。该融合蛋白质的表达可以受控于可操作地与所述多核苷酸连接的启动子。该启动子可以是组成型启动子或诱导型启动子。在某些优选的实施方式中,该启动子是诱导型启动子,如cbh1诱导型启动子。

[0202] 在另一个方面,本发明提供的丝状真菌细胞与野生型丝状真菌细胞中的alg3基因的表达水平相比具有降低的本发明alg3基因的表达水平,其中该丝状真菌细胞还包含编码重组N-乙酰葡萄糖胺基转移酶I催化结构域的第一多核苷酸和编码重组N-乙酰葡萄糖胺基转移酶II催化结构域的第二多核苷酸。在这样的实施方式中,重组N-乙酰葡萄糖胺基转移酶I催化结构域的表达受可操作地与第一多核苷酸连接的启动子的控制,且重组N-乙酰葡萄糖胺基转移酶II催化结构域的表达受可操作地与第二多核苷酸连接的启动子的控制。启动子可以是组成型启动子或诱导型启动子。在某些优选的实施方式中,启动子是诱导型启动子,如cbh1诱导型启动子。

[0203] 在其它实施方式中,单一多核苷酸可以同时编码重组N-乙酰葡萄糖胺基转移酶I催化结构域和重组N-乙酰葡萄糖胺基转移酶II催化结构域,使得它们表达为独立的多肽。在这样的实施方式中,所述多核苷酸可以含有允许从该多核苷酸单独翻译各个催化结构域的内部核糖体进入位点。在这样的实施方式中,重组N-乙酰葡萄糖胺基转移酶I催化结构域的表达受可操作地与编码N-乙酰葡萄糖胺基转移酶I催化结构域的多核苷酸部分连接的启动子的控制,且重组N-乙酰葡萄糖胺基转移酶II催化结构域的表达受可操作地与编码N-乙酰葡萄糖胺基转移酶II催化结构域的多核苷酸部分连接的启动子的控制。启动子可以是组成型启动子或诱导型启动子。在某些优选的实施方式中,启动子是诱导型启动子,如cbh1诱导型启动子。

[0204] 如本文所公开的,N-乙酰葡萄糖胺基转移酶I(GlcNAc-TI;GnTI;EC 2.4.1.101)催化

反应 $\text{UDP-N-乙酰基-D-葡萄糖胺} + 3-(\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R} \rightleftharpoons \text{UDP} + 3-(2-(\text{N-乙酰基}-\beta\text{-D-葡萄糖胺基})-\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R}$, 其中R代表聚糖受体中N-连接的寡糖的其余部分。N-乙酰葡萄糖胺基转移酶I催化结构域是能够催化该反应的N-乙酰葡萄糖胺基转移酶I酶的任何部分。来自各种生物体的N-乙酰葡萄糖胺基转移酶I酶的氨基酸序列在SEQ ID NO:1-19中列出。另外的GnTI酶在CAZy数据库中的糖基转移酶家族13(cazy.org/GT13_all)中列出。酶学特征性种类包括拟南芥AAR78757.1(US6 653 459)、秀丽隐杆线虫AAD03023.1(Chen S.等*J. Biol. Chem* 1999; 274(1):288-97)、黑腹果蝇AAF57454.1(Sarkar和Schachter *Biol Chem.* 2001 Feb; 382(2):209-17)、中国仓鼠AAC52872.1(Puthalakath H.等*J. Biol. Chem* 1996 271(44):27818-22)、智人AAA52563.1(Kumar R.等*Proc Natl Acad Sci U S A.* 1990 Dec; 87(24):9948-52)、金黄地鼠AAD04130.1(Opat As等*Biochem J.* 1998 Dec 15; 336(Pt 3):593-8)(包括失活突变体的例子)、兔, 穴兔AAA31493.1(Sarkar M等.*Proc Natl Acad Sci U S A.* 1991 Jan 1; 88(1):234-8)。可以在cazy.org/GT13_characterized中找到特征性活性酶的其他例子。在Unligil UM等*EMBO J.* 2000 Oct 16; 19(20):5269-80中通过X-射线晶体学描述了兔GnTI的催化结构域的三维结构。GnTI的蛋白质数据库(PDB)结构是1F08、1F09、1F0A、2AM3、2AM4、2AM5和2APC。在某些实施方式中, N-乙酰葡萄糖胺基转移酶I催化结构域是来自人N-乙酰葡萄糖胺基转移酶I酶(SEQ ID NO:1)或其变异体。在某些实施方式中, N-乙酰葡萄糖胺基转移酶I催化结构域包含与SEQ ID NO:1的氨基酸残基84-445至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。在一些实施方式中, 可以使用较短的序列作为催化结构域(如人类酶的氨基酸残基105-445或兔子的酶的氨基酸残基107-447; Sarkar等.(1998)*Glycoconjugate J* 15:193-197)。可以用作GnTI催化结构域的另外的序列包括来自人类酶的大致氨基酸30至445的氨基酸残基或者在氨基酸残基30至105之间开始和延续至人类酶的大致氨基酸445的任何C-末端非催化结构域, 或者另一个GnTI的相应同源序列或其催化活性的变异体或突变体。催化结构域可以包括例如非催化结构域、跨膜结构域或胞质结构域的全部或部分的酶的N-末端部分。

[0205] 如本文所公开的, N-乙酰葡萄糖胺基转移酶II(GlcNAc-TII; GnTII; EC 2.4.1.143)催化反应 $\text{UDP-N-乙酰基-D-葡萄糖胺} + 6-(\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R} \rightleftharpoons \text{UDP} + 6-(2-(\text{N-乙酰基}-\beta\text{-D-葡萄糖胺基})-\alpha\text{-D-甘露糖基})-\beta\text{-D-甘露糖基-R}$, 其中R代表聚糖受体中N-连接的寡糖的其余部分。N-乙酰葡萄糖胺基转移酶II催化结构域是能够催化该反应的N-乙酰葡萄糖胺基转移酶II酶的任何部分。SEQ ID NO:20-33中列出了来自各种生物体的N-乙酰葡萄糖胺基转移酶II酶的氨基酸序列。在某些实施方式中, N-乙酰葡萄糖胺基转移酶II催化结构域是来自人N-乙酰葡萄糖胺基转移酶II酶(SEQ ID NO:20)或其变异体。另外的GnTII种类在CAZy数据库中的糖基转移酶家族16(cazy.org/GT16_all)中列出。酶学特征性种类包括, 秀丽隐杆线虫、黑腹果蝇、智人、褐家鼠、野猪(cazy.org/GT16_characterized)的GnTII。在某些实施方式中, N-乙酰葡萄糖胺基转移酶II催化结构域包含与SEQ ID NO:21的氨基酸残基大致30至大致447的至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%相同的序列。催化结构域可以包括例如非催化结构域、跨膜结构域或胞质结构域的全部或部分的酶的N-末端部分。

[0206] 在其中丝状真菌细胞含有本发明的融合蛋白质的实施方式中, 该融合蛋白质可进

一步包含在N-乙酰葡萄糖胺基转移酶I催化结构域和N-乙酰葡萄糖胺基转移酶II催化结构域之间的间隔区。可以使用如题为“间隔区”的一节中所描述的任何本发明的间隔区。在某些优选的实施方式中,间隔区是EGIV间隔区、2xG4S间隔区、3xG4S间隔区或CBHI间隔区。在其它实施方式中,间隔区包含来自非催化结构域的序列。

[0207] 对于ER/高尔基体表达,N-乙酰葡萄糖胺基转移酶I和/或N-乙酰葡萄糖胺基转移酶II催化结构域通常与靶向肽或者ER或早期高尔基蛋白质的部分融合,或与动物或植物N-乙酰葡萄糖胺基转移酶的内源性ER靶向结构一起表达。在某些优选的实施方式中,N-乙酰葡萄糖胺基转移酶I和/或N-乙酰葡萄糖胺基转移酶II催化结构域包含如题为“靶向肽”的章节中所描述的任何本发明的靶向肽。优选地,靶向肽被连接至催化结构域的N末端。在一些实施方式中,靶向肽包含如题为“靶向肽”的章节中所描述的任何本发明的非催化结构域。在某些优选的实施方式中,靶向肽是Kre2靶向肽。在其它实施方式中,靶向肽进一步包含连接到非催化结构域的N-末端的跨膜结构域或连接到非催化结构域N-末端的胞质结构域。在靶向肽还包含跨膜结构域的实施方式中,靶向肽可以进一步含有连接到跨膜结构域的N末端的胞质结构域。

[0208] 可以通过本领域中已知的任何适宜的方法降低本发明的alg3基因的表达水平,包括但不限于突变alg3基因。可通过例如,点突变或删除整个alg3基因来突变alg3。优选地,通过alg3的突变降低或消除alg3蛋白质的功能。alg3基因编码多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基 α -1,3-甘露糖基转移酶。如本文所公开的,本发明的多萜醇基-P-Man:Man(5)GlcNAc(2)-PP-多萜醇基甘露糖基转移酶将 α -D-甘露糖残基从多萜醇-磷酸D-甘露糖移到膜脂连接的寡糖中。

[0209] 在某些实施方式中,丝状真菌细胞可以含有编码UDP-GlcNAc转运蛋白的多核苷酸。编码UDP-GlcNAc转运蛋白的多核苷酸可以是丝状真菌细胞中内源性的(即,天然存在的),或者它可以对于丝状真菌细胞是异源的。

[0210] 在其它实施方式中,丝状真菌细胞也可以包含编码如题为“宿主细胞”的章节中描述的本发明的 α -1,2-甘露糖苷酶的多核苷酸。编码 α -1,2-甘露糖苷酶的多核苷酸在丝状真菌细胞中可以是内源性的,或者它可以对于丝状真菌细胞是异源的。这些多核苷酸特别可用于表达从高尔基体转移到内质网的高甘露糖聚糖而没有有效的外切- α -2-甘露糖苷酶切割的丝状真菌细胞。对于细胞质表达,甘露糖苷酶的催化结构域通常与靶向肽融合,如HDEL、KDEL或者ER或早期高尔基蛋白的部分,或与动物或植物甘露糖苷酶I酶的内源性ER靶向结构一起表达。

[0211] 在进一步的实施方式中,丝状真菌细胞也可以包含编码如题为“宿主细胞”的章节中描述的本发明的半乳糖基转移酶的多核苷酸。半乳糖基转移酶将 β -连接的半乳糖残基转移到末端N-乙酰葡萄糖胺残基。在某些实施方式中,半乳糖基转移酶是 β -4-半乳糖基转移酶。半乳糖基转移酶可以在丝状真菌的细胞质中表达。可以使用异源靶向肽,如Schwientek J.Biol.Chem 1996 3398中描述的Kre2肽。可用于表达半乳糖基转移酶的启动子包括组成型启动子,如gpd,在高尔基体或ER中合成N-聚糖的内源性糖基化酶和糖基转移酶(如甘露糖基转移酶)的启动子,和高产率的内源性蛋白质的诱导型启动子,如cbh1启动子。在其中宿主细胞包含编码半乳糖基转移酶的多核苷酸的本发明的实施方式中,该宿主细胞还包含编码UDP-Gal和/或UDP-Gal转运蛋白的多核苷酸。在其中该丝状真菌细胞包含编码半乳糖

基转移酶的多核苷酸的本发明的某些实施方式中,在培养丝状真菌细胞时,乳糖可以代替葡萄糖用作碳源。培养基可以是在pH 4.5和7.0之间,或在5.0和6.5之间。在其中丝状真菌细胞含有编码半乳糖基转移酶的多核苷酸和编码UDP-Gal和/或UDP-Gal转运蛋白的多核苷酸的本发明的某些实施方式中,二价阳离子如 Mn^{2+} 、 Ca^{2+} 或 Mg^{2+} 可被加入到细胞培养基中。

[0212] 在其它实施方式中,丝状真菌细胞也可以包含编码如题为“宿主细胞”的章节中所描述的本发明的唾液酸转移酶的多核苷酸。唾液酸转移酶将 $\alpha 3$ -或 $\alpha 6$ -连接的唾液酸,如Neu5 Ac,转移到半乳糖基化复合聚糖的末端Gal。编码 $\alpha 3$ -或 $\alpha 6$ -连接的唾液酸转移酶的多核苷酸可以对于丝状真菌细胞是内源性的,或者它可以对于丝状真菌细胞是异源的。丝状真菌细胞中的唾液酸化可能需要表达合成供体CMP-唾液酸如CMP-Neu5Ac的酶,特别是在真菌、植物、线虫/寄生虫或昆虫细胞中。

[0213] 此外,丝状真菌细胞可以具有增加的或降低的各种另外的内源性酶的活性水平。活性水平降低可以通过用抑制剂、抗体等抑制内源性酶的活性来提供。在某些实施方式中,丝状真菌细胞被遗传修饰来增加或降低一种或多种内源性酶的活性。对丝状真菌细胞进行遗传修饰以增加或降低一种或多种内源性酶的活性的方法是本领域中公知的,并包括但不限于,如题为“宿主细胞”的章节中所描述的那些。在某些实施方式中,与野生型丝状真菌细胞中的活性水平相比,该丝状真菌细胞具有降低的 $\alpha 1,6$ -甘露糖基转移酶的活性水平。 $\alpha 1,6$ -甘露糖基转移酶(EC 2.4.1.232)在高尔基器中将引发GDP-甘露糖的 α -D-甘露糖残基的伸长转移到蛋白质连接的N-聚糖寡糖中,从而形成 $\alpha(1\rightarrow6)$ -D-甘露糖基-D-甘露糖连接。通常情况下, $\alpha 1,6$ -甘露糖基转移酶由och1基因编码。在某些实施方式中,与野生型丝状真菌细胞中的表达水平相比,丝状真菌细胞具有降低的och1基因的表达水平。在某些实施方式中,从丝状真菌细胞中删除och1基因。

[0214] 丝状真菌细胞可以是,例如,枝顶孢属、曲霉属、短梗霉属、隐球菌属、金孢子菌属、Chrysosporium lucknowense、线黑粉酵母属、镰孢菌属、赤霉属、腐质霉属、梨孢菌属、毛霉属、毁丝霉属、漆斑菌属、新考马脂霉属、脉孢霉属、拟青霉属、青霉属、瘤胃壶菌属、裂褶菌属、柱霉属、踝节菌属、嗜热子囊菌属、梭孢壳属、弯颈霉属或木霉属细胞。在某些实施方式中,丝状真菌细胞是里氏木霉细胞。

[0215] 含有由本发明的方法制备的复合N-聚糖的药物组合物

[0216] 在另一个方面中,本发明提供了组合物,例如,药物组合物,其含有一种或多种与通过本发明的方法所产生的异源分子连接的复合N-聚糖,与药学上可接受的载体一起配制。本发明的药物组合物还可以在联合疗法中施用,即与其它药剂组合。例如,联合治疗可以包括与至少一种其它治疗剂相结合的与根据本发明的异源分子连接的复合N-聚糖。

[0217] 如本文所使用的“药学上可接受的载体”包括任何和所有的生理上相容的溶剂、分散介质、涂层、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂等。优选地,载体适合于静脉内、肌内、皮下、肠胃外、脊柱或表皮施用(例如,通过注射或输注)。取决于给药的途径,活性化合物,即连接到根据本发明的异源分子的复合N-聚糖,可以包被在材料中以保护该化合物免于酸和其它可能使化合物失活的自然条件的影响。

[0218] 本发明的药物组合物可包含一种或多种药学上可接受的盐。“药学上可接受的盐”是指保留母体化合物的所需生物活性并且不产生任何不需要的毒理学效应的盐(见例如, Berge, S.M. 等人(1977) J. Pharm. Sci. 66:1-19)。这类盐的例子包括酸加成盐和碱加成盐。

酸加成盐包括由无毒的无机酸如盐酸、硝酸、磷酸、硫酸、氢溴酸、氢碘酸、亚磷酸等以及由无毒的有机酸如脂肪族单羧酸和二羧酸、苯基取代的链烷酸、羟基链烷酸、芳族酸、脂族和芳族磺酸等衍生的那些。碱加成盐包括由碱土金属如钠、钾、镁、钙等以及由无毒的有机胺如N,N'-二苄基乙二胺、N-甲基葡萄糖胺、氯普鲁卡因、胆碱、二乙醇胺、乙二胺、普鲁卡因等衍生的那些。

[0219] 本发明的药物组合物也可包括药学上可接受的抗氧化剂。药学上可接受的抗氧化剂的实例包括：(1)水溶性抗氧化剂，如抗坏血酸、半胱氨酸盐酸盐、硫酸氢钠、偏亚硫酸氢钠、亚硫酸钠等；(2)油溶性抗氧化剂，如抗坏血酸棕榈酸酯、丁基羟基苯甲醚(BHA)、丁基化的羟基甲苯(BHT)、卵磷脂、没食子酸丙酯、 α -生育酚等；以及(3)金属螯合剂，如柠檬酸、乙二胺四乙酸(EDTA)、山梨醇、酒石酸、磷酸等等。

[0220] 在本发明的药物组合物中可以采用的合适的水性和非水性载体的例子包括水、乙醇、多元醇(如甘油、丙二醇、聚乙二醇等)及其合适的混合物、植物油如橄榄油和可注射的有机酯如油酸乙酯。可以，例如，通过使用包衣材料如卵磷脂，在分散体的情况下通过维持所需的颗粒尺寸，和通过使用表面活性剂保持适当的流动性。

[0221] 这些组合物也可以包含辅剂如防腐剂、润湿剂、乳化剂和分散剂。可以通过消毒程序和通过包含各种抗细菌剂和抗真菌剂，例如对羟基苯甲酸酯、氯丁醇、苯酚山梨酸等来确保防止微生物的存在。组合物中包括等渗剂如糖、氯化钠等也可能是需要的。另外，可通过包括延迟吸收的试剂如单硬脂酸铝和明胶导致可注射药物形式的延长吸收。

[0222] 药学上可接受的载体包括无菌水性溶液或分散液以及用于临时配制成无菌注射溶液或分散液的无菌粉末。使用这些用于药物活性物质的介质和试剂是本领域已知的。只要不是与活性化合物不相容的任何常规介质或试剂，就可以考虑将其用于本发明的药物组合物中。补充的活性化合物也可以加入组合物中。

[0223] 治疗组合物通常在制备和贮存条件下必须是无菌的和稳定的。该组合物可以配制成溶液、微乳液、脂质体或其他适合于高药物浓度的有序结构。载体可以是溶剂或分散介质，其包含例如，水、乙醇、多元醇(例如，甘油、丙二醇和液体聚乙二醇等)及其合适的混合物。可以例如，通过使用包衣如卵磷脂，在分散体的情况下通过维持所需颗粒大小以及通过使用表面活性剂保持适当的流动性。在许多情况下，优选在组合物中包括等渗剂例如糖、多元醇如甘露醇、山梨醇或氯化钠。另外，可通过在组合物中包括延迟吸收的试剂如单硬脂酸铝和明胶产生可注射药物形式的延长吸收。

[0224] 可通过按照需要将所需量的活性化合物掺入具有上面列举的一种成分或其组合的溶剂中然后无菌微滤制得无菌注射溶液。一般情况下，分散体的制备是通过将活性化合物加入包含基本分散介质和上面所列举的那些的所需其它成分的无菌媒介中。在用于制备无菌注射溶液的无菌粉末的情况下，特定的制备方法是从预先无菌过滤的溶液产生活性成分的粉末加任何附加的所需成分的真空干燥和冷冻干燥(冻干)。

[0225] 可以与载体材料结合以产生单一剂型的活性成分的量根据被治疗的主体和特定的给药方式而变化。通常，可以与载体材料结合以产生单一剂型的活性成分的量是产生治疗效果的组合物的量。一般来说，100%中，这个量的范围是与药学上可接受的载体组合的从约0.01%到约99%的活性成分，优选地从约0.1%至约70%，最优选地从约1%至约30%的活性成分。

[0226] 调整给药方案以提供最佳的所需反应(例如,治疗反应)。例如,可以施用单一大剂量,可以随着时间施用分割的数个剂量或可以按治疗情况的紧急状况所指示的按比例减少或增加该剂量。特别有利的是,为了便于给药的方便性和剂量均匀性,以单位剂量形式配制肠胃外组合物。本文所用的单位剂量形式是指适合作为用于待治疗受试者的单一剂量的物理分散单元,各个单元包含与所需的药物载体结合的经计算以产生所需的治疗效果的预定量的活性化合物。本发明的单位剂量形式的规格决定于并直接依赖于(a)该活性化合物的独特特性和所要达到的特定治疗效果,以及(b)合成这样的用于个体的治疗敏感性的活性化合物的领域中的固有局限性。

[0227] 对于施用连接到异源分子的复合N-聚糖,特别是其中异源分子是抗体的情况中,其剂量范围从bolus 0.0001至100mg/kg宿主体重,更通常为0.01至5mg/kg宿主体重。例如,剂量可以是0.3mg/kg体重、1mg/kg体重、3mg/kg体重、5mg/kg体重或10mg/kg体重或者在1-10mg/kg范围内。示例性的治疗方案要求每周给药一次、每两周给药一次、每三周给药一次、每四周给药一次、每月给药一次、每3个月给药一次或每3至6个月给药一次。连接到异源抗体的复合N-聚糖的某些剂量方案包括通过静脉内施用的1mg/kg体重或3mg/kg体重,其中抗体使用以下给药时间表之一施用:(i)每四周6个剂量,然后每隔3个月,(ii)每三个星期,(iii)3mg/kg体重一次,然后每三个星期1mg/kg体重。

[0228] 或者,根据本发明的连接到异源分子的复合N-聚糖可以作为缓释制剂施用,在这种情况下,需要较低频率的施用。剂量和频率根据给药物质在患者中的半衰期变化。通常,人类抗体显示出最长的半衰期,其次是人源化抗体、嵌合抗体和非人抗体。施用剂量和频率可以根据治疗是预防性还是治疗性的而变化。在预防性应用中,在长的一段时间内以相对不频繁的间隔施用相对低的剂量。有些患者他们的余下生命中持续接受治疗。在治疗性应用中,有时需要以相对短的时间间隔的相对高剂量直到病情发展减轻或终止,且优选直到病人显示部分或完全的疾病症状的改善。此后,患者可以施用预防方案。

[0229] 本发明的药物组合物中活性成分的实际剂量水平可以变化以便获得有效地实现对特定患者、组合物和给药方式的所需治疗反应而对病人无毒的活性组分的量。所选择的剂量水平取决于各种各样的药代动力学因素,包括所使用的本发明特定组合物或其酯、盐或酰胺的活性、给药途径、给药时间、施用的特定化合物排泄速率、治疗的持续时间、与使用的特定组合物结合的其它药物、化合物和/或材料、接受治疗的病人的年龄、性别、体重、状况、一般健康及过往病史,以及医学领域中公知的类似因素。

[0230] 本发明的免疫球蛋白的“治疗有效剂量”优选地导致疾病症状严重程度的减轻、无疾病症状期频率和持续时间增加或预防由于疾病的痛苦带来的障碍或残疾。例如,对于肿瘤的治疗,“治疗有效剂量”优选地相对于未治疗的受试者抑制细胞生长或肿瘤生长至少约20%、更优选至少约40%、甚至更优选至少约60%和仍更优选至少约80%。可以预测人类肿瘤中的功效的动物模型系统中评估化合物抑制肿瘤生长的能力。或者,可以通过检验化合物的抑制能力评估组合物的这一性能,这种抑制在体外通过熟练技术人员熟知的分析评估。治疗有效量的治疗化合物可降低肿瘤大小,或以其他方式改善在受试者的症状。本技术领域的普通技术人员能够基于如对象的大小、受试者症状的严重程度和选择的特定组合物或给药途径的因素确定这些量。

[0231] 可以使用本领域中已知的各种方法中的一种或多种,通过一个或多个给药途径施

用本发明的组合物。如本领域技术人员将理解的,给药途径和/或方式根据所期望的结果而变化。本发明的结合部分的某些给药途径包括静脉内、肌内、皮内、腹膜内、皮下、脊柱或其它肠胃外给药途径,例如通过注射或输注。本文所用的短语“肠胃外给药”是指肠道和局部给药以外的施用方式,通常是通过注射,且包括,但不限于,静脉内、肌肉内、动脉内、鞘内、囊内、眶内、心脏内、皮内、腹膜内、经气管、皮下、表皮下、关节内、包膜下、蛛网膜下腔、椎管内、硬膜外和胸骨内注射和输注。

[0232] 或者,根据本发明的与异源分子连接的复合N-聚糖可以通过非肠胃外途径给药,如局部、表皮或粘膜给药途径,例如鼻内、口腔、阴道、直肠、舌下或局部。

[0233] 活性化合物可以与防止化合物快速释放的载体一起制备,例如控释制剂,包括植入物、透皮贴片和微囊化递送系统。可以使用可生物降解的、生物相容的聚合物,如乙烯乙酸乙烯酯、聚酐、聚乙醇酸、胶原、聚原酸酯和聚乳酸。许多用来制备这些制剂的方法是获得专利的或通常是本领域技术人员已知的。见,例如,Sustained and Controlled Release Drug Delivery Systems, J.R. Robinson, ed., Marcel Dekker, Inc., New York, 1978。

[0234] 治疗组合物可以用本领域中已知的医疗设备施用。例如,在某个实施方式中,本发明的治疗组合物可以用无针皮下注射装置施用,如在第5,399,163、5,383,851、5,312,335、5,064,413、4,941,880、4,790,824或4,596,556号美国专利中公开的装置。可以用于本发明中的众所周知的植入物和模块的实例包括:4,487,603号美国专利,其公开了用于按控制速率分配药物的可植入微型输液泵;4,486,194号美国专利,其公开了通过皮肤施用药物的治疗装置;4,447,233号美国专利,其公开了用于以精确输注速率递送药物的药物输液泵;4,447,224号美国专利,其公开了用于连续药物递送的可变流量可植入输注装置;4,439,196号美国专利,其公开了具有多腔室隔室的渗透药物递送系统;和4,475,196号美国专利,其公开了渗透药物递送系统。

[0235] 在某些实施方式中,使用根据本发明的与异源分子连接的复合N-聚糖来治疗可以用治疗性抗体治疗的任何疾病。

[0236] 应当理解,虽然已经结合了本发明的某些特定实施方式来描述本发明,但前面的描述旨在说明而不是限制本发明的范围。本发明的范围内的其他方面、优点和修改对发明所属领域的技术人员是显而易见的。

[0237] 尽管已经描述了本发明,但提供下列实施例以说明的方式而不是以限制的方式来说明本发明。

实施例

[0238] 实施例1-用于糖工程的宿主菌株的选择

[0239] 这个实施例的目的是确定用于糖工程的最佳的里氏木霉菌株。最佳的菌株产生大量的Man5 N-聚糖和少量的酸性聚糖。

[0240] 样品

[0241] 对不同的里氏木霉菌株包括M44(VTT-D-00775; Selinheimo等, FEBSJ. 2006, 273 (18): 4322-35)、M81、M84、M109、M110、M131、M132、M133、M134和M124(M44的mus53缺失菌株)进行了分析。这10个菌株中的每一个都在摇瓶培养中生长。在三个不同的时间点取样:3天、5天和7天。收集上清液(分泌的蛋白质)和细胞沉淀,并在-20℃下冷冻保存直至进行聚糖分

析。

[0242] 在指定的时间点从分泌的蛋白质分离N-聚糖,随后是基质辅助激光解析电离飞行时间(MALDI-TOF)聚糖谱分析。对来自5天的时间点的细胞沉淀进行N-聚糖谱分析。对共80个样品(每个中性-和酸性上清部分30个,和每个中性-和酸性颗粒部分10个)进行分析。

[0243] 还对菌株M44进行分批和分批补料发酵培养以评估摇瓶和发酵罐培养之间的聚糖谱差异。对于聚糖分析,分析来自三个不同时间点的样品,总共12个样品(6个中性和6个酸性部分)。作为对照,对培养液进行了分析。

[0244] 质谱方法

[0245] 采用Bruker Ultraflex TOF/TOF仪(Bruker Daltonics,Germany)进行MALDI-TOF质谱分析。在作为 $[M+Na]^+$ 离子的正离子反射模式中检测中性N-聚糖和在作为 $[M-H]^-$ 离子的负离子线性模式中检测酸性N-聚糖。基于它们在光谱中的相对信号强度对中性N-聚糖组分的相对摩尔丰度赋值。将在存在的聚糖图谱中的所得聚糖信号归一化为100%以便于进行样本之间的比较。

[0246] 蛋白质特异性糖基化方法

[0247] 用SDS-PAGE分离来自发酵罐培养样品的蛋白质,并印迹到PVDF膜上。切除目的蛋白条带,并通过用PNGase F酶解来释放N-聚糖。

[0248] 里氏木霉菌株的中性N-聚糖谱

[0249] 所需的Man5结构在图1所示的质谱中可以被观察为m/z值1257.4处的 $[M+Na]^+$ 信号。发现分析的里氏木霉菌株的中性糖组具有Man5或Man8作为主要中性聚糖种类(在表2中的H5N2和H8N2)。

[0250] 表2:分析的里氏木霉菌株的不同的中性N-聚糖信号的百分比。

[0251]

菌株											
组合物	m/z	M44	M81	M84	M109	M110	M131	M132	M133	M134	M124
		%	%	%	%	%	%	%	%	%	%
H3N2	933	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
H4N2	1095	1.2	0.0	0.0	1.8	0.9	0.0	2.3	0.0	2.3	4.1
H5N2	1257	81.0	70.8	4.0	78.9	5.8	78.8	84.1	10.7	73.2	77.9
H6N2	1419	5.8	5.3	0.0	5.3	0.9	4.8	4.6	0.9	6.0	7.3
H7N2	1581	4.8	7.3	1.5	4.7	3.0	4.8	3.9	3.8	5.8	4.8
H8N2	1743	3.7	8.6	81.5	5.1	68.2	5.9	2.6	68.1	6.3	3.3
H9N2	1905	2.9	8.0	9.0	3.4	16.0	4.6	2.0	12.8	5.7	2.3
H10N2	2067	0.5	0.0	2.5	0.8	3.7	1.1	0.4	2.5	0.7	0.4
H11N2	2229	0.0	0.0	1.5	0.0	1.4	0.0	0.0	1.2	0.0	0.0
H12N2	2391	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[0252] 在中性N-糖基部分中观察到一些酸性N-聚糖。这可能是由于磷酸化聚糖的特定性质,如存在磷酸二酯结构,或可能导致在本研究中使用的实验条件下酸性物质泄漏至中性部分中的磷酸聚糖的其它性质。为检查相应的结构,对感兴趣的信号进行MS/MS分析。在MS/MS分析模式中使用Bruker Ultraflex TOF/TOF进行聚糖的质谱断裂化(图2)。由于聚糖没有全甲基化,不能得到基于MS/MS数据的明确结构指定。

[0253] 里氏木霉菌株的酸性N-聚糖谱

[0254] 对于糖工程目的,有最小量的酸性N-聚糖的菌株是有用的。因此,从用于筛选的菌株中分析酸性N-聚糖谱。分析的菌株的酸性N-聚糖谱示于图3和下表3中。

[0255] 表3:分析的里氏木霉菌株的不同酸性N-聚糖信号的百分比。

[0256]

		M44	M81	M84	M109	M110	M131	M132	M133	M134	M124
	m/z	%	%	%	%	%	%	%	%	%	%
Hex3HexNAc2SP	989	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex4HexNAc2SP	1151	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex5HexNAc2SP	1313	4.0	5.2	0.0	3.7	0.0	2.8	7.4	0.0	5.2	2.8
Hex5HexNAc2SP2	1393	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.9	0.0
Hex6HexNAc2SP	1475	23.7	27.3	2.2	18.1	2.1	22.4	21.0	3.9	24.9	26.3
Hex6HexNAc2SP2	1555	0.0	2.8	0.0	2.4	0.0	3.2	1.1	0.0	3.6	1.7
Hex7HexNAc2SP	1637	30.3	18.8	1.1	16.2	2.0	14.9	24.7	0.0	17.2	23.3
Hex7HexNAc2SP2	1717	0.0	7.7	0.0	8.6	0.0	10.7	2.5	0.0	10.4	7.0
Hex8HexNAc2SP	1799	18.4	11.8	17.9	12.8	9.7	9.1	19.7	14.5	8.8	11.2
Hex8HexNAc2SP2	1879	5.1	8.8	0.0	11.0	0.0	14.8	4.0	0.0	12.4	10.0
Hex9HexNAc2SP	1961	7.3	6.4	49.1	9.5	37.9	5.9	6.1	53.9	4.1	3.5
Hex9HexNAc2SP2	2041	4.2	5.0	0.0	5.7	0.0	7.3	5.1	0.0	5.9	7.2
Hex10HexNAc2SP	2123	2.8	2.9	19.7	4.5	28.1	2.6	2.3	19.3	2.1	1.6
Hex10HexNAc2SP2	2203	2.8	2.1	0.0	2.2	0.0	2.7	3.6	0.0	1.9	3.3
Hex11HexNAc2SP	2285	1.5	1.3	3.7	2.1	9.5	1.2	0.9	5.0	1.0	0.8
Hex11HexNAc2SP2	2365	0.0	0.0	0.0	0.9	0.0	1.3	1.5	0.0	0.8	1.3
Hex12HexNAc2SP	2447	0.0	0.0	1.3	1.0	1.6	1.0	0.0	0.0	0.5	0.0
Hex12HexNAc2SP2	2527	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.3	0.0
Hex13HexNAc2SP	2609	0.0	0.0	1.2	0.4	1.1	0.0	0.0	0.0	0.0	0.0
Hex14HexNAc2SP	2771	0.0	0.0	0.6	0.0	0.9	0.0	0.0	0.0	0.0	0.0

[0257] 来自发酵罐培养的M44菌株的N-聚糖谱

[0258] 在发酵罐中培养M44菌株以找出不同培养条件下是否可以导致其聚糖谱的变化。对发酵罐中培养样品进行N-聚糖分析(分批;41:10、88:45和112:50小时和分批补料;45:50、131:40和217:20小时)并与摇瓶培养的样品相比较。图4中显示了在发酵罐中培养的里氏木霉菌株M44分泌蛋白质的中性和酸性N-聚糖。下面的表4中给出了来自烧瓶和发酵罐培养物的N-聚糖百分比之间的对比。

[0259] 表4:在烧瓶中和发酵罐中培养的里氏木霉菌株M44的N-聚糖信号的百分比。

组成	m/z	烧瓶	发酵罐
		%	%
[0260] H3N2	933	0.0	0.0
H4N2	1095	1.2	0.0
H5N2	1257	81.0	91.3

[0261]

H6N2	1419	5.8	4.5
H7N2	1581	4.8	4.2
H8N2	1743	3.7	0.0
H9N2	1905	2.9	0.0
H10N2	2067	0.5	0.0
H11N2	2229	0.0	0.0
H12N2	2391	0.0	0.0

[0262] 摇瓶培养基的N-聚糖分析

[0263] 作为对照实验,对里氏木霉的培养基(不接触真菌)进行了分析。图5a显示了中性N-聚糖分析,其中没有观察到N-聚糖。仅在基线以上可见轻微的己糖低聚物的信号(最有可能是来源于培养基中使用的植物材料)。在图5b(酸性聚糖)中,没有观察到对应于N-聚糖的信号。

[0264] 分泌蛋白质的N-糖基化

[0265] 要检查单个分泌蛋白质之间是否有糖基化的变异,用SDS-PAGE分离来自发酵培养物上清液的样品,并印迹到PVDF膜上。然后用膜上酶促释放解离选定条带的N-聚糖。结果示于图6和7中。

[0266] 结论:中性聚糖

[0267] 本研究的目的是确定具有最高量的Man5N-聚糖和最低量的酸性聚糖的用于糖工程的里氏木霉菌株。在质谱分析中以Man5作为主峰的菌株可以有较高的内源性 α -1,2-甘露糖苷酶活性。基于里氏木霉N-糖基化的背景信息,Man5的可能结构是Man α 3[Man α 3(Man α 6)Man α 6]Man β 4GlcNAc β 4GlcNAc(Salovuori等,1987;Stals等,Glycobiology 14,2004,第725页)。

[0268] 一些菌株包含H8N2作为主要的中性糖形。基于文献记载,该糖形最有可能是Glc α 3Man α 2Man α 2Man5结构(Stals等,Glycobiology 14,2004,第725页)。有可能在这些菌株中的葡萄糖苷酶缺陷阻止了聚糖修剪为较小的糖形。

[0269] 在一些菌株中,在中性谱中观察到酸性N-聚糖。这种情况可能是由于酸性的N-聚糖的较高比例或在将中性聚糖与酸性聚糖分离的过程中特定的结构泄漏到中性部分中。

[0270] 在发酵罐中培养的菌株的聚糖谱与在摇瓶中培养的菌株的聚糖谱相比,对于糖工程稍微更有利。来自发酵罐培养的样品的单个蛋白质的糖基化与一般糖基化没有显著差异。所有分析的蛋白质含有Man5作为主要的糖形。这一观察结果表明所有分泌蛋白质经过了类似的聚糖加工。因此,表现为大多数分泌蛋白质类似地被里氏木霉宿主细胞糖基化,对于哺乳动物细胞并不总是这种情况。

[0271] 酸性聚糖

[0272] N-聚糖的磷酸化一般不是糖工程所期望的,因为常规的治疗性蛋白质(包括抗体)中不存在末端磷酸残基。这一规则的一些例外情况是用于溶酶体糖基化储积障碍一些特殊化的蛋白质。N-聚糖的磷酸化在真菌中可能是蛋白特异性的。在动物中,甘露糖磷酸化是保

守的溶酶体靶向信号。

[0273] 至今为止,还没有过里氏木霉N-聚糖硫酸盐化的报告。因此,本报告中提及的酸性结构可能是磷酸化的聚糖。

[0274] 如在烧瓶培养的情况下,当里氏木霉被培养在低pH值下时,磷酸化是更常见的,这可能与低pH值应激和菌丝体破坏有关(Stals等,2004,Glycobiology 14:713-724)。在这项研究中,在烧瓶和发酵罐培养的样品之间观察到明显的差异。在摇瓶培养样品中观察到酸性N-聚糖(其全部磷酸化)。发酵罐样品中酸性N-聚糖的量可能在检测限以下,或者因为较高的pH值,可能已没有明显的聚糖磷酸化。酸性N-聚糖与N-聚糖总量的比例不能用本研究中所使用的方法验证,这是由于中性和酸性聚糖物质之间不同的电离效率。

[0275] 为了测定磷酸化水平,用N-聚糖酶从分批和分批补料发酵罐中培养的10 μ g里氏木霉分泌蛋白质释放N-聚糖。使用基于Bradford的方法用BSA作为标准测定蛋白质浓度。在MALDI-TOF分析之前将1pmol的标准分子NeuAcHex4HexNAc2加入酸性N-聚糖样品中。主要糖形(对于发酵罐为Hex7HexNAc2P和对于烧瓶培养为Hex6-8HexNAc2P)的量为0.9pmol/10 μ g分批培养的分泌蛋白质,0.6pmol/10 μ g分批补料培养的分泌蛋白质,和160pmol/10 μ g烧瓶培养的分泌蛋白质(当使得培养物的pH值下降时)。在MALDI-TOF分析之前,用10pmol的标准聚糖Hex2HexNAc4加入中性N-聚糖样品中来测量中性N-聚糖的量。主要糖形Hex5HexNAc2的量是87pmol/10 μ g分批和分批补料培养的分泌蛋白质以及145pmol/10 μ g烧瓶培养的分泌蛋白质。因此,酸性N-聚糖与N-聚糖总量的比例在分批培养中为1%,在分批补料培养中为0.7%和在烧瓶培养中为52%。定量是仅基于使用MALDI-TOF数据的信号强度比较。

[0276] 在酸性部分中N-聚糖也较大。这可能是由于其中磷与一个己糖单元连接到聚糖主链上的磷酸-甘露糖基化反应。在酸性谱中看到一些双磷酸化结构。这种解释与此前公布的关于里氏木霉中发现的磷酸化聚糖的数据一致(Stals等2004,Glycobiology 14:725-737)。在在发酵罐中培养时,酸性N-聚糖的比例是很低的,在检测限以下。

[0277] 里氏木霉培养基的N-聚糖谱没有显示里氏木霉的N-糖组被源自含植物材料培养基的聚糖污染。

[0278] 总之,不同里氏木霉菌株的N-聚糖分析表明,菌株M44、M109、M131、M132和M124中的主要糖形是Man5或Man α 3[Man α 3(Man α 6)Man α 6]Man β 4GlcNAc β 4GlcNAc。考虑到产生Man5的菌株中可能存在葡萄糖,包括H8N2作为次要组分。两个菌株(M109和M131)包含的H8N2的量比H7N2大。H8N2的富集可能表明部分葡萄糖苷酶缺陷。

[0279] M44菌株几乎不含磷酸化聚糖。来自菌株M131、M109、M132和M124的样品中观察到在中性聚糖组分中看到的泄漏酸性聚糖作为在m/z 1521和m/z 1683处的信号,这表明较高的磷酸化水平和存在潜在的磷酸二酯结构。

[0280] 本研究的目的是要找到具有最大的Man5Gn2结构产量和低水平的酸性(磷酸化)的N-聚糖产量的菌株。最好的菌株在pH控制的摇瓶培养条件下有80%以上的Man5。最好的菌株也具有降低的双磷酸化的聚糖产量和/或较大的磷酸化结构(见表3)。

[0281] 实施例2-*alg3*缺陷木霉菌株的产生

[0282] 载体构建和菌株产生

[0283] 在里氏木霉的基因组序列中鉴定了编码ALG3甘露糖基转移酶的基因。设计了破坏构建体(disruption construct)以在*alg3*基因的1000bp的5'和3'侧翼区片段之间插入乙

酰胺酶选择标记。通过PCR扩增侧翼区片段,并通过在酿酒酵母中的同源重组克隆制备构建体。通过消化从其骨架载体释放破坏盒(disruption cassette)和转化到里氏木霉菌株M124中。在乙酰胺酶培养基上选择转化体,并通过PCR用构建体5'侧翼区片段外的正向引物和AmdS选择标记内的反向引物筛选。

[0284] 转化体的筛选

[0285] 62个筛选的转化体中的58个给出了构建体整合到alg3基因座预期的大小的PCR产物。通过单孢子培养将9个PCR阳性转化体纯化为单核克隆,并从中制得孢子悬浮液。通过Southern杂交对这9个克隆进行破坏盒的正确整合的分析。在标准的杂交条件下来自亲本菌株和来自9个克隆的EcoRI-消化的基因组DNA与alg3探针杂交。探针与来自亲本菌株的DNA杂交但不与来自任何克隆的DNA杂交表示成功删除了alg3(图8)。

[0286] 通过用AmdS探针的Southern杂交进行进一步分析。AmdS基因被包含在删除盒中且预计可在来自转化体的DNA中但不在来自亲本菌株的DNA中检测到。用EcoRI+PvuI(E+P)和KpnI+NheI(K+N)消化亲本菌株M124和9个转化体的基因组DNA。携带alg3-AmdS删除盒的NotI消化的质粒被用作阳性对照。探针识别来自阳性对照的预期约2.7kb的片段(AmdS),但不与亲本菌株杂交。全部转化体给出了预期的信号(对于E+P为1.6+2.8kb和对于K+N为1.7+3.4kb,图9B中用箭头表示),表明正确整合了删除盒。克隆11A和15A也表现出一些额外片段的杂交,表明删除盒非特异性地整合到基因组(图9B)。

[0287] N-聚糖分析学

[0288] 对五个不同Alg3敲除菌株(4A、5A、6A、10A和16A)和亲本菌株M124的摇瓶培养物进行N-聚糖分析。从时间点3、5、7和9天采集样品。所有培养物重复地进行生长。

[0289] 使用基于Bradford的分析针对BSA标准曲线测定来自所有时间点的随机选择的基因敲除菌株(4A)的分泌蛋白质的蛋白质浓度。第5天检测到最高的蛋白浓度。因此,第5天的样品用于所有五个基因敲除菌株的N-聚糖分析。一式三份地分析所有样品(包括重复的培养物)。将10μg用于N-聚糖分析。通过MALDI-TOF对中性 and 酸性N-聚糖进行分析。

[0290] 在亲本菌株M124中的主要糖形是Man5Gn2。在所有的Alg3敲除菌株中的主要糖形是Man3(图10)。在亲本菌株M124中没有发现Man3。在不同的Alg3敲除菌株中Man3的量的范围在摇瓶培养物中为49.7%至55.2%之间(使得pH值下降)。亲本菌株中Hex6Gn2增加。作为观察到的中性N-聚糖信号的百分比的信号强度列于下面的表5中。

[0291] 表5:Alg3敲除菌株的中性N-聚糖含量。

[0292]

		菌株											
		亲本 M124		4A		5A		6A		10A		16A	
组成	m/z	平均	STDEV	平均	STDEV	平均	STDEV	平均	STDEV	平均	STDEV	平均	STDEV
Hex3HexNAc2	983.31	0.0	0.0	53.6	0.2	55.2	4.2	49.7	0.3	53.3	0.9	53.4	0.9
Hex4HexNAc2	1095.37	1.6	0.1	2.7	0.0	2.9	0.7	3.4	0.1	3.2	0.4	3.4	0.4
Hex5HexNAc2	1257.42	70.2	3.3	8.5	0.2	7.3	1.1	10.4	0.5	8.6	0.9	9.7	0.9
Hex6HexNAc2	1419.48	7.9	1.1	35.0	0.3	34.4	1.9	36.1	0.6	34.9	0.5	33.2	0.7
Hex7HexNAc2	1581.53	7.8	0.6	0.3	0.4	0.3	0.4	0.3	0.4	0.0	0.0	0.3	0.4
Hex8HexNAc2	1743.58	5.9	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex9HexNAc2	1905.63	6.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex10HexNAc2	2067.69	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[0293] 通过MALDI MS分析无法观察到各糖形的不同异构体的存在,因此进一步进行串联质谱研究。首先,研究Man3和Hex5Gn2结构。对于Man3,探询Man3结构是分支的还是线性的。

对于此分析,将含有这两种结构的样品全甲基化和按照制造商的说明使用Bruker Ultraflex III TOF/TOF仪通过质谱片断化进行分析(图11和12)。

[0294] 接下来,确定Hex6Gn2结构的非还原性末端上的己糖单元是甘露糖还是葡萄糖。用 α -甘露糖苷酶消化对所有敲除菌株和亲本菌株进行(图13)。使用切割 α -甘露糖和保留主链的 β -甘露糖不发生变化的杰克豆甘露糖苷酶。预期所得到的结构是Man1Gn2。

[0295] 由于MALDI的低分子量范围效应,Man1GlcNAc2聚糖的相对强度可能有所降低,这解释了Hex6相对量的小幅增加。 α -甘露糖苷酶消化后,Man3和Man4糖形消失了。没有观察到Man2结构。然而,Hex6(M/Z 1419)不被消化(表6),这表明在该结构的非还原性末端上有葡萄糖单元。也存在一些不可消化的Hex5,可能是通过去除Hex6的空间位阻Man6分支的弱反应产生的。

[0296] 表6: α -甘露糖苷酶消化前(天然)和消化后Alg3敲除菌株4A的中性N-聚糖。

		4A	
		天然的	α -甘露糖苷
组成	m/z	平均	%
Hex1HexNAc2	609.21	0.0	53.2
Hex2HexNAc2	771.26	0.0	0.0
Hex3HexNAc2	933.31	47.5	0.0
[0297] Hex4HexNAc2	1095.37	3.8	0.0
Hex5HexNAc2	1257.42	11.7	5.0
Hex6HexNAc2	1419.48	36.8	41.0
Hex7HexNAc2	1581.53	0.2	0.8
Hex8HexNAc2	1743.58	0.0	0.0
Hex9HexNAc2	1905.63	0.0	0.0
Hex10HexNAc2	2067.69	0.0	0.0

[0298] 对于Alg3敲除菌株中发现的不同结构的最终分析,对Alg3基因敲除菌株4A进行大规模的PNGase F消化。用HPLC纯化两种主要的聚糖(图14),并通过NMR进行分析(图15)。

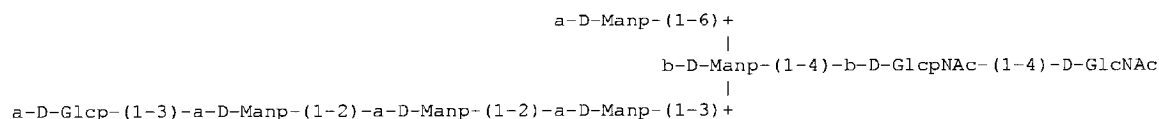
[0299] 基于图15中给出的数据,Hex3HexNAc2物质被明确鉴定为Man α 1-3(Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAc。Man α 3和Man α 6 H-1单元分别在5.105和4.914ppm处共振。在4.245ppm处观察到Man β 4H-2单元。这个信号是非常特征性的,因为相邻的Man α 3-OH取代。在2.038和2.075处观察到核心GlcNAc单元的N-乙酰基-CH₃信号。这些值与Sugabase数据库(www.boc.chem.uu.nl/sugabase/sugabase.html)中对于该戊多糖所报道的完全一致。此外,在相同的实验条件下,对于商业化生产的Man α 1-3(Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAc(Glycoseparations, Inc.)测定质子-NMR谱,并得到几乎相同的化学位移。

[0300] 图15B中显示了Hex6HexNAc2组分的NMR谱。该数据暗示此组分代表八糖Glc α 1-3Man α 1-2Man α 1-2Man α 1-3(Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAc。显示典型的 α Glc 2.4Hz耦合的5.255信号证实了葡萄糖单元的存在。由于平伏的H-2构型,所有Man信号通常显示<1赫兹耦合。与Sugabase数据相比观察到小的差异(表7),这可能是由于本NMR测量中使用的

不同温度(40℃对26℃)。

[0301] 表7:Glcα1-3Manα1-2Manα1-2Manα1-3(Manα1-6)Manβ1-4GlcNAcβ1-4GlcNAc的公开NMR数据。数据从Sugabase获得(在**boc.chem.uu.nl/sugabase/sugabase**上找到)。

[0302]



残基	连接	质子	PPM	J	Hz
D-GlcNAc		H-1a	5.189		
		H-1b	4.694		
		H-2a	3.867		
		H-2b	3.692		
		NAc	2.038		
b-D-GlcpNAc	4	H-1	4.606		
		H-2	3.792		
		NAc	2.077		
b-D-Manp	4, 4	H-1	4.773		
		H-2	4.237		
a-D-Manp	6, 4, 4	H-1	4.913		
		H-2	3.964		
a-D-Manp	3, 4, 4	H-1	5.346		
		H-2	4.080		
a-D-Manp	2, 3, 4, 4	H-1	5.304		
		H-2	4.103		
a-D-Manp	2, 2, 3, 4, 4	H-1	5.038		
		H-2	4.224		
a-D-Glcp	3, 2, 2, 3, 4, 4	H-1	5.247		
		H-2	3.544		

[0303] 最后,在不同时间点(第3天、第5天、第7天和第9天)分析随机选择的敲除菌株4A的N-聚糖谱。摇瓶培养的pH在开始时间点为4.8和在结束时间点为2.6。对重复培养的每个时间点的一式三份样品进行了分析。在全部重复中观察到,由于pH值的降低,Man3Gn2信号的相对量随着生长时间的变化下降。然而,Hex6Gn2信号的量随着生长时间的变化增高(表8)。

[0304] 表8:来自Alg3敲除菌株4A的观察的中性聚糖信号的信号强度的百分比。分析了四个不同时间点(第3天、第5天、第7天和第9天)的重复培养物(3A和4A)。

[0305] Alg3敲除菌株4A(烧瓶3A)

[0306]

组成	m/z	第3天, 3A		第5天, 3A		第7天, 3A		第9天, 3A	
		平均	stdev	平均	stdev	平均	stdev	平均	stdev
Hex3HexNAc	730.24	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex2HexNAc2	771.26	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Hex3HexNAc2	933.31	61.7	3.7	61.3	0.8	61.1	1.9	52.7	7.7
Hex4HexNAc2	1095.37	2.6	0.2	2.5	0.1	2.1	0.4	3.7	1.0
Hex5HexNAc2	1257.42	4.3	0.6	6.5	0.4	5.7	0.6	6.4	1.0
Hex6HexNAc2	1419.48	31.4	3.5	29.8	0.4	31.1	1.6	37.2	5.7

[0307] Alg3敲除菌株4A(烧瓶4A)

[0308]

组成	m/z	第3天, 4A		第5天, 4A		第7天, 4A		第9天, 4A	
		平均	stdev	平均	stdev	平均	stdev	平均	stdev
Hex3HexNAc	730.24	0.0	0.0	0.0	0.0	0.0	0.0	0.7	1.2
Hex2HexNAc2	771.26	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.5
Hex3HexNAc2	933.31	61.7	3.2	58.6	1.1	55.6	1.9	54.8	5.9
Hex4HexNAc2	1095.37	3.4	1.0	2.6	0.2	3.1	0.2	2.6	0.5
Hex5HexNAc2	1257.42	5.2	1.5	6.7	0.4	7.1	0.4	7.6	3.7
Hex6HexNAc2	1419.48	29.7	0.9	32.1	0.8	34.3	1.5	34.0	3.6

[0309] 注意到这两个分析(表4和7)之间关于在克隆4A(第5天)中Man3的百分比的差异。这种差异可能是由于分析方法中的差异。冻融循环后观察到某些异源培养基蛋白质制剂的不稳定性,这可能是由于聚糖和/或蛋白质的降解,从而导致较大的聚糖的量减少。表5中的数据的生产包括额外的冻融循环。

[0310] 还通过MALDI分析了酸性N-聚糖部分(图16)。在亲本菌株M124中不同酸性化合物的丰度与所有Alg3敲除菌株(其中酸性部分似乎是非常相似的)都不同。

[0311] 在亲本菌株中的三种主要聚糖是H6N2P1、H7N2P1和H8N2P1。在Alg3敲除菌株中,聚糖大小移向较小的聚糖:H5N2P1、H6N2P1和H4N2P1。此外,亲本菌株中双磷酸化聚糖更丰富。这可能是由于缺少将磷酸化的甘露糖连接到聚糖上的特定的酶的合适底物。磷酸化的甘露糖可以由其他的甘露糖残基进一步延长。在发酵条件下产生的亲本M124菌株的聚糖中实质上不存在磷酸化。

[0312] 发酵罐和摇瓶生长样品的比较

[0313] 将一个Alg3敲除菌株(转化体4A)以分批发酵在乳糖和酒糟提取物培养基上生长。该培养基是60克/升乳糖,加20克/升酒糟提取物,接种后体积为7升(发酵罐运行bio01616)。其他培养基组分为磷酸二氢钾和(NH₄)₂SO₄。培养物的pH值控制在5.5和5.8之间。在运行的过程中提取生物质和培养上清液样品,并储存于-20℃。还收集菌丝体样品用于可能的RNA分析并立即冷冻在液氮中和转移至-70℃。分析从这些发酵的全过程中收集的样品的N-聚糖组成。对于发酵罐运行(bio01616)和转化体4A的摇瓶培养的5天时间点样品(图17和图18)进行N-聚糖分析。在摇瓶培养中的主要信号是Man3(59%)。在发酵罐培养中,主要信号是Man3(85%),且Hex6比例下降到8%。

[0314] 结论

[0315] Alg3敲除成功地产生50%或更多的预期的Man3糖形。通过碎片质谱和核磁共振谱验证了Man α 3(Man α 6)Man β -的所需分支结构。

[0316] Alg3敲除的其他产物包括Man4(含甘露糖的次要产物)、Hex5(如图13所示的Hex6的降解产物)和Hex6(这是第二大组分)。Hex6组分由甘露糖苷酶抗性和特定的核磁共振信号(包括G1ca3Man末端)表征为包含末端G1c。据认为通过用于减少末端G1c的方法可以进一步优化聚糖结构,末端G1c可能造成葡萄糖苷酶II对分子Man α 6臂上缺乏甘露糖的聚糖的效率不佳。对发酵条件的进一步优化可能降低末端G1c的量。

[0317] 这一数据表明,与较早的在曲霉属(Kainz等,Appl Environ Microb.2008 1076-

86)和毕赤酵母(Davidson等Glycobiology 2004,399-407)中的Alg3敲除的数据相比,里氏木霉中的Alg3基因敲除产生更好的糖基化结果。在Kainz等和Davidson等的工作中报道了相似或更高的Hex6相应产物水平。这些研究还报告了与 α 2-甘露糖,0CH1产物和较大的尺寸,及由毕赤酵母中产生的细胞类型特异性聚糖相关的额外问题。总之,里氏木霉Alg3敲除的N-聚糖分析表明敲除菌株中的主要糖形是Man3Gn2,哺乳动物型N-聚糖的有效产生的理想起点。

[0318] 实施例3-单个GnTI和GnTII酶的纯化和活性

[0319] 为了研究其受体特异性和活性,人GnTI和GnTII(N-乙酰葡萄糖胺基转移酶I和N-乙酰葡萄糖胺基转移酶II)在毕赤酵母中表达为水溶性的分泌蛋白质。

[0320] 用于在毕赤酵母中生产的GnTI构建体的产生

[0321] 获得人GnTI(P26572)序列的全长序列并亚克隆到里氏木霉过表达载体中。将编码人GnTI可溶性部分的蛋白质编码序列(CDS)克隆到pBLARG-SX表达载体以便在毕赤酵母中产生分泌形式的蛋白质用于酶学研究。在克隆过程中,将His标签编码序列添加到框的5'端以在截短的蛋白质的N-末端得到标记。通过测序分析验证了该序列。产生的载体pTTg5线性化并通过电穿孔转化至毕赤酵母GY190细胞中以产生菌株GY4。挑取Arg⁺转化体,并通过PCR筛选。测试含有整合的质粒的GY4克隆的蛋白质表达。

[0322] 可溶性GnTI的表达和纯化

[0323] 首先将表达可溶性GnTI的毕赤酵母菌株GY4在+30℃下在BMGY培养基(1%酵母提取物,2%蛋白胨,100mM磷酸钾pH 6.0,1.34%酵母氮基质,4×10⁻⁵%生物素,1%甘油)中振荡过夜进行生长到OD₆₀₀2-6。然后通过离心收获细胞,并在BMMY培养基(与BMGY一样,但用0.5%的甲醇代替1%甘油)中重悬浮到OD₆₀₀为1。将培养物放置在带挡板的烧瓶中,并返回到在+16℃下的振摇孵育器中。每24小时加入100%甲醇至终浓度为0.5%以保持诱导。诱导0、24、48和72小时后获取1毫升的表达培养物样品,并存储细胞沉淀和上清液用于分析。3天的诱导之后,通过离心收获整个培养物的细胞,并收集上清液用于进一步纯化GnTI。

[0324] 制备粗GnTI样品用于活性测定

[0325] 处理含有可溶性His标记的GnTI的毕赤酵母细胞培养物用于通过浓度和缓冲交换的活性测定。简单地说,在用MeOH诱导后的第3天,通过将细胞沉淀在50ml的Falcon管(Eppendorf 5810R,3220rcf,在4℃下5分钟)并收集上清液来收获摇瓶培养物的40毫升毕赤酵母上清液。然后用Millipore Amicon Ultracel 30K浓缩器通过顺序离心(Eppendorf 5810R或相当的,3220rcf,在4℃下10分钟)将上清液浓缩至<2.5毫升。用100mM MES,pH 6.1将浓缩物的体积调切至2.5毫升。浓缩物用PD-10凝胶过滤柱(GE Healthcare 17-0851-01)进行缓冲液交换。首先用100mM的MES,pH 6.1平衡柱,然后加入样品(2.5毫升),丢弃穿流液,收集用2.25毫升MES缓冲液进行的洗脱。最后,用Millipore Biomax 30K浓缩器(Eppendorf 5417,12 000rcf,5分钟4℃下)将500μl的洗脱液浓缩到100μl,并直接用于活性测定。

[0326] GnTI酶的活性测定

[0327] Man α 1-6(Man α 1-3)Man β 1-4GlcNAc(Man₃Gn)在GnTI活性测定中被用作GnTI的受体。通过在室温下孵育总体积为10μl的含有0.1mM的受体Man₃GlcNAc、20mM的UDP-GlcNAc、50mM的GlcNAc、100mM MnCl₂、0.5%BSA和8μl GnTI在100mM MES,pH 6.1中的反应混合物过

夜进行GnTI反应。通过在100℃下孵育反应物5分钟使反应停止。

[0328] 在GnTI活性测定的同时,控制粗酶制剂中可能的HexNAc'ase活性。GlcNAc β 1-2Man α 1-6(GlcNAc β 1-2Man α 1-3)Man β 1-4GlcNAc β 1-4GlcNAc-Asn(=Gn₂Man₃Gn₂-Asn)被用作HexNAc'ase的底物。GnTI以类似的方式进行反应,除了加入100pmol的Gn₂Man₃Gn₂-Asn而不是Man₃Gn和UDP-GlcNAc。没有检测到HexNAc'ase活性。

[0329] 纯化反应混合物用于通过顺序Hypersep C₁₈(100mg,Thermo Scientific,cat no: 60300-428)和在HyperSep 96孔真空管(Thermo Scientific)上的Hypercarb(10mg/96孔板/1PKG,cat no 60302-606)层析的MALDI分析。用300 μ l的EtOH和300 μ l的MQ水制备Hypersep C₁₈,然后把收集板放在下面,并上样和用150 μ l MQ水洗脱。用300 μ l的MeOH和300 μ l的MQ水制备Hypercarb。加载来自Hypersep C₁₈的洗脱液,用150 μ l 0.5M的NH₄Ac移除盐,并用2 $\times$ 300 μ l的MQ水洗脱孔。用150 μ l 25%的ACN洗脱GnTI反应产物,和用25%的ACN和0.05%的TFA洗脱HexNAc'ase反应产物。在Speedvac中干燥样品。

[0330] 用Bruker Ultraflex TOF/TOF仪(Bruker Daltonics,Germany)进行基质辅助激光解析电离飞行时间(MALDI-TOF)质谱(MS)分析。作为[M+Na]⁺离子的正离子反射模式中检测受体糖和产物。Hex₃HexNAc₁和Hex₃HexNAc₂的[M+Na]⁺信号的计算m/z值分别为733.238和933.318。从Hex₃HexNAc₁和Hex₃HexNAc₂的对应信号计算受体和产物的百分比(图19)。

[0331] 用于毕赤酵母中生产的GnTII构建体的产生

[0332] 用分别含有KpnI和EcoRI限制性位点的引物GP3和GP13PCR-扩增编码人GnTII的核苷酸序列。将EcoRI/KpnI消化的PCR片段接合到类似消化的pBLARG-SX克隆载体中。验证序列后,将最终的构建体转化到毕赤酵母菌株GS190中以产生菌株GY22。通过PCR筛选阳性酵母转化体。研究了在16℃和30℃下两个克隆(其中只有一个在图20中示出)在甲醇诱导AOX1启动子控制下的GnTII表达。

[0333] 可溶性GnTII的表达

[0334] 根据Western印迹分析(图20),毕赤酵母菌株GY22产生可溶性重组GnTII酶。GnTII有49049.0Da的计算分子量和两个预测的N-糖基化位点。重组GnTII在16℃下分泌到培养基中(泳道9)。当生长在+30℃下时,重组GnTII滞留在细胞内(泳道4)。

[0335] 可溶性GnTII的活性测定

[0336] 按如上对于GnTI所述,处理含有可溶性His标记GnTII的毕赤酵母细胞培养物用于活性测定。对细胞培养物进行离心,收集上清液并浓缩,缓冲液交换至100mM MES,pH 6.1中,并在活性测试之前将所得样品进行进一步浓缩。

[0337] 活性测定与GnTI类似地进行。使用GnMan₃Gn作为GnTII受体。

[0338] 在存在0.1mM的受体GnMan₃Gn,20mM UDP-GlcNAc,50mMGlcNAc,100mM MnCl₂,0.5%的BSA和在100mM的MES,pH 6.1中的GnTII的情况下进行GnTII反应。如以上对于GnTI所述的,通过顺序Hypersep C₁₈和在真空歧管上在96-孔板上的Hypercarb色谱进行用于MALDI-TOF MS分析的反应混合物的纯化。

[0339] 用Bruker Ultraflex TOF/TOF仪(Bruker Daltonics,Germany)进行MALDI-TOF MS。作为[M+Na]⁺离子的正离子反射模式中检测受体糖和产物。从它们的信号强度计算反应结束时产物和受体的比率(GnMan₃Gn受体和具有一个GlcNAc添加的产物的[M+Na]⁺信号的计算m/z值分别为933.318和1136.397)。

[0340] 重复产生GnTII的毕赤酵母的培养,并制备来自上清液的GnTII浓缩物(60倍)和根据上面描述的方法测定其活性。在2.5小时、5小时和过夜的时间点样品的MALDI谱分别表明80%、83%和82%的受体转化为产物。在2.5小时内达到接近最大反应。

[0341] 另外,制备粗GnTII样品,并如上面对于粗GnTI样品所述的进行活性分析。将反应混合物孵育过夜,纯化,并进行MALDI分析。MALDI谱显示GnTII活性(图21)。粗GnTII样品中没有检测到HexNAc'ase活性。

[0342] 用于合成在上述GnTII活性测定中使用的GnTII受体的方法如下所述。从如上所述的毕赤酵母培养基制备GnTI样品。此GnTI样品显示出高GnTI活性,且因此它可以用于将约40nmol的Man3Gn转化成GnMan3Gn。在0.5mM的Man3Gn、20mM的UDP-GlcNAc、50mM的GlcNAc、100mM的MnCl₂、0.5%的BSA和GnTI样品存在下进行反应。在室温下将反应混合物孵育3天。用Hypercarb色谱和MALDI分析将约1%的样品进行纯化。根据MALDI谱,GnTI反应将几乎所有Man3Gn受体转化为GnMan3Gn产物。只有2.8%的受体没有被转化。

[0343] 实施例4-GnTI/GnTII融合蛋白质

[0344] GnTI/GnTII表达构建体的产生

[0345] 通过用5'-末端的65-mer融合引物(其包含同框融合位点(来自GnTI的短序列,含有天然发生的A1eI限制性位点,移除终止密码子并与GnTII序列重叠))和含有SpeI或NdeI限制性位点的与GnTII同源的3'-端引物扩增1313bp的GnTII片段来构建重组GnTI/II融合蛋白质。这一融合位点允许融合片段直接克隆到具有cbh1启动子(用A1eI/NdeI克隆)控制下的野生型GnTI或具有在gpd启动子(用A1eI/SpeI克隆)控制下的野生型GnTI的里氏木霉过表达载体中。使用高保真Phusion聚合酶(Finnzymes)和标准的扩增和克隆程序。通过直接的表达载体测序验证该序列。将得到的载体用于表达作为里氏木霉中的跨膜蛋白的融合体。

[0346] 为了获得融合蛋白质的功能性的更多信息,融合GnTI/II蛋白质也在毕赤酵母中表达为可溶性蛋白质。将编码蛋白质的可溶性部分的GnTI/II融合体的CDS克隆到pBLARG-SX表达载体中以产生用于酶学研究的蛋白质。在克隆过程中,将His标签编码序列添加到框的5'-端以在截短的蛋白质的N-末端获得标记。通过测序分析验证了该序列。将得到的载体线性化,并通过电穿孔转化至毕赤酵母菌株GS 190中以产生菌株GY6。挑取Arg⁺转化体,并通过PCR筛选。测试含有整合的质粒的毕赤酵母克隆的蛋白质表达。

[0347] 毕赤酵母中产生的可溶性GnTI/II的纯化

[0348] 毕赤酵母中的表达和纯化过程按以上对于重组GnTI蛋白所述的进行。

[0349] GnTI/II融合蛋白质的酶活性测试

[0350] 如上对GnTI分析所述的使用Man3Gn寡糖作为受体和UDP-GlcNAc作为供体进行活性分析。通过MALDI-TOF质谱分析反应的产物。对于GnTI/GnTII融合蛋白质只观察到GnTI活性(图22)。

[0351] 通过随机整合用GnTI/GnTII构建体转化里氏木霉

[0352] 利用随机整合将带有gpdA启动子的嵌合人GnTI/GnTII质粒共转化至里氏木霉M124菌株中。通过含有乙酰胺酶标记基因的质粒的共转化实现选择。将20个PCR阳性转化体纯化至单核克隆并在摇瓶培养中生长以用于聚糖分析。将所有的转化体和亲本菌株M124培养在补充有4%的乳糖和2%的酒糟提取物的木霉基本培养基(TrMM),pH 4.8中。在第3天、

第5天和第7天收集上清液和菌丝体样品,并冷冻贮存直至分析。此外,作为对照,通过随机整合用GnTII构建体转化里氏木霉。

[0353] 通过随机整合获得的里氏木霉GnTI/GnTII菌株的聚糖分析

[0354] 对三个不同时间点(第3天、第5天和第7天)来自里氏木霉菌株M124的GnTI/GnTII转化体的20个不同克隆的样品进行了分析。分析来自两个亲本M124菌株的样品用于对照。在96孔板中一式三份对5 μ g的上清液蛋白质进行N-聚糖酶反应而无SDS变性。通过基于Bradford的分析(Bio-Rad Quick Start Bradford蛋白质分析)使用BSA作为标准测定上清液的蛋白质浓度。通过MALDI-TOF MS对中性 and 酸性N-聚糖进行分析。使用GnTI/GnTII构建体在任何时间点在任何克隆中以及具有gpdA启动子的GnTI转化体的克隆中没有检出G0产物。

[0355] 通过靶向整合用GnTI/GnTII构建体转化里氏木霉

[0356] 将嵌合的GnTI/GnTII序列亚克隆到pTTv38骨架中,pTTv38为包含乙酰胺酶标记基因和用于alg3基因座整合的5'-和3'-侧翼序列位点的载体。将该载体作为消化的片段转化到里氏木霉M124菌株中。通过这一转化,检测到18个PCR阳性转化体,其产生表明正确整合到alg3基因座的PCR片段。在单孢纯化步骤后在摇瓶中培养这些转化体,并如下文所述进行分析。

[0357] 通过靶向于alg3基因座获得的里氏木霉GnTI/GnTII菌株的聚糖分析

[0358] 在三个不同的时间点(第3天、第5天和第7天)获得 Δ alg3里氏木霉GnTI/GnTII转化体的10个不同克隆的上清液样品。在摇瓶中用两种不同的培养基组合培养克隆。TrMM,pH 5.5,具有2%的酒糟提取物,4%的乳糖和K-邻苯二甲酸盐缓冲用于所有克隆,且同时地,TrMM,pH 5.5,具有2%的酒糟提取物,4%的乳糖,1%酪蛋白氨基酸和K-邻苯二甲酸盐缓冲用于其中5个克隆。继续培养7天:在+28 $^{\circ}$ C下5天和+24 $^{\circ}$ C下第6和7天。

[0359] 在96孔板中一式三份地对5 μ g上清液蛋白质进行N-聚糖分析。从第3天、第5天和第7天对样品进行分析。使用BSA作为标准,通过基于布Bradford的分析(Bio-Rad Quick Start Bradford protein Assay)测定上清液的蛋白质浓度。通过MALDI-TOF MS对中性 and 酸性N-聚糖进行分析。

[0360] 在每一个克隆中发现可检测量的糖形G0。克隆201A包含最多的1.2%的Gn2Man3(图23和表9)。此外,在这个特定的克隆中Hex6量是最低的。具有1%酪蛋白氨基酸的第二种培养基没有给出任何额外的G0/G1cNAc β 2Man α 3(G1cNAc β 2Man α 6)Man β 4G1cNAc β 4G1cNAc β 的产生。第3天和第7天的样品的结果与第5天样品基本上是相同的。

[0361] 表9:观察到的里氏木霉GnTI/II转化体(GnTI/II整合到alg3基因座)的分泌蛋白质的N-聚糖的信号强度百分比。在其名称中带有字母A的克隆在培养基A)中培养,带有B的在培养基B)中培养,培养基B)相对于培养基A)有额外的1%的酪蛋白氨基酸。

[0362]

		克隆 201A, 第 5 天						克隆 202A, 第 5 天						克隆 208A, 第 5 天						克隆 210A, 第 5 天					
组成	m/z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX				
Man2	771.3	0.6	0.5	86.8	0.0	1.0	0.4	0.7	173.2	0.0	1.1	0.6	0.6	92.3	0.0	1.1	0.1	0.2	173.2	0.0	0.4				
Man3	933.3	47.9	14.5	30.2	39.0	64.6	41.3	0.2	0.4	41.1	41.5	38.2	1.1	2.8	37.0	38.9	38.2	1.1	3.0	37.5	39.5				
Man4	1095.4	7.9	2.9	36.5	5.9	11.3	6.4	0.6	8.7	6.0	7.0	5.3	0.2	4.0	5.0	5.5	6.0	0.4	6.6	5.5	6.2				
GnMan3	1136.4	1.4	0.7	46.9	1.0	2.2	1.1	0.3	23.5	0.8	1.3	1.0	0.2	17.0	0.9	1.2	1.1	0.1	8.9	1.0	1.2				
Man5	1257.4	10.5	2.5	23.5	8.7	13.3	8.6	0.8	9.7	7.7	9.4	8.2	0.3	4.0	7.8	8.5	8.9	0.3	3.7	8.6	9.3				
Gn2Man3	1339.5	1.2	0.8	69.1	0.6	2.2	0.6	0.1	21.0	0.5	0.8	0.6	0.1	21.5	0.5	0.7	0.6	0.1	17.5	0.6	0.8				
Hex6	1419.5	27.3	23.7	86.7	0.0	42.0	40.5	0.6	1.5	39.9	41.1	44.7	0.7	1.6	43.9	45.2	43.2	0.7	1.6	42.7	44.0				
Hex7	1581.5	2.9	3.0	103.3	1.1	6.4	1.0	0.1	11.0	1.0	1.2	1.1	0.1	11.7	1.0	1.2	1.2	0.0	3.7	1.2	1.2				
Hex8	1743.6	0.1	0.2	173.2	0.0	0.4	0.2	0.3	173.2	0.0	0.5	0.3	0.2	87.0	0.0	0.4	0.6	0.3	57.0	0.3	1.0				
		克隆 212A, 第 5 天						克隆 213A, 第 5 天						克隆 215A, 第 5 天						克隆 216A, 第 5 天					
组成	m/z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX				
Man2	771.3	0.4	0.4	86.8	0.0	0.7	0.6	0.6	94.4	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0				
Man3	933.3	45.6	1.3	2.8	44.2	46.8	40.0	2.8	7.0	37.3	42.9	43.4	1.9	4.4	41.3	45.1	42.6	2.0	4.6	40.5	44.4				
Man4	1095.4	5.6	0.3	5.1	5.4	5.9	6.5	0.6	8.8	6.0	7.1	6.3	0.5	8.5	5.7	6.8	6.1	0.6	10.3	5.4	6.7				
GnMan3	1136.4	0.9	0.2	22.4	0.7	1.1	0.9	0.1	8.5	0.8	1.0	1.1	0.1	6.9	1.0	1.2	1.1	0.2	14.1	0.9	1.2				
Man5	1257.4	7.2	0.5	7.0	6.8	7.7	9.5	0.4	3.8	9.1	9.8	8.5	0.4	4.2	8.2	8.9	7.7	0.6	8.4	7.0	8.3				
Gn2Man3	1339.5	0.5	0.1	11.9	0.5	0.6	0.6	0.1	18.3	0.5	0.7	0.7	0.2	29.3	0.6	1.0	0.7	0.2	26.4	0.5	0.9				
Hex6	1419.5	38.6	1.2	3.0	37.4	39.7	40.7	2.5	6.1	38.2	43.2	38.5	1.8	4.6	37.4	40.5	40.5	1.7	4.2	39.0	42.4				
Hex7	1581.5	0.8	0.0	4.1	0.8	0.8	1.0	0.1	10.8	0.9	1.2	1.1	0.1	4.5	1.1	1.2	1.0	0.1	6.4	0.9	1.0				
Hex8	1743.6	0.4	0.1	34.8	0.3	0.5	0.1	0.2	173.2	0.0	0.3	0.4	0.3	88.5	0.0	0.6	0.4	0.3	87.6	0.0	0.6				
		克隆 217A, 第 5 天						克隆 219A, 第 5 天						克隆 201B, 第 5 天						克隆 202B, 第 5 天					
组成	m/z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX				
Man2	771.3	0.9	0.0	5.0	0.9	1.0	0.5	0.4	96.7	0.0	0.9	0.4	0.7	173.2	0.0	1.1	0.6	1.1	173.2	0.0	1.8				
Man3	933.3	54.1	1.1	1.9	53.0	55.0	44.0	1.8	4.1	42.4	45.9	46.9	0.2	0.5	46.6	47.1	40.6	1.7	4.3	38.6	41.8				
Man4	1095.4	5.2	0.3	6.5	4.9	5.5	5.7	0.1	1.5	5.6	5.8	6.9	0.9	12.7	6.0	7.8	8.5	0.9	10.0	7.7	9.4				
GnMan3	1136.4	0.9	0.2	17.4	0.7	1.0	1.0	0.2	16.6	0.9	1.2	1.2	0.4	32.1	0.9	1.6	1.3	0.4	0.0	0.9	1.8				
Man5	1257.4	5.8	0.1	2.6	5.6	5.9	8.0	1.2	15.6	6.7	9.2	8.1	0.5	5.7	7.8	8.6	10.0	0.6	6.2	9.5	10.6				
Gn2Man3	1339.5	0.7	0.1	14.7	0.6	0.7	0.9	0.1	14.2	0.8	1.0	0.8	0.1	7.1	0.8	0.9	0.7	0.5	70.8	0.3	1.3				
Hex6	1419.5	31.5	1.5	4.7	30.5	33.3	38.5	1.1	2.8	37.3	39.2	34.2	0.7	2.1	33.8	35.1	37.5	1.1	2.8	36.7	38.7				
Hex7	1581.5	0.9	0.1	12.9	0.8	1.0	1.0	0.2	15.4	0.8	1.1	1.1	0.1	5.2	1.0	1.2	0.8	0.7	86.9	0.0	1.2				
Hex8	1743.6	0.0	0.0	0.0	0.0	0.0	0.4	0.1	17.9	0.3	0.5	0.4	0.3	90.7	0.0	0.7	0.0	0.0	0.0	0.0	0.0				
		克隆 208B, 第 5 天						克隆 210B, 第 5 天						克隆 219B, 第 5 天											
组成	m/z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX									
Man2	771.3	0.9	0.8	87.1	0.0	1.5	0.8	0.7	86.7	0.0	1.2	1.2	0.1	10.3	1.0	1.3									
Man3	933.3	48.4	1.2	2.4	47.3	49.6	39.6	1.1	2.7	38.6	40.8	34.9	1.8	5.2	33.2	36.8									
Man4	1095.4	7.2	0.2	2.2	7.0	7.3	7.9	0.6	8.0	7.3	8.5	8.1	0.3	4.1	7.8	8.4									
GnMan3	1136.4	0.6	0.6	92.1	0.0	1.1	1.0	0.1	12.7	0.9	1.1	1.1	0.1	12.1	1.0	1.2									
Man5	1257.4	8.7	0.7	7.6	7.9	9.1	9.6	0.2	2.0	9.4	9.8	11.3	0.8	7.5	10.7	12.3									
Gn2Man3	1339.5	0.4	0.2	44.3	0.2	0.6	0.6	0.2	32.4	0.4	0.8	0.6	0.1	13.9	0.5	0.6									
Hex6	1419.5	32.4	0.4	1.4	32.1	32.9	38.5	0.3	0.8	38.3	38.9	40.6	0.7	1.8	39.8	41.1									
Hex7	1581.5	1.0	0.2	15.5	0.8	1.1	1.5	0.1	8.2	1.4	1.6	1.4	0.2	13.5	1.2	1.5									
Hex8	1743.6	0.4	0.4	87.7	0.0	0.7	0.5	0.5	92.4	0.0	0.9	0.8	0.1	16.3	0.7	0.9									

[0363] 实施例5-GnTII/GnTI融合蛋白质

[0364] GnTII/GnTI表达构建体的产生

[0365] 通过应用PCR重叠技术产生GnTII/GnTI融合表达构建体。用在融合位点处包含50bp同框重叠的引物单独地从GnTII和GnTI模板扩增融合片段。从琼脂糖凝胶纯化片段并根据标准程序用作用于融合构建体扩增的PCR模板。将融合构建体克隆到带有ApaI/SpeI限制性位点的载体中。通过测序分析验证得到的构建体。生成用于在毕赤酵母中表达在目标蛋白质的N-末端带有His标签的可溶性形式GnTII/GnTI的载体。与如上对于GnTI/II融合构建体所述类似的方式产生该载体。

[0366] 毕赤酵母中产生的可溶性GnTII/GnTI的纯化

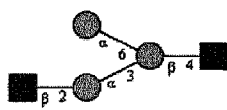
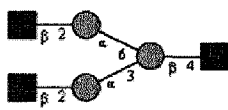
[0367] 按以上对于重组GnTI蛋白质所述的进行毕赤酵母中的表达和纯化过程。

[0368] GnTII/GnTI融合蛋白质的酶活性测试

[0369] 使用Man3Gn寡糖作为受体,按如上对于GnTI所述的进行活性分析。来自GnTII/GnTI反应的纯化的反应混合物的MALDI谱表明两个GlcNAcP残基被转移到受体上(图24)。

[0370] 表10:GnTII/GnTI融合蛋白质活性的总结。

[0371]

		形成的产物	
GnTII/GnTI 转化体	受体浓度		
转化体 1	0.5 mM	47%	5%
转化体 1	0.1 mM	-	11%
转化体 2	0.5 mM	3%	2.4%

[0372] 通过 β -N-乙酰葡萄糖胺糖苷酶表征

[0373] 用来自肺炎链球菌的 β 1-2,3,4,6-N-乙酰葡萄糖胺糖苷酶处理在GnTII/GnTI活性反应中形成的混合物。用MALDI MS分析来确定转移的 β -连接GlcNAc残基都被切割了(图25)。

[0374] 通过 β 1-4GalT半乳糖基化

[0375] 用来自牛奶的 β 1-4GalT处理在GnTII/GnTI活性反应中形成的混合物。预期 β 1-4GalT使产物混合物中的末端GlcNAc残基半乳糖基化。根据 β 1-4GalT反应混合物的MALDI谱,两种产物均被半乳糖基化。两个半乳糖转移到Gn2Man3Gn产物上,这表明GlcNAc残基与单独的甘露糖分支连接(图26)。

[0376] 通过随机整合用GnTII/GnTI构建体转化里氏木霉

[0377] 设计嵌合GnTII/GnTI序列,并克隆到含有gpdA启动子的载体中。质粒序列验证后,共转化至带有潮霉素标记基因的里氏木霉M124菌株中。对十三个PCR阳性转化体进行了鉴定。将所有阳性转化体和亲本菌株M124培养在TrMM,pH 4.8(补充有4%的乳糖和2%的酒糟提取物)中。此外,7个转化体和亲本菌株培养在TrMM,pH 5.5(具有4%的乳糖,2%的酒糟提取物和1%酪蛋白氨基酸,用100mMPIPPS(哌嗪1,4-二丙基磺酸)缓冲)中。用pH测量来监测菌株的增长率。在第3天、第5天和第7天收集上清液和菌丝体样品,冷冻保存,并分析聚糖结构。GnTII/GnTI序列也被克隆至含有cbh1启动子的质粒中。此外,作为对照,通过随机整合用GnTI构建体转化里氏木霉。

[0378] 通过随机整合获得的里氏木霉GnTII/GnTI菌株的聚糖分析

[0379] 对在两种不同培养基中培养的里氏木霉菌株M124GnTII/GnTI转化体和亲本M124株的156个上清液样品进行了分析。第一种培养基是TrMM,pH 4.8,补充有2%的酒糟提取物和4%的乳糖,和第二种培养基是TrMM,pH 5.5,补充有2%酒糟提取物,4%的乳糖,100mM的PIPPS和1%的酪蛋白氨基酸。细胞在这两种类型的培养基中生长3、5和7天。

[0380] 对于用于来自第3天和第5天时间点的样品的5 μ g上清液蛋白,在96孔板中一式三份地进行N-聚糖酶反应而无SDS变性。使用BSA作为标准,由基于Bradford的分析(Bio Rad Quick Start Bradford Protein Assay)测定上清液的蛋白质浓度。通过MALDI-TOF MS进行中性和酸性N-聚糖的分析。

[0381] 在3天和5天时间点的任何克隆中没有看见预期的GnTII/GnTI产物的征兆。此外，从通过随机整合所产生的带有gpdA启动子的GnTI和GnTI/II转化体没有观察到产物。

[0382] 通过靶向整合用GnTII/GnTI构建体转化里氏木霉

[0383] 构建具有在cbh1启动子控制下的嵌合GnTII/GnTI序列的带有pyr4基因环路标记(loopout marker)的载体，并亚克隆至alg3侧翼区片段之间的骨架载体中进行靶向整合。将PmeI消化的表达盒转化至里氏木霉菌株M127(M124的pyr4⁻菌株)中。在板选择后，克隆经PCR筛选和通过单孢子纯化。为了获得用于聚糖分析的材料，如所描述的进行摇瓶培养。将表明正确整合至M127转化体的alg3基因座的五个PCR阳性转化体在28℃下在300毫升体积中培养7天，培养基含有TrMM, pH 5.5, 补充40克/升的乳糖、20克/升酒糟提取物和100mM的PIPPS。为了避免细菌污染，在接种时将100毫克/升氨苄青霉素加入到烧瓶中。第3天、第5天和第7天收集用于聚糖分析的样品。

[0384] 通过靶向于alg3基因座获得的里氏木霉GnTII/GnTI菌株的聚糖分析

[0385] 在96孔板上对于5μg上清液蛋白质一式三份地制备里氏木霉菌株M124(对照)、M127GnTII/GnTI转化体的5个不同克隆和对照培养基样品的上清液样品。使用牛BSA作为标准，通过基于Bradford的分析(Bio-Rad Quick Start Bradford Protein Assay)测定上清液的蛋白质浓度。如所描述地进行PNGase F反应而无SDS变性。首先用Hypersep C-18纯化释放的N-聚糖，然后用Hypersep Hypercarb(都来自Thermo Scientific)纯化，其中中性和酸性聚糖分离。这两种纯化在96孔形式中进行。通过MALDI-TOF MS分析中性N-聚糖。

[0386] 对比来自里氏木霉M127 GnTII/GnTI转化体的中性N-聚糖的比例与来自菌株M124(其除了pyr4阳性外其它都与菌株M127相同)的比例。五个GnTII/GnTI转化体中的四个在所有时间点(3、5和7天)产生G0作为主要糖形。只有克隆46A是G0阴性的(图27)。在所有时间点每一个克隆的Man3Gn比例都是小的，但Hex6的比例还是相当大的。在第7天，与其它克隆相比，克隆17A产生最多的G0和最少的Hex6(图27)。第5天在摇瓶条件下GnTII/GnTI转化体中的四个克隆产生40%左右的糖形G0(图27)。具有受控pH值的发酵条件可以增加alg3敲除体中的G0产物的量和减少Hex6的量。

[0387] 在培养基样品中，观察到了一系列植物型的N-聚糖，但没有观察到G0相对应的信号。

[0388] 通过靶向整合用GnTII/GnTI构建体转化产生利妥昔单抗的里氏木霉

[0389] 将题为“通过靶向整合用GnTII/GnTI构建体转化里氏木霉”的章节中所述的表达盒转化到里氏木霉菌株M279(菌株M202的pyr4⁻株)中。通过删除M124中的pep1蛋白酶和引入利妥昔单抗重链和轻链(带有Kex2切割位点)获得M202。板选择后，克隆经PCR筛选和通过单孢子纯化。为了获得用于聚糖分析的材料，按题为“通过靶向整合用GnTII/GnTI构建体转化里氏木霉”的章节中所述的进行摇瓶培养，和此外，某些培养基补充有0.3毫克/毫升大豆胰蛋白酶抑制剂(SBTI)和1%酪蛋白氨基酸。首先接种时加入SBTI，然后在第3-6天每天添加。在冷冻前向所有样品添加PMSF和胃蛋白酶抑制剂A。

[0390] 通过靶向于alg3基因座获得的产生利妥昔单抗的里氏木霉GnTII/GnTI菌株的聚糖分析

[0391] 从具有SBTI的第5天的上清液样品和从没有SBTI的第5天和第7天样品用蛋白G亲和层析纯化利妥昔单抗。对约10μg的变性蛋白质进行PNGase F反应。首先用Hypersep C-18

纯化释放的N-聚糖,然后用Hypersep Hypercarb(都来自Thermo Scientific)纯化,其中中性和酸性聚糖分离。以96孔形式进行纯化步骤。通过MALDI-TOF MS分析中性和酸性N-聚糖。GnTII/GnTI转化体克隆中的两个,9A-1和31A-1,分别产生约30%和约24%的G0糖形。然而,合理数量的Hex6和GnMan3仍观察到(图28)。来自其他克隆的利妥昔单抗含有很少或没有G0。

[0392] 间隔区的优化

[0393] 构建了一系列用于GnTII/GnTI融合蛋白质的间隔区修饰。这些变异体在毕赤酵母中产生并在体外研究酶稳定性和活性。

[0394] 在此描述了用于克隆GnTII/GnTI融合蛋白质的材料和方法。通过采用PCR重叠策略在两部分中扩增T45序列。首先,片段用GP13 5'引物和GP93 3'引物扩增,且第二片段用GP92 5'引物和GP2 3'引物扩增。在由供应商提供的标准条件下用Phusion高保真PCR聚合酶(Finnzymes)进行扩增。循环条件如下:98℃下初始变性30秒,在98℃下变性5秒,在65℃下退火30秒,在72℃下延伸45秒,重复20次,并在72℃下最后延伸20分钟。用Fermentas GeneJET凝胶提取试剂盒从琼脂糖凝胶纯化得到的PCR产物。在无引物的标准条件下将这些具有重叠的经修饰的序列的片段在相同的反应混合物中合并。十个退火/延伸循环如下进行:在98℃下进行初始变性30秒,在98℃下变性5秒,在65℃下退火30秒,在72℃延伸45秒,重复10次,并在72℃下进行20分钟的最后延伸。加入引物GP13(5')和GP2(3'),和如上20个扩增循环所述继续循环。用Fermentas GeneJET PCR纯化试剂盒纯化扩增的T45片段,按照标准方法用EcoRI/KpnI(New England Biolabs)进行消化,并克隆到EcoRI/KpnI消化的酵母表达载体pBLARG-SX中。用引物3' AOX、5' AOX、GP9、GP37、GP38和GP122测序得到的载体。发现该序列是正确的。

[0395] 将这种得到的质粒作为模板用于3xG4S间隔区修饰。按如上对于T45所述进行T46序列的克隆,GP13 5'-引物和GP95 3'-引物用于第一片段的合成,和GP94 5'-引物和GP2 3'-引物用于第二片段的合成。将片段组合,并加入引物GP13(5')和GP2(3')用于扩增。然后用EcoRI/KpnI消化扩增的片段T46,并克隆到酵母表达载体pBLARG-SX中。用上述引物对得到的载体进行测序,发现该序列是正确的。

[0396] 用类似的PCR重叠方法构建纤维素酶相关的天然间隔区。利用CBHI相关间隔区,用GP13 5'-引物和GP107 3'-引物扩增第一片段。用GP108 5'-引物和GP2 3'-引物扩增第二片段(表11)。利用EGIV相关间隔区,用GP13 5'引物和GP109 3'引物扩增第一片段。用GP1105'-引物和GP2 3'-引物扩增第二片段(表11)。在这两种情况下,从琼脂糖凝胶纯化PCR产物,组合,并用作下一PCR反应的模板以扩增序列T50和T51。然后分别用EcoRI/KpnI消化T50和T51 PCR产物并克隆到酵母表达载体pBLARG-SX中。

[0397] 用高保真Phusion聚合酶(Finnzymes)进行所有的PCR扩增。从MWG Operon订购引物(表11)。测序是由Helsinki大学生物技术研究所的DNA测序实验室作为商业服务进行的。

[0398] 表11:引物序列。

[0399] 引物 序列5'-3'

[0400]

3'AOX	GCAAATGGCATTTCTGACATCC (SEQ ID NO: 99)
5'AOX	GACTGGTTCCAATTGACAAGC (SEQ ID NO: 100)
GP2	CAGTGGTACCCTAATTCCAGCTAGGATCATAGCCCTCCCACG (SEQ ID NO: 101)
GP9	CGGACCACCGCAAGTTCC (SEQ ID NO: 102)
GP13	ATGCGGAATTCTGCATCATCATCATCATCGCCAGCGTAAGAACGAGGCCCT (62113) (SEQ ID NO: 103)
GP37	CCTTTCTCTATCCAACCTCTACC (SEQ ID NO: 104)
GP38	GGAACCTGCGGTGGTCCG (SEQ ID NO: 105)
GP92	CCGCCGGCTCCAGGGAGGTGGGGGCAGTGGAGGTGGCGGCAGTGGGAGGGTGCCCACC GCCGCCCC (SEQ ID NO: 106)
GP93	GCGGTGGGCACCCTCCCCTGCGCCACCTCCACTGCCCCCACCTCCCTGGAGCCGGCGG TAAGAC (SEQ ID NO: 107)
GP94	AGGTGGGGGCAGTGGAGGTGGCGGCAGTGGCGGCGGTGGAAGTGGGAGGGTGCCCACC GCCGCCC (SEQ ID NO: 108)
GP95	CGGTGGGCACCCTCCCCTGCGCCACCTCCACTGCGCCACCTCCACTGCCCCCACCTC CCTG (SEQ ID NO: 109)
GP107	GTTTCCGCCGGGAGGGTTGCCGCCGCTAGGGTTGCCGGTGCTCTGGAGCCGGCGGTAAG ACTTGC (SEQ ID NO: 110)
GP108	GCAACCCCTCCCGCGGAAACCCGCCTGGCAGCACCGGGAGGGTGCCCCACCGCCGCCCT CCCGCCC (SEQ ID NO: 111)
GP109	CCGCTCCAGGAACAGTGGCGCTGGCGGTGGCCGTCGCGGCGGAGCTCTGGAGCCGGCG GTAAGACTTGC (SEQ ID NO: 112)
GP110	CGCCACTGTTCTGGAGGCGGTAGCGCCCCACCAGCGGGAGGGTGCCCCACCGCCGCCCT CCCCCGCCAGC (SEQ ID NO: 113)
GP122	CATTAGCGAGAAGTTTACGG (SEQ ID NO: 114)

[0401] 按实施例3中对于GnTI所描述的类似方式,通过浓缩和缓冲液交换处理间隔区修饰的(3xG4S和2xG4S)GnTII/GnTI融合酶用于活性测定。用Man3Gn受体进行活性测定,和如在GnTI活性测定中所描述的进行反应混合物的纯化。还如对于GnTI反应混合物所描述的进行MALDI分析,但是此外,接着形成GnTII产物,Hex3HexNAc3。Hex3HexNAc3的[M+Na]⁺信号的计算m/z值是1136.318(图29)。

[0402] 间隔区变异体

[0403] GnTII/I间隔区变异体从GnTII/I融合蛋白质的野生型间隔区序列进行改变。修饰的间隔区被列于表12中。在+16℃下用蛋白酶抑制剂表达所有四个间隔区变异菌株(GY32、GY33、GY49和GY50)、野生型GnTII/I融合菌株(GY7-2)和模拟菌株(GY3)。菌株在+30℃, 220rpm下接种在60毫升BMY培养基中过夜(0/N)。沉淀过夜培养物,并将细胞重悬浮于60毫升BMY培养基中。在开始甲醇诱导的同时并在此后每天一次在培养基中加入蛋白酶抑制剂,1mMEDTA,1.5μM胃蛋白酶抑制剂A(Sigma)和1完全无EDTA的蛋白酶抑制剂混合片(Roche)。在第3天和第4天从培养物取25毫升样品,并使用浓缩管(Millipore)浓缩上清液样品,在PD-10柱中缓冲交换到100mM MES pH 6.1中和浓缩至最终50X。细胞沉淀重悬浮于500μl的1xPBS中,除了野生型的细胞沉淀(第三)(其重悬浮在500μl的100mM MES pH 6.1和完全(无EDTA)抑制剂混合物中)。

[0404] 包含3xG4S间隔区的GnTII/GnTI融合蛋白质的氨基酸序列示于SEQ ID NO:119中。包含3xG4S间隔区的GnTII/GnTI融合蛋白质的核苷酸序列示于SEQ ID NO:141中。包含2xG4S间隔区的GnTII/GnTI融合蛋白质的氨基酸序列示于SEQ ID NO:121中。包含2xG4S间隔区的GnTII/GnTI融合蛋白质的核苷酸序列示于SEQ ID NO:139中。包含CBHI间隔区的

GnTII/GnTI融合蛋白质的氨基酸序列示于SEQ ID NO:123中。包含CBHI间隔区的GnTII/GnTI融合蛋白质的核苷酸序列示于SEQ ID NO:143中。包含EGIV间隔区的GnTII/GnTI融合蛋白质的氨基酸序列示于SEQ ID NO:125中。包含EGIV间隔区的GnTII/GnTI融合蛋白质的核苷酸序列示于SEQ ID NO:145中。

[0405] 通过重复离心和重悬细胞在具有完全(无EDTA)抑制剂混合物的100mM的MES pH 6.1中洗涤200 μ l细胞悬浮液样品。通过取得200 μ l洗涤的细胞样品,加入50 μ l的玻璃珠和2 μ l的Triton X-100并置于珠磨器中6分钟来制备细胞裂解液。如上进行50倍浓缩的毕赤酵母培养物上清液、细胞样品和细胞裂解液的GnTI活性分析。

[0406] 表12:酵母菌株的说明

[0407]

酵母菌株	说明	间隔区变异体序列
GY3	模拟菌株	
GY7-2	野生型 GnTII/I 融合体	
GY32-5	GnTII/I 融合 3xG4S 间隔区变异体	SEQ ID NO: 118
GY32-9		
GY33-7	GnTII/I 融合 2xG4S 间隔区变异体	SEQ ID NO: 120
GY33-8		
GY49-3	GnTII/I 融合 CBHI 间隔区变异体	SEQ ID NO: 122
GY50-7	GnTII/I 融合 EGIV 间隔区变异体	SEQ ID NO: 124
GY50-10		

[0408] 来自第3天的细胞沉淀和50倍浓缩的培养基上清液的Western印迹分析如图30所示。CBHI间隔区变异体(GY49)从细胞沉淀样品而非上清液给出强信号。从上清液检测到EGIV间隔区变异体(GY50),但只获得微弱的信号。从野生型GnTII/I融合菌株(GY7-2)和2xG4S间隔区变异菌株GY33-7和GY33-8的上清液样品中也获得了微弱的信号(图30)。

[0409] 然后将含有间隔区变异体的GnTII/I融合蛋白质的活性与包含野生型间隔区的GnTII/I融合蛋白质的活性进行比较。

[0410] 上清液中的融合GnTII/I活性。提供了GnTI底物Man3Gn,和反应产物GnMan3Gn作为用于融合蛋白质的GnTII活性的受体。在第3天和第4天表达阶段后获取用于活性分析的样品。图31显示了包含野生型间隔区或间隔区变异体的GnTII/I融合蛋白质的培养物活性分析结果。在抑制剂(1.5 μ M胃蛋白抑制剂A,1mM EDTA,1片/50毫升完全无EDTA的蛋白酶抑制剂混合片)存在下进行样品培养。为简单起见,一起加入GnTI和GnTII反应产物。所有活性测定样品中只含有少量(<5%)的GnTI产物GnMan3Gn,说明GnTII主动地将GnMan3Gn转化为Gn2Man3Gn。

[0411] 所有四个间隔区变异体显示GnT活性,虽然克隆和培养天数之间有一些变异性。包含2xG4S(克隆_1)、3xG4S(克隆_1和克隆_2)或EGIV间隔区变异体的GnTII/I融合蛋白质比具有野生型间隔区的酶显示出更高的活性(图31)。包含CBHI间隔区变异体的GnTII/I融合蛋白质表现出与具有野生型间隔区的酶相当的活性(图31)。包含2xG4S变异体(克隆_2)的GnTII/I融合蛋白质比具有野生型间隔区的酶的活性低(图31)。第4天的样品比第3天的样品具有较高的活性,例外的是包含3xG4S间隔区变异体的GnTII/I融合蛋白质(克隆_1和克隆_2),其在第3天显示出较高的活性(图31)。包含EGIV间隔区变异体的GnTII/I融合蛋白质在第4天有最高的活性(图31)。

[0412] 细胞和细胞裂解液中的融合GnTII/I活性。细胞、细胞裂解液和来自包含具有野生型间隔区的GnTII/I融合蛋白质的细胞的上清液样品的活性试验表明,裂解液样品含有最高的活性(图32)。在细胞表面上的活性第二高,且上清液样品中观察到的活性最低(图32)。因此,表现为大多数GnTII/I融合蛋白质定位于细胞内或在细胞表面上,只有少量被分泌。

[0413] 包含具有野生型间隔区或间隔区变异体的GnTII/I融合蛋白质的细胞的GnT活性如图33中所示。将细胞重悬浮在500 μ l的100mM MES, pH 6.1与完全无EDTA抑制剂混合物中,且按上述方法制备于500 μ l的PBS中的间隔区变异体及用于活性测试的细胞和裂解液。

[0414] 如图33所示,包含间隔区变异体的GnTII/I融合蛋白质在细胞中比在上清液中有显着高得多的GnTII/I活性。在裂解液中,酶呈现为失活的。据认为,这种活性的缺乏是由于释放的蛋白酶的作用。包含CBHI间隔区变异体的GnTII/I融合蛋白质在细胞和裂解液中显示高活性(图33),这与在细胞沉淀样品中显示较高信号的Western印迹分析相关(图30)。

[0415] 讨论。在上清液中,包含2xG4S和3xG4S间隔区变异体的GnTII/I融合蛋白质比包含野生型间隔区的GnTII/I融合蛋白质具有更高的活性,而CBHI间隔区变异体与包含野生型间隔区的GnTII/I融合蛋白质具有相当的活性。此外,含EGIV间隔区变异体的GnTII/I融合蛋白质显示最高的GnT活性。第3天样品的Western印迹分析与第4天活性的结果有一定的相关性。Western印迹分析显示野生型,两个克隆2xG4S和EGIV的上清液样品的微弱条带。检测的活性顺序如下:EGIV > 2xG4S(克隆_1) > 3xG4S(克隆_2) > 3xG4S(克隆_1) > CBHI = 野生型 = 2xG4S(克隆_2)。

[0416] 包含野生型间隔区的GnTII/I融合蛋白质的上清液、细胞和细胞裂解液样品中GnTII/I融合蛋白质活性的测定表明,大部分的活性关联在细胞内和较低的量分泌。据认为,这解释了为什么Western印迹分析中观察到His标记的GnTII/I的信号在细胞部分中比在上清液部分中好得多。

[0417] 完全无EDTA抑制剂片对丝氨酸和半胱氨酸蛋白酶、EDTA对金属蛋白酶和胃蛋白酶抑制剂A对天冬氨酸蛋白酶的抑制提高了GnTII/I融合蛋白质的产量。对于使用丝氨酸蛋白酶抑制剂的这个观察与Salamon等(*Appl. Environ. Microbiol.*, 76(2010)4269-4276)的工作一致,这表明毕赤酵母培养基中的丝氨酸型蛋白酶活性完全被PMSF抑制。此外,Vad等(*J. Biotechnol.* 116(2005)251-260)报道在毕赤酵母中存在与酿酒酵母蛋白质二硫键异构酶共表达结合的10mM EDTA的情况下完整人甲状旁腺激素的高产量,超过300毫克/升。

[0418] 包含四个间隔区变异体中的每一个的所有GnTII/I融合蛋白质具有GnTII/I的活性,并且具有2xG4S和EGIV间隔区变异体的酶的活性比包含野生型间隔区的GnTII/I融合蛋白质的活性更高。

[0419] 实施例6-Man5作为受体聚糖的融合蛋白质的用途

[0420] 具有Man5型N-糖基化的利妥昔单抗表达里氏木霉菌株的构建

[0421] 原始的利妥昔单抗序列是统一的密码子协和的。产生了含有合成的利妥昔单抗轻链和重链的原始质粒。抗体链和CBHI融合蛋白质被设计有40个核苷酸的重叠序列,如用于重链的表达载体pHH01(乙酰胺酶选择标记,用于整合到cbh1基因座中的cbh1侧翼)或用于轻链的pHH02(潮霉素选择标记,用于整合到egl1基因座中的egl1侧翼),以确保使用酵母同源重组的克隆。

[0422] 将所获得的基因质粒转化入大肠杆菌中。制备DNA,并消化合成的基因和从质粒骨架分离。通过对带有CBHI融合蛋白质和重链或轻链任一的里氏木霉表达载体进行酵母同源重组来构建表达载体。从酵母援救重组质粒并转化至大肠杆菌中。经过PCR筛选后,分离正确的克隆并测序。消化表达盒片段并从质粒主链分离,产生大约10.2kb的片段用于重链构建体和10.8kb的片段用于轻链构建体。将重链和轻链片段共转化到里氏木霉菌株M124中。筛选转化体的潮霉素抗性和以乙酰胺作为唯一氮源生长的能力。连续两轮将转化体在双重选择培养基上划线,并通过PCR测试表达构建体到基因组中的整合。

[0423] GnTII/I串联酶和甘露糖苷酶II引入表达利妥昔单抗抗体的里氏木霉菌株

[0424] 除了将重组GnTII/I引入产生Man5的菌株,如M124外,还需要甘露糖苷酶II活性以从GlcNAcMan5聚糖结构移除两个甘露糖,以使GnTII/I可以使用GlcNAcMan3作为受体分子。

[0425] 在前面的实施例中所描述的GnTII/I的表达盒可以使用基本如上所述的方法靶向于,例如,里氏木霉的cbh2基因座。为生成GlcNAcMan3受体分子用于GnTII/I融合蛋白质,然后使用如上所述的转化方法将甘露糖苷酶II活性引入菌株中。

[0426] 通过设计带有用于驱动甘露糖苷酶表达的启动子的所需含甘露糖苷酶的表达盒,将甘露糖苷酶II活性引入到表达利妥昔单抗抗体的M124菌株中。有用的启动子为来自gpdA或cbh1的那些。可通过随机整合转化甘露糖苷酶II活性,然后筛选具有最合适的表达水平的菌株。表达盒与专有的选择标记基因连接,或选择标记作为单独的表达盒共转化。根据上面描述的方法进行转化。

[0427] 甘露糖苷酶II融合构建体可以源自里氏木霉的胞质结构域、跨膜结构域和非催化结构域,或KRE2的靶向肽,并与人甘露糖苷酶II的N-末端氨基酸缺失同框接合。通过KRE2靶向肽序列将编码的融合蛋白质定位在ER/高尔基体内同时保留其甘露糖苷酶催化结构域活性,且能够将GlcNAcMan5GlcNAc2水解成GlcNAcMan3GlcNAc2。在某些实施方式中,可以在M124菌株中表达全长人甘露糖苷酶II。

[0428] KRE2靶向肽包含KRE2的从约1至约106或从约1至约83的氨基酸。

[0429] Kre2 aa 1-106

[0430]

MASTNARYVRYLLIAFFTILVFYFVSNSKYEGVDLNKGTFTAPDSTKTTTPKPPATGDAKDFPLALTPNDPGFN DLVG
IAPGPRMNATFVTLARNSDVWDIARSIRQ(SEQ ID NO:115)

[0431] Kre2 aa 1-83

[0432]

MASTNARYVRYLLIAFFTILVFYFVSNSKYEGVDLNKGTFTAPDSTKTTTPKPPATGDAKDFPLALTPNDPGFN DLVG
IAPGPR(SEQ ID NO:116)

[0433] 如上所述用甘露糖苷酶II构建体转化木霉后,选择木霉菌株,连续两轮在选择性培养基上划线,并通过PCR测试表达构建体到基因组中的整合。然后将选择的产生Man5和表达GnTII/I融合蛋白质、甘露糖苷酶II和利妥昔单抗抗体的木霉菌株的转化体培养在摇瓶或发酵罐条件中并如上所述分析聚糖含量。

[0434] 实施例7-GnTI和GnTII在里氏木霉中的表达

[0435] 通过随机整合用GnTI构建体转化里氏木霉M124

[0436] 将密码子优化的人GntI转化至里氏木霉M124株中。GntI基因被克隆到两个不同启动子的控制下的载体中:(1)cbh1基因的诱导型启动子,和(2)gpdA基因的组成型表达启动子。在这两个启动子任一控制下的含GntI的载体用含乙酰胺或潮霉素抗性标记基因的质粒各自共转化至里氏木霉M124菌株中。

[0437] 通过PCR筛选具有gpdA启动子下和乙酰胺选择下的GntI的34个转化体,且均为GntI阳性。对于具有cbh1启动子和乙酰胺选择下的GntI的转化体,26个中的19个对于GntI构建体为PCR阳性。此外,对具有在cbh1启动子和潮霉素选择下的GntI的五个菌株进行最初DNA提取。所有这些菌株为PCR阳性。纯化二十五个gpdA启动子转化体,和所有的cbh1启动子的转化体(14+5)纯化至单核克隆并制备孢子悬浮液。

[0438] 为了初始分析的目的,将23个gpdA启动子转化体和19个cbh1启动子转化体(14个从乙酰胺和5个从潮霉素选择生长),以及亲本菌株M124培养在具有50ml木霉基本培养基(补充有2%的酒糟提取物和4%的乳糖)的250ml摇瓶中。通过pH测量监测菌株的生长。在第3天、第5天和第7天收集样品(上清液和菌丝体),冷冻保存直至用于聚糖结构分析。

[0439] 通过随机整合获得的里氏木霉GnTI菌株的聚糖分析

[0440] 通过基于Bradford的分析(Bio-Rad Quick Start Bradford Protein Assay),使用BSA作为标准测定所有上清液样品的蛋白质浓度。将进行N-聚糖分析的样品的分泌蛋白质含量调整到5 μ g或10 μ g。在96孔板上对5 μ g上清液蛋白或在1.5ml管中对10 μ g上清液蛋白进行N-聚糖分析。一式三份进行所有的N-聚糖分析。用MALDI-TOF MS进行中性和酸性N-聚糖的分析。

[0441] 为了得到更精确的四个GnT1转化体中产生的GnT1产物Gn1Hex5的量(从第3和5天)和产生的酸性N-聚糖的量的测量,用已知的聚糖掺入MALDI谱。对于中性和酸性N-聚糖,分别使用2pmol/谱的1177Da质量值的Hex2HexNAc4和0.5pmol的1362Da质量值的甘露糖基苷Hex4HexNAc2的内部标准。一式三份进行分析。

[0442] 在任何gpdA启动子转化体中没有观察到GnT1产物。但是,8个cbhI启动子转化体产生了GnT1产物Gn1Man5(图34和图35,及表13);五个具有潮霉素选择,三个具有乙酰胺选择。

[0443] 表13:与内部标准Hex2HexNAc4相比第3天和第5天的四个阳性GnT1转化体和亲本M124菌株中Man5和Gn1Man5的信号强度百分比。Man5是亲本M124菌株中的主要糖形。

[0444]

		M124 1., 第 3 天						M124 1., 第 5 天						M124/GNT1,克隆 HM1,第 3 天						M124/GNT1, 克隆 HM1,第 5 天					
组成	m\z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX				
Hex2HexNAc4	1177.42	97.7	0.5	0.5	97.1	98.0	36.5	0.8	2.3	35.9	37.1	78.5	14.5	18.4	68.2	88.7	50.1	10.6	21.2	42.6	57.6				
Hex5HexNAc2	1257.42	2.3	0.5	22.5	2.0	2.9	63.5	0.8	1.3	62.9	64.1	14.5	9.9	68.0	7.5	21.5	44.0	9.6	21.9	37.2	50.8				
Hex5HexNAc3	1460.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.1	4.6	65.6	3.8	10.3	5.9	1.0	16.7	5.2	6.6				
		M124/GNT1,克隆 8,第 3 天						M124/GNT1,克隆 8,第 5 天						M124/GNT1,克隆 39,第 3 天						M124/GNT1, 克隆 39,第 5 天					
组成	m\z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX				
Hex2HexNAc4	1177.42	77.3	7.6	9.8	72.0	82.7	67.3	10.0	14.9	56.5	76.3	83.7	1.5	1.8	82.7	84.8	40.0	1.9	4.6	37.9	41.1				
Hex5HexNAc2	1257.42	15.0	5.2	34.4	11.4	18.7	18.9	6.2	32.5	12.8	25.1	8.3	1.0	11.7	7.6	8.9	46.9	1.8	3.8	45.6	49.0				
Hex5HexNAc3	1460.5	7.6	2.4	31.6	5.9	9.3	13.8	4.0	29.1	10.8	18.3	8.0	0.6	6.9	7.6	8.4	13.1	0.3	2.1	12.7	13.3				
		M124/GNT1,克隆 90,第 3 天						M124/GNT1,克隆 90,第 5 天																	
组成	m\z	平均	SD	RSD	MIN	MAX	平均	SD	RSD	MIN	MAX														
Hex2HexNAc4	1177.42	93.8	1.6	1.7	92.4	95.6	92.6	2.7	2.9	89.8	95.3														
Hex5HexNAc2	1257.42	3.7	1.0	25.9	2.6	4.5	4.7	1.4	30.9	3.2	6.0														
Hex5HexNAc3	1460.5	2.5	0.7	26.2	1.8	3.1	2.7	1.3	47.8	1.5	4.1														

[0445] 在所有阳性转化体的磷酸化N-聚糖中也发现了GnT1产物Gn1Man6P1、Gn1Man7P1和Gn1Man8P1。GnT1转化体中磷酸化N-聚糖的量增加,图谱偏向较大的N-聚糖,其中Man7P1或Man8P1具有最强的信号(Man6P1在亲本M124中具有最强的信号)(图36)。

[0446] 八个GnTI转化全产生Gn1Man5结构。Gn1Man5在克隆39中是最丰富的。然而,最好的克隆似乎是克隆8,其产生第二高水平的Gn1Man5,但具有高比例的Man5和Gn1Man5(图35)。克隆8,其包含是cbh1启动子控制下的GnTI,被命名为菌株M198,并被选择用于继续分析。

[0447] 通过靶向整合用GnTII构建体转化里氏木霉M198菌株

[0448] 建立了五个携带GnTII的载体(表14)。其中的两个载体含有在GNTII中的天然哺乳动物高尔基体靶向肽。在其他三个载体中,哺乳动物靶向肽被里氏木霉MNT1(α -1,2-甘露糖转移酶)靶向肽替代。所有五个载体含有cbh1启动子或gpdA启动子和pyr4环标记。此外,所有五个载体靶向于整合入alg3基因座中,从而删除alg3基因。在cbh1启动子下的MNT1/GnTII结构中,对两种不同大小的GnTII序列缺失进行了测试。

[0449] 表14:构建的GNT2载体。

[0450]

质粒名称	启动子	靶向肽	N-末端缺失(GnTII)
pTTv140	cbh1	哺乳动物	N/A
pTTv141	gpdA	哺乳动物	N/A
pTTv142	cbh1	里氏木霉MNT1	74个氨基酸

[0451]

pTTv143	cbh1	里氏木霉MNT1	104个氨基酸
pTTv144	gpdA	里氏木霉MNT1	74个氨基酸

[0452] 除了pTTv144载体外,将这些载体分别转化至产生最佳py4-阴性GnTI的菌株M198(M319)中作为PmeI片段。将转化体纯化至单核克隆并进行PCR筛选。显示两端的正确整合的克隆然后被选择用于继续分析。

[0453] 为了研究产生的表达GNTII的菌株的生长特性,制备了大摇瓶培养物。以两个单独批次制备摇瓶培养物。第一批包含pTTv140、pTTv142和pTTv143。第二批包含pTTv141。将亲本菌株M198用作对照菌株。细胞生长在TrMM培养基(补充有40克/升乳糖,20克/升酒糟提取物和100mM的PIPPS),pH 5.5中。培养每个构建体的五个转化体。pTTv140、pTTv142和pTTv143培养物在第3、5、7和9天采样。pTTv141培养物在第3、5、7和10天进行采样。测定各个

样品的pH值和细胞干重,并将培养物上清液样品用于聚糖结构分析。

[0454] 通过将GnTII靶向于里氏木霉M198菌株的alg3基因座获得的里氏木霉菌株的多糖分析

[0455] 对含有pTTv140载体(含天然靶向肽和cbh1启动子)、pTTv142载体(含MNT1靶向肽、GNTII 74 aa N-末端缺失和cbh1启动子)、pTTv143载体(含MNT1靶向肽、GNTII 110 aa N-末端缺失和cbh1启动子)和pTTv141载体(含靶向肽和gpdA启动子)的五种不同的克隆进行了分析。

[0456] 在96孔板上对于5 μ g的上清液蛋白对第5天的样品一式三份地制备N-聚糖分析物,并对第3天和第7天的样品一式两份地制备N-聚糖分析物。上清液的蛋白质浓度通过使用BSA作为标准通过基于Bradford的分析(BioRad Quickstart Bradford Protein Assay)测定。如所述的进行PNGase F反应。首先用Hypersep C-18 100毫克和然后用Hypersep Hypercarb 10毫克(都来自Thermo Scientific)纯化释放的N-聚糖,其中中性和酸性聚糖分离。这两种纯化均以96孔形式进行。通过MALDI-TOF MS分析中性N-聚糖。

[0457] 对用GnTII转化的四个不同菌株的N-聚糖进行分析。用pTTv140载体转化并因此包含天然靶向肽和cbhI启动子的克隆1-117A产生约40%的G0和约13%的Hex6(图37A)。用pTTv143载体转化并因此包含MNT1靶向肽、GnTII 110 aa的N-末端缺失和cbhI启动子的克隆产生约10%的G0(图37C)。克隆3B(其包含gbdA启动子)产生约28%的G0和约19%的Hex6(图37D)。

[0458] 含pTTv140、pTTv141和pTTv142载体的代表性克隆的糖基化模式也显示为随时间变化而稳定的(图38)。

[0459] 蛋白质特异性糖基化

[0460] 为了分析糖基化中的蛋白质特异性变化,用SDS-PAGE分离来自含pTTv142载体的克隆3-17A和来自亲本菌株M198的样品并印迹至PVDF膜上。切除目的蛋白质条带(M198的四个条带和3-17A克隆的四个条带),并用PNGase F通过膜上酶促释放来释放N-聚糖(图39)。

[0461] 利用MALDI-TOF MS分析脱离和纯化的中性N-聚糖。总分泌蛋白质的糖基化模式与M198亲本菌株的分离的50kDa蛋白质类似(图40)。最小尺寸的蛋白质条带为未糖基化的。

[0462] 在GnTII克隆3-17A中,大多数的非典型信号消失了,证实它们来自培养基中。此外,克隆3-17A的糖基化模式与总分泌蛋白质的聚糖模式不同(图40B)。来自克隆3-17A的G0量约为35%至36%(图40B)。

[0463] GnTII菌株的发酵罐培养

[0464] GnTII菌株1-117A M329(其包含pTTv140载体)的发酵罐培养是在TrMM pH 5.5+2%的酒糟提取物+6%的乳糖+0.5%磷酸二氢钾+0.5%硫酸铵在+28 $^{\circ}$ C(pH 5.5)中发酵。对第3天获取的样品如上面在“蛋白质特异性糖基化”一节中描述的对5 μ g分泌蛋白质一式三份地进行N-聚糖分析。在第3天G0的量为约48%和Hex6的量为约19%(图41)。

[0465] 实施例8-里氏木霉ALG3同源物

[0466] 通过随机整合用GnTI构建体转化里氏木霉M124

[0467] 从其他生物体鉴定了里氏木霉ALG3同源物。这些同源物可以被用来设计ALG3缺失构建体用于里氏木霉以外的丝状真菌细胞。ALG3同源物被列于表15中。里氏木霉ALG3和ALG3同源物的多氨基酸序列比对如图42中所示。

[0468] 表15:ALG3同源物。

[0469]

参考序列	生物体	SEQ ID NO:
Trire2 104121 fgenes5_pg.C_scaffold_3000076	Trichoderma reesei	126
Triat2 270085 fgenes1_pg.contig_14#_149	Trichoderma atroviride	127
TriviGv29_8_2 194462 fgenes1_pm.87_#_115	Trichoderma virens	128
EGU81920.1	Fusarium oxysporum Fo5176	129
XP_389829.1	Gibberella zeae PH-1	130
AE060805.1	Myceliophthora thermophila	131
XP_962259.1	Neurospora crassa OR74A	132
XP_001824044.1	Aspergillus oryzae R1B40	133
XP_001259497.1	Neosartorya fischeri NRRL 181	134
XP_001398696.2	Aspergillus niger CBS 513.88	135
XP_362427.2	Magnaporthe oryzae 70-15	136
NP_593853.1	Schizosaccharomyces pombe 972h	137

[0470] 实施例9-GnTII/GnTI融合蛋白质变异体

[0471] GnTII/GnTI表达构建体的产生

[0472] 如实施例4和5中所描述的构建在诱导型启动子cbh1控制下和包含4个间隔区变异体之一的重组GnTI/II融合蛋白质。这四种间隔区变异体是2xG4S间隔区、3xG4S间隔区、CBHI间隔区以及EGIV间隔区。

[0473] 简单地说,用在融合位点含有50bp的同框重叠的引物单独地从GnTII和GnTI模板扩增融合片段。从琼脂糖凝胶纯化片段,并根据标准程序用作用于融合构建体扩增的PCR模板。将该融合构建体克隆到在诱导型启动子cbh1的控制下的具有ApaI/SpeI限制性位点的载体中。此外,GnTII域中的天然哺乳动物高尔基体靶向肽被里氏木霉MNT1(α -1,2-甘露糖基转移酶)靶向肽取代。

[0474] 为了将2xG4S间隔区变异体引入到融合蛋白质中,通过用PCR重叠策略以两部分扩增T45序列。首先,用AKT1-6-1 5'引物(GGTACCGGGCCCACTGCGCATCATGCGCTCCGAATCTACAAGCG (SEQ ID NO:146))和GP933'引物扩增片段,和用GP92 5'引物和AKT1-6-4 3'引物(GGCGCGCCACTAGTCTAATTCCAGCTGGGATCATAGCC (SEQ ID NO:147))扩增第二片段。用Phusion高保真PCR聚合酶(Finnzymes)在供应商提供的标准条件下进行扩增。循环条件如实施例5中所描述。从琼脂糖凝胶纯化所得的PCR产物,且在无引物的标准条件下,具有重叠的、经修饰的序列的片段在相同的反应混合物中合并。如实施例5中所描述地进行十个退火/延伸循环。加入引物AKT1-6-1(5')和AKT1-6-4(3'),和如实施例5中所描述地继续进行20个扩增循环。然后纯化扩增的T45片段,根据标准方案用ApaI/SpeI(New England Biolabs)消化,并克隆到里氏木霉表达载体中。然后用合适的引物组通过测序来验证克隆的片段,和将产生的序列用于构建具有2xG4S启动子和alg3靶向的里氏木霉表达载体。

[0475] 将得到的质粒用作模板用于3xG4S间隔区修饰。如上对于T45所述的进行T46序列的克隆。将AKT1-6-1 5'引物和GP95 3'-引物用于第一片段的合成,和GP94 5'-引物和AKT1-6-4 3'引物用于第二片段的合成。将片段合并,并添加引物AKT1-6-1(5')和AKT1-6-4(3')用于扩增。然后用ApaI/SpeI消化扩增的片段T46,并克隆到里氏木霉表达载体中。然后通过用适当的引物组测序来验证克隆的片段和将产生的序列用于构建具有3xG4S启动子和

alg3靶向的里氏木霉表达载体。

[0476] 用类似的PCR重叠方法构建CBHI和EGIV间隔区。对于CBHI间隔区,用AKT1-6-1 5'引物和GP107 3'-引物扩增第一片段。用GP1085'-引物和AKT1-6-4 3'-引物(表11)扩增第二片段。对于EGIV间隔区,用AKT1-6-1 5'引物和GP109 3'引物扩增第一片段。用GP110 5'引物和AKT1-6-4 3'引物(表11)扩增第二片段。在这两种情况下,从琼脂糖凝胶纯化PCR产物,合并并用作下一PCR反应的模板来扩增序列T50和T51。然后用ApaI/SpeI消化T50和T51PCR产物,并克隆到里氏木霉表达载体中。然后通过用适当的引物组测序来验证克隆的片段和将产生的序列用于构建具有CBHI或EGIV启动子和alg3靶向的里氏木霉表达载体。

[0477] 用Phusion高保真PCR聚合酶(Finnzymes)进行所有的PCR扩增。引物(表11)从MWG Operon订购。测序是由Helsinki大学生物技术研究所的DNA测序实验室作为商业服务完成。

[0478] 然后构建所述的带有间隔区变异体(2xG4S、3xG4S、CBHI和EGIV)的嵌合GnTII/GnTI序列亚克隆到cbh1启动子的控制下的里氏木霉表达载体,其中pyr4基因环路标记和alg3侧翼区片段用于靶向整合到骨架中。表达盒转化至里氏木霉菌株M279(M202的pyr4⁻菌株)。板选择后,克隆经PCR筛选和通过单孢子纯化。为了获得用于聚糖分析的材料,按所述进行摇瓶培养。

[0479] GnTII/I融合蛋白质变异体引入表达利妥昔单抗抗体的里氏木霉菌株按实施例5中所描述的将重组GnTII/I融合蛋白质变异体引入表达利妥昔单抗的里氏木霉菌株M279中。

[0480] 简单地说,将具有cbh1启动子控制下的GnTII/GnTI融合蛋白质、MNTI靶向肽、pyr4环路标记及4个间隔区变异体中的各变异体的载体各自亚克隆到用于靶向整合的alg3侧翼区片段之间的骨架载体中,从而删除alg3基因。将PmeI消化的表达盒转化至里氏木霉菌株M279(pyr4⁻菌株)中。板选择后,克隆经PCR筛选和通过单孢子纯化。

[0481] 通过靶向于alg3基因座获得的产生利妥昔单抗的里氏木霉GnTII/GnTI变异体菌株的多糖分析

[0482] 为了获得用于聚糖分析的材料,按照实施例5中所描述的进行摇瓶培养,和此外,一些培养基补充有0.3毫克/毫升大豆胰蛋白酶抑制剂(SBTI)和1%酪蛋白氨基酸。接种时首次添加SBTI,然后在第3-6天每天添加。在冷冻前向所有样品加入PMSF和胃蛋白酶抑制剂A。

[0483] 用蛋白G亲和层析法从具有SBTI的第5天的上清液样品和没有SBTI的第5天和第7天的样品纯化利妥昔单抗。对约10μg的变性蛋白质进行PNGase F反应。首先用Hypersep C-18然后用Hypersep Hypercarb(均来自Thermo Scientific)纯化N-聚糖,其中中性和酸性聚糖分离。以96孔形式进行纯化步骤。通过MALDI-TOF MS分析中性和酸性N-聚糖来测试利妥昔单抗抗体上G0糖形的存在。

序列表

<110> 诺瓦提斯国际药品有限公司
 格利科斯芬兰公司
 NATUNEN, Jari
 KANERVA, Anne
 HILTUNEN, Jukka
 SALOHEIMO, Markku
 VISKARI, Heli
 HUUSKONEN, Anne

<120> 融合酶

<130> 61967-20010.40

<140> 未指明

<141> 与本申请同时

<150> US 61/417,144

<151> 2010-11-24

<160> 147

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 445

<212> PRT

<213> 智人

<400> 1

[0001]

```

Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
1      5      10      15
Phe Val Ala Trp Asn Ala Leu Leu Leu Leu Phe Phe Trp Thr Arg Pro
20     25     30
Ala Pro Gly Arg Pro Pro Ser Val Ser Ala Leu Asp Gly Asp Pro Ala
35     40     45
Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
50     55     60
Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Gly Asp Ala Leu Ser
65     70     75     80
Ser Gln Arg Gly Arg Val Pro Thr Ala Ala Pro Pro Ala Gln Pro Arg
85     90     95
Val Pro Val Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala
100    105    110
Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr
115    120    125
Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp Cys Gly
130    135    140
His Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr
145    150    155    160
His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His
165    170    175
Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala
180    185    190
Leu Gly Gln Val Phe Arg Gln Phe Arg Phe Pro Ala Ala Val Val Val
195    200    205
Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe Arg Ala
210    215    220
Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala
225    230    235    240
Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Ser Arg Pro Glu
245    250    255
Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu
260    265    270
Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp
275    280    285
Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys Ile
290    295    300
Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser
305    310    315    320
His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln
325    330    335
Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu

```

Met	Leu	Lys	Lys	Gln	Ser	Ala	Gly	Leu	Val	Leu	Trp	Gly	Ala	Ile	Leu
1				5					10					15	
Phe	Val	Ala	Trp	Asn	Ala	Leu	Leu	Leu	Phe	Phe	Trp	Gly	Thr	Arg	Pro
			20					25					30		
Ala	Pro	Gly	Arg	Pro	Pro	Ser	Val	Ser	Ala	Leu	Asp	Asp	Asp	Pro	Ala
		35					40				45				
Ser	Leu	Thr	Arg	Glu	Val	Ile	Arg	Leu	Ala	Gln	Asp	Ala	Glu	Val	Glu
	50					55					60				
Leu	Glu	Arg	Gln	Arg	Gly	Leu	Leu	Gln	Gln	Ile	Gly	Asp	Ala	Leu	Ser
65					70					75				80	
Ser	Gln	Arg	Gly	Arg	Val	Pro	Thr	Ala	Ala	Pro	Pro	Ala	Gln	Pro	Arg
				85					90					95	
Val	Pro	Val	Thr	Pro	Ala	Pro	Ala	Val	Ile	Pro	Ile	Leu	Val	Ile	Ala
			100					105					110		
Cys	Asp	Arg	Ser	Thr	Val	Arg	Arg	Cys	Leu	Asp	Lys	Leu	Leu	His	Tyr
		115					120					125			
Arg	Pro	Ser	Ala	Glu	Leu	Phe	Pro	Ile	Ile	Val	Ser	Gln	Asp	Cys	Gly
		130				135					140				
His	Glu	Glu	Thr	Ala	Gln	Ala	Ile	Ala	Ser	Tyr	Gly	Ser	Ala	Val	Thr
145					150					155				160	
His	Ile	Arg	Gln	Pro	Asp	Leu	Ser	Ser	Ile	Ala	Val	Pro	Pro	Asp	His
				165					170					175	
Arg	Lys	Phe	Gln	Gly	Tyr	Tyr	Lys	Ile	Ala	Arg	His	Tyr	Arg	Trp	Ala
			180					185					190		
Leu	Gly	Gln	Val	Phe	Arg	Gln	Phe	Gly	Phe	Pro	Ala	Ala	Val	Val	Val
		195					200					205			
Glu	Asp	Asp	Leu	Glu	Val	Ala	Pro	Asp	Phe	Phe	Glu	Tyr	Phe	Gln	Ala
	210					215					220				
Thr	Tyr	Pro	Leu	Leu	Lys	Ala	Asp	Pro	Ser	Leu	Trp	Cys	Val	Ser	Ala
225					230					235				240	
Trp	Asn	Asp	Asn	Gly	Lys	Glu	Gln	Met	Val	Asp	Ala	Ser	Arg	Pro	Glu
				245					250					255	
Leu	Leu	Tyr	Arg	Thr	Asp	Phe	Phe	Pro	Gly	Leu	Gly	Trp	Leu	Leu	
			260				265						270		
Ala	Glu	Leu	Trp	Ala	Glu	Leu	Glu	Pro	Lys	Trp	Pro	Lys	Ala	Phe	Trp
		275					280					285			
Asp	Asp	Trp	Met	Arg	Arg	Pro	Glu	Gln	Arg	Gln	Gly	Arg	Ala	Cys	Ile
	290					295					300				
Arg	Pro	Glu	Ile	Ser	Arg	Thr	Met	Thr	Phe	Gly	Arg	Lys	Gly	Val	Ser
305				310						315				320	
His	Gly	Gln	Phe	Phe	Asp	Gln	His	Leu	Lys	Phe	Ile	Lys	Leu	Asn	Gln
				325					330					335	
Gln	Phe	Val	His	Phe	Thr	Gln	Leu	Asp	Leu	Ser	Tyr	Leu	Gln	Arg	Glu
			340					345					350		
Ala	Tyr	Asp	Arg												

435

440

445

<210> 3
 <211> 445
 <212> PRT
 <213> 苏门答腊猩猩

<400> 3
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ala Pro Gly Arg Pro Pro Ser Val Ser Ala Leu Asp Asp Asp Pro Ala
 35 40 45
 Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Gly Asp Ala Leu Trp
 65 70 75 80
 Ser Gln Arg Gly Arg Val Pro Thr Pro Ala Leu Pro Ala Gln Pro Arg
 85 90 95
 Val Pro Ala Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala
 100 105 110
 Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu Gln Tyr
 115 120 125
 Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp Cys Gly
 130 135 140
 His Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr
 145 150 155 160
 His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His
 165 170 175
 Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala
 180 185 190
 Leu Gly Gln Ile Phe Gln Arg Phe Arg Phe Pro Ala Ala Val Val Val
 195 200 205
 Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe Gln Ala
 210 215 220
 Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala
 225 230 235 240
 Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Ser Lys Pro Glu
 245 250 255
 Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu
 260 265 270
 Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp
 275 280 285
 Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala Cys Ile
 290 295 300
 Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser
 305 310 315 320
 His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln
 325 330 335
 Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu
 340 345 350
 Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro Gln Leu
 355 360 365
 Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu Val
 370 375 380
 Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala
 385 390 395 400
 Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr
 405 410 415
 Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His Leu Ala
 420 425 430
 Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
 435 440 445

[0003]

<210> 4
 <211> 445
 <212> PRT
 <213> 恒河猴

<400> 4
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Leu Phe Phe Trp Thr Arg Pro

[0004]

```

      20      25      30
Ala Pro Gly Arg Pro Pro Ser Val Ser Ala Leu Asn Asp Asp Pro Ala
      35      40      45
Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
      50      55      60
Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Gly Asp Ala Leu Trp
      65      70      75      80
Ser Gln Arg Gly Arg Val Pro Thr Ala Gly Pro Pro Ala Gln Pro His
      85      90      95
Val Pro Val Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala
      100      105      110
Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr
      115      120      125
Arg Pro Ser Ala Glu Arg Phe Pro Ile Ile Val Ser Gln Asp Cys Gly
      130      135      140
His Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr
      145      150      155      160
His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His
      165      170      175
Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala
      180      185      190
Leu Gly Gln Val Phe His Arg Phe Arg Phe Pro Ala Ala Val Val Val
      195      200      205
Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe Gln Ala
      210      215      220
Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala
      225      230      235      240
Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Gly Lys Pro Glu
      245      250      255
Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu
      260      265      270
Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp
      275      280      285
Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala Cys Ile
      290      295      300
Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser
      305      310      315      320
His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln
      325      330      335
Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu
      340      345      350
Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Ala Ala Pro Gln Leu
      355      360      365
Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu Val
      370      375      380
Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala
      385      390      395      400
Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr
      405      410      415
Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His Leu Ala
      420      425      430
Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
      435      440      445

```

<210> 5
 <211> 447
 <212> PRT
 <213> 中国仓鼠

```

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

```

115 120 125
 His Tyr Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp
 130 135 140
 Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala
 145 150 155 160
 Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Ala Val Pro Pro
 165 170 175
 Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
 180 185 190
 Trp Ala Leu Gly Gln Ile Phe Asn Lys Phe Lys Phe Pro Ala Ala Val
 195 200 205
 Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
 210 215 220
 Gln Ala Thr Tyr Pro Leu Arg Thr Asp Pro Ser Leu Trp Cys Val
 225 230 235 240
 Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
 245 250 255
 Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
 260 265 270
 Leu Met Ala Glu Leu Trp Thr Glu Leu Glu Pro Lys Trp Pro Lys Ala
 275 280 285
 Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala
 290 295 300
 Cys Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
 305 310 315 320
 Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
 325 330 335
 Asn Gln Gln Phe Val Ser Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln
 340 345 350
 Arg Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Ser Ala Pro
 355 360 365
 Leu Leu Gln Val Glu Lys Val Arg Thr Asn Asp Gln Lys Glu Leu Gly
 370 375 380
 Glu Val Arg Val Gln Tyr Thr Ser Arg Asp Ser Phe Lys Ala Phe Ala
 385 390 395 400
 Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
 405 410 415
 Gly Tyr Arg Gly Val Val Thr Phe Gln Phe Arg Gly Arg Arg Val His
 420 425 430
 Leu Ala Pro Pro Gln Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
 435 440 445

[0005]

<210> 6
 <211> 447
 <212> PRT
 <213> 褐家鼠

<400> 6
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Ile
 1 5 10 15
 Phe Val Gly Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ala Pro Gly Arg Leu Pro Ser Asp Ser Ala Leu Gly Asp Asp Pro Ala
 35 40 45
 Ser Leu Thr Arg Glu Val Ile His Leu Ala Glu Asp Ala Glu Ala Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Lys Glu His Tyr Ser
 65 70 75 80
 Leu Trp Arg Gln Arg Trp Arg Val Pro Thr Val Ala Pro Pro Ala Trp
 85 90 95
 Pro Arg Val Pro Gly Thr Pro Ser Pro Ala Val Ile Pro Ile Leu Val
 100 105 110
 Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
 115 120 125
 His Tyr Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp
 130 135 140
 Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Thr Ala
 145 150 155 160
 Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Ala Val Gln Pro
 165 170 175
 Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
 180 185 190
 Trp Ala Leu Gly Gln Ile Phe Asn Lys Phe Lys Phe Pro Ala Ala Val
 195 200 205
 Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe

```

      210      215      220
Gln Ala Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val
225      230      235      240
Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
      245      250      255
Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
      260      265      270
Leu Leu Ala Asp Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
      275      280      285
Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala
      290      295      300
Cys Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
305      310      315      320
Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
      325      330      335
Asn Gln Gln Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln
      340      345      350
Arg Glu Ala Tyr Asp Arg Asp Phe Leu Ala Gln Val Tyr Gly Ala Pro
      355      360      365
Gln Leu Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly
      370      375      380
Glu Val Arg Val Gln Tyr Thr Ser Arg Asp Ser Phe Lys Ala Phe Ala
385      390      395      400
Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
      405      410      415
Gly Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His
      420      425      430
Leu Ala Pro Pro Glu Thr Trp Asn Gly Tyr Asp Pro Ser Trp Asn
      435      440      445

```

<210> 7
 <211> 447
 <212> PRT
 <213> 小家鼠

[0006]

```

      <400> 7
Met Leu Lys Lys Gln Thr Ala Gly Leu Val Leu Trp Gly Ala Ile Ile
  1      5      10      15
Phe Val Gly Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
      20      25      30
Ala Pro Gly Arg Leu Pro Ser Asp Ser Ala Leu Gly Asp Asp Pro Ala
      35      40      45
Ser Leu Thr Arg Glu Val Ile His Leu Ala Glu Asp Ala Glu Ala Glu
      50      55      60
Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Lys Glu His Tyr Ala
      65      70      75      80
Leu Trp Arg Gln Arg Trp Arg Val Pro Thr Val Ala Pro Pro Ala Trp
      85      90      95
Pro Arg Val Pro Val Thr Pro Ser Pro Val Gln Ile Pro Ile Leu Val
      100      105      110
Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
      115      120      125
His Tyr Arg Pro Ser Ala Glu Arg Phe Pro Ile Ile Val Ser Gln Asp
      130      135      140
Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Thr Ala
      145      150      155      160
Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Ala Val Gln Pro
      165      170      175
Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
      180      185      190
Trp Ala Leu Gly Gln Ile Phe Asn Lys Phe Lys Phe Pro Ala Ala Val
      195      200      205
Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
      210      215      220
Gln Ala Thr Tyr Pro Leu Leu Arg Thr Asp Pro Ser Leu Trp Cys Val
      225      230      235      240
Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
      245      250      255
Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
      260      265      270
Leu Leu Ala Asp Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
      275      280      285
Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala
      290      295      300
Cys Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly

```

```

305          310          315          320
Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
          325          330          335
Asn Gln Gln Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln
          340          345          350
Gln Glu Ala Tyr Asp Arg Asp Phe Leu Ala Gln Val Tyr Gly Ala Pro
          355          360          365
Gln Leu Gln Val Glu Lys Val Arg Thr Asn Asp Gln Lys Glu Leu Gly
          370          375          380
Glu Val Arg Val Gln Tyr Thr Ser Arg Asp Ser Phe Lys Ala Phe Ala
          385          390          395          400
Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
          405          410          415
Gly Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His
          420          425          430
Leu Ala Pro Pro Gln Thr Trp Thr Gly Tyr Asp Pro Ser Trp Asn
          435          440          445

```

<210> 8
 <211> 447
 <212> PRT
 <213> 穴兔

[0007]

```

<400> 8
Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
  1          5          10          15
Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
          20          25          30
Val Pro Ser Arg Leu Pro Ser Asp Asn Ala Leu Asp Asp Asp Pro Ala
          35          40          45
Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
          50          55          60
Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Arg Glu His His Ala
          65          70          75          80
Leu Trp Ser Gln Arg Trp Lys Val Pro Thr Ala Ala Pro Pro Ala Gln
          85          90          95
Pro His Val Pro Val Thr Pro Pro Pro Ala Val Ile Pro Ile Leu Val
          100          105          110
Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
          115          120          125
His Tyr Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp
          130          135          140
Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala
          145          150          155          160
Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Ala Val Gln Pro
          165          170          175
Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
          180          185          190
Trp Ala Leu Gly Gln Ile Phe His Asn Phe Asn Tyr Pro Ala Ala Val
          195          200          205
Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Glu Tyr Phe
          210          215          220
Gln Ala Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val
          225          230          235          240
Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
          245          250          255
Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
          260          265          270
Leu Leu Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
          275          280          285
Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala
          290          295          300
Cys Val Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
          305          310          315          320
Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
          325          330          335
Asn Gln Gln Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln
          340          345          350
Gln Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro
          355          360          365
Gln Leu Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly
          370          375          380
Glu Val Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala
          385          390          395          400
Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala

```

405 410 415
 Gly Tyr Arg Gly Ile Val Thr Phe Leu Phe Arg Gly Arg Arg Val His
 420 425 430
 Leu Ala Pro Pro Gln Thr Trp Asp Gly Tyr Asp Pro Ser Trp Thr
 435 440 445

<210> 9
 <211> 447
 <212> PRT
 <213> 金黄地鼠

<400> 9
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Ile
 1 5 10 15
 Phe Val Gly Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ala Pro Gly Arg Pro Pro Leu Asp Ser Ala Leu Asp Asp Pro Ala
 35 40 45
 Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Glu Asp Ala Glu Val Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Arg Glu His His Thr
 65 70 75 80
 Leu Trp Asn Gln Arg Trp Lys Val Pro Thr Val Ala Pro Pro Ala Trp
 85 90 95
 Pro Arg Val Pro Val Thr Pro Ser Pro Pro Val Ile Pro Ile Leu Val
 100 105 110
 Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
 115 120 125
 His Tyr Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp
 130 135 140
 Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala
 145 150 155 160
 Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Ala Val Gln Pro
 165 170 175
 Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
 180 185 190
 Trp Ala Leu Gly Gln Ile Phe Asn Lys Phe Lys Phe Pro Ala Ala Val
 195 200 205
 Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
 210 215 220
 Gln Ala Thr Tyr Pro Leu Leu Arg Thr Asp Pro Ser Leu Trp Cys Val
 225 230 235 240
 Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
 245 250 255
 Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
 260 265 270
 Leu Leu Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
 275 280 285
 Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Lys Gly Arg Ala
 290 295 300
 Cys Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
 305 310 315 320
 Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
 325 330 335
 Asn Gln Gln Phe Val Ser Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln
 340 345 350
 Arg Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro
 355 360 365
 Leu Leu Gln Val Glu Lys Val Arg Thr Asn Asp Gln Lys Glu Leu Gly
 370 375 380
 Glu Val Arg Val Gln Tyr Thr Ser Arg Asp Ser Phe Lys Ala Phe Ala
 385 390 395 400
 Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
 405 410 415
 Gly Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His
 420 425 430
 Leu Ala Pro Pro Arg Ser Trp Glu Gly Tyr Asp Pro Ser Trp Thr
 435 440 445

[0008]

<210> 10
 <211> 445
 <212> PRT
 <213> 大熊猫

<400> 10
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ser Pro Gly Arg Leu Pro Ser Glu Ser Ala Leu Asp Asp Asp Pro Ala
 35 40 45
 Val Leu Thr Arg Glu Val Ile Arg Leu Ala Glu Asp Ala Glu Val Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Arg Glu His His Ala
 65 70 75 80
 Arg Trp Ser Gln Arg Trp Arg Ala Pro Thr Ala Thr Val Pro Ala Pro
 85 90 95
 Ala Pro Ala Ser Asn Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala
 100 105 110
 Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr
 115 120 125
 Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp Cys Gly
 130 135 140
 His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala Val Thr
 145 150 155 160
 His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His
 165 170 175
 Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala
 180 185 190
 Leu Gly Gln Val Phe His Arg Phe Lys Phe Pro Ala Ala Val Val Val
 195 200 205
 Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe Gln Ala
 210 215 220
 Thr Tyr Pro Leu Leu Arg Ala Asp Pro Ser Leu Trp Cys Val Ser Ala
 225 230 235 240
 Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys Pro Glu
 245 250 255
 Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu
 260 265 270
 Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Arg Ala Phe Trp
 275 280 285
 Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys Val
 290 295 300
 Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser
 305 310 315 320
 His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln
 325 330 335
 His Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Arg Gln Glu
 340 345 350
 Thr Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro Leu Leu
 355 360 365
 Gln Val Glu Lys Val Arg Thr Ser Glu Arg Asn Glu Leu Gly Glu Val
 370 375 380
 Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala
 385 390 395 400
 Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr
 405 410 415
 Arg Gly Ile Val Ser Phe Leu Phe Arg Gly Arg Arg Val His Leu Ala
 420 425 430
 Pro Pro Gln Thr Trp Asp Gly Tyr Asp Pro Ser Trp Asn
 435 440 445

[0009]

<210> 11
 <211> 447
 <212> PRT
 <213> 野猪

<400> 11
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ala Pro Gly Arg Leu Pro Ser Asp Ser Ala Leu Asp Asp Asp Pro Ala
 35 40 45
 Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala Glu Val Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Arg Glu His His Ala
 65 70 75 80
 Arg Trp Ser Gln Arg Trp Arg Val Pro Thr Val Ala Pro Pro Val Pro

85 90 95
 Pro Arg Val Pro Val Thr Ser Ala Pro Thr Val Ile Pro Ile Leu Val
 100 105 110
 Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
 115 120 125
 His Tyr Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp
 130 135 140
 Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala
 145 150 155 160
 Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Val Val Pro Pro
 165 170 175
 Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
 180 185 190
 Trp Ala Leu Gly Gln Val Phe Glu Lys Phe Lys Phe Ser Ala Ala Val
 195 200 205
 Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
 210 215 220
 Gln Ala Thr Tyr Pro Leu Leu Arg Ala Asp Pro Ser Leu Trp Cys Val
 225 230 235 240
 Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
 245 250 255
 Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
 260 265 270
 Leu Leu Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
 275 280 285
 Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala
 290 295 300
 Cys Val Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
 305 310 315 320
 Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
 325 330 335
 Asn Gln His Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Arg
 340 345 350
 Arg Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro
 355 360 365
 Leu Leu Gln Val Glu Lys Val Arg Thr Ser Glu Arg Ser Glu Leu Gly
 370 375 380
 Glu Val Arg Val Gln Tyr Thr Ser Arg Asp Ser Phe Lys Ala Phe Ala
 385 390 395 400
 Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
 405 410 415
 Gly Tyr Arg Gly Ile Val Ser Phe Leu Phe Arg Gly Arg Arg Val Tyr
 420 425 430
 Leu Ala Pro Pro Glu Thr Trp Asp Gly Tyr Asp Pro Ser Trp Asn
 435 440 445

[0010]

<210> 12
 <211> 447
 <212> PRT
 <213> 家犬

<400> 12
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Trp Thr Arg Pro
 20 25 30
 Ser Pro Ser Arg Leu Pro Ser Asp Ser Ala Leu Asp Asp Asp Pro Ala
 35 40 45
 Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Glu Asp Ala Glu Val Glu
 50 55 60
 Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Arg Glu His His Ala
 65 70 75 80
 Arg Trp Ser Gln Arg Trp Arg Val Pro Thr Ala Ala Pro Pro Ala Pro
 85 90 95
 Pro Arg Val Pro Val Ser Ser Pro Pro Ala Val Ile Pro Ile Leu Val
 100 105 110
 Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
 115 120 125
 His Tyr Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp
 130 135 140
 Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala
 145 150 155 160
 Ile Thr His Ile Arg Gln Pro Asp Leu Ser Ser Ile Thr Val Pro Pro
 165 170 175
 Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg

Trp	Ala	Leu	Gly	Gln	Val	Phe	His	Lys	Phe	Lys	Phe	Pro	Ala	Ala	Val
180	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Val	Val	Glu	Asp	Asp	Leu	Glu	Val	Ala	Pro	Asp	Phe	Phe	Glu	Tyr	Phe
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Gln	Ala	Thr	Tyr	Pro	Leu	Leu	Arg	Ala	Asp	Pro	Ser	Leu	Trp	Cys	Val
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Ser	Ala	Trp	Asn	Asp	Asn	Gly	Lys	Glu	Gln	Met	Val	Asp	Ser	Ser	Lys
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Pro	Glu	Leu	Leu	Tyr	Arg	Thr	Asp	Phe	Phe	Pro	Gly	Leu	Gly	Trp	Leu
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Leu	Leu	Ala	Glu	Leu	Trp	Ala	Glu	Leu	Glu	Pro	Lys	Trp	Pro	Arg	Ala
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Phe	Trp	Asp	Asp	Trp	Met	Arg	Arg	Pro	Glu	Gln	Arg	Gln	Gly	Arg	Ala
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Cys	Val	Arg	Pro	Glu	Ile	Ser	Arg	Thr	Met	Thr	Phe	Gly	Arg	Lys	Gly
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Val	Ser	His	Gly	Gln	Phe	Phe	Asp	Gln	His	Leu	Lys	Phe	Ile	Lys	Leu
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Asn	Gln	His	Phe	Val	Pro	Phe	Thr	Gln	Leu	Asp	Leu	Ser	Tyr	Leu	Arg
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Gln	Glu	Thr	Tyr	Asp	Arg	Asp	Phe	Leu	Ala	Arg	Val	Tyr	Gly	Ala	Pro
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Leu	Leu	Gln	Val	Glu	Lys	Val	Arg	Thr	Ser	Glu	Arg	Ser	Glu	Leu	Gly
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Glu	Val	Arg	Val	Gln	Tyr	Thr	Gly	Arg	Asp	Ser	Phe	Lys	Ala	Phe	Ala
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Lys	Ala	Leu	Gly	Val	Met	Asp	Asp	Leu	Lys	Ser	Gly	Val	Pro	Arg	Ala
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Gly	Tyr	Arg	Gly	Ile	Val	Ser	Phe	Leu	Phe	Arg	Gly	Arg	Arg	Val	His
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405
Leu	Ala	Pro	Pro	Gln	Thr	Trp	Asp	Gly	Tyr	Asp	Pro	Ser	Trp	Asn	
185	195	210	225	240	255	270	285	300	315	330	345	360	375	390	405

[0011]

<210> 13
 <211> 447
 <212> PRT
 <213> 牛

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

```

      275      280      285
Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala
290      295      300
Cys Val Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
305      310      315      320
Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
      325      330      335
Asn Gln His Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Arg
      340      345      350
Gln Glu Thr Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro
      355      360      365
Leu Leu Gln Val Glu Lys Val Arg Thr Ser Glu Arg Ser Glu Leu Gln
      370      375      380
Glu Val Arg Val Gln Tyr Thr Ser Arg Asp Ser Phe Lys Ala Phe Ala
385      390      395      400
Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
      405      410      415
Gly Tyr Arg Gly Ile Val Ser Phe Leu Tyr Arg Gly Arg Arg Val His
      420      425      430
Leu Ala Pro Pro Gln Thr Trp Asp Gly Tyr Asp Pro Ser Trp Asn
      435      440      445

```

```

<210> 14
<211> 447
<212> PRT
<213> 马

```

[0012]

```

<400> 14
Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
1      5      10      15
Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Trp Met Arg Pro
      20      25      30
Ser Pro Ser Arg Leu Pro Ser Asp Gly Thr Leu Asp Asp Asp Pro Thr
      35      40      45
Gly Leu Thr Arg Lys Val Ile His Leu Ala Gln Asp Val Glu Val Glu
      50      55      60
Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Arg Glu His His Ala
65      70      75      80
Arg Trp Ser Gln Trp Trp Arg Val Pro Thr Val Pro Pro Pro Val Pro
      85      90      95
Pro His Val Ser Val Thr Ser Leu Pro Ala Val Ile Pro Ile Leu Val
      100      105      110
Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
      115      120      125
His Tyr Arg Pro Ser Ala Glu His Phe Pro Ile Ile Val Ser Gln Asp
      130      135      140
Cys Gly His Glu Glu Thr Ala Gln Val Ile Ala Ser Tyr Gly Ser Ala
145      150      155      160
Val Thr His Ile Arg Gln Pro Asp Leu Ser Asn Ile Ala Val Pro Pro
      165      170      175
Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
      180      185      190
Trp Ala Leu Ala Gln Val Phe His Arg Phe Lys Phe Pro Ala Ala Val
      195      200      205
Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
      210      215      220
Gln Ala Thr Tyr Pro Leu Leu Arg Ala Asp Pro Ser Leu Trp Cys Val
      225      230      235
Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ser Ser Lys
      245      250      255
Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
      260      265      270
Leu Leu Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
      275      280      285
Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala
      290      295      300
Cys Val Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Ile Gly
305      310      315      320
Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
      325      330      335
Asn Gln His Phe Val Pro Phe Thr Gln Leu Asp Leu Ser Tyr Leu Arg
      340      345      350
Gln Glu Ala Tyr Asp Lys Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro
      355      360      365
Leu Leu Gln Val Glu Lys Val Arg Thr Gly Glu Arg Ser Glu Leu Gly

```

370	375	380
Glu Val Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala		
385	390	395
Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala		400
	405	410
Gly Tyr Arg Gly Ile Val Ser Phe Leu Phe Arg Gly Arg Arg Val His		415
	420	425
Leu Ala Pro Pro Gln Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn		430
435	440	445

<210> 15
 <211> 447
 <212> PRT
 <213> 灰短尾负鼠

[0013]

<400> 15
 Met Leu Lys Lys Gln Ser Ala Gly Leu Val Leu Trp Gly Ala Ile Leu
 1 5 10 15
 Phe Val Ala Trp Asn Ala Leu Leu Leu Phe Phe Phe Trp Ala Arg Pro
 20 25 30
 Leu Pro Gly Gly Pro Ser Ser Glu Asp Pro Phe Ala Asn Asp Pro Ala
 35 40 45
 Ser Leu Ser Arg Arg Val Ile Arg Leu Ala Gln Glu Ala Glu Ile Glu
 50 55 60
 Leu Glu Arg Gln His Val Leu Leu Gln Gln Ile Gln Lys His Ser Val
 65 70 75 80
 Leu Trp Asn Gln Arg Gln Gln Val Ala Thr Ala Gly Pro Pro Ala Val
 85 90 95
 Ser His Pro Thr Val Ala Pro Thr Thr Phe Val Leu Pro Ile Leu Val
 100 105 110
 Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
 115 120 125
 His Tyr Arg Pro Ser Ala Glu Arg Phe Pro Ile Ile Val Ser Gln Asp
 130 135 140
 Cys Gly His Lys Val Thr Ala Gln Val Ile Ala Ser Tyr Gly Asn Ala
 145 150 155 160
 Ile Met His Ile Lys Gln Pro Asp Leu Ser Ser Ile Pro Val Pro Thr
 165 170 175
 Glu His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
 180 185 190
 Trp Ala Leu Asn Gln Val Phe Arg Thr Phe Lys Tyr Gln Ala Ala Val
 195 200 205
 Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
 210 215 220
 Gln Ala Thr Tyr Pro Leu Leu Arg Thr Asp Pro Ser Leu Trp Cys Val
 225 230 235 240
 Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Lys Arg
 245 250 255
 Pro Asp Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
 260 265 270
 Leu Leu Ala Glu Leu Trp Asp Glu Leu Glu Pro Lys Trp Pro Lys Ala
 275 280 285
 Phe Trp Asp Asp Trp Met Arg Gln Pro Glu Gln Arg Arg Asp Arg Ala
 290 295 300
 Cys Leu Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
 305 310 315 320
 Val Ser Gln Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
 325 330 335
 Asn Gln Gly Phe Val Phe Phe Thr Gln Leu Asp Leu Ser Tyr Leu Lys
 340 345 350
 Gln Glu Ala Tyr Asp Arg Asp Phe Ser Ala Arg Val Tyr Ala Ala Pro
 355 360 365
 Gln Val Gln Val Glu Glu Leu Lys Ser Asn Gln Lys Gln Glu Leu Gly
 370 375 380
 Glu Val Arg Val Gln Tyr Arg Gly Arg Asp Ser Phe Arg Ala Phe Ala
 385 390 395 400
 Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
 405 410 415
 Ser Tyr Arg Gly Ile Val Ser Phe Leu Phe Arg Gly Arg Arg Val Tyr
 420 425 430
 Leu Ala Pro Pro Gln Asp Trp Thr Gly Tyr Asp Pro Ser Trp Ser
 435 440 445

<210> 16

<211> 438
 <212> PRT
 <213> 大西洋鲑

<400> 16

Met Leu Arg Lys Arg Gly Ser Ala Ile Leu Cys Gly Ala Phe Leu Phe
 1 5 10 15
 Val Ala Trp Asn Ala Val Val Val Leu Tyr Leu Trp Gly Arg Pro Leu
 20 25 30
 Ser Gly Arg Glu Glu Arg Glu Met Asp Gly Gly Arg Gly Ala Asp
 35 40 45
 Leu Ala Gly Asp Val Ile His Met Ala Glu Ala Phe Glu Ala Glu Leu
 50 55 60
 Glu Met Gln Arg Lys Ile Leu Leu Gln Ile Gln Gly His Arg Ser Leu
 65 70 75 80
 Trp Glu Gln Pro Asn Glu Asn Gly Ala Ser Arg Ile Gly Pro Pro Gln
 85 90 95
 Val Val Ile Pro Ile Leu Val Ile Ala Cys Asn Arg Val Thr Val Lys
 100 105 110
 Arg Cys Leu Asp Lys Leu Leu Glu Tyr Arg Pro Ser Ala Glu Leu Tyr
 115 120 125
 Pro Ile Ile Val Ser Gln Asp Cys Gly His Ala Glu Thr Ala Gln Val
 130 135 140
 Ile Gly Ser Tyr Gly Ser Gln Val Thr His Leu Lys Gln Pro Asp Leu
 145 150 155 160
 Ser Asp Ile Ala Val Arg Pro Glu His Lys Lys Phe Gln Gly Tyr Tyr
 165 170 175
 Lys Ile Ser Arg His Tyr Arg Trp Ala Leu Asn Gln Val Phe Asn Ser
 180 185 190
 Leu Ser His Ser Ser Val Val Ile Val Glu Asp Asp Leu Glu Val Ala
 195 200 205
 Pro Asp Phe Phe Glu Tyr Phe Arg Ser Leu His Pro Ile Leu Lys Ser
 210 215 220
 Asp Leu Ser Leu Trp Cys Val Ser Ala Trp Asn Asp Asn Gly Arg Asp
 225 230 235 240
 Gly Tyr Val Asp Pro Ala Lys Ala Asp Leu Leu Tyr Arg Thr Asp Phe
 245 250 255
 Phe Pro Gly Leu Gly Trp Met Met Leu Lys Glu Leu Trp Val Glu Leu
 260 265 270
 Glu Pro Lys Trp Pro Gly Ala Phe Trp Asp Asp Trp Met Arg Gln Pro
 275 280 285
 Asp Gln Arg Arg Asp Arg Ala Cys Ile Arg Pro Glu Ile Scr Arg Thr
 290 295 300
 Leu Thr Phe Gly Arg Lys Gly Val Ser Leu Gly Gln Phe Tyr Asp Lys
 305 310 315 320
 Tyr Leu Arg Tyr Ile Lys Leu Asn Ser Glu Phe Val Pro Phe Thr Lys
 325 330 335
 Leu Asp Leu Ala Tyr Leu Lys Glu Glu Lys Tyr Lys Glu Ile Phe Glu
 340 345 350
 Lys Gln Val Tyr Ser Ala Pro Leu Val Lys Tyr Glu Glu Val Gln Arg
 355 360 365
 Gly Gln Leu Lys Gly Ala Gly Pro Phe Cys Leu His Tyr Leu Ser Lys
 370 375 380
 Asp Gly Phe Lys Val Leu Ala Lys Asn Leu Gly Val Met Glu Asp Leu
 385 390 395 400
 Lys Ser Gly Val Pro Arg Thr Gly Tyr Arg Gly Val Val Ser Phe Leu
 405 410 415
 Ser Arg Gly Arg Arg Ile Phe Leu Ala Pro Pro Pro Gly Trp Ser Lys
 420 425 430
 Tyr Asp Pro Thr Trp Ser
 435

[0014]

<210> 17
 <211> 452
 <212> PRT
 <213> 斑马鱼

<400> 17

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

50 55 60
 Val Ala Asp Ala Phe Glu Ala Glu Leu Ala Ala Gln Lys Lys Ile Leu
 65 70 75 80
 Gln Gln Ile Gln Ser His Trp Ser Val Trp Asp Ser Lys Asp Gly Val
 85 90 95
 Ile Pro Glu Lys Ser Lys Ser Glu Val Glu His Thr Ala Pro Val Val
 100 105 110
 Ile Pro Ile Leu Val Ile Ala Cys Asn Arg Val Thr Val Lys Arg Cys
 115 120 125
 Leu Asp Lys Leu Ile Glu His Arg Pro Ser Ala Glu Leu His Pro Ile
 130 135 140
 Ile Val Ser Gln Asp Cys Gly His Arg Glu Thr Ser Asp Val Ile Gly
 145 150 155 160
 Ser Tyr Gly Ser Gln Leu Thr His Ile Lys Gln Pro Asp Leu Ser Asp
 165 170 175
 Val Ala Val Pro Pro Gln His Lys Lys Phe Gln Gly Tyr Tyr Lys Ile
 180 185 190
 Ser Arg His Tyr Lys Trp Ala Leu Ser Gln Val Phe Asn Thr Phe Ser
 195 200 205
 Tyr Ser Ser Val Val Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp
 210 215 220
 Phe Phe Glu Tyr Phe Arg Ala Leu His Pro Met Leu Lys Ser Asp Pro
 225 230 235 240
 Thr Leu Trp Cys Val Ser Ala Trp Asn Asp Asn Gly Arg Asp Gly Phe
 245 250 255
 Val Asp Pro Gly Lys Ala Ser Leu Leu Tyr Arg Thr Asp Phe Phe Pro
 260 265 270
 Gly Leu Gly Trp Met Leu Thr Lys Asp Leu Trp Ala Glu Leu Glu Pro
 275 280 285
 Lys Trp Pro Ala Ser Phe Trp Asp Asp Trp Met Arg His Pro Asp Gln
 290 295 300
 Arg Lys Asp Arg Ser Cys Ile Arg Pro Glu Ile Ser Arg Thr Leu Thr
 305 310 315 320
 Phe Gly Arg Lys Gly Val Ser Leu Gly Gln Phe Tyr Asp Lys Tyr Leu
 325 330 335
 Arg Phe Ile Lys Leu Asn Thr Glu Phe Val Pro Phe Thr Lys Met Asp
 340 345 350
 [0015] Leu Ser Tyr Leu Glu Lys Glu Lys Tyr Asp Glu Ser Phe Glu Lys Glu
 355 360 365
 Val Tyr Ala Ala Ser Leu Val Thr Leu Glu Asp Leu Lys Ser Gly Lys
 370 375 380
 Leu Ser Gly Ser Gly Pro Phe Arg Val Gln Tyr Ser Ser Pro Asp Ser
 385 390 395 400
 Phe Lys Ser Leu Ala Arg Asn Leu Gly Val Met Asp Asp Leu Lys Ser
 405 410 415
 Gly Val Pro Arg Ala Gly Tyr Arg Gly Ala Val Ser Phe Leu Leu Arg
 420 425 430
 Gly Lys Arg Val Tyr Leu Ala Pro Pro Ala Gly Trp Ser Arg Tyr Asp
 435 440 445
 Pro Ser Trp Ser
 450

<210> 18

<211> 448

<212> PRT

<213> 非洲爪蟾

<400> 18

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

130 135 140
 Asp Cys Gly His Glu Glu Thr Gly Lys Val Ile Asp Ser Tyr Gly Asp
 145 150 155 160
 Ala Val Thr His Ile Lys Gln Pro Asp Leu Ser Glu Val Ala Val Pro
 165 170 175
 Pro Glu His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ser Arg His Tyr
 180 185 190
 Arg Trp Ala Leu Asn Gln Ile Phe Lys Ser Met Gly Tyr Lys Ala Ala
 195 200 205
 Ile Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Tyr Glu Tyr
 210 215 220
 Phe Gln Ala Thr Leu Pro Leu Leu Gln Lys Asp Arg Met Leu Trp Cys
 225 230 235 240
 Val Ser Ala Trp Asn Asp Asn Gly Lys Glu Ala Leu Ile Asp Pro Gly
 245 250 255
 Gly Thr Ser Leu Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp
 260 265 270
 Leu Leu Leu Arg Glu Leu Trp Glu Glu Leu Glu Pro Lys Trp Pro Ser
 275 280 285
 Ala Phe Trp Asp Asp Trp Val Arg Arg Pro Glu Gln Arg Leu Asp Arg
 290 295 300
 Ala Cys Val Arg Pro Glu Leu Ser Arg Thr Arg Thr Phe Gly Arg Lys
 305 310 315 320
 Gly Val Ser Gln Gly Gln Phe Phe Asp Gln His Leu Arg Phe Ile Lys
 325 330 335
 Leu Asn Gln Asp Leu Val Pro Phe Thr Lys Met Asp Leu Ser Tyr Leu
 340 345 350
 Leu Lys Asp Thr Tyr Asp Pro Trp Phe Leu Glu Gln Val Tyr Gly Ala
 355 360 365
 Pro Lys Ala Arg Ala Glu Glu Val Leu His Gly Gln Val Pro Gly Gly
 370 375 380
 Arg Thr Val Arg Val Glu Tyr Thr Thr Lys Asp Thr Phe Lys Ala Met
 385 390 395 400
 Ala Arg Ala Phe Gly Val Met Glu Asp Leu Lys Ser Gly Val Ala Arg
 405 410 415
 Ala Ala Tyr Lys Gly Val Val Ser Phe Ser His Arg Gly Arg Arg Val
 420 425 430
 Phe Leu Ala Pro Pro Lys Asp Trp Thr Gly Tyr Asp Pro Leu Trp Asn
 435 440 445

[0016]

<210> 19
 <211> 458
 <212> PRT
 <213> 黑腹果蝇

<400> 19
 Met Arg Thr Arg Lys Val Leu Leu Val Ile Gly Phe Leu Val Thr Trp
 1 5 10 15
 Thr Tyr Ala Thr Tyr Tyr Leu Leu Leu Arg Gln Thr Gly Ile His Thr
 20 25 30
 Ser Arg His Gln Ser Leu Gln Ala Tyr Lys Leu Asn Ser Gln Ala Arg
 35 40 45
 Asp Ala Asn Met Gln Ser His His Leu Ala Lys Asn Val Phe Glu Phe
 50 55 60
 Val Lys Leu Lys Tyr Leu Glu Lys Gln Pro Pro Ser Val Ala Ser Thr
 65 70 75 80
 Pro Gln Ile Ser Ile Ile Ala Ala Glu Ile Ser Ala Glu Leu Pro Glu
 85 90 95
 Gln His Val Ala Lys Ser Ala Thr Ala Arg Ile Pro Thr Lys Thr Tyr
 100 105 110
 Leu Ala Asn Gly Glu Pro Val Phe Pro Val Val Val Phe Ala Cys Asn
 115 120 125
 Arg Val Ser Val Lys Lys Cys Ile Asp Asn Leu Val Gln Tyr Arg Pro
 130 135 140
 Ser Val Glu Gln Phe Pro Ile Ile Val Ser Gln Asp Cys Gly Asp Glu
 145 150 155 160
 Pro Thr Lys Glu Ala Ile Leu Ser Tyr Gly Lys Gln Val Thr Leu Ile
 165 170 175
 Glu Gln Pro Asp Leu Ser Asp Ile Thr Val Leu Pro Lys Glu Lys Lys
 180 185 190
 Phe Lys Gly Tyr Tyr Lys Ile Ala Arg His Tyr Gly Trp Ala Leu Asn
 195 200 205
 Thr Thr Phe Ala Val Gly Phe Glu Phe Val Ile Ile Val Glu Asp Asp
 210 215 220
 Leu Asn Val Ala Pro Asp Phe Phe Glu Tyr Phe Leu Gly Thr His Lys

```

225          230          235          240
Leu Leu Lys Gln Asp Pro Ser Leu Trp Cys Val Ser Ala Trp Asn Asp
          245          250          255
Asn Gly Lys Ala Ala Val Val Asp Ala Ala Gln Pro Glu Leu Tyr
          260          265          270
Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Thr Lys Asp Leu
          275          280          285
Trp Ala Glu Leu Ser Val Lys Trp Pro Lys Ser Phe Trp Asp Asp Trp
          290          295          300
Ile Arg His Pro Ala Gln Arg Lys Asp Arg Val Cys Ile Arg Pro Glu
305          310          315          320
Ile Ser Arg Thr Arg Thr Phe Gly Lys Ile Gly Val Ser Asn Gly Leu
          325          330          335
Phe Phe Asp Lys Tyr Leu Lys His Ile Lys Leu Ser Glu Asp Phe Val
          340          345          350
Gln Phe Thr Lys Ile Asn Met Ser Tyr Leu Leu Lys Asp Asn Tyr Asp
          355          360          365
Asn Thr Phe Leu Arg Arg Val Tyr Thr Tyr Pro Ile Val Thr Tyr Asp
          370          375          380
Glu Leu Arg Arg Asn Leu Ile Arg Ile Glu Gly Pro Val Arg Ile Gln
385          390          395          400
Tyr Thr Thr Arg Glu Gln Tyr Lys Arg Thr Thr Lys Met Leu Gly Leu
          405          410          415
Met Asp Asp Phe Lys Ser Gly Val Pro Arg Thr Ala Tyr His Gly Ile
          420          425          430
Val Ser Phe Tyr Tyr Asn Lys Arg Arg Val His Leu Ala Pro Asn Ala
          435          440          445
Asn Trp Lys Gly Tyr Glu Leu Ser Trp Ser
          450          455

```

<210> 20
 <211> 447
 <212> PRT
 <213> 智人

[0017]

```

<400> 20
Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
1          5          10          15
Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
          20          25          30
Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
          35          40          45
Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
          50          55          60
Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
65          70          75          80
Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
          85          90          95
Gln Leu Asn Phe Asp Asn Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
          100          105          110
Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
          115          120          125
Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
          130          135          140
Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
145          150          155          160
Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
          165          170          175
Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
          180          185          190
Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
          195          200          205
Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
          210          215          220
Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Lys Leu His Phe
225          230          235          240
Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
          245          250          255
Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
          260          265          270
Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
          275          280          285
Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
          290          295          300
Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly

```

```

305          310          315          320
Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
          325          330          335
Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
          340          345          350
Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
          355          360          365
Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
          370          375          380
Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
385          390          395          400
Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
          405          410          415
Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
          420          425          430
Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
          435          440          445

```

<210> 21
 <211> 447
 <212> PRT
 <213> 黑猩猩

[0018]

```

<400> 21
Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
1          5          10          15
Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
          20          25          30
Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
          35          40          45
Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
50          55          60
Arg Arg Val Ser Asn Val Ser Ala Ala Pro Leu Val Pro Ala Val Pro
65          70          75          80
Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
          85          90          95
Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
100          105          110
Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
115          120          125
Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
130          135          140
Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
145          150          155          160
Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
          165          170          175
Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
180          185          190
Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
195          200          205
Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
210          215          220
Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
225          230          235          240
Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Val Leu
          245          250          255
Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
260          265          270
Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
275          280          285
Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
290          295          300
Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
305          310          315          320
Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
          325          330          335
Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
          340          345          350
Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
          355          360          365
Val Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
          370          375          380
Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
385          390          395          400
Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys

```

405 410 415
 Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430
 Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

<210> 22
 <211> 447
 <212> PRT
 <213> 穴兔

<400> 22
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Ala Ala Gly Gly Arg Gly Gly Asp His Pro Ala Val Ser Val Gly Ile
 50 55 60
 Arg Arg Val Ser Asn Glu Ser Ala Ala Pro Leu Val Pro Ala Ala Ala
 65 70 75 80
 Gln Pro Glu Ala Asp Asn Gln Thr Leu Arg Tyr Arg Ser Leu Val Tyr
 85 90 95
 Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Ser
 100 105 110
 Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Leu
 115 120 125
 Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140
 Glu Asp Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Pro Glu Ile
 145 150 155 160
 Asn Gln Leu Ile Ala Gly Val Asp Phe Cys Pro Ile Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Gln Lys Asn Ala Ala Leu Lys Leu
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220
 Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Val Leu Gln Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
 275 280 285
 Leu Ser Leu Gly Thr Tyr Thr Ala Ser Arg Ser Phe His Gly Ile Ala
 290 295 300
 His Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Leu Trp Arg Val Leu Val Pro Gln
 355 360 365
 Val Pro Arg Val Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380
 Thr Cys Arg Pro Phe Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Ser Asn Arg Gln Tyr Met Phe Pro Glu Thr Leu Ile Ile Ser Glu Lys
 405 410 415
 Ser Pro Val Val Ser Ile Ala Ser Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430
 Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

[0019]

<210> 23
 <211> 446
 <212> PRT
 <213> 大熊猫

<400> 23
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Gly Ala Gly Gly Arg Gly Gly Asp His Ser Ala Val Ser Ala Gly Ile
 50 55 60
 Arg Arg Val Ser Asn Asp Ser Ala Ala Pro Leu Val Pro Ala Ala Pro
 65 70 75 80
 Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
 85 90 95
 Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Ser
 100 105 110
 Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
 115 120 125
 Asp Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140
 Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
 145 150 155 160
 Asn Gln Leu Ile Ala Gly Val Asp Phe Cys Pro Val Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Glu Lys Asn Ala Ala Leu Lys Met
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220
 Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Val Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Glu Glu Cys Thr Glu Cys Asp Val
 275 280 285
 [0020] Leu Ser Leu Gly Thr Tyr Thr Ala Val Arg Ser Phe His Gly Ile Ala
 290 295 300
 Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asp Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365
 Val Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380
 Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Asn Asn Lys Gln Tyr Leu Phe Pro Glu Thr Leu Ile Ile Ser Glu Lys
 405 410 415
 Phe Val Ala Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly Asp
 420 425 430
 Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

<210> 24
 <211> 446
 <212> PRT
 <213> 家犬

<400> 24
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Gly Ala Gly Gly Arg Gly Gly Asp His Ser Ala Val Ser Val Gly Ile
 50 55 60
 Arg Arg Gly Ser Asn Glu Ser Ala Ala Pro Leu Val Pro Ala Ala Pro
 65 70 75 80
 Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr

85 90 95
 Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Ser
 100 105 110
 Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
 115 120 125
 Asp Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140
 Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
 145 150 155 160
 Asn Gln Leu Ile Ala Gly Val Asp Phe Cys Pro Val Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Glu Lys Asn Ala Ala Leu Lys Met
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220
 Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Val Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Glu Glu Cys Pro Glu Cys Asp Val
 275 280 285
 Leu Ser Leu Gly Thr Tyr Thr Ala Ile Arg Ser Phe His Gly Ile Ala
 290 295 300
 Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asp Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365
 Val Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His Lys Lys
 370 375 380
 Thr Cys Lys Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Ser Asn Lys Gln Tyr Leu Phe Pro Glu Thr Leu Ile Ile Ser Glu Lys
 405 410 415
 Phe Val Ala Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly Asp
 420 425 430
 Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440 445

[0021]

<210> 25
 <211> 446
 <212> PRT
 <213> 野猪

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

```

      180      185      190
Pro Arg Asp Cys Pro Arg Asp Leu Glu Lys Asn Ala Ala Leu Lys Met
      195      200      205
Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
      210      215      220
Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
      225      230      235      240
Val Trp Glu Arg Val Lys Val Leu Arg Asp Tyr Ala Gly Leu Ile Leu
      245      250      255
Phe Leu Glu Glu Asp His Tyr Val Ala Pro Asp Phe Tyr His Val Phe
      260      265      270
Lys Lys Met Trp Asn Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
      275      280      285
Leu Ser Leu Gly Thr Tyr Thr Thr Val Arg Ser Phe Arg Asp Val Ala
      290      295      300
Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
      305      310      315      320
Leu Ala Leu Thr Arg Asp Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
      325      330      335
Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
      340      345      350
Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
      355      360      365
Val Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
      370      375      380
Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
      385      390      395      400
Ser Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
      405      410      415
Leu Thr Ala Ala Leu Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly Asp
      420      425      430
Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
      435      440      445

```

<210> 26
 <211> 446
 <212> PRT
 <213> 牛

[0022]

```

<400> 26
Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Met Leu Val
  1      5      10      15
Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
      20      25      30
Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Asp Pro Val Arg
      35      40      45
Gly Ala Gly Ala Arg Ala Gly Asp His Pro Ala Val Ser Val Gly Ile
      50      55      60
Arg Arg Gly Ser Asn Glu Ser Ala Ala Pro Leu Val Ala Ala Ala Pro
      65      70      75      80
Gln Pro Glu Val Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
      85      90      95
Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Ala Ser
      100      105      110
Trp Thr Pro Arg Glu Leu Ala Leu Val Val Gln Val His Asn Arg Pro
      115      120      125
Glu Tyr Leu Lys Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
      130      135      140
Asp Asp Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
      145      150      155      160
Asn Gln Leu Ile Ala Gly Val Asp Phe Cys Pro Val Leu Gln Val Phe
      165      170      175
Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Thr Asp
      180      185      190
Pro Arg Asp Cys Pro Arg Asp Met Glu Lys Asn Ala Ala Leu Arg Met
      195      200      205
Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
      210      215      220
Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
      225      230      235      240
Val Trp Glu Arg Val Lys Val Leu Arg Asp Tyr Ala Gly Leu Ile Leu
      245      250      255
Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
      260      265      270
Lys Lys Met Trp Lys Leu Lys Gln Leu Glu Cys Pro Glu Cys Asp Val

```

275	280	285
Leu Ser Leu Gly Thr Tyr Thr	Ala Ile Arg Asn Phe Tyr	Asp Val Ala
290	295	300
Asp Lys Val Asp Val Lys Thr	Trp Lys Ser Thr Glu His	Asn Met Gly
305	310	315
Leu Ala Leu Thr Arg Glu Ala	Tyr Gln Lys Leu Ile Glu Cys	Thr Asp
325	330	335
Thr Phe Cys Thr Tyr Asp Asp	Tyr Asn Trp Asp Trp Thr	Leu Gln Tyr
340	345	350
Leu Thr Val Ser Cys Leu Pro	Lys Phe Trp Lys Val Leu Val	Pro Gln
355	360	365
Val Pro Arg Ile Phe His Ala	Gly Asp Cys Gly Met His His	Gln Lys
370	375	380
Thr Cys Arg Pro Ala Thr Gln	Ser Ala Gln Leu Glu Ser	Leu Leu Asn
385	390	395
Asn Asn Lys Gln Tyr Leu Phe	Pro Glu Thr Leu Thr Ile Ser	Glu Lys
405	410	415
Phe Met Thr Ser Leu Ser Pro	Pro Arg Lys Asn Gly Gly Trp	Gly Asp
420	425	430
Ile Arg Asp His Glu Leu Cys	Lys Ser Tyr Arg Arg Leu Gln	
435	440	445

<210> 27
 <211> 442
 <212> PRT
 <213> 褐家鼠

[0023]

<400> 27
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Asp Ala Leu Ala Pro Pro Leu Leu Asp Ser Glu Pro Leu Arg
 35 40 45
 Gly Ala Gly His Phe Ala Ala Ser Val Gly Ile Arg Arg Val Ser Asn
 50 55 60
 Asp Ser Ala Ala Pro Leu Val Pro Ala Val Pro Arg Pro Glu Val Asp
 65 70 75 80
 Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr Gln Leu Asn Phe Asp
 85 90 95
 Gln Met Leu Arg Asn Val Asp Lys Asp Gly Thr Trp Ser Pro Gly Glu
 100 105 110
 Leu Val Leu Val Val Gln Val His Asn Arg Pro Glu Tyr Leu Arg Leu
 115 120 125
 Leu Ile Asp Ser Leu Arg Lys Ala Gln Gly Ile Arg Glu Val Leu Val
 130 135 140
 Ile Phe Ser His Asp Phe Trp Ser Ala Glu Ile Asn Ser Leu Ile Ser
 145 150 155 160
 Ser Val Asp Phe Cys Pro Val Leu Gln Val Phe Phe Pro Phe Ser Ile
 165 170 175
 Gln Leu Tyr Pro Ser Glu Phe Pro Gly Ser Asp Pro Arg Asp Cys Pro
 180 185 190
 Arg Asp Leu Lys Lys Asn Ala Ala Leu Lys Leu Gly Cys Ile Asn Ala
 195 200 205
 Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu Ala Lys Phe Ser Gln
 210 215 220
 Thr Lys His His Trp Trp Trp Lys Leu His Phe Val Trp Glu Arg Val
 225 230 235 240
 Lys Val Leu Gln Asp Tyr Thr Gly Leu Ile Leu Phe Leu Glu Glu Asp
 245 250 255
 His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe Lys Lys Met Trp Lys
 260 265 270
 Leu Lys Gln Gln Glu Cys Pro Gly Cys Asp Val Leu Ser Leu Gly Thr
 275 280 285
 Tyr Thr Thr Ile Arg Ser Phe Tyr Gly Ile Ala Asp Lys Val Asp Val
 290 295 300
 Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly Leu Ala Leu Thr Arg
 305 310 315 320
 Asp Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp Thr Phe Cys Thr Tyr
 325 330 335
 Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr Leu Thr Leu Ala Cys
 340 345 350
 Leu Pro Lys Val Trp Lys Val Leu Val Pro Gln Ala Pro Arg Ile Phe
 355 360 365
 His Ala Gly Asp Cys Gly Met His His Lys Lys Thr Cys Arg Pro Ser

```

      370      375      380
Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn Asn Asn Lys Gln Tyr
385      390      395      400
Leu Phe Pro Glu Thr Leu Val Ile Gly Glu Lys Phe Pro Met Ala Ala
      405      410      415
Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly Asp Ile Arg Asp His
      420      425      430
Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
      435      440

```

<210> 28
 <211> 442
 <212> PRT
 <213> 小家鼠

[0024]

```

<400> 28
Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
1      5      10      15
Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
      20      25      30
Lys Ser Asp Ala Leu Gly Pro Pro Leu Leu Asp Ala Glu Pro Val Arg
      35      40      45
Gly Ala Gly His Leu Ala Val Ser Val Gly Ile Arg Arg Val Ser Asn
      50      55      60
Glu Ser Ala Ala Pro Leu Val Pro Ala Val Pro Arg Pro Glu Val Asp
      65      70      75      80
Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr Gln Leu Asn Phe Asp
      85      90      95
Gln Met Leu Arg Asn Val Gly Asn Asp Gly Thr Trp Ser Pro Gly Glu
      100      105      110
Leu Val Leu Val Val Gln Val His Asn Arg Pro Glu Tyr Leu Arg Leu
      115      120      125
Leu Ile Asp Ser Leu Arg Lys Ala Gln Gly Ile Gln Glu Val Leu Val
      130      135      140
Ile Phe Ser His Asp Phe Trp Ser Ala Glu Ile Asn Ser Leu Ile Ser
      145      150      155      160
Arg Val Asp Phe Cys Pro Val Leu Gln Val Phe Phe Pro Phe Ser Ile
      165      170      175
Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp Pro Arg Asp Cys Pro
      180      185      190
Arg Asp Leu Lys Lys Asn Ala Ala Leu Lys Leu Gly Cys Ile Asn Ala
      195      200      205
Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu Ala Lys Phe Ser Gln
      210      215      220
Thr Lys His His Trp Trp Trp Lys Leu His Phe Val Trp Glu Arg Val
      225      230      235      240
Lys Val Leu Gln Asp Tyr Thr Gly Leu Ile Leu Phe Leu Glu Glu Asp
      245      250      255
His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe Lys Lys Met Trp Lys
      260      265      270
Leu Lys Gln Gln Glu Cys Pro Gly Cys Asp Val Leu Ser Leu Gly Thr
      275      280      285
Tyr Thr Thr Ile Arg Ser Phe Tyr Gly Ile Ala Asp Lys Val Asp Val
      290      295      300
Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly Leu Ala Leu Thr Arg
      305      310      315      320
Asp Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp Thr Phe Cys Thr Tyr
      325      330      335
Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr Leu Thr Leu Ala Cys
      340      345      350
Leu Pro Lys Ile Trp Lys Val Leu Val Pro Gln Ala Pro Arg Ile Phe
      355      360      365
His Ala Gly Asp Cys Gly Met His His Lys Lys Thr Cys Arg Pro Ser
      370      375      380
Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn Ser Asn Lys Gln Tyr
      385      390      395      400
Leu Phe Pro Glu Thr Leu Val Ile Gly Glu Lys Phe Pro Met Ala Ala
      405      410      415
Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly Asp Ile Arg Asp His
      420      425      430
Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
      435      440

```

<210> 29

<211> 444
 <212> PRT
 <213> 灰短尾负鼠

<400> 29

Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Phe Trp Asn Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Phe Ala Gly Ser Val Pro Ala Pro Val Arg Ala Val
 35 40 45
 Gly Pro Gly Asp Leu Arg Arg Phe Pro Asn Gly Ser Ala Ala Pro Pro
 50 55 60
 Pro Glu Val Asp Asn Met Thr Leu Val Tyr Arg Ser Leu Val Tyr Gln
 65 70 75 80
 Val Asn Phe Asp Gln Thr Leu Lys Asn Ala Leu Ala Ala Ala Val
 85 90 95
 Gly Ala Gly Gly Ala Gly Gly Gly Gly Gly Pro Ala Gln Leu Glu
 100 105 110
 Leu Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro Asp Tyr Leu
 115 120 125
 Lys Leu Leu Leu Asp Ser Leu Arg Lys Val Gln Gly Ile Gly Asn Leu
 130 135 140
 Leu Val Ile Phe Ser His Asp Phe Trp Ser Ala Glu Ile Asn Gln Leu
 145 150 155 160
 Ile Ala Gly Val Asp Phe Cys Pro Val Leu Gln Val Phe Phe Pro Phe
 165 170 175
 Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Asn Asp Pro Lys Asp
 180 185 190
 Cys Pro Arg Asp Leu Gln Lys Lys Ala Ala Leu Lys Met Gly Cys Ile
 195 200 205
 Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu Ala Lys Phe
 210 215 220
 Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe Ala Trp Glu
 225 230 235 240
 Arg Val Lys Ile Leu Arg Asn Tyr Ala Gly Leu Met Val Phe Leu Glu
 245 250 255
 Glu Asp His Tyr Leu Ala Pro Asp Phe Phe His Val Leu Lys Lys Met
 260 265 270
 Trp Lys Leu Lys Leu Gln Glu Cys Pro Asp Cys Asp Val Leu Ser Leu
 275 280 285
 Gly Ser Tyr Ala Val Ser Arg Ser Phe Phe Gly Lys Ala Asp Lys Val
 290 295 300
 Glu Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly Leu Ala Leu
 305 310 315 320
 Thr Arg Asp Thr Tyr Gln Lys Leu Ile Glu Cys Thr Asp Thr Phe Cys
 325 330 335
 Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr Leu Thr Thr
 340 345 350
 Thr Cys Leu Lys Asn Phe Trp Lys Val Met Val Pro Glu Val Pro Arg
 355 360 365
 Ile Tyr His Ala Gly Asp Cys Gly Met His His Lys Asp Pro Cys Arg
 370 375 380
 Pro Ser Thr Gln Ser Ala Gln Ile Glu Leu Leu Leu Asn Lys Asn Lys
 385 390 395 400
 Gln Tyr Leu Phe Pro Lys Thr Leu Ser Ile Ser Lys Lys Tyr Ser Met
 405 410 415
 Val Pro Leu Leu Pro His Gly Lys Asn Gly Gly Trp Gly Asp Ile Arg
 420 425 430
 Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln
 435 440

[0025]

<210> 30
 <211> 432
 <212> PRT
 <213> 非洲爪蟾

<400> 30

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

50 55 60
 Lys Glu Gln Lys Thr Arg Val Asp Asn Met Thr Leu Val Tyr Arg Ser
 65 70 75 80
 Val Val Phe Gln Trp Asn Phe Asp Gln Ala Ile Arg Asn Val Asp Lys
 85 90 95
 Ile Asn Arg Pro Gln Asp Asp Val Val Val Val Gln Val His Asn
 100 105 110
 Arg Pro Glu Phe Leu Arg Arg Leu Leu Asp Ser Leu Gly Lys Ala Lys
 115 120 125
 Gly Ile Glu Asn Val Leu Leu Val Phe Ser His Asp Tyr Trp Ser Pro
 130 135 140
 Glu Ile Asn Gln Ile Ile Ala Ser Val Asp Phe Cys Gln Val Leu Gln
 145 150 155 160
 Ile Phe Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly
 165 170 175
 His Asp Pro Lys Asp Cys Pro Arg Asp Ile Lys Lys Lys Asp Ala Val
 180 185 190
 Glu Leu Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr
 195 200 205
 Arg Glu Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu
 210 215 220
 Gln Phe Val Trp Asp Lys Leu Lys Val Leu Lys Glu His Asn Gly Leu
 225 230 235 240
 Val Leu Phe Ile Glu Glu Asp His Tyr Leu Ser Pro Asp Phe Tyr Tyr
 245 250 255
 Thr Leu Lys Lys Met Trp Ser Lys Lys Asn Glu Glu Cys Pro Asp Cys
 260 265 270
 Asp Met Leu Cys Leu Gly Thr Tyr Ala His Thr Pro Phe Ala Asp Lys
 275 280 285
 Ala Gly Lys Val Glu Val Lys Thr Trp Lys Ser Thr Glu His Asn Met
 290 295 300
 Gly Met Ala Met Asn Arg Glu Thr Tyr Lys Lys Leu Val Ala Cys Ser
 305 310 315 320
 Glu Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln
 325 330 335
 Tyr Leu Thr Val Asn Cys Leu Pro Lys Phe Trp Lys Val Met Val Pro
 340 345 350
 [0026] Glu Val Pro Arg Ile Tyr His Ile Gly Asp Cys Gly Met His His Asn
 355 360 365
 Lys Pro Cys Arg Pro Thr Thr Glu Ser Ala Lys Leu Glu Ala Leu Phe
 370 375 380
 Thr Ser Asn Gln Arg Asp Leu Phe Pro Glu Lys Ile Asp Ile Ser Arg
 385 390 395 400
 Arg Tyr Thr Met Ala Ala Leu Ser Pro His Val Lys Asn Gly Gly Trp
 405 410 415
 Gly Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr His Arg Leu Gln
 420 425 430

<210> 31
 <211> 450
 <212> PRT
 <213> 斑马鱼

<400> 31
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Val Ile Leu Thr Leu Val
 1 5 10 15
 Val Ile Ile Cys Gly Phe Ala Val Trp Asn Ser Gly Lys Pro Lys Lys
 20 25 30
 Ala Ser Thr Val Phe Pro Lys Glu Val Glu Thr Val Lys Arg Ser Ser
 35 40 45
 Val Gly Ser Gln Ile Gln Ala Thr Ile Pro Val Thr Arg Lys Pro Ile
 50 55 60
 Asn Glu Ser Ile Pro Glu Lys Gln Gln Gln Gln Gln Pro Val Ala Lys
 65 70 75 80
 Ser Glu Ala Asp Asn Thr Thr Leu Val Tyr Arg Gly Ile Val Phe Gln
 85 90 95
 Leu Asn Phe Asp Gln Asn Leu Lys Asn Glu Glu Lys Phe Arg Ala Val
 100 105 110
 Arg Gln Lys Asp Asp Leu Val Ile Val Val Gln Val His Asn Arg Pro
 115 120 125
 Glu Tyr Leu Arg Leu Leu Val Asp Ser Leu Arg Lys Ser Lys Gly Ile
 130 135 140
 Glu Asn Ile Leu Leu Ile Phe Ser His Asp Phe Trp Ser Pro Glu Ile
 145 150 155 160
 Asn Gln Ile Val Ala Ser Val Asp Phe Cys Leu Val Leu Gln Ile Phe

```

165      170      175
Phe Pro Phe Ser Ile Gln Leu Tyr Pro Gln Glu Phe Pro Gly Asn Asp
180      185      190
Pro Arg Asp Cys Pro Arg Asp Ile Pro Lys Lys Glu Ala Leu Thr Leu
195      200      205
Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
210      215      220
Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
225      230      235      240
Val Trp Asp Arg Val Arg Val Leu Lys Asp His Lys Gly Leu Val Leu
245      250      255
Leu Ile Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Leu Leu
260      265      270
Lys Leu Met Ala Ser Leu Lys Lys Glu Gln Cys Pro Asp Cys Asp Ile
275      280      285
Leu Ser Leu Gly Ser Tyr Gly His Ile Gly Tyr Ser Ser Lys Ala Asn
290      295      300
Lys Val Glu Val Lys Ala Trp Lys Ser Thr Glu His Asn Met Gly Met
305      310      315      320
Ala Leu Asn Arg Asp Ala Tyr Gln Lys Leu Leu Arg Cys Thr Asp Ala
325      330      335
Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Ser Leu Gln His Leu
340      345      350
Thr Val Thr Cys Leu Pro Ala Phe Leu Lys Val Met Val Ser Glu Ala
355      360      365
Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys Ser
370      375      380
Ala Cys Met Pro Ser Gly Gln Lys Thr Lys Ile Glu Asn Val Leu Gln
385      390      395      400
Asn Ser Gly Asn Gln Leu Phe Pro Lys Gln Leu Leu Ile Thr Lys Arg
405      410      415
Leu Pro Ala Ser Gly Ala Lys Gly Val Ala Pro His Val Lys Asn Gly
420      425      430
Gly Trp Gly Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Leu Arg
435      440      445
Leu Gln
450

```

[0027]

```

<210> 32
<211> 465
<212> PRT
<213> 大西洋鲑

```

```

<400> 32
Met Arg Phe Arg Val Tyr Lys Arg Lys Val Val Ile Leu Thr Leu Val
1      5      10      15
Val Val Val Cys Gly Leu Ala Phe Trp Thr Ser Gly Lys Gln Lys Lys
20      25      30
Ser Ser Gly Val Val Val Leu Lys Glu Ala Glu Gly Ala Arg Arg Ser
35      40      45
Ser Ser Ser Gln Val Gln Pro Gln Pro Gln Ala Thr Pro Glu Val Ser
50      55      60
Arg Ile Pro Asn Val Pro Pro Ile Ala Pro Val Asn Glu Thr His Pro
65      70      75      80
Lys Asn Gln Pro Glu Lys His Leu Glu Lys Glu Glu Val Val Lys Pro
85      90      95
Glu Val Asp Asn Thr Thr Gln Val Tyr Arg Gly Ile Val Phe Gln Leu
100      105      110
Asn Phe Asp Gln Thr Val Arg His Glu Glu Lys Phe Arg Ala Ala Arg
115      120      125
Lys Lys Asp Asp Leu Val Val Val Gln Val His Asn Arg Pro Asp
130      135      140
Tyr Leu Arg Leu Leu Val Glu Ser Leu Arg Lys Ala Arg Gly Val Glu
145      150      155      160
Ser Ile Leu Leu Ile Phe Ser His Asp Phe Trp Ser Pro Glu Ile Asn
165      170      175
Gln Val Val Ala Ser Val Asp Phe Cys Gln Val Leu Gln Ile Phe Phe
180      185      190
Pro Phe Ser Ile Gln Leu Tyr Pro Gln Glu Phe Pro Gly His Asp Pro
195      200      205
Arg Asp Cys Pro Arg Asp Ile Ser Lys Ile Asp Ala Leu Lys Leu Gly
210      215      220
Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu Ala
225      230      235      240
Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe Val

```

245 250 255
 Trp Asp Arg Val Arg Ala Leu Lys Asp His Arg Gly Leu Val Leu Leu
 260 265 270
 Ile Glu Glu Asp His Phe Leu Ser Pro Asp Phe Leu His Phe Leu Lys
 275 280 285
 Leu Met Ser Ile Leu Lys Arg Glu Asn Cys Pro Asp Cys Asp Ile Leu
 290 295 300
 Ser Leu Gly Ser Tyr Gly His Ile Ser Tyr Pro Ser Lys Ala Asn Lys
 305 310 315 320
 Val Glu Val Lys Ala Trp Lys Ser Thr Glu His Asn Met Gly Met Ala
 325 330 335
 Leu Ser Arg Glu Thr Tyr Gln Lys Leu Ile Gln Cys Thr Asp Ala Phe
 340 345 350
 Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Ser Leu Gln His Leu Thr
 355 360 365
 Val Thr Cys Leu Pro Ser Tyr Trp Lys Val Met Val Ser Glu Ala Pro
 370 375 380
 Arg Val Phe His Ala Gly Asp Cys Gly Met His Lys Lys Ser Val
 385 390 395 400
 Cys Met Pro Ser Ser Gln Lys Ser Lys Ile Asp Thr Ile Leu Gln Ser
 405 410 415
 Ser Ser Asn Gln Leu Phe Pro Lys Asn Leu Leu Ile Thr Lys Arg Leu
 420 425 430
 Pro Ala Asn Gly Ala Gly Gly Val Ala Pro His Val Lys Asn Gly Gly
 435 440 445
 Trp Gly Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Pro Arg Leu
 450 455 460
 Gln
 465

<210> 33
 <211> 471
 <212> PRT
 <213> 黑腹果蝇

[0028]

<400> 33
 Met Gly Arg Lys Arg Asn Asn Phe Tyr Met Arg Ser Leu Phe Leu Leu
 1 5 10 15
 Ala Leu Gly Ile Phe Gly Leu Leu Gln Tyr Asn Asn Phe Asn Tyr Leu
 20 25 30
 Asp Ser Arg Asp Asn Val Leu Gly Asp Ala Val Thr Asn Asp Ser Asp
 35 40 45
 Asp Ala Ile Leu Ala Met Val Pro Ala Thr Leu His Lys Tyr Leu Thr
 50 55 60
 Pro His Ser Arg Asn His Ser Ala Ser Gly Ala Gly Ala Leu Asn Gly
 65 70 75 80
 Ala Ala Leu Leu Leu Asn Ala Ser Ser Pro Gly Ala Ala Thr Ala Ser
 85 90 95
 Thr Ile Ser Phe Asp Val Tyr His Pro Pro Asn Ile Thr Glu Ile Lys
 100 105 110
 Arg Gln Ile Val Arg Tyr Asn Asp Met Gln Met Val Leu Asn Glu Asp
 115 120 125
 Val Phe Gly Pro Leu Gln Asn Asp Ser Val Ile Ile Val Val Gln Val
 130 135 140
 His Thr Arg Ile Thr Tyr Leu Arg His Leu Ile Val Ser Leu Ala Gln
 145 150 155 160
 Ala Arg Asp Ile Ser Lys Val Leu Leu Val Phe Ser His Asp Tyr Tyr
 165 170 175
 Asp Asp Asp Ile Asn Asp Leu Val Gln Gln Ile Asp Phe Cys Lys Val
 180 185 190
 Met Gln Ile Phe Tyr Pro Tyr Ser Ile Gln Thr His Pro Asn Glu Tyr
 195 200 205
 Pro Gly Val Asp Pro Asn Asp Cys Pro Arg Asn Ile Lys Lys Glu Gln
 210 215 220
 Ala Leu Ile Thr Asn Cys Asn Asn Ala Met Tyr Pro Asp Leu Tyr Gly
 225 230 235 240
 His Tyr Arg Glu Ala Lys Phe Thr Gln Thr Lys His His Trp Ile Trp
 245 250 255
 Lys Ala Asn Arg Val Phe Asn Glu Leu Glu Val Thr Arg Tyr His Thr
 260 265 270
 Gly Leu Val Leu Phe Leu Glu Glu Asp His Tyr Val Ala Glu Asp Phe
 275 280 285
 Leu Tyr Leu Leu Ala Met Met Gln Gln Arg Thr Lys Asp Leu Cys Pro
 290 295 300
 Gln Cys Asn Val Leu Ser Leu Gly Thr Tyr Leu Lys Thr Phe Asn Tyr

```

305          310          315          320
Tyr Thr Tyr His Ser Lys Val Glu Val Met Pro Trp Val Ser Ser Lys
          325          330          335
His Asn Met Gly Phe Ala Phe Asn Arg Thr Thr Trp Ser Asn Ile Arg
          340          345          350
Lys Cys Ala Arg His Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp
          355          360          365
Ser Leu Gln His Val Ser Gln Gln Cys Leu Arg Arg Lys Leu His Ala
          370          375          380
Met Ile Val Lys Gly Pro Arg Val Phe His Ile Gly Glu Cys Gly Val
385          390          395          400
His His Lys Asn Lys Asn Cys Glu Ser Asn Gln Val Ile Ser Lys Val
          405          410          415
Gln His Val Leu Arg Ile Ala Arg Asn Ser His Gln Leu Phe Pro Arg
          420          425          430
Ser Leu Thr Leu Thr Val Pro Ser Leu Met Lys Lys Ser Lys Leu Arg
          435          440          445
Lys Gly Asn Gly Gly Trp Gly Asp Met Arg Asp His Glu Leu Cys Leu
          450          455          460
Asn Met Thr Leu Ala Thr Arg
465          470

```

<210> 34
 <211> 54
 <212> PRT
 <213> 智人

```

<400> 34
Thr Arg Pro Ala Pro Gly Arg Pro Pro Ser Val Ser Ala Leu Asp Gly
1          5          10          15
Asp Pro Ala Ser Leu Thr Arg Glu Val Ile Arg Leu Ala Gln Asp Ala
          20          25          30
Glu Val Glu Leu Glu Arg Gln Arg Gly Leu Leu Gln Gln Ile Gly Asp
          35          40          45
Ala Leu Ser Ser Gln Arg
50

```

[0029]

<210> 35
 <211> 15
 <212> PRT
 <213> 人工序列

<220>
 <223> Synthesized Construct

```

<400> 35
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1          5          10          15

```

<210> 36
 <211> 10
 <212> PRT
 <213> 人工序列

<220>
 <223> Synthesized Construct

```

<400> 36
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1          5          10

```

<210> 37
 <211> 21
 <212> PRT
 <213> 里氏木霉

```

<400> 37
Ser Ser Ala Ala Thr Ala Thr Ala Ser Ala Thr Val Pro Gly Gly Gly
1          5          10          15
Ser Gly Pro Thr Ser
20

```

<210> 38
 <211> 19
 <212> PRT
 <213> 里氏木霉

<400> 38
 Ser Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro
 1 5 10 15
 Gly Ser Thr

<210> 39
 <211> 10
 <212> PRT
 <213> 里氏木霉

<400> 39
 Met Ala Ser Thr Asn Ala Arg Tyr Val Arg
 1 5 10

<210> 40
 <211> 17
 <212> PRT
 <213> 里氏木霉

<400> 40
 Tyr Leu Leu Ile Ala Phe Phe Thr Ile Leu Val Phe Tyr Phe Val Ser
 1 5 10 15
 Asn

<210> 41
 <211> 379
 <212> PRT
 <213> 里氏木霉

[0030]

<400> 41
 Ser Lys Tyr Glu Gly Val Asp Leu Asn Lys Gly Thr Phe Thr Ala Pro
 1 5 10 15
 Asp Ser Thr Lys Thr Thr Pro Lys Pro Pro Ala Thr Gly Asp Ala Lys
 20 25 30
 Asp Phe Pro Leu Ala Leu Thr Pro Asn Asp Pro Gly Phe Asn Asp Leu
 35 40 45
 Val Gly Ile Ala Pro Gly Pro Arg Met Asn Ala Thr Phe Val Thr Leu
 50 55 60
 Ala Arg Asn Ser Asp Val Trp Asp Ile Ala Arg Ser Ile Arg Gln Val
 65 70 75 80
 Glu Asp Arg Phe Asn Arg Arg Tyr Asn Tyr Asp Trp Val Phe Leu Asn
 85 90 95
 Asp Lys Pro Phe Asp Asn Thr Phe Lys Lys Val Thr Thr Ser Leu Val
 100 105 110
 Ser Gly Lys Thr His Tyr Gly Glu Ile Ala Pro Glu His Trp Ser Phe
 115 120 125
 Pro Asp Trp Ile Asp Gln Asp Lys Ala Lys Lys Val Arg Glu Asp Met
 130 135 140
 Ala Glu Arg Lys Ile Ile Tyr Gly Asp Ser Val Ser Tyr Arg His Met
 145 150 155 160
 Cys Arg Phe Glu Ser Gly Phe Phe Phe Arg Gln Pro Leu Met Met Asn
 165 170 175
 Tyr Glu Tyr Tyr Trp Arg Val Glu Pro Ser Ile Glu Leu Tyr Cys Asp
 180 185 190
 Ile His Tyr Asp Pro Phe Arg Leu Met Val Glu Gln Gly Lys Lys Tyr
 195 200 205
 Ser Phe Val Ile Ser Leu Tyr Glu Tyr Pro Ala Thr Ile Ala Thr Leu
 210 215 220
 Trp Glu Ser Thr Lys Lys Phe Met Lys Asn His Pro Glu His Ile Ala
 225 230 235 240
 Pro Asp Asn Ser Met Arg Phe Leu Ser Asp Asp Gly Gly Glu Thr Tyr
 245 250 255
 Asn Asn Cys His Phe Trp Ser Asn Phe Glu Ile Gly Ser Leu Glu Trp
 260 265 270
 Leu Arg Ser Lys Gln Tyr Ile Asp Phe Phe Glu Ser Leu Asp Lys Asp

275 280 285
 Gly Gly Phe Phe Tyr Glu Arg Trp Gly Asp Ala Pro Val His Ser Ile
 290 295 300
 Ala Ala Gly Leu Met Leu Asn Arg Ser Glu Ile His Phe Phe Asn Asp
 305 310 315 320
 Ile Ala Tyr Trp His Val Pro Phe Thr His Cys Pro Thr Gly Glu Lys
 325 330 335
 Thr Arg Leu Asp Leu Lys Cys His Cys Asp Pro Lys Glu Asn Phe Asp
 340 345 350
 Trp Lys Gly Tyr Ser Cys Thr Ser Arg Phe Phe Glu Met Asn Gly Met
 355 360 365
 Asp Lys Pro Glu Gly Trp Glu Asn Gln Gln Asp
 370 375

<210> 42
 <211> 8
 <212> PRT
 <213> 里氏木霉

<400> 42
 Met Ala Ile Ala Arg Pro Val Arg
 1 5

<210> 43
 <211> 15
 <212> PRT
 <213> 里氏木霉

<400> 43
 Ala Leu Gly Gly Leu Ala Ala Ile Leu Trp Cys Phe Phe Leu Tyr
 1 5 10 15

[0031]

<210> 44
 <211> 371
 <212> PRT
 <213> 里氏木霉

<400> 44
 Gln Leu Leu Arg Pro Ser Ser Ser Tyr Asn Ser Pro Gly Asp Arg Tyr
 1 5 10 15
 Ile Asn Phe Glu Arg Asp Pro Asn Leu Asp Pro Thr Gly Glu Pro Glu
 20 25 30
 Gly Ile Leu Val Arg Thr Ser Asp Arg Tyr Ala Pro Asp Ala Lys Asp
 35 40 45
 Thr Asp Arg Ala Ser Ala Thr Leu Leu Ala Leu Val Arg Asn Glu Glu
 50 55 60
 Val Asp Asp Met Val Ala Ser Met Val Asp Leu Glu Arg Thr Trp Asn
 65 70 75 80
 Ser Lys Phe Asn Tyr Pro Trp Thr Phe Phe Asn Asp Lys Pro Phe Ser
 85 90 95
 Glu Glu Phe Lys Lys Lys Thr Ser Ala Val Thr Asn Ala Thr Cys Asn
 100 105 110
 Tyr Glu Leu Ile Pro Lys Glu His Trp Asp Ala Pro Ser Trp Ile Asp
 115 120 125
 Pro Ala Ile Phe Glu Glu Ser Ala Ala Val Leu Lys Lys Asn Gly Val
 130 135 140
 Gln Tyr Ala Asn Met Met Ser Tyr His Gln Met Cys Arg Trp Asn Ser
 145 150 155 160
 Gly Met Phe Tyr Lys His Pro Ala Leu Lys Asp Val Arg Tyr Tyr Trp
 165 170 175
 Arg Val Glu Pro Lys Val His Phe Phe Cys Asp Val Asp Tyr Asp Val
 180 185 190
 Phe Arg Tyr Met Gln Asp Asn Asn Lys Thr Tyr Gly Phe Thr Ile Asn
 195 200 205
 Leu Tyr Asp Asp Pro His Thr Leu Pro Thr Leu Trp Pro Gln Thr Ala
 210 215 220
 Lys Phe Leu Ala Asp His Pro Asn Tyr Leu His Glu His Ser Ala Ile
 225 230 235 240
 Lys Trp Val Ile Asp Asp Ala Arg Arg Pro Gln His Asn Arg Glu Ala
 245 250 255
 Gln Gly Phe Ser Thr Cys His Phe Trp Ser Asn Phe Glu Val Ala Asp
 260 265 270
 Met Glu Phe Trp Arg Ser Lys Val Tyr Glu Asp Tyr Phe Glu His Leu

```

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

```

<210> 45
 <211> 6
 <212> PRT
 <213> 里氏木霉

<400> 45
 Met Leu Asn Pro Arg Arg
 1 5

<210> 46
 <211> 15
 <212> PRT
 <213> 里氏木霉

<400> 46
 Ala Leu Ile Ala Ala Ala Phe Ile Leu Thr Val Phe Phe Leu Ile
 1 5 10 15

[0032]

<210> 47
 <211> 352
 <212> PRT
 <213> 里氏木霉

<400> 47
 Ser Arg Ser His Asn Ser Glu Ser Ala Ser Thr Ser Glu Pro Lys Asp
 1 5 10 15
 Ala Glu Ala Glu Ala Leu Ser Ala Ala Asn Ala Gln Gln Arg Ala Ala
 20 25 30
 Pro Pro Pro Pro Gln Lys Pro Met Ile Asp Met Ser Gly Met Ser
 35 40 45
 Thr Tyr Asp Lys Leu Ala Tyr Ala Tyr Glu Tyr Asp Ile Glu Ser Lys
 50 55 60
 Phe Pro Ala Tyr Ile Trp Gln Thr Trp Arg Lys Thr Pro Ser Glu Gly
 65 70 75 80
 Asp Phe Glu Phe Arg Glu Gln Glu Ala Ser Trp Ser Ile Glu His Pro
 85 90 95
 Gly Phe Ile His Glu Val Ile Thr Asp Ser Val Ala Asp Thr Leu Leu
 100 105 110
 Gln Leu Leu Tyr Gly Ser Ile Pro Glu Val Leu Glu Ala Tyr His Ala
 115 120 125
 Leu Pro Leu Pro Val Leu Lys Ala Asp Leu Phe Arg Tyr Leu Ile Leu
 130 135 140
 Tyr Ala Arg Gly Gly Ile Tyr Ser Asp Ile Asp Thr Tyr Ala Ile Arg
 145 150 155 160
 Ser Ala Leu Glu Trp Ile Pro Pro Gln Ile Pro Lys Glu Thr Val Gly
 165 170 175
 Leu Val Ile Gly Ile Glu Ala Asp Pro Asp Arg Pro Asp Trp Ala Asp
 180 185 190
 Trp Tyr Ser Arg Arg Ile Gln Phe Cys Gln Trp Thr Ile Gln Ser Lys
 195 200 205
 Pro Gly His Pro Val Leu Arg Asp Ile Ile Ser Arg Ile Thr Asn Gln
 210 215 220
 Thr Leu Glu Met Lys Lys Ser Gly Lys Leu Ser Ala Phe Gln Gly Asn
 225 230 235 240
 Arg Val Val Asp Leu Thr Gly Pro Ala Val Trp Thr Asp Thr Ile Met
 245 250 255
 Asp Tyr Phe Asn Asp Glu Arg Tyr Phe Asp Met Glu Asn Ser Lys Gly
 260 265 270
 Arg Ile Asp Tyr Arg Asn Phe Thr Gly Met Glu Thr Ser Lys Arg Val

<210> 50
 <211> 362
 <212> PRT
 <213> 里氏木霉

<400>	50															
Asn	Phe	Asp	Trp	Ile	Leu	Ala	Ile	Pro	Asn	Pro	Ala	Scr	Val	Leu	Arg	
1				5					10					15		
Arg	Glu	Pro	Lys	Ala	Pro	Pro	Leu	Pro	Gly	Ser	Thr	Phe	Pro	Gln	Lys	
			20					25					30			
Ile	Trp	Gln	Thr	Trp	Lys	Val	Asp	Pro	Leu	Asn	Phe	Asp	Glu	Arg	Asp	
			35				40					45				
Leu	Val	Thr	Ala	Arg	Thr	Trp	Thr	Thr	Ile	Asn	Pro	Gly	Met	Arg	Tyr	
			50			55					60					
Glu	Val	Val	Thr	Asp	Ala	Asn	Glu	Met	Ala	Tyr	Ile	Glu	Asp	Arg	Tyr	
65					70					75					80	
Gly	Pro	Asn	Gly	Phe	Asp	Arg	Pro	Asp	Ile	Val	Glu	Phe	Tyr	Lys	Met	
				85					90					95		
Ile	Asn	Leu	Pro	Ile	Ile	Lys	Ala	Asp	Leu	Leu	Arg	Tyr	Met	Ile	Met	
			100					105					110			
Tyr	Ala	Glu	Gly	Gly	Ile	Tyr	Ala	Asp	Ile	Asp	Val	Glu	Thr	Met	Lys	
			115				120					125				
Pro	Phe	His	Arg	Phe	Ile	Pro	Asp	Arg	Tyr	Asp	Glu	Lys	Asp	Ile	Asp	
						135					140					
Ile	Ile	Ile	Gly	Val	Glu	Ile	Asp	Gln	Pro	Asp	Phe	Lys	Asp	His	Pro	
145					150					155					160	
Ile	Leu	Gly	Lys	Lys	Ser	Met	Ser	Phe	Cys	Gln	Trp	Thr	Phe	Val	Ala	
				165					170					175		
Arg	Pro	Gln	Gln	Pro	Val	Met	Met	Arg	Leu	Ile	Glu	Asn	Ile	Met	Lys	
			180					185					190			
Trp	Phe	Lys	Thr	Val	Ala	Arg	Asp	Gln	Gly	Val	Pro	Leu	Gly	Glu	Val	
			195				200					205				
Gln	Leu	Asp	Phe	Asp	Gln	Val	Ile	Ser	Gly	Thr	Gly	Pro	Ser	Ala	Phe	
						215					220					
Thr	Lys	Ala	Met	Leu	Glu	Glu	Met	Asn	Arg	Lys	Thr	Lys	Gly	Pro	Lys	
225					230					235					240	
Val	Thr	Trp	Asp	Ala	Phe	His	Asn	Leu	Asp	Glu	Ser	Lys	Leu	Val	Gly	
				245					250					255		
Gly	Val	Leu	Val	Leu	Thr	Val	Glu	Ala	Phe	Cys	Ala	Gly	Gln	Gly	His	
			260					265					270			
Ser	Asp	Ser	Gly	Asn	His	Asn	Ala	Arg	Asn	Ala	Leu	Val	Lys	His	His	
		275					280					285				
Phe	His	Ala	Ser	Asn	Trp	Pro	Ser	Arg	His	Pro	Arg	Tyr	Lys	His	Pro	

290 295 300
 Ala Tyr Gly Gln Val Glu Asp Cys Asn Trp Val Pro Glu Cys Val Arg
 305 310 315 320
 Lys Trp Asp Glu Asp Thr Ser Asn Trp Asp Lys Tyr Ser Glu Asn Glu
 325 330 335
 Gln Lys Lys Ile Leu Gln Asp Ile Glu Asn Ala Arg Leu Glu Arg Glu
 340 345 350
 Arg Gln Gln Gln Ala Leu Ala Ala Leu Pro
 355 360

<210> 51
 <211> 17
 <212> PRT
 <213> 里氏木霉

<400> 51
 Met Ala Arg Pro Met Gly Ser Val Arg Leu Lys Lys Ala Asn Pro Ser
 1 5 10 15
 Thr

<210> 52
 <211> 16
 <212> PRT
 <213> 里氏木霉

<400> 52
 Leu Ile Leu Gly Ala Val Leu Cys Ile Phe Ile Ile Ile Phe Leu Val
 1 5 10 15

<210> 53
 <211> 339
 <212> PRT
 <213> 里氏木霉

[0034]

<400> 53
 Ser Pro Ser Ser Pro Ala Ser Ala Ser Arg Leu Ser Ile Val Ser Ala
 1 5 10 15
 Gln His His Leu Ser Pro Pro Thr Ser Pro Tyr Gln Ser Pro Arg Ser
 20 25 30
 Gly Ala Val Gln Gly Pro Pro Pro Val Thr Arg Tyr Asn Leu Asn Lys
 35 40 45
 Val Thr Val Thr Ser Asp Pro Val Arg Asn Gln Glu His Ile Leu Ile
 50 55 60
 Leu Thr Pro Met Ala Arg Phe Tyr Gln Glu Tyr Trp Asp Asn Leu Leu
 65 70 75 80
 Arg Leu Asn Tyr Pro His Glu Leu Ile Thr Leu Gly Phe Ile Leu Pro
 85 90 95
 Lys Thr Lys Glu Gly Asn Gln Ala Thr Ser Met Leu Gln Lys Gln Ile
 100 105 110
 Gln Lys Thr Gln Asn Tyr Gly Pro Glu Lys Asp Arg Phe Lys Ser Ile
 115 120 125
 Ile Ile Leu Arg Gln Asp Phe Asp Pro Ala Val Val Ser Gln Asp Glu
 130 135 140
 Ser Glu Arg His Lys Leu Ala Asn Gln Lys Ala Arg Arg Glu Val Met
 145 150 155 160
 Ala Lys Ala Arg Asn Ser Leu Leu Phe Thr Thr Leu Gly Pro Ser Thr
 165 170 175
 Ser Trp Val Leu Trp Leu Asp Ala Asp Ile Thr Glu Thr Ala Pro Thr
 180 185 190
 Leu Ile Gln Asp Leu Ala Ser His Asp Lys Pro Ile Ile Val Ala Asn
 195 200 205
 Cys Phe Gln Lys Tyr Tyr Asp Pro Glu Ser Lys Lys Met Ala Glu Arg
 210 215 220
 Pro Tyr Asp Phe Asn Ser Trp Gln Asp Ser Glu Thr Ala Leu Lys Met
 225 230 235 240
 Ala Glu Gln Met Gly Pro Asp Asp Ile Leu Leu Glu Gly Tyr Ala Glu
 245 250 255
 Met Ala Thr Tyr Arg Thr Leu Leu Ala Tyr Met Ser Thr Pro Gly Gly
 260 265 270
 Ser Lys Asp Leu Val Val Pro Leu Asp Gly Val Gly Gly Thr Ala Leu
 275 280 285
 Leu Val Lys Ala Asp Val His Arg Asp Gly Ala Met Phe Pro Pro Phe

290 295 300
 Ala Phe Tyr His Leu Ile Glu Ser Glu Gly Phe Ala Lys Met Ala Lys
 305 310 315 320
 Arg Leu Gly Trp Gln Pro Tyr Gly Leu Pro Asn Tyr Lys Val Tyr His
 325 330 335
 Tyr Asn Glu

<210> 54
 <211> 31
 <212> PRT
 <213> 里氏木霉

<400> 54
 Met Leu Leu Pro Lys Gly Gly Leu Asp Trp Arg Ser Ala Arg Ala Gln
 1 5 10 15
 Ile Pro Pro Thr Arg Ala Leu Trp Asn Ala Val Thr Arg Thr Arg
 20 25 30

<210> 55
 <211> 15
 <212> PRT
 <213> 里氏木霉

<400> 55
 Phe Ile Leu Leu Val Gly Ile Thr Gly Leu Ile Leu Leu Leu Trp
 1 5 10 15

<210> 56
 <211> 358
 <212> PRT
 <213> 里氏木霉

[0035]

<400> 56
 Arg Gly Val Ser Thr Ser Ala Ser Glu Met Gln Ser Phe Tyr Cys Trp
 1 5 10 15
 Gly Pro Ala Lys Pro Pro Met Glu Met Ser Pro Asn Glu His Asn Arg
 20 25 30
 Trp Asn Gly His Leu Gln Thr Pro Val Ile Phe Asn His His Ala Pro
 35 40 45
 Val Glu Val Asn Ser Ser Thr Ile Glu His Val Asp Leu Asn Pro Ile
 50 55 60
 Asn Ser Thr Lys Gln Ala Val Thr Lys Glu Glu Arg Ile Leu Ile Leu
 65 70 75 80
 Thr Pro Leu Lys Asp Ala Ala Pro Tyr Leu Ser Lys Tyr Phe Glu Leu
 85 90 95
 Leu Ala Glu Leu Thr Tyr Pro His Arg Leu Ile Asp Leu Ala Phe Leu
 100 105 110
 Val Ser Asp Ser Thr Asp Asp Thr Leu Ala Val Leu Ala Ser Glu Leu
 115 120 125
 Asp Arg Ile Gln Lys Arg Pro Asp Gln Ile Pro Phe His Ser Ala Thr
 130 135 140
 Val Ile Glu Lys Asp Phe Gly Phe Lys Leu Ser Gln Asn Val Glu Glu
 145 150 155 160
 Arg His Ser Phe Glu Ala Gln Gly Pro Arg Arg Lys Ala Met Gly Arg
 165 170 175
 Ala Arg Asn Tyr Leu Leu Tyr Thr Ala Leu Lys Pro Glu His Ser Trp
 180 185 190
 Val Tyr Trp Arg Asp Val Asp Ile Val Asp Ser Pro Thr Gly Ile Leu
 195 200 205
 Glu Asp Phe Ile Ala His Asp Arg Asp Ile Leu Val Pro Asn Ile Trp
 210 215 220
 Phe His Arg Tyr Arg Asp Gly Val Asp Ile Glu Gly Arg Phe Asp Tyr
 225 230 235 240
 Asn Ser Trp Val Glu Ser Asp Lys Gly Arg Lys Leu Ala Asn Ser Leu
 245 250 255
 Asp Lys Asp Val Val Leu Ala Glu Gly Tyr Lys Gln Tyr Asp Thr Gly
 260 265 270
 Arg Thr Tyr Met Ala Lys Met Gly Asp Trp Arg Glu Asn Lys Asp Val
 275 280 285
 Glu Leu Glu Leu Asp Gly Ile Gly Gly Val Asn Ile Leu Val Lys Ala
 290 295 300
 Asp Val His Arg Ser Gly Ile Asn Phe Pro Cys Tyr Ala Phe Glu Asn

305 310 315 320
 Gln Ala Glu Thr Glu Gly Phe Ala Lys Met Ala Lys Arg Ala Gly Tyr
 325 330 335
 Glu Val Tyr Gly Leu Pro Asn Tyr Val Val Trp His Ile Asp Thr Glu
 340 345 350
 Glu Lys Gly Gly Asn Ala
 355

<210> 57
 <211> 45
 <212> PRT
 <213> 里氏木霉

<400> 57
 Met Met Pro Arg His His Ser Ser Gly Phe Ser Asn Gly Tyr Pro Arg
 1 5 10 15
 Ala Asp Thr Phe Glu Ile Ser Pro His Arg Phe Gln Pro Arg Ala Thr
 20 25 30
 Leu Pro Pro His Arg Lys Arg Lys Arg Thr Ala Ile Arg
 35 40 45

<210> 58
 <211> 16
 <212> PRT
 <213> 里氏木霉

<400> 58
 Val Gly Ile Ala Val Val Val Ile Leu Val Leu Val Leu Trp Phe Gly
 1 5 10 15

<210> 59
 <211> 407
 <212> PRT
 <213> 里氏木霉

[0036]

<400> 59
 Gln Pro Arg Ser Val Ala Ser Leu Ile Ser Leu Gly Ile Leu Ser Gly
 1 5 10 15
 Tyr Asp Asp Leu Lys Leu Glu Thr Val Arg Tyr Tyr Asp Leu Ser Asn
 20 25 30
 Val Gln Gly Thr Ala Arg Gly Trp Glu Arg Glu Glu Arg Ile Leu Leu
 35 40 45
 Cys Val Pro Leu Arg Asp Ala Glu Gln His Leu Pro Met Phe Phe Ser
 50 55 60
 His Leu Lys Asn Phe Thr Tyr Pro His Asn Leu Ile Asp Leu Ala Phe
 65 70 75 80
 Leu Val Ser Asp Ser Lys Asp His Thr Leu Glu Ser Leu Thr Glu His
 85 90 95
 Leu Glu Ala Ile Gln Ala Asp Pro Asp Pro Lys Gln Pro Tyr Gly Glu
 100 105 110
 Ile Ser Ile Ile Glu Lys Asp Phe Gly Gln Lys Val Asn Gln Asp Val
 115 120 125
 Glu Ser Arg His Gly Phe Ala Ala Gln Ala Ser Arg Arg Lys Leu Met
 130 135 140
 Ala Gln Ala Arg Asn Trp Leu Leu Ser Ala Ala Leu Arg Pro Tyr His
 145 150 155 160
 Ser Trp Val Tyr Trp Arg Asp Val Asp Val Glu Thr Ala Pro Phe Thr
 165 170 175
 Ile Leu Glu Asp Leu Met Arg His Asn Lys Asp Val Ile Val Pro Asn
 180 185 190
 Val Trp Arg Pro Leu Pro Asp Trp Leu Gly Gly Glu Gln Pro Tyr Asp
 195 200 205
 Leu Asn Ser Trp Gln Glu Ser Glu Thr Ala Leu Ala Leu Ala Asp Thr
 210 215 220
 Leu Asp Glu Asp Ala Val Ile Val Glu Gly Tyr Ala Glu Tyr Ala Thr
 225 230 235 240
 Trp Arg Pro His Leu Ala Tyr Leu Arg Asp Pro Tyr Gly Asp Pro Asp
 245 250 255
 Met Glu Met Glu Ile Asp Gly Val Gly Gly Val Ser Ile Leu Ala Lys
 260 265 270
 Ala Lys Val Phe Arg Ala Gly Val His Phe Pro Ala Phe Ser Phe Glu
 275 280 285
 Lys His Ala Glu Thr Glu Gly Phe Gly Lys Met Ala Lys Arg Met His

```

      290              295              300
Phe Ser Val Val Gly Leu Pro His Tyr Thr Ile Trp His Leu Tyr Glu
305              310              315              320
Pro Ser Val Asp Asp Ile Lys His Met Glu Glu Met Glu Arg Glu Arg
      325              330              335
Ile Ala Arg Glu Lys Glu Glu Glu Glu Arg Lys Lys Lys Glu Ala Gln
      340              345              350
Ile Lys Glu Glu Phe Gly Asp Ala Asn Ser Gln Trp Glu Gln Asp Lys
      355              360              365
Gln Gln Met Gln Asp Leu Lys Leu Gln Asp Arg Gly Gly Asp Lys Glu
      370              375              380
Ala Ala Ala Ala Gly Val Asn Gln Gly Ala Ala Ala Lys Ala Ala Gly
385              390              395              400
Ala Met Glu Gly Gln Lys Asn
      405

```

<210> 60
 <211> 119
 <212> PRT
 <213> 里氏木霉

```

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

```

[0037]

<210> 61
 <211> 15
 <212> PRT
 <213> 里氏木霉

```

<400> 61
Ile Leu Leu Pro Leu Ile Ile Ile Cys Thr Ile Val Ala Tyr Tyr
1              5              10              15

```

<210> 62
 <211> 325
 <212> PRT
 <213> 里氏木霉

```

<400> 62
Gly Thr His Glu Ala Pro Gly Phe Val His Trp Trp Arg Arg Ile Ser
1              5              10              15
Met Gly Gly Gly Gly Glu Lys Phe Val Ile Ile Leu Gly Ala Asn Val
      20              25              30
Gly Gly Gly Val Met Glu Trp Lys Gly Ala Arg Glu Trp Ala Ile Glu
      35              40              45
Arg Asp Ser Val Arg Asn Lys Arg Lys Tyr Ala Thr Arg Trp Gly Tyr
      50              55              60
Asp Leu Glu Ile Val Asp Met Lys Thr Lys Lys Arg Tyr Ala His Glu
65              70              75              80
Trp Arg Glu Ser Trp Glu Lys Val Asp Phe Ile Arg Ala Ala Met Arg
      85              90              95
Lys Tyr Pro Lys Ala Glu Trp Phe Trp Trp Leu Asp Leu Asn Thr Tyr
      100              105              110
Val Met Glu Pro Ser Tyr Ser Leu Gln Arg His Leu Phe Asn His Leu
      115              120              125
Asp Arg His Val Tyr Arg Asp Ile Asn Val Phe Asn Pro Leu Asn Ile
      130              135              140
Thr His Pro Pro Thr Glu Glu Tyr Leu Asp Ala Glu Ala Arg Ser Pro

```

```

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

```

<210> 63
 <211> 33
 <212> PRT
 <213> 里氏木霉

```

<400> 63
Met His Phe Ala Tyr Pro Ser Arg Lys Ser Ser Asn Pro Pro Pro Phe
  1          5          10          15
Arg Pro Arg Ser Thr Arg Leu Pro Gly Leu Arg Arg Ser Arg Ile Lys
          20          25          30
Thr

```

[0038]

<210> 64
 <211> 15
 <212> PRT
 <213> 里氏木霉

```

<400> 64
Ile Gly Ile Val Leu Phe Leu Val Leu Ala Thr Leu Trp Phe Phe
  1          5          10          15

```

<210> 65
 <211> 262
 <212> PRT
 <213> 里氏木霉

```

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

```

```

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

```

<210> 66
 <211> 90
 <212> PRT
 <213> 里氏木霉

```

<400> 66
Met Ile Arg Asp Pro Phe Gly Ile His Ser Lys Asn Ala Phe Lys Ala
  1          5          10          15
Thr Ala Leu Arg Ala Ala Arg Asp Ile Lys Glu Ala Ala Thr Gln Ala
          20          25          30
Gly Ala Asn Ala Leu Glu Met Ser Phe Ser Leu Pro Lys His Val Pro
          35          40          45
Asp Phe Gly Asp Pro Ser Arg Ala Leu Glu Asp Arg Ala Trp Ala Ala
          50          55          60
Leu Leu Pro Met Tyr Lys Asp Lys Pro Tyr Ala Tyr Ala Pro Ser Met
          65          70          75          80
Arg Leu Arg Pro Trp Trp Arg Arg Arg Lys
          85          90

```

[0039]

<210> 67
 <211> 19
 <212> PRT
 <213> 里氏木霉

```

<400> 67
Val Leu Gly Met Ile Ala Ala Ala Val Met Phe Val Leu Tyr Val Thr
  1          5          10          15
Gly Phe Phe

```

<210> 68
 <211> 565
 <212> PRT
 <213> 里氏木霉

```

<400> 68
Ser Ser Gly Gln Thr Glu Glu Ala Lys Lys Lys Ala Ser Gly Ser Ala
  1          5          10          15
Phe Ser Trp Leu Gly Leu Ser Gln Glu Arg Gly Gly Val Asp Trp Asp
          20          25          30
Glu Arg Arg Lys Ser Val Val Glu Ala Phe Glu Val Trp Asp Ala Tyr
          35          40          45
Glu Arg Tyr Ala Trp Gly Lys Asp Glu Phe His Pro Ile Ser Lys Asn
          50          55          60
Gly Arg Asn Met Ala Pro Lys Gly Leu Gly Trp Ile Ile Ile Asp Ser
          65          70          75          80
Leu Asp Thr Met Met Leu Met Asn Gln Thr Thr Arg Leu Gln His Ala
          85          90          95
Arg Glu Trp Ile Ser Thr Ser Leu Thr Trp Asp Gln Asp Gln Asp Val
          100          105          110
Asn Thr Phe Glu Thr Thr Ile Arg Met Leu Gly Gly Leu Leu Ser Ala
          115          120          125
His Tyr Leu Ser Thr Glu Phe Pro Glu Leu Ala Pro Leu Thr Glu Asp
          130          135          140
Asp Glu Gly Ala Pro Gly Glu Asp Leu Tyr Leu Glu Lys Ala Lys Asp
          145          150          155          160
Leu Ala Asp Arg Leu Leu Ser Ala Phe Glu Ser Glu Ser Gly Ile Pro
          165          170          175
Tyr Ala Ser Val Asn Ile Gly Glu Tyr Lys Gly Pro Ser His Ser Asp

```

[0040]

```

      180      185      190
Asn Gly Ala Ser Ser Thr Ala Glu Ala Thr Thr Leu Gln Leu Glu Phe
      195      200      205
Lys Tyr Leu Ala Lys Leu Thr Gly Glu Lys Asn Phe Trp Asp Lys Val
      210      215      220
Glu Lys Val Met Glu Val Val Asp Asp Asn Gln Pro Glu Asp Gly Leu
      225      230      235      240
Val Pro Ile Tyr Ile Tyr Ala Thr Thr Gly Glu Phe Arg Gly Gln Asn
      245      250      255
Ile Arg Leu Gly Ser Arg Gly Asp Ser Tyr Tyr Glu Tyr Leu Ile Lys
      260      265      270
Gln Tyr Leu Gln Thr Asn Lys Gln Glu Pro Ile Tyr Glu Glu Met Trp
      275      280      285
Asp Glu Ala Leu Ala Gly Val Arg Lys His Leu Val Thr Tyr Thr Glu
      290      295      300
Pro Ser Glu Phe Thr Ile Ile Ala Glu Arg Pro Asp Gly Leu Glu His
      305      310      315      320
Pro Met Ser Pro Lys Met Asp His Leu Val Cys Phe Met Pro Gly Thr
      325      330      335
Ile Ala Leu Ala Ala Thr Gly Gly Leu Thr Glu Ala Glu Ala Arg Lys
      340      345      350
Leu Ser Thr Trp Asn Lys Lys Lys Asp Asp Asp Met Gln Leu Ala Arg
      355      360      365
Glu Leu Met His Thr Cys Trp Gly Met Tyr Lys Tyr Met Lys Thr Gly
      370      375      380
Leu Ala Pro Glu Ile Met Tyr Phe Asn Ile Pro Asn Pro Pro Pro Glu
      385      390      395      400
Ser Ser Ala Pro His Gln Ala Pro Ala Ala Phe Asp Glu Asp Pro His
      405      410      415
Ala Glu Trp Arg Lys Asp Phe Val Val His Ser Asn Asp Val His Asn
      420      425      430
Leu Gln Arg Pro Glu Thr Val Glu Ser Leu Phe Tyr Met Trp Arg Ile
      435      440      445
Thr Gly Asp Val Lys Tyr Arg Glu Trp Gly Trp Asp Met Phe Lys Ser
      450      455      460
Phe Val Asn Tyr Thr Ala Val Glu Asp Gln Gly Gly Phe Thr Ser Leu
      465      470      475      480
Leu Asp Ala Asn Ser Ile Pro Pro Thr Pro Lys Asp Asn Met Glu Ser
      485      490      495
Phe Trp Leu Ala Glu Thr Leu Lys Tyr Met Tyr Leu Leu Phe Ser Pro
      500      505      510
Asn Asp Val Leu Pro Leu His Lys Ile Val Leu Asn Thr Glu Ala His
      515      520      525
Pro Phe Pro Arg Phe Asp Met Gly Pro Leu Phe Ser Thr Gly Trp Lys
      530      535      540
Arg Lys Pro Arg Asp Gly Ser Ala Lys Lys Lys Ala Thr Thr Ala Ala
      545      550      555      560
Thr Thr Asp Ala Glu
      565

```

<210> 69
 <211> 7
 <212> PRT
 <213> 里氏木霉

<400> 69
 Met Ala Arg Arg Arg Tyr Arg
 1 5

<210> 70
 <211> 15
 <212> PRT
 <213> 里氏木霉

<400> 70
 Leu Phe Met Ile Cys Ala Ala Val Ile Leu Phe Leu Leu Tyr Arg
 1 5 10 15

<210> 71
 <211> 871
 <212> PRT
 <213> 里氏木霉

<400> 71
 Val Ser Gln Asn Thr Trp Asp Asp Ser Ala His Tyr Ala Thr Leu Arg
 1 5 10 15
 His Pro Pro Ala Ser Asn Pro Pro Ala Gly Gly Glu Ser Pro Leu
 20 25 30
 Lys Pro Ala Ala Lys Pro Glu His Glu His Glu His Glu Asn Gly Tyr
 35 40 45
 Ala Pro Glu Ser Lys Pro Lys Pro Gln Ser Glu Pro Lys Pro Glu Ser
 50 55 60
 Lys Pro Ala Pro Glu His Ala Ala Gly Gly Gln Lys Ser Gln Gly Lys
 65 70 75 80
 Pro Ser Tyr Glu Asp Asp Glu Glu Thr Gly Lys Asn Pro Pro Lys Ser
 85 90 95
 Ala Val Ile Pro Ser Asp Thr Arg Leu Pro Pro Asp Asn Lys Val His
 100 105 110
 Trp Arg Pro Val Lys Glu His Phe Pro Val Pro Ser Glu Ser Val Ile
 115 120 125
 Ser Leu Pro Thr Gly Lys Pro Leu Lys Val Pro Arg Val Gln His Glu
 130 135 140
 Phe Gly Val Glu Ser Pro Glu Ala Lys Ser Arg Arg Val Ala Arg Gln
 145 150 155 160
 Glu Arg Val Gly Lys Glu Ile Glu Arg Ala Trp Ser Gly Tyr Lys Lys
 165 170 175
 Phe Ala Trp Met His Asp Glu Leu Ser Pro Val Ser Ala Lys His Arg
 180 185 190
 Asp Pro Phe Cys Gly Trp Ala Ala Thr Leu Val Asp Ser Leu Asp Thr
 195 200 205
 Leu Trp Ile Ala Gly Leu Lys Glu Gln Phe Asp Glu Ala Ala Arg Ala
 210 215 220
 Val Glu Gln Ile Asp Phe Thr Thr Thr Pro Arg Asn Asn Ile Pro Val
 225 230 235 240
 Phe Glu Thr Thr Ile Arg Tyr Leu Gly Gly Leu Leu Gly Ala Phe Asp
 245 250 255
 Val Ser Gly Gly His Asp Gly Gly Tyr Pro Met Leu Leu Thr Lys Ala
 260 265 270
 Val Glu Leu Ala Glu Ile Leu Met Gly Ile Phe Asp Thr Pro Asn Arg
 275 280 285
 Met Pro Ile Leu Tyr Tyr Gln Trp Gln Pro Glu Tyr Ala Ser Gln Pro
 290 295 300
 His Arg Ala Gly Ser Val Gly Ile Ala Glu Leu Gly Thr Leu Ser Met
 305 310 315 320
 Glu Phe Thr Arg Leu Ala Gln Leu Thr Ser Gln Tyr Lys Tyr Tyr Asp
 325 330 335
 Ala Val Asp Arg Ile Thr Asp Ala Leu Ile Glu Leu Gln Lys Gln Gly
 340 345 350
 Thr Ser Ile Pro Gly Leu Phe Pro Glu Asn Leu Asp Ala Ser Gly Cys
 355 360 365
 Asn His Thr Ala Thr Ala Leu Arg Ser Ser Leu Ser Glu Ala Ala Gln
 370 375 380
 Lys Gln Met Asp Glu Asp Leu Ser Asn Lys Pro Glu Asn Tyr Arg Pro
 385 390 395 400
 Gly Lys Asn Ser Lys Ala Asp Pro Gln Thr Val Glu Lys Gln Pro Ala
 405 410 415
 Lys Lys Gln Asn Glu Pro Val Glu Lys Ala Lys Gln Val Pro Thr Gln
 420 425 430
 Gln Thr Ala Lys Arg Gly Lys Pro Pro Phe Gly Ala Asn Gly Phe Thr
 435 440 445
 Ala Asn Trp Asp Cys Val Pro Gln Gly Leu Val Val Gly Gly Tyr Gly
 450 455 460
 Phe Gln Gln Tyr His Met Gly Gly Gln Asp Ser Ala Tyr Glu Tyr
 465 470 475 480
 Phe Pro Lys Glu Tyr Leu Leu Leu Gly Gly Leu Glu Ser Lys Tyr Gln
 485 490 495
 Lys Leu Tyr Val Asp Ala Val Glu Ala Ile Asn Glu Trp Leu Leu Tyr
 500 505 510
 Arg Pro Met Thr Asp Gly Asp Trp Asp Ile Leu Phe Pro Ala Lys Val
 515 520 525
 Ser Thr Ala Gly Asn Pro Ser Gln Asp Leu Val Ala Thr Phe Glu Val
 530 535 540
 Thr His Leu Thr Cys Phe Ile Gly Gly Met Tyr Gly Leu Gly Gly Lys
 545 550 555 560
 Ile Phe Gly Arg Glu Lys Asp Leu Glu Thr Ala Lys Arg Leu Thr Asp
 565 570 575
 Gly Cys Val Trp Ala Tyr Gln Ser Thr Val Ser Gly Ile Met Pro Glu
 580 585 590
 Gly Ser Gln Val Leu Ala Cys Pro Thr Leu Glu Lys Cys Asp Phe Asn

```

      595              600              605
Glu Thr Leu Trp Trp Glu Lys Leu Asp Pro Ala Lys Asp Trp Arg Asp
      610              615              620
Lys Gln Val Ala Asp Asp Lys Asp Lys Ala Thr Val Gly Glu Ala Leu
625              630              635              640
Lys Glu Thr Ala Asn Ser His Asp Ala Ala Gly Gly Ser Lys Ala Val
      645              650              655
His Lys Arg Ala Ala Val Pro Leu Pro Lys Pro Gly Ala Asp Asp Asp
      660              665              670
Val Gly Ser Glu Leu Pro Gln Ser Leu Lys Asp Lys Ile Gly Phe Lys
      675              680              685
Asn Gly Glu Gln Lys Lys Pro Thr Gly Ser Ser Val Gly Ile Gln Arg
      690              695              700
Asp Pro Asp Ala Pro Val Asp Ser Val Leu Glu Ala His Arg Leu Pro
705              710              715              720
Pro Gln Glu Pro Glu Glu Gln Gln Val Ile Leu Pro Asp Lys Pro Gln
      725              730              735
Thr His Glu Glu Phe Val Lys Gln Arg Ile Ala Glu Met Gly Phe Ala
      740              745              750
Pro Gly Val Val His Ile Gln Ser Arg Gln Tyr Ile Leu Arg Pro Glu
      755              760              765
Ala Ile Glu Ser Val Trp Tyr Met Tyr Arg Ile Thr Gly Asp Pro Ile
770              775              780
Trp Met Glu Lys Gly Trp Lys Met Phe Glu Ala Thr Ile Arg Ala Thr
785              790              795              800
Arg Thr Glu Ile Ala Asn Ser Ala Ile Asp Asp Val Asn Ser Glu Glu
      805              810              815
Pro Gly Leu Lys Asp Glu Met Glu Ser Phe Trp Leu Ala Glu Thr Leu
      820              825              830
Lys Tyr Tyr Tyr Leu Leu Phe Ser Glu Pro Ser Val Ile Ser Leu Asp
      835              840              845
Glu Trp Val Leu Asn Thr Glu Ala His Pro Phe Lys Arg Pro Gly Gly
      850              855              860
Ser Val Ile Gly His Ser Ile
865              870

```

[0042]

```

<210> 72
<211> 13
<212> PRT
<213> 里氏木霉

```

```

<400> 72
Met Leu Asn Gln Leu Gln Gly Arg Val Pro Arg Arg Tyr
  1              5              10

```

```

<210> 73
<211> 15
<212> PRT
<213> 里氏木霉

```

```

<400> 73
Ile Ala Leu Val Ala Phe Ala Phe Phe Val Ala Phe Leu Leu Trp
  1              5              10              15

```

```

<210> 74
<211> 542
<212> PRT
<213> 里氏木霉

```

```

<400> 74
Ser Gly Tyr Asp Phe Val Pro Arg Thr Ala Thr Val Gly Arg Phe Lys
  1              5              10              15
Tyr Val Pro Ser Tyr Asp Trp Ser Lys Ala Lys Val Tyr Tyr Pro
      20              25              30
Val Lys Asp Met Lys Thr Leu Pro Gln Gly Thr Pro Val Thr Phe Pro
      35              40              45
Arg Leu Gln Leu Arg Asn Gln Ser Glu Ala Gln Asp Asp Thr Thr Lys
      50              55              60
Ala Arg Lys Gln Ala Val Lys Asp Ala Phe Val Lys Ser Trp Glu Ala
      65              70              75              80
Tyr Lys Thr Tyr Ala Trp Thr Lys Asp Gln Leu Gln Pro Leu Ser Leu
      85              90              95
Ser Gly Lys Glu Thr Phe Ser Gly Trp Ser Ala Gln Leu Val Asp Ala

```

100 105 110
 Leu Asp Thr Leu Trp Ile Met Asp Leu Lys Asp Asp Phe Phe Leu Ala
 115 120 125
 Val Lys Glu Val Ala Val Ile Asp Trp Ser Lys Thr Lys Asp Asn Lys
 130 135 140
 Val Ile Asn Leu Phe Glu Val Thr Ile Arg Tyr Leu Gly Gly Leu Ile
 145 150 155 160
 Ala Ala Tyr Asp Leu Ser Gln Glu Pro Val Leu Arg Ala Lys Ala Ile
 165 170 175
 Glu Leu Gly Asp Thr Leu Tyr Ala Thr Phe Asp Thr Pro Asn Arg Leu
 180 185 190
 Pro Ser His Trp Leu Asp Tyr Ser Lys Ala Lys Lys Gly Thr Gln Arg
 195 200 205
 Ala Asp Asp Ser Met Ser Gly Ala Ala Gly Gly Thr Leu Cys Met Glu
 210 215 220
 Phe Thr Arg Leu Ser Gln Ile Thr Gly Asp Pro Lys Tyr Tyr Asp Ala
 225 230 235 240
 Thr Glu Arg Ile Lys Gln Phe Phe Tyr Arg Phe Gln Asn Glu Thr Thr
 245 250 255
 Leu Pro Gly Met Trp Pro Val Met Met Asn Tyr Arg Glu Glu Thr Met
 260 265 270
 Val Glu Ser Arg Tyr Ser Met Gly Gly Ser Ala Asp Ser Leu Tyr Glu
 275 280 285
 Tyr Leu Val Lys Met Pro Ala Leu Leu Gly Gly Leu Asp Pro Gln Tyr
 290 295 300
 Pro Glu Met Ala Ile Arg Ala Leu Asp Thr Ala Arg Asp Asn Leu Leu
 305 310 315 320
 Phe Arg Pro Met Thr Glu Lys Gly Asp Asn Ile Leu Ala Leu Gly Asn
 325 330 335
 Ala Leu Val Asp His Gly Asn Val Gln Arg Ile Thr Glu Met Gln His
 340 345 350
 Leu Thr Cys Phe Ala Gly Gly Met Tyr Ala Met Ala Gly Lys Leu Phe
 355 360 365
 Lys Arg Asp Asp Tyr Val Asp Leu Gly Ser Arg Ile Ser Ser Gly Cys
 370 375 380
 Val Trp Ala Tyr Asp Ser Phe Pro Ser Gly Ile Met Pro Glu Ser Ala
 385 390 395 400
 Asp Met Ala Ala Cys Ala Lys Leu Asp Gly Pro Cys Pro Tyr Asp Glu
 405 410 415
 Val Lys Ala Pro Val Asp Pro Asp Gly Arg Arg Pro His Gly Phe Ile
 420 425 430
 His Val Lys Ser Arg His Tyr Leu Leu Arg Pro Glu Ala Ile Glu Ser
 435 440 445
 Val Phe Tyr Met Trp Arg Ile Thr Gly Asp Gln Val Trp Arg Asp Thr
 450 455 460
 Ala Trp Arg Met Trp Glu Asn Ile Val Arg Glu Ala Glu Thr Glu His
 465 470 475 480
 Ala Phe Ala Ile Val Glu Asp Val Thr Arg Thr Ala Ser Lys Leu Thr
 485 490 495
 Asn Asn Tyr Leu Leu Gln Thr Phe Trp Leu Ala Glu Thr Leu Lys Tyr
 500 505 510
 Phe Tyr Leu Ile Phe Asp Asp Glu Ser Ala Ile Asp Leu Asp Lys Trp
 515 520 525
 Val Phe Asn Thr Glu Ala His Pro Phe Lys Arg Pro Ala Val
 530 535 540

[0043]

<210> 75
 <211> 13
 <212> PRT
 <213> 里氏木霉

<400> 75
 Met Leu Val Val Gly Arg Pro Arg Leu Val Arg Asn Ser
 1 5 10

<210> 76
 <211> 16
 <212> PRT
 <213> 里氏木霉

<400> 76
 Ile Ile Leu Thr Leu Ala Ile Leu Ser Ile Trp His Leu Gly Leu Leu
 1 5 10 15

<210> 77
 <211> 576
 <212> PRT
 <213> 里氏木霉

<400> 77

Ser Arg Thr Pro Thr Ser Ala Ser Ala Leu Val Ser Ala Ser Val Ser
 1 5 10 15
 Ala Ser Ser Glu Trp Ser Arg Leu Glu Arg Leu Met Asn Arg Gly Ala
 20 25 30
 Pro Leu Thr Pro Tyr Pro Asp Ser Asn Ser Ser Phe Asp Trp Ser Ala
 35 40 45
 Ile Pro Phe Arg Tyr Pro Pro His Asn Thr Thr His Leu Pro Pro Arg
 50 55 60
 His Lys Gln Pro Pro Leu Pro Arg Ile Gln His Arg Phe Gly Pro Glu
 65 70 75 80
 Ser Pro Ala Ala Ala Lys Glu Arg Ile Lys Arg Leu Lys Ala Val Lys
 85 90 95
 Gln Val Phe Leu Arg Ala Trp Gln Ala Tyr Lys Gly Tyr Ala Trp Lys
 100 105 110
 Gln Asp Ala Leu Leu Pro Ile Ser Gly Gly Arg Glu Gln Phe Ser
 115 120 125
 Gly Trp Ala Ala Thr Leu Val Asp Ala Leu Asp Thr Leu Trp Ile Met
 130 135 140
 Gly Leu Arg Glu Glu Phe Asp Glu Ala Val Ala Ala Val Ala Glu Ile
 145 150 155 160
 Asp Phe Gly Ser Ser Thr Ser Ser Arg Val Asn Ile Phe Glu Thr Asn
 165 170 175
 Ile Arg Tyr Leu Gly Gly Leu Leu Ala Ala Tyr Asp Leu Ser Gly Arg
 180 185 190
 Glu Val Leu Leu Lys Lys Ala Val Glu Leu Gly Asp Leu Ile Tyr Ala
 195 200 205
 Gly Phe Asn Thr Glu Asn Gly Met Pro Val Asp Phe Leu Asn Phe Tyr
 210 215 220
 Ser Ala Lys Ser Gly Glu Gly Leu Val Val Glu Ser Ser Val Val Ser
 225 230 235 240
 Ala Ser Pro Gly Thr Leu Ser Leu Glu Leu Ala His Leu Ser Gln Val
 245 250 255
 Thr Gly Asp Asp Lys Tyr Tyr Ser Ala Val Ser Gln Val Met Asp Val
 260 265 270
 Phe Tyr Gln Gly Gln Asn Lys Thr Arg Leu Pro Gly Val Trp Pro Ile
 275 280 285
 Asp Val Asn Met Arg Ala Lys Asp Val Val Ser Gly Ser Arg Phe Thr
 290 295 300
 Leu Gly Gly Cys Ala Asp Ser Leu Tyr Glu Tyr Leu Pro Lys Met His
 305 310 315 320
 Gln Leu Leu Gly Gly Glu Pro Lys Tyr Glu Thr Met Ser Arg Thr
 325 330 335
 Phe Leu Gln Ala Ala Asp Arg His Phe Val Phe Arg Pro Met Leu Pro
 340 345 350
 Gly Ala Glu Glu Asp Val Leu Met Pro Gly Asn Val Asn Val Asp Glu
 355 360 365
 Asp Ser Gly Glu Ala Val Leu Asp Pro Glu Thr Glu His Leu Ala Cys
 370 375 380
 Phe Val Gly Gly Met Phe Gly Leu Ala Gly Arg Leu Phe Ser Arg Pro
 385 390 395 400
 Asp Asp Val Glu Thr Gly Val Arg Leu Thr Asn Gly Cys Val Tyr Ala
 405 410 415
 Tyr Arg Ala Phe Pro Thr Gly Met Met Pro Glu Arg Leu Asp Leu Ala
 420 425 430
 Pro Cys Arg Asp Arg Ser Ser Arg Cys Pro Trp Asp Glu Glu His Trp
 435 440 445
 Leu Glu Glu Arg Ala Lys Arg Pro Glu Trp Glu Pro His Leu Pro Arg
 450 455 460
 Gly Phe Thr Ser Ala Lys Asp Pro Arg Tyr Leu Leu Arg Pro Glu Ala
 465 470 475 480
 Ile Glu Ser Val Phe Tyr Ser Tyr Arg Ile Thr Gly Arg Gln Glu Phe
 485 490 495
 Gln Thr Ala Ala Trp Asp Met Phe Thr Ala Val Glu Lys Gly Thr Arg
 500 505 510
 Thr Gln Phe Ala Asn Ala Ala Val Leu Asp Val Thr Arg Ala Ala Asp
 515 520 525
 Glu Leu Pro Gln Glu Asp Tyr Met Glu Ser Phe Trp Leu Ala Glu Thr
 530 535 540
 Leu Lys Tyr Phe Tyr Leu Met Phe Thr Thr Pro Asp Ile Ile Ser Leu

[0044]

545 550 555 560
 Asp Asp Tyr Val Leu Asn Thr Glu Ala His Pro Phe Lys Leu Val Gly
 565 570 575

<210> 78
 <211> 17
 <212> PRT
 <213> 里氏木霉

<400> 78
 Met Val Met Leu Val Ala Ile Ala Leu Ala Trp Leu Gly Cys Ser Leu
 1 5 10 15
 Leu

<210> 79
 <211> 1053
 <212> PRT
 <213> 里氏木霉

<400> 79
 Arg Pro Val Asp Ala Met Arg Ala Asp Tyr Leu Ala Gln Leu Arg Gln
 1 5 10 15
 Glu Thr Val Asp Met Phe Tyr His Gly Tyr Ser Asn Tyr Met Glu His
 20 25 30
 Ala Phe Pro Glu Asp Glu Leu Arg Pro Ile Ser Cys Thr Pro Leu Thr
 35 40 45
 Arg Asp Arg Asp Asn Pro Gly Arg Ile Ser Leu Asn Asp Ala Leu Gly
 50 55 60
 Asn Tyr Ser Leu Thr Leu Ile Asp Ser Leu Ser Thr Leu Ala Ile Leu
 65 70 75 80
 Ala Gly Gly Pro Gln Asn Gly Pro Tyr Thr Gly Pro Gln Ala Leu Ser
 85 90 95
 Asp Phe Gln Asp Gly Val Ala Glu Phe Val Arg His Tyr Gly Asp Gly
 100 105 110
 Arg Ser Gly Pro Ser Gly Ala Gly Ile Arg Ala Arg Gly Phe Asp Leu
 115 120 125
 Asp Ser Lys Val Gln Val Phe Glu Thr Val Ile Arg Gly Val Gly Gly
 130 135 140
 Leu Leu Ser Ala His Leu Phe Ala Ile Gly Glu Leu Pro Ile Thr Gly
 145 150 155 160
 Tyr Val Pro Arg Pro Glu Gly Val Ala Gly Asp Asp Pro Leu Glu Leu
 165 170 175
 Ala Pro Ile Pro Trp Pro Asn Gly Phe Arg Tyr Asp Gly Gln Leu Leu
 180 185 190
 Arg Leu Ala Leu Asp Leu Ser Glu Arg Leu Leu Pro Ala Phe Tyr Thr
 195 200 205
 Pro Thr Gly Ile Pro Tyr Pro Arg Val Asn Leu Arg Ser Gly Ile Pro
 210 215 220
 Phe Tyr Val Asn Ser Pro Leu His Gln Asn Leu Gly Glu Ala Val Glu
 225 230 235 240
 Glu Gln Ser Gly Arg Pro Glu Ile Thr Glu Thr Cys Ser Ala Gly Ala
 245 250 255
 Gly Ser Leu Val Leu Glu Phe Thr Val Leu Ser Arg Leu Thr Gly Asp
 260 265 270
 Ala Arg Phe Glu Gln Ala Ala Lys Arg Ala Phe Trp Glu Val Trp His
 275 280 285
 Arg Arg Ser Glu Ile Gly Leu Ile Gly Asn Gly Ile Asp Ala Glu Arg
 290 295 300
 Gly Leu Trp Ile Gly Pro His Ala Gly Ile Gly Ala Gly Met Asp Ser
 305 310 315 320
 Phe Phe Glu Tyr Ala Leu Lys Ser His Ile Leu Leu Ser Gly Leu Gly
 325 330 335
 Met Pro Asn Ala Ser Thr Ser Arg Arg Gln Ser Thr Thr Ser Trp Leu
 340 345 350
 Asp Pro Asn Ser Leu His Pro Pro Leu Pro Pro Glu Met His Thr Ser
 355 360 365
 Asp Ala Phe Leu Gln Ala Trp His Gln Ala His Ala Ser Val Lys Arg
 370 375 380
 Tyr Leu Tyr Thr Asp Arg Ser His Phe Pro Tyr Tyr Ser Asn Asn His
 385 390 395 400
 Arg Ala Thr Gly Gln Pro Tyr Ala Met Trp Ile Asp Ser Leu Gly Ala
 405 410 415
 Phe Tyr Pro Gly Leu Leu Ala Leu Ala Gly Glu Val Glu Glu Ala Ile

[0045]

117

[illegible]

```

      325      330      335
Ile Ala Arg Glu Lys Glu Glu Glu Glu Arg Lys Lys Lys Glu Ala Gln
      340      345      350
Ile Lys Glu Glu Phe Gly Asp Ala Asn Ser Gln Trp Glu Gln Asp Lys
      355      360      365
Gln Gln Met Gln Asp Leu Lys Leu Gln Asp Arg Gly Gly Asp Lys Glu
      370      375      380
Ala Ala Ala Ala Gly Val Asn Gln Gly Ala Ala Ala Lys Ala Ala Gly
385      390      395      400
Ala Met Glu Gly Gln Lys Asn
      405

```

<210> 83
 <211> 31
 <212> PRT
 <213> 里氏木霉

```

<400> 83
Met Leu Leu Pro Lys Gly Gly Leu Asp Trp Arg Ser Ala Arg Ala Gln
  1      5      10      15
Ile Pro Pro Thr Arg Ala Leu Trp Asn Ala Val Thr Arg Thr Arg
      20      25      30

```

<210> 84
 <211> 15
 <212> PRT
 <213> 里氏木霉

```

<400> 84
Phe Ile Leu Leu Val Gly Ile Thr Gly Leu Ile Leu Leu Leu Trp
  1      5      10      15

```

[0048]

<210> 85
 <211> 358
 <212> PRT
 <213> 里氏木霉

```

<400> 85
Arg Gly Val Ser Thr Ser Ala Ser Glu Met Gln Ser Phe Tyr Cys Trp
  1      5      10      15
Gly Pro Ala Lys Pro Pro Met Glu Met Ser Pro Asn Glu His Asn Arg
      20      25      30
Trp Asn Gly His Leu Gln Thr Pro Val Ile Phe Asn His His Ala Pro
      35      40      45
Val Glu Val Asn Ser Ser Thr Ile Glu His Val Asp Leu Asn Pro Ile
      50      55      60
Asn Ser Thr Lys Gln Ala Val Thr Lys Glu Glu Arg Ile Leu Ile Leu
      65      70      75      80
Thr Pro Leu Lys Asp Ala Ala Pro Tyr Leu Ser Lys Tyr Phe Glu Leu
      85      90      95
Leu Ala Glu Leu Thr Tyr Pro His Arg Leu Ile Asp Leu Ala Phe Leu
      100      105      110
Val Ser Asp Ser Thr Asp Asp Thr Leu Ala Val Leu Ala Ser Glu Leu
      115      120      125
Asp Arg Ile Gln Lys Arg Pro Asp Gln Ile Pro Phe His Ser Ala Thr
      130      135      140
Val Ile Glu Lys Asp Phe Gly Phe Lys Leu Ser Gln Asn Val Glu Glu
      145      150      155      160
Arg His Ser Phe Glu Ala Gln Gly Pro Arg Lys Ala Met Gly Arg
      165      170      175
Ala Arg Asn Tyr Leu Leu Tyr Thr Ala Leu Lys Pro Glu His Ser Trp
      180      185      190
Val Tyr Trp Arg Asp Val Asp Ile Val Asp Ser Pro Thr Gly Ile Leu
      195      200      205
Glu Asp Phe Ile Ala His Asp Arg Asp Ile Leu Val Pro Asn Ile Trp
      210      215      220
Phe His Arg Tyr Arg Asp Gly Val Asp Ile Glu Gly Arg Phe Asp Tyr
      225      230      235      240
Asn Ser Trp Val Glu Ser Asp Lys Gly Arg Lys Leu Ala Asn Ser Leu
      245      250      255
Asp Lys Asp Val Val Leu Ala Glu Gly Tyr Lys Gln Tyr Asp Thr Gly
      260      265      270
Arg Thr Tyr Met Ala Lys Met Gly Asp Trp Arg Glu Asn Lys Asp Val

```

275 280 285
 Glu Leu Glu Leu Asp Gly Ile Gly Gly Val Asn Ile Leu Val Lys Ala
 290 295 300
 Asp Val His Arg Ser Gly Ile Asn Phe Pro Cys Tyr Ala Phe Glu Asn
 305 310 315 320
 Gln Ala Glu Thr Glu Gly Phe Ala Lys Met Ala Lys Arg Ala Gly Tyr
 325 330 335
 Glu Val Tyr Gly Leu Pro Asn Tyr Val Val Trp His Ile Asp Thr Glu
 340 345 350
 Glu Lys Gly Gly Asn Ala
 355

<210> 86
 <211> 17
 <212> PRT
 <213> 里氏木霉

<400> 86
 Met Ala Arg Pro Met Gly Ser Val Arg Leu Lys Lys Ala Asn Pro Ser
 1 5 10 15
 Thr

<210> 87
 <211> 16
 <212> PRT
 <213> 里氏木霉

<400> 87
 Leu Ile Leu Gly Ala Val Leu Cys Ile Phe Ile Ile Ile Phe Leu Val
 1 5 10 15

[0049]

<210> 88
 <211> 339
 <212> PRT
 <213> 里氏木霉

<400> 88
 Ser Pro Ser Ser Pro Ala Ser Ala Ser Arg Leu Ser Ile Val Ser Ala
 1 5 10 15
 Gln His His Leu Ser Pro Pro Thr Ser Pro Tyr Gln Ser Pro Arg Ser
 20 25 30
 Gly Ala Val Gln Gly Pro Pro Val Thr Arg Tyr Asn Leu Asn Lys
 35 40 45
 Val Thr Val Thr Ser Asp Pro Val Arg Asn Gln Glu His Ile Leu Ile
 50 55 60
 Leu Thr Pro Met Ala Arg Phe Tyr Gln Glu Tyr Trp Asp Asn Leu Leu
 65 70 75 80
 Arg Leu Asn Tyr Pro His Glu Leu Ile Thr Leu Gly Phe Ile Leu Pro
 85 90 95
 Lys Thr Lys Glu Gly Asn Gln Ala Thr Ser Met Leu Gln Lys Gln Ile
 100 105 110
 Gln Lys Thr Gln Asn Tyr Gly Pro Glu Lys Asp Arg Phe Lys Ser Ile
 115 120 125
 Ile Ile Leu Arg Gln Asp Phe Asp Pro Ala Val Val Ser Gln Asp Glu
 130 135 140
 Ser Glu Arg His Lys Leu Ala Asn Gln Lys Ala Arg Arg Glu Val Met
 145 150 155 160
 Ala Lys Ala Arg Asn Ser Leu Leu Phe Thr Thr Leu Gly Pro Ser Thr
 165 170 175
 Ser Trp Val Leu Trp Leu Asp Ala Asp Ile Thr Glu Thr Ala Pro Thr
 180 185 190
 Leu Ile Gln Asp Leu Ala Ser His Asp Lys Pro Ile Ile Val Ala Asn
 195 200 205
 Cys Phe Gln Lys Tyr Tyr Asp Pro Glu Ser Lys Lys Met Ala Glu Arg
 210 215 220
 Pro Tyr Asp Phe Asn Ser Trp Gln Asp Ser Glu Thr Ala Leu Lys Met
 225 230 235 240
 Ala Glu Gln Met Gly Pro Asp Asp Ile Leu Leu Glu Gly Tyr Ala Glu
 245 250 255
 Met Ala Thr Tyr Arg Thr Leu Leu Ala Tyr Met Ser Thr Pro Gly Gly
 260 265 270
 Ser Lys Asp Leu Val Val Pro Leu Asp Gly Val Gly Gly Thr Ala Leu

275 280 285
 Leu Val Lys Ala Asp Val His Arg Asp Gly Ala Met Phe Pro Pro Phe
 290 295 300
 Ala Phe Tyr His Leu Ile Glu Ser Glu Gly Phe Ala Lys Met Ala Lys
 305 310 315 320
 Arg Leu Gly Trp Gln Pro Tyr Gly Leu Pro Asn Tyr Lys Val Tyr His
 325 330 335
 Tyr Asn Glu

<210> 89
 <211> 33
 <212> PRT
 <213> 里氏木霉

<400> 89
 Met His Phe Ala Tyr Pro Ser Arg Lys Ser Ser Asn Pro Pro Pro Phe
 1 5 10 15
 Arg Pro Arg Ser Thr Arg Leu Pro Gly Leu Arg Arg Ser Arg Ile Lys
 20 25 30
 Thr

<210> 90
 <211> 15
 <212> PRT
 <213> 里氏木霉

<400> 90
 Ile Gly Ile Val Leu Phe Leu Val Leu Ala Thr Leu Trp Phe Phe
 1 5 10 15

[0050]

<210> 91
 <211> 262
 <212> PRT
 <213> 里氏木霉

<400> 91
 Ser Asn Pro Arg Val Pro Arg Pro Asp Pro Glu Arg Val Pro Ser Gly
 1 5 10 15
 Arg Pro Pro Val Val Leu Val Thr Val Ile Asp Pro Thr Gln Tyr Pro
 20 25 30
 Asn Ala Tyr Leu Lys Thr Ile Lys Glu Asn Arg Glu Gln Tyr Ala Ala
 35 40 45
 Lys His Gly Tyr Glu Ala Phe Ile Val Lys Ala Tyr Asp Tyr Asp Thr
 50 55 60
 Gln Gly Ala Pro Gln Ser Trp Ser Lys Leu Met Ala Met Arg His Ala
 65 70 75 80
 Leu Thr Lys Phe Pro Glu Cys Arg Phe Val Trp Tyr Leu Asp Gln Asp
 85 90 95
 Ala Tyr Ile Met Asp Met Ser Lys Ser Leu Glu Glu Gln Leu Leu Asn
 100 105 110
 Arg Gln Lys Leu Glu Ser Leu Met Ile Lys Asn Tyr Pro Val Val Pro
 115 120 125
 Pro Asp Ser Ile Ile Lys Thr Phe Ser His Leu Arg Pro Asp Glu Val
 130 135 140
 Asp Leu Ile Val Ser Gln Asp Ser Ser Gly Leu Val Ala Gly Ser Val
 145 150 155 160
 Val Val Arg Asn Ser Gln Trp Ser Lys Phe Leu Leu Glu Thr Trp Met
 165 170 175
 Asp Pro Leu Tyr Arg Ser Tyr Asn Phe Gln Lys Ala Glu Arg His Ala
 180 185 190
 Leu Glu His Ile Val Gln Trp His Pro Thr Ile Leu Ser Lys Leu Ala
 195 200 205
 Leu Val Pro Gln Arg Thr Leu Gly Pro Tyr Thr Arg Thr Asp Gln Gly
 210 215 220
 Asp Ala Tyr Gln Asp Gly Asp Phe Val Val Met Phe Thr Gly Cys Thr
 225 230 235 240
 Lys Ser Gly Glu Gln Ser Cys Glu Thr Val Ser Ala Ser Tyr Tyr Gln
 245 250 255
 Lys Trp Ser Ser Ser Leu
 260

<210> 92
 <211> 119
 <212> PRT
 <213> 里氏木霉

<400> 92
 Met Ser Leu Ser Arg Ser Pro Ser Pro Val Pro Gly Gly Gly Trp Ser
 1 5 10 15
 Ser Pro Gly Leu Asn Ile Asn Ser Gly Arg Ser Ser Pro Ser Asn Ala
 20 25 30
 Ala Gly Ser Ser Val Ser Trp Glu Ser Ala Lys Met Arg Lys Gln Gly
 35 40 45
 Ala Asn Gly Tyr Pro Ser Phe Ser Thr Gln Asn Gln Gly Phe Phe Thr
 50 55 60
 Arg His Met Arg Arg Ile Ser Ser Ser Leu Pro Arg Phe Ala Ala Gly
 65 70 75 80
 Pro Gly Asn Thr Tyr Ala Glu Arg Glu Lys Tyr Glu Arg Gly Gly His
 85 90 95
 Ser Pro His Ala Gly Gly Arg Leu Arg Ala Phe Leu Ala Arg Ile
 100 105 110
 Gly Arg Arg Leu Lys Trp Arg
 115

<210> 93
 <211> 16
 <212> PRT
 <213> 里氏木霉

<400> 93
 Ile Leu Leu Pro Leu Ile Ile Ile Cys Thr Ile Val Ala Tyr Tyr Gly
 1 5 10 15

<210> 94
 <211> 324
 <212> PRT
 <213> 里氏木霉

[0051]

<400> 94
 Thr His Glu Ala Pro Gly Phe Val His Trp Trp Arg Arg Ile Ser Met
 1 5 10 15
 Gly Gly Gly Gly Glu Lys Phe Val Ile Ile Leu Gly Ala Asn Val Gly
 20 25 30
 Gly Gly Val Met Glu Trp Lys Gly Ala Arg Glu Trp Ala Ile Glu Arg
 35 40 45
 Asp Ser Val Arg Asn Lys Arg Lys Tyr Ala Thr Arg Trp Gly Tyr Asp
 50 55 60
 Leu Glu Ile Val Asp Met Lys Thr Lys Lys Arg Tyr Ala His Glu Trp
 65 70 75 80
 Arg Glu Ser Trp Glu Lys Val Asp Phe Ile Arg Ala Ala Met Arg Lys
 85 90 95
 Tyr Pro Lys Ala Glu Trp Phe Trp Trp Leu Asp Leu Asn Thr Tyr Val
 100 105 110
 Met Glu Pro Ser Tyr Ser Leu Gln Arg His Leu Phe Asn His Leu Asp
 115 120 125
 Arg His Val Tyr Arg Asp Ile Asn Val Phe Asn Pro Leu Asn Ile Thr
 130 135 140
 His Pro Pro Thr Glu Glu Tyr Leu Asp Ala Glu Ala Arg Ser Pro Val
 145 150 155 160
 Gly Asp Gly Asn Ile Asn Ser Val Asn Leu Met Leu Thr Gln Asp Cys
 165 170 175
 Ser Gly Phe Asn Leu Gly Ser Phe Phe Ile Arg Arg Ser Ala Trp Thr
 180 185 190
 Glu Gln Leu Leu Asp Ile Trp Trp Asp Pro Val Leu Tyr Glu Gln Lys
 195 200 205
 His Met Glu Trp Glu His Lys Glu Gln Asp Ala Leu Glu Gln Leu Tyr
 210 215 220
 Arg Thr Gln Pro Trp Ile Arg Gln His Thr Gly Phe Leu Pro Gln Arg
 225 230 235 240
 Leu Ile Asn Ser Phe Pro Pro Ala Ala Cys Ala Asp Glu Ser Gly Leu
 245 250 255
 Asn Asn Thr Arg Ile His Tyr Asn Glu Lys Asp Arg Asp Phe Val Val
 260 265 270
 Asn Met Ala Gly Cys Glu Trp Gly Arg Asp Cys Trp Gly Glu Met Tyr

<210>	95
<211>	863
<212>	PRT
<213>	智人

[0052]

485 490 495
 Leu Ser Ser Gln Arg Gly Arg Val Pro Thr Ala Ala Pro Pro Ala Gln
 500 505 510
 Pro Arg Val Pro Val Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val
 515 520 525
 Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu
 530 535 540
 His Tyr Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp
 545 550 555 560
 Cys Gly His Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala
 565 570 575
 Val Thr His Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro
 580 585 590
 Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala Arg His Tyr Arg
 595 600 605
 Trp Ala Leu Gly Gln Val Phe Arg Gln Phe Arg Phe Pro Ala Ala Val
 610 615 620
 Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe Phe Glu Tyr Phe
 625 630 635 640
 Arg Ala Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val
 645 650 655
 Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Ser Arg
 660 665 670
 Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu
 675 680 685
 Leu Leu Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala
 690 695 700
 Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala
 705 710 715 720
 Cys Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly
 725 730 735
 Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu
 740 745 750
 Asn Gln Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln
 755 760 765
 Arg Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro
 770 775 780
 Gln Leu Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly
 785 790 795 800
 Glu Val Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala
 805 810 815
 Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala
 820 825 830
 Gly Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His
 835 840 845
 Leu Ala Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
 850 855 860

[0053]

<210> 96
 <211> 2592
 <212> DNA
 <213> 智人

<400> 96
 atgcgcttcc gaatctacaa gcggaaggtc ctcattctga cccttgctgt ggccgcttgc 60
 ggctttgttc tctggtccag caacggctgc cagcgtaaga acgaggccct ggccgcttccc 120
 ctcttgacg ccgaaccggc cagaggcgca gglggcaggg gaggggatca cccctcggtc 180
 gctgtcgga tccgccgct cagcaatgtg tccgccgect ctctggtccc ggccggttccg 240
 cagcctgagg cagacaacct cagctgcgc taccgatcac tcgtgtatca acttaacttc 300
 gaccagactc tgcggaacgt cgacaaggcc ggaacctggg ctccgcgtga gttggtcttc 360
 gtcgttcagg tgcacaacag gcccgagtac ctccgcttc tcttgattc gcttcgaaag 420
 gccagggcca tcgacaacgt cctggtgatt ttcagccatc acttttggtc cacagagatc 480
 aatcagctca ttgcgggtgt caacttttgc cccgtcttgc aagtttctt ccttttctt 540
 atccaactct accccaacga gttcccgggc agtgaccccc gcgactgtcc tcgggatctg 600
 ccaaaaaacg ccgctctcaa gctgggtcgc atcaacgccg aataccccga cagcttttgc 660
 cactatcgcg aggccaagtt ctgcgacagc aagcaccact ggtggtggaa gctccatttt 720
 gtctgggagc gactgaagat ccttcgtgat tacgcaggac tcattctgtt ctggaagag 780
 gaccactacc tggccccgga cttctaccac gtctttaaga agatgtggaa gctcaagcag 840
 caggaatgcc ccgagtgcga cgttctgtcc ctggcactc atagcgcgtc ccgctcgttc 900
 tacggtatgg ttgcacaaggc cgtatgtgaaa acctggaagt caactgagca caatatgggc 960
 ctgccccgta cgaggaaacgc ctaccagaaa ctcatcgagt gtaccgacac cttctgcacg 1020
 tacgacgact ataactggga ttggacactg cagtacttga ctgtcagctg cctccctaag 1080
 ttttgaagg tccttgttcc ccagatcccg agaattttcc atgctggcga ctgcgggatg 1140
 caccacaaga aaacctgtcg cccatccacg cagtctgccc aaatcgagtc gctcctgaac 1200
 aacaacaagc agtacatggt ccccgagaca ctgaccatta gcgagaagtt tacggtcgtg 1260

```

gcgactctccc cgcctcga aaatggcggc tgggggtgaca tccgcgatca cgagctgtgc 1320
aagtcttacc gccggctcca gacgcgcccc gcacctggca ggccaccctc agtcagcgct 1380
ctcgatggcg accccgccag cctcaccggg gaagtgatcc gcctggcccc agacgccgag 1440
gtggagctgg agcggcagcg tgggctgctg cagcagatcg gggatgccct gtcgagccag 1500
cgggggaggg tgcaccgcgc cgcctctccc gccagccgcg gtgtgcctgt gaccccgcg 1560
ccggcggtga ttccatcct ggtcatcgcc tgtgaccgca gcactgttcg gcgctgcctg 1620
gacaagctgc tgcattatcg gccctcggtt gactcttccc ccatcatcgt cagccaggac 1680
tgccggcacg aggagacggc ccaggccatc gcctcctacg gcagcgcggt cagccacatc 1740
cggcagcccg acctgagcag cattgcggtg ccgcgggacc accgcaagtt ccagggttac 1800
tacaagatcg cgcgccacta ccgctggggc ctggggccagg tcttcgggca gtttcgcttc 1860
cccggccgcg tgggtgtgga ggatgacctg gaggtggccc cggacttctt cgagtacttt 1920
cgggcacact atccgctgct gaaggccgac ccctccctgt ggtgcgtctc ggcttgggat 1980
gacaacggca aggagcagat ggtggacgcc agcagccctg agctgtctta ccgaccgac 2040
tttttccctg gcttggctg gctgctgttg gccagctctt gggctgagct ggagcccaag 2100
tgcccaaggc ctttctggga cgactggatg cggcggccgg agcagcggca gggcggggccc 2160
tgcatccgccc ctgagatctc aagaacgatg accttggccc gcaagggtgt gagccacggg 2220
cagttctttg accagcactt caagtctatc aagctgaacc agcagtttgt gcacttcacc 2280
cagctggacc tgtcttacct gcagcgggag gcctatgacc gagatttctt cggccgcttc 2340
tacggtgctc ccagctgca ggtggagaaa gtgaggacca atgaccggaa ggagctgggg 2400
gaggtgcggg tgcagtacac gggcagggac agcttcaagg ctttcgcaa ggccttgggt 2460
gtcatggatg acctcaagtc ggggggttccg agagctggct accggggcat tgtcaccttc 2520
cagttccggg gccgcggtgt ccacttggcg ccccccacga cgtgggaggg ctatgatccc 2580
agctggaatt ag 2592

```

<210> 97
 <211> 1263
 <212> DNA
 <213> 里氏木霉

```

<400> 97
atggcgctcac tcatcaaaac tgcggtggac attgccaaag gccgcatgc gctgtccaga 60
tatgtcatct tgggctctg gcttgcggat gcggtgctgt gcgggctgat tatctggaaa 120
gtgcttata cggaaatcga ctgggtcgcc tacatggagc aagtcacca gttctgccac 180
ggagagcgag actaccccaa gatggagggc ggcacagggc ccctggtgta tcccgcgccc 240
catgtgtaca tctacacagg gctctactac ctgacgaaca agggcaccga catcctgtg 300
gcgcagcagc tctttgccgt gctctacatg gctactctgg cggctgctat gacatgctac 360
tccaaggcca aggtcccgcc gtacatcttc ccgcttctca tecttccaa aagaattcac 420
acgctcttcc tctgagatg cttcaacgac tgccttcgccc ccttcttctt ctggctctgc 480
atcttcttct tccagaggcg agagtggacc atcgagctc tgcataacag catcgccctg 540
ggcgtaaaaa tgcgctgctt actggttctc ccgcgctgg tcatcgctct ctacctggc 600
cgcggttcca agggcgccct gcggtgctgc tggctcatgg tgcaggteca gctcctctc 660
gccalacctt tcatcacgac aaattggcgc ggctacctg gccgtgcatt cgagctctcg 720
aggcagttca agtttgaatg gacagtcact tggcgcatgc tggcgaggga tctgttctc 780
agccggggct tctctatcac gctactgcca ttacacgcca tcttctctct cgcctttatc 840
ctcggccggt ggcgaagat tagggaacgg accgtactcg ggatgatecc ctatgtcatc 900
cgattcagat cgccttttac cgagcaggaa gagcgcgcca tctccaaccg cgtcgctcac 960
cccggctatg tcatgtccac catcttctcg gccaacgtgg tgggactgct gtttggcccg 1020
tctctgcaat accagttcta tgcatactcg gcgtgggcca cccctatct cctgtggagc 1080
gcctgcccca atcttttggg ggtggccccc ctctggcgcg cgcaagaatg ggcttgggac 1140
gtcttcccca gcacgctct tagctcgagc gtcgtggtga gcgtgctggc cgtgacggtg 1200
gcatgagcgt ttgcaggttc aaatccgcag ccacgtgaaa catcgaagcc gaagcagcac 1260
taa 1263

```

<210> 98
 <211> 420
 <212> PRT
 <213> 里氏木霉

```

<400> 98
Met Ala Ser Leu Ile Lys Thr Ala Val Asp Ile Ala Asn Gly Arg His
1           5           10          15
Ala Leu Ser Arg Tyr Val Ile Phe Gly Leu Trp Leu Ala Asp Ala Val
20          25          30
Leu Cys Gly Leu Ile Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp
35          40          45
Val Ala Tyr Met Glu Gln Val Thr Gln Phe Val His Gly Glu Arg Asp
50          55          60
Tyr Pro Lys Met Glu Gly Gly Thr Gly Pro Leu Val Tyr Pro Ala Ala
65          70          75          80
His Val Tyr Ile Tyr Thr Gly Leu Tyr Tyr Leu Thr Asn Lys Gly Thr
85          90          95
Asp Ile Leu Leu Ala Gln Gln Leu Phe Ala Val Leu Tyr Met Ala Thr
100         105         110
Leu Ala Val Val Met Thr Cys Tyr Ser Lys Ala Lys Val Pro Pro Tyr
115         120         125
Ile Phe Pro Leu Leu Ile Leu Ser Lys Arg Leu His Ser Val Phe Val

```

```

130          135          140
Leu Arg Cys Phe Asn Asp Cys Phe Ala Ala Phe Phe Leu Trp Leu Cys
145          150          155          160
Ile Phe Phe Phe Gln Arg Arg Glu Trp Thr Ile Gly Ala Leu Ala Tyr
          165          170          175
Ser Ile Gly Leu Gly Val Lys Met Ser Leu Leu Leu Val Leu Pro Ala
          180          185          190
Val Val Ile Val Leu Tyr Leu Gly Arg Gly Phe Lys Gly Ala Leu Arg
          195          200          205
Leu Leu Trp Leu Met Val Gln Val Gln Leu Leu Leu Ala Ile Pro Phe
          210          215          220
Ile Thr Thr Asn Trp Arg Gly Tyr Leu Gly Arg Ala Phe Glu Leu Ser
225          230          235          240
Arg Gln Phe Lys Phe Glu Trp Thr Val Asn Trp Arg Met Leu Gly Glu
          245          250          255
Asp Leu Phe Leu Ser Arg Gly Phe Ser Ile Thr Leu Leu Ala Phe His
          260          265          270
Ala Ile Phe Leu Leu Ala Phe Ile Leu Gly Arg Trp Leu Lys Ile Arg
          275          280          285
Glu Arg Thr Val Leu Gly Met Ile Pro Tyr Val Ile Arg Phe Arg Ser
290          295          300
Pro Phe Thr Glu Gln Glu Arg Ala Ile Ser Asn Arg Val Val Thr
305          310          315          320
Pro Gly Tyr Val Met Ser Thr Ile Leu Ser Ala Asn Val Val Gly Leu
          325          330          335
Leu Phe Ala Arg Ser Leu His Tyr Gln Phe Tyr Ala Tyr Leu Ala Trp
          340          345          350
Ala Thr Pro Tyr Leu Leu Trp Thr Ala Cys Pro Asn Leu Leu Val Val
          355          360          365
Ala Pro Leu Trp Ala Ala Gln Glu Trp Ala Trp Asn Val Phe Pro Ser
          370          375          380
Thr Pro Leu Ser Ser Ser Val Val Val Ser Val Leu Ala Val Thr Val
385          390          395          400
Ala Met Ala Phe Ala Gly Ser Asn Pro Gln Pro Arg Glu Thr Ser Lys
          405          410          415
Pro Lys Gln His
          420

```

[0055]

```

<210> 99
<211> 21
<212> DNA
<213> 人工序列

```

```

<220>
<223> 合成构建体

```

```

<400> 99
gcaaattggca ttctgacatc c 21

```

```

<210> 100
<211> 21
<212> DNA
<213> 人工序列

```

```

<220>
<223> 合成构建体

```

```

<400> 100
gactgggttcc aattgacaag c 21

```

```

<210> 101
<211> 42
<212> DNA
<213> 人工序列

```

```

<220>
<223> 合成构建体

```

```

<400> 101
cagtgggtacc ctaattccag ctaggatcat agccctccca cg 42

```

```

<210> 102
<211> 18
<212> DNA
<213> 人工序列

```

	<220>		
	<223> 合成构建体		
	<400> 102		
	cggaccaccg caagtcc	18	
	<210> 103		
	<211> 54		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 103		
	atgcggaatt ctgcatcatc atcatcatca tcgccagcgt aagaacgagg ccct	54	
	<210> 104		
	<211> 22		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 104		
	cctttctcta tccaactcta cc	22	
	<210> 105		
	<211> 18		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
[0056]	<400> 105		
	ggaacttgcg gtggtccg	18	
	<210> 106		
	<211> 66		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 106		
	ccgcccgcctc caggaggtg ggggcagtgg aggtggcggc agtgggaggg tgcccaccgc	60	
	cgcccc	66	
	<210> 107		
	<211> 66		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 107		
	gcggtgggca ccctcccact gccgccacct ccactgcccc cacctccctg gagccggcgg	60	
	taagac	66	
	<210> 108		
	<211> 65		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 合成构建体		
	<400> 108		
	aggtgggggc agtggaggtg gcggcagtgg cgcggtgga agtgggaggg tgcccaccgc	60	
	cgccc	65	

	<210> 109	
	<211> 64	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 109	
	cggtgggcac cctccactt ccaccgcgc cactgccgc acctccactg cccccacctc 60	
	cctg	64
	<210> 110	
	<211> 65	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 110	
	gtttccgccg ggagggttgc cgccgctagg gttgccggtg ctctggagcc ggcggtaaga 60	
	cttgc	65
	<210> 111	
	<211> 66	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 111	
	gcaaccctcc cgcggaac ccgcctggca gcaccgggag ggtgcccacc gccgcccctc 60	
	ccgccc	66
[0057]	<210> 112	
	<211> 70	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 112	
	ccgcctccag gaacagtggc gctggcggtg gccgtcgcgg cggagctctg gagccggcgg 60	
	taagacttgc	70
	<210> 113	
	<211> 71	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 113	
	cgccactgtt cctggaggcg gtagcgccc caccagcggg aggggtgccc cgcgcgccc 60	
	tcccgcccag c	71
	<210> 114	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 114	
	cattagcgag aagtttacgg	20
	<210> 115	
	<211> 106	
	<212> PRT	
	<213> 里氏木霉	

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

<210> 116
 <211> 83
 <212> PRT
 <213> 里氏木霉

<400> 116
 Met Ala Ser Thr Asn Ala Arg Tyr Val Arg Tyr Leu Leu Ile Ala Phe
 1 5 10 15
 Phe Thr Ile Leu Val Phe Tyr Phe Val Ser Asn Ser Lys Tyr Glu Gly
 20 25 30
 Val Asp Leu Asn Lys Gly Thr Phe Thr Ala Pro Asp Ser Thr Lys Thr
 35 40 45
 Thr Pro Lys Pro Pro Ala Thr Gly Asp Ala Lys Asp Phe Pro Leu Ala
 50 55 60
 Leu Thr Pro Asn Asp Pro Gly Phe Asn Asp Leu Val Gly Ile Ala Pro
 65 70 75 80
 Gly Pro Arg

[0058]

<210> 117
 <211> 8
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 117
 Asp Tyr Lys Asp Asp Asp Lys
 1 5

<210> 118
 <211> 15
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 118
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15

<210> 119
 <211> 824
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 119
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30

[0059]

Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
 50 55 60
 Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
 65 70 75 80
 Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
 85 90 95
 Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
 100 105 110
 Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
 115 120 125
 Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140
 Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
 145 150 155 160
 Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220
 Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
 275 280 285
 Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
 290 295 300
 Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365
 Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380
 Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
 405 410 415
 Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430
 Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln Gly
 435 440 445
 Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Arg
 450 455 460
 Val Pro Thr Ala Ala Pro Pro Ala Gln Pro Arg Val Pro Val Thr Pro
 465 470 475 480
 Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala Cys Asp Arg Ser Thr
 485 490 495
 Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr Arg Pro Ser Ala Glu
 500 505 510
 Leu Phe Pro Ile Ile Val Ser Gln Asp Cys Gly His Glu Glu Thr Ala
 515 520 525
 Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr His Ile Arg Gln Pro
 530 535 540
 Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His Arg Lys Phe Gln Gly
 545 550 555 560
 Tyr Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala Leu Gly Gln Val Phe
 565 570 575
 Arg Gln Phe Arg Phe Pro Ala Ala Val Val Val Glu Asp Asp Leu Glu
 580 585 590
 Val Ala Pro Asp Phe Phe Glu Tyr Phe Arg Ala Thr Tyr Pro Leu Leu
 595 600 605
 Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala Trp Asn Asp Asn Gly
 610 615 620
 Lys Glu Gln Met Val Asp Ala Ser Arg Pro Glu Leu Leu Tyr Arg Thr
 625 630 635 640

Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu Ala Glu Leu Trp Ala
 645 650 655
 Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp Asp Asp Trp Met Arg
 660 665 670
 Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys Ile Arg Pro Glu Ile Ser
 675 680 685
 Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser His Gly Gln Phe Phe
 690 695 700
 Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln Gln Phe Val His Phe
 705 710 715 720
 Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu Ala Tyr Asp Arg Asp
 725 730 735
 Phe Leu Ala Arg Val Tyr Gly Ala Pro Gln Leu Gln Val Glu Lys Val
 740 745 750
 Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu Val Arg Val Gln Tyr Thr
 755 760 765
 Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala Leu Gly Val Met Asp
 770 775 780
 Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr Arg Gly Ile Val Thr
 785 790 795 800
 Phe Gln Phe Arg Gly Arg Arg Val His Leu Ala Pro Pro Pro Thr Trp
 805 810 815
 Glu Gly Tyr Asp Pro Ser Trp Asn
 820

<210> 120
 <211> 10
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 120
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10

[0060]

<210> 121
 <211> 819
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 121
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
 50 55 60
 Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
 65 70 75 80
 Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
 85 90 95
 Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
 100 105 110
 Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
 115 120 125
 Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140
 Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
 145 150 155 160
 Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220

[0061]

Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
 275 280 285
 Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
 290 295 300
 Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365
 Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380
 Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
 405 410 415
 Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430
 Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln Gly
 435 440 445
 Gly Gly Gly Ser Gly Gly Gly Ser Gly Arg Val Pro Thr Ala Ala
 450 455 460
 Pro Pro Ala Gln Pro Arg Val Pro Val Thr Pro Ala Pro Ala Val Ile
 465 470 475 480
 Pro Ile Leu Val Ile Ala Cys Asp Arg Ser Thr Val Arg Arg Cys Leu
 485 490 495
 Asp Lys Leu Leu His Tyr Arg Pro Ser Ala Glu Leu Phe Pro Ile Ile
 500 505 510
 Val Ser Gln Asp Cys Gly His Glu Thr Ala Gln Ala Ile Ala Ser
 515 520 525
 Tyr Gly Ser Ala Val Thr His Ile Arg Gln Pro Asp Leu Ser Ser Ile
 530 535 540
 Ala Val Pro Pro Asp His Arg Lys Phe Gln Gly Tyr Tyr Lys Ile Ala
 545 550 555 560
 Arg His Tyr Arg Trp Ala Leu Gly Gln Val Phe Arg Gln Phe Arg Phe
 565 570 575
 Pro Ala Ala Val Val Val Glu Asp Asp Leu Glu Val Ala Pro Asp Phe
 580 585 590
 Phe Glu Tyr Phe Arg Ala Thr Tyr Pro Leu Leu Lys Ala Asp Pro Ser
 595 600 605
 Leu Trp Cys Val Ser Ala Trp Asn Asp Asn Gly Lys Glu Gln Met Val
 610 615 620
 Asp Ala Ser Arg Pro Glu Leu Leu Tyr Arg Thr Asp Phe Phe Pro Gly
 625 630 635 640
 Leu Gly Trp Leu Leu Leu Ala Glu Leu Trp Ala Glu Leu Glu Pro Lys
 645 650 655
 Trp Pro Lys Ala Phe Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg
 660 665 670
 Gln Gly Arg Ala Cys Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe
 675 680 685
 Gly Arg Lys Gly Val Ser His Gly Gln Phe Phe Asp Gln His Leu Lys
 690 695 700
 Phe Ile Lys Leu Asn Gln Gln Phe Val His Phe Thr Gln Leu Asp Leu
 705 710 715 720
 Ser Tyr Leu Gln Arg Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val
 725 730 735
 Tyr Gly Ala Pro Gln Leu Gln Val Glu Lys Val Arg Thr Asn Asp Arg
 740 745 750
 Lys Glu Leu Gly Glu Val Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe
 755 760 765
 Lys Ala Phe Ala Lys Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly
 770 775 780
 Val Pro Arg Ala Gly Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly
 785 790 795 800
 Arg Arg Val His Leu Ala Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro
 805 810 815
 Ser Trp Asn

<210> 122
 <211> 19
 <212> PRT
 <213> 里氏木霉

<400> 122
 Ser Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro
 1 5 10 15
 Gly Ser Thr

<210> 123
 <211> 828
 <212> PRT
 <213> 里氏木霉

<400> 123
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
 50 55 60
 Arg Arg Val Ser Asn Val Ser Ala Ala Ser Leu Val Pro Ala Val Pro
 65 70 75 80
 Gln Pro Glu Ala Asp Asn Leu Thr Leu Arg Tyr Arg Ser Leu Val Tyr
 85 90 95
 Gln Leu Asn Phe Asp Gln Thr Leu Arg Asn Val Asp Lys Ala Gly Thr
 100 105 110
 Trp Ala Pro Arg Glu Leu Val Leu Val Val Gln Val His Asn Arg Pro
 115 120 125
 Glu Tyr Leu Arg Leu Leu Leu Asp Ser Leu Arg Lys Ala Gln Gly Ile
 130 135 140
 Asp Asn Val Leu Val Ile Phe Ser His Asp Phe Trp Ser Thr Glu Ile
 145 150 155 160
 Asn Gln Leu Ile Ala Gly Val Asn Phe Cys Pro Val Leu Gln Val Phe
 165 170 175
 Phe Pro Phe Ser Ile Gln Leu Tyr Pro Asn Glu Phe Pro Gly Ser Asp
 180 185 190
 Pro Arg Asp Cys Pro Arg Asp Leu Pro Lys Asn Ala Ala Leu Lys Leu
 195 200 205
 Gly Cys Ile Asn Ala Glu Tyr Pro Asp Ser Phe Gly His Tyr Arg Glu
 210 215 220
 Ala Lys Phe Ser Gln Thr Lys His His Trp Trp Trp Lys Leu His Phe
 225 230 235 240
 Val Trp Glu Arg Val Lys Ile Leu Arg Asp Tyr Ala Gly Leu Ile Leu
 245 250 255
 Phe Leu Glu Glu Asp His Tyr Leu Ala Pro Asp Phe Tyr His Val Phe
 260 265 270
 Lys Lys Met Trp Lys Leu Lys Gln Gln Glu Cys Pro Glu Cys Asp Val
 275 280 285
 Leu Ser Leu Gly Thr Tyr Ser Ala Ser Arg Ser Phe Tyr Gly Met Ala
 290 295 300
 Asp Lys Val Asp Val Lys Thr Trp Lys Ser Thr Glu His Asn Met Gly
 305 310 315 320
 Leu Ala Leu Thr Arg Asn Ala Tyr Gln Lys Leu Ile Glu Cys Thr Asp
 325 330 335
 Thr Phe Cys Thr Tyr Asp Asp Tyr Asn Trp Asp Trp Thr Leu Gln Tyr
 340 345 350
 Leu Thr Val Ser Cys Leu Pro Lys Phe Trp Lys Val Leu Val Pro Gln
 355 360 365
 Ile Pro Arg Ile Phe His Ala Gly Asp Cys Gly Met His His Lys Lys
 370 375 380
 Thr Cys Arg Pro Ser Thr Gln Ser Ala Gln Ile Glu Ser Leu Leu Asn
 385 390 395 400
 Asn Asn Lys Gln Tyr Met Phe Pro Glu Thr Leu Thr Ile Ser Glu Lys
 405 410 415
 Phe Thr Val Val Ala Ile Ser Pro Pro Arg Lys Asn Gly Gly Trp Gly
 420 425 430
 Asp Ile Arg Asp His Glu Leu Cys Lys Ser Tyr Arg Arg Leu Gln Ser
 435 440 445

[0062]

[0063]

Thr Gly Asn Pro Ser Gly Gly Asn Pro Pro Gly Gly Asn Pro Pro Gly
 450 455 460
 Ser Thr Gly Arg Val Pro Thr Ala Ala Pro Pro Ala Gln Pro Arg Val
 465 470 475 480
 Pro Val Thr Pro Ala Pro Ala Val Ile Pro Ile Leu Val Ile Ala Cys
 485 490 495
 Asp Arg Ser Thr Val Arg Arg Cys Leu Asp Lys Leu Leu His Tyr Arg
 500 505 510
 Pro Ser Ala Glu Leu Phe Pro Ile Ile Val Ser Gln Asp Cys Gly His
 515 520 525
 Glu Glu Thr Ala Gln Ala Ile Ala Ser Tyr Gly Ser Ala Val Thr His
 530 535 540
 Ile Arg Gln Pro Asp Leu Ser Ser Ile Ala Val Pro Pro Asp His Arg
 545 550 555 560
 Lys Phe Gln Gly Tyr Lys Ile Ala Arg His Tyr Arg Trp Ala Leu
 565 570 575
 Gly Gln Val Phe Arg Gln Phe Arg Phe Pro Ala Ala Val Val Val Glu
 580 585 590
 Asp Asp Leu Glu Val Ala Pro Asp Phe Glu Tyr Phe Arg Ala Thr
 595 600 605
 Tyr Pro Leu Leu Lys Ala Asp Pro Ser Leu Trp Cys Val Ser Ala Trp
 610 615 620
 Asn Asp Asn Gly Lys Glu Gln Met Val Asp Ala Ser Arg Pro Glu Leu
 625 630 635 640
 Leu Tyr Arg Thr Asp Phe Phe Pro Gly Leu Gly Trp Leu Leu Leu Ala
 645 650 655
 Glu Leu Trp Ala Glu Leu Glu Pro Lys Trp Pro Lys Ala Phe Trp Asp
 660 665 670
 Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys Ile Arg
 675 680 685
 Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val Ser His
 690 695 700
 Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn Gln Gln
 705 710 715 720
 Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg Glu Ala
 725 730 735
 Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro Gln Leu Gln
 740 745 750
 Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu Val Arg
 755 760 765
 Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys Ala Leu
 770 775 780
 Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly Tyr Arg
 785 790 795 800
 Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Val His Leu Ala Pro
 805 810 815
 Pro Pro Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
 820 825

<210> 124
 <211> 21
 <212> PRT
 <213> 里氏木霉

<400> 124
 Ser Ser Ala Ala Thr Ala Thr Ala Ser Ala Thr Val Pro Gly Gly Gly
 1 5 10 15
 Ser Gly Pro Thr Ser
 20

<210> 125
 <211> 830
 <212> PRT
 <213> 里氏木霉

<400> 125
 Met Arg Phe Arg Ile Tyr Lys Arg Lys Val Leu Ile Leu Thr Leu Val
 1 5 10 15
 Val Ala Ala Cys Gly Phe Val Leu Trp Ser Ser Asn Gly Arg Gln Arg
 20 25 30
 Lys Asn Glu Ala Leu Ala Pro Pro Leu Leu Asp Ala Glu Pro Ala Arg
 35 40 45
 Gly Ala Gly Gly Arg Gly Gly Asp His Pro Ser Val Ala Val Gly Ile
 50 55 60

[0064]

Arg	Arg	Val	Ser	Asn	Val	Ser	Ala	Ala	Ser	Leu	Val	Pro	Ala	Val	Pro
65					70					75					80
Gln	Pro	Glu	Ala	Asp	Asn	Leu	Thr	Leu	Arg	Tyr	Arg	Ser	Leu	Val	Tyr
				85					90					95	
Gln	Leu	Asn	Phe	Asp	Gln	Thr	Leu	Arg	Asn	Val	Asp	Lys	Ala	Gly	Thr
			100					105					110		
Trp	Ala	Pro	Arg	Glu	Leu	Val	Leu	Val	Val	Gln	Val	His	Asn	Arg	Pro
		115					120					125			
Glu	Tyr	Leu	Arg	Leu	Leu	Leu	Asp	Ser	Leu	Arg	Lys	Ala	Gln	Gly	Ile
	130					135					140				
Asp	Asn	Val	Leu	Val	Ile	Phe	Ser	His	Asp	Phe	Trp	Ser	Thr	Glu	Ile
145					150					155					160
Asn	Gln	Leu	Ile	Ala	Gly	Val	Asn	Phe	Cys	Pro	Val	Leu	Gln	Val	Phe
			165						170					175	
Phe	Pro	Phe	Ser	Ile	Gln	Leu	Tyr	Pro	Asn	Glu	Phe	Pro	Gly	Ser	Asp
			180					185					190		
Pro	Arg	Asp	Cys	Pro	Arg	Asp	Leu	Pro	Lys	Asn	Ala	Ala	Leu	Lys	Leu
		195					200					205			
Gly	Cys	Ile	Asn	Ala	Glu	Tyr	Pro	Asp	Ser	Phe	Gly	His	Tyr	Arg	Glu
	210					215					220				
Ala	Lys	Phe	Ser	Gln	Thr	Lys	His	His	Trp	Trp	Trp	Lys	Leu	His	Phe
225					230					235					240
Val	Trp	Glu	Arg	Val	Lys	Ile	Leu	Arg	Asp	Tyr	Ala	Gly	Leu	Ile	Leu
			245						250					255	
Phe	Leu	Glu	Glu	Asp	His	Tyr	Leu	Ala	Pro	Asp	Phe	Tyr	His	Val	Phe
		260						265					270		
Lys	Lys	Met	Trp	Lys	Leu	Lys	Gln	Gln	Glu	Cys	Pro	Glu	Cys	Asp	Val
		275					280					285			
Leu	Ser	Leu	Gly	Thr	Tyr	Ser	Ala	Ser	Arg	Ser	Phe	Tyr	Gly	Met	Ala
	290					295					300				
Asp	Lys	Val	Asp	Val	Lys	Thr	Trp	Lys	Ser	Thr	Glu	His	Asn	Met	Gly
305					310					315					320
Leu	Ala	Leu	Thr	Arg	Asn	Ala	Tyr	Gln	Lys	Leu	Ile	Glu	Cys	Thr	Asp
			325						330					335	
Thr	Phe	Cys	Thr	Tyr	Asp	Asp	Tyr	Asn	Trp	Asp	Trp	Thr	Leu	Gln	Tyr
		340						345					350		
Leu	Thr	Val	Ser	Cys	Leu	Pro	Lys	Phe	Trp	Lys	Val	Leu	Val	Pro	Gln
		355					360					365			
Ile	Pro	Arg	Ile	Phe	His	Ala	Gly	Asp	Cys	Gly	Met	His	His	Lys	Lys
		370				375					380				
Thr	Cys	Arg	Pro	Ser	Thr	Gln	Ser	Ala	Gln	Ile	Glu	Ser	Leu	Leu	Asn
385					390					395					400
Asn	Asn	Lys	Gln	Tyr	Met	Phe	Pro	Glu	Thr	Leu	Thr	Ile	Ser	Glu	Lys
			405						410					415	
Phe	Thr	Val	Val	Ala	Ile	Ser	Pro	Pro	Arg	Lys	Asn	Gly	Gly	Trp	Gly
		420						425					430		
Asp	Ile	Arg	Asp	His	Glu	Leu	Cys	Lys	Ser	Tyr	Arg	Arg	Leu	Gln	Ser
		435					440					445			
Ser	Ala	Ala	Thr	Ala	Thr	Ala	Ser	Ala	Thr	Val	Pro	Gly	Gly	Gly	Ser
		450				455					460				
Gly	Pro	Thr	Ser	Gly	Arg	Val	Pro	Thr	Ala	Ala	Pro	Pro	Ala	Gln	Pro
465					470					475					480
Arg	Val	Pro	Val	Thr	Pro	Ala	Pro	Ala	Val	Ile	Pro	Ile	Leu	Val	Ile
			485						490					495	
Ala	Cys	Asp	Arg	Ser	Thr	Val	Arg	Arg	Cys	Leu	Asp	Lys	Leu	Leu	His
		500						505					510		
Tyr	Arg	Pro	Ser	Ala	Glu	Leu	Phe	Pro	Ile	Ile	Val	Ser	Gln	Asp	Cys
		515					520					525			
Gly	His	Glu	Glu	Thr	Ala	Gln	Ala	Ile	Ala	Ser	Tyr	Gly	Ser	Ala	Val
		530				535					540				
Thr	His	Ile	Arg	Gln	Pro	Asp	Leu	Ser	Ser	Ile	Ala	Val	Pro	Pro	Asp
545					550					555					560
His	Arg	Lys	Phe	Gln	Gly	Tyr	Tyr	Lys	Ile	Ala	Arg	His	Tyr	Arg	Trp
			565						570					575	
Ala	Leu	Gly	Gln	Val	Phe	Arg	Gln	Phe	Arg	Phe	Pro	Ala	Ala	Val	Val
			580					585					590		
Val	Glu	Asp	Asp	Leu	Glu	Val	Ala	Pro	Asp	Phe	Phe	Glu	Tyr	Phe	Arg
		595					600					605			
Ala	Thr	Tyr	Pro	Leu	Leu	Lys	Ala	Asp	Pro	Ser	Leu	Trp	Cys	Val	Ser
		610				615					620				
Ala	Trp	Asn	Asp	Asn	Gly	Lys	Glu	Gln	Mct	Val	Asp	Ala	Ser	Arg	Pro
625					630					635					640
Glu	Leu	Leu	Tyr	Arg	Thr	Asp	Phe	Phe	Pro	Gly	Leu	Gly	Trp	Leu	Leu
			645						650					655	
Leu	Ala	Glu	Leu	Trp	Ala	Glu	Leu	Glu	Pro	Lys	Trp	Pro	Lys	Ala	Phe
			660					665						670	

Trp Asp Asp Trp Met Arg Arg Pro Glu Gln Arg Gln Gly Arg Ala Cys
 675 680 685
 Ile Arg Pro Glu Ile Ser Arg Thr Met Thr Phe Gly Arg Lys Gly Val
 690 695 700
 Ser His Gly Gln Phe Phe Asp Gln His Leu Lys Phe Ile Lys Leu Asn
 705 710 715 720
 Gln Gln Phe Val His Phe Thr Gln Leu Asp Leu Ser Tyr Leu Gln Arg
 725 730 735
 Glu Ala Tyr Asp Arg Asp Phe Leu Ala Arg Val Tyr Gly Ala Pro Gln
 740 745 750
 Leu Gln Val Glu Lys Val Arg Thr Asn Asp Arg Lys Glu Leu Gly Glu
 755 760 765
 Val Arg Val Gln Tyr Thr Gly Arg Asp Ser Phe Lys Ala Phe Ala Lys
 770 775 780
 Ala Leu Gly Val Met Asp Asp Leu Lys Ser Gly Val Pro Arg Ala Gly
 785 790 795 800
 Tyr Arg Gly Ile Val Thr Phe Gln Phe Arg Gly Arg Arg Val His Leu
 805 810 815
 Ala Pro Pro Pro Thr Trp Glu Gly Tyr Asp Pro Ser Trp Asn
 820 825 830

<210> 126
 <211> 420
 <212> PRT
 <213> 里氏木霉

[0065]

<400> 126
 Met Ala Ser Leu Ile Lys Thr Ala Val Asp Ile Ala Asn Gly Arg His
 1 5 10 15
 Ala Leu Ser Arg Tyr Val Ile Phe Gly Leu Trp Leu Ala Asp Ala Val
 20 25 30
 Leu Cys Gly Leu Ile Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp
 35 40 45
 Val Ala Tyr Met Glu Gln Val Thr Gln Phe Val His Gly Glu Arg Asp
 50 55 60
 Tyr Pro Lys Met Glu Gly Gly Thr Gly Pro Leu Val Tyr Pro Ala Ala
 65 70 75 80
 His Val Tyr Ile Tyr Thr Gly Leu Tyr Tyr Leu Thr Asn Lys Gly Thr
 85 90 95
 Asp Ile Leu Leu Ala Gln Gln Leu Phe Ala Val Leu Tyr Met Ala Thr
 100 105 110
 Leu Ala Val Val Met Thr Cys Tyr Ser Lys Ala Lys Val Pro Pro Tyr
 115 120 125
 Ile Phe Pro Leu Leu Ile Leu Ser Lys Arg Leu His Ser Val Phe Val
 130 135 140
 Leu Arg Cys Phe Asn Asp Cys Phe Ala Ala Phe Phe Leu Trp Leu Cys
 145 150 155 160
 Ile Phe Phe Phe Gln Arg Arg Glu Trp Thr Ile Gly Ala Leu Ala Tyr
 165 170 175
 Ser Ile Gly Leu Gly Val Lys Met Ser Leu Leu Leu Val Leu Pro Ala
 180 185 190
 Val Val Ile Val Leu Tyr Leu Gly Arg Gly Phe Lys Gly Ala Leu Arg
 195 200 205
 Leu Leu Trp Leu Met Val Gln Val Gln Leu Leu Leu Ala Ile Pro Phe
 210 215 220
 Ile Thr Thr Asn Trp Arg Gly Tyr Leu Gly Arg Ala Phe Glu Leu Ser
 225 230 235 240
 Arg Gln Phe Lys Phe Glu Trp Thr Val Asn Trp Arg Met Leu Gly Glu
 245 250 255
 Asp Leu Phe Leu Ser Arg Gly Phe Ser Ile Thr Leu Leu Ala Phe His
 260 265 270
 Ala Ile Phe Leu Leu Ala Phe Ile Leu Gly Arg Trp Leu Lys Ile Arg
 275 280 285
 Glu Arg Thr Val Leu Gly Met Ile Pro Tyr Val Ile Arg Phe Arg Ser
 290 295 300
 Pro Phe Thr Glu Gln Glu Arg Ala Ile Ser Asn Arg Val Val Thr
 305 310 315 320
 Pro Gly Tyr Val Met Ser Thr Ile Leu Ser Ala Asn Val Val Gly Leu
 325 330 335
 Leu Phe Ala Arg Ser Leu His Tyr Gln Phe Tyr Ala Tyr Leu Ala Trp
 340 345 350
 Ala Thr Pro Tyr Leu Leu Trp Thr Ala Cys Pro Asn Leu Leu Val Val
 355 360 365
 Ala Pro Leu Trp Ala Ala Gln Glu Trp Ala Trp Asn Val Phe Pro Ser
 370 375 380

Thr Pro Leu Ser Ser Ser Val Val Val Ser Val Leu Ala Val Thr Val
 385 390 395 400
 Ala Met Ala Phe Ala Gly Ser Asn Pro Gln Pro Arg Glu Thr Ser Lys
 405 410 415
 Pro Lys Gln His
 420

<210> 127
 <211> 525
 <212> PRT
 <213> 深绿木霉

[0066]

<400> 127
 Met Ala Ser Leu Ile Lys Phe Ala Ser Asp Val Ala Thr Gly Arg His
 1 5 10 15
 Ala Leu Ser Lys Leu Ile Pro Val Gly Leu Phe Leu Ala Asp Ala Ile
 20 25 30
 Leu Cys Gly Leu Val Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp
 35 40 45
 Thr Ala Tyr Met Glu Gln Val Thr Gln Phe Val Asn Gly Glu Arg Asp
 50 55 60
 Tyr Pro Lys Met Glu Gly Gly Thr Gly Pro Leu Val Tyr Pro Ala Ala
 65 70 75 80
 His Val Tyr Ile Tyr Thr Gly Leu Tyr Tyr Leu Thr Asn Arg Gly Thr
 85 90 95
 Asp Ile Leu Leu Ala Gln Gln Leu Phe Ala Val Leu Tyr Met Ala Thr
 100 105 110
 Leu Gly Val Val Met Leu Ser Tyr Trp Lys Ala Arg Val Pro Pro Tyr
 115 120 125
 Ile Phe Pro Leu Leu Ile Leu Ser Lys Arg Leu His Ser Val Phe Val
 130 135 140
 Leu Arg Cys Phe Asn Asp Cys Phe Ala Ala Phe Phe Leu Trp Leu Cys
 145 150 155 160
 Ile Tyr Ser Phe Gln Asn Arg Ala Trp Thr Phe Gly Ala Leu Ala Tyr
 165 170 175
 Thr Leu Gly Leu Gly Val Lys Met Ser Leu Leu Leu Val Leu Pro Ala
 180 185 190
 Val Val Ile Ile Leu Phe Leu Gly Arg Gly Phe Lys Gly Ala Leu Arg
 195 200 205
 Leu Val Trp Leu Met Ala Gln Val Gln Leu Val Leu Ala Ile Pro Phe
 210 215 220
 Ile Thr Thr Asn Trp Ala Gly Tyr Leu Gly Arg Ala Phe Glu Leu Ser
 225 230 235 240
 Arg Gln Phe Lys Phe Glu Trp Thr Val Asn Trp Arg Met Met Gly Glu
 245 250 255
 Glu Thr Phe Leu Ser Arg Gly Phe Ser Ile Thr Leu Leu Thr Phe His
 260 265 270
 Val Val Thr Leu Leu Val Phe Ile Ala Ala Arg Trp Leu Lys Leu Gln
 275 280 285
 Glu Arg Ser Leu Leu Gly Ile Ile Thr Tyr Ala Val Arg Phe Gln Ser
 290 295 300
 Pro Phe Thr Glu Gln Glu Glu Ala Lys Val Ser Lys Lys Val Val Thr
 305 310 315 320
 Pro Arg Tyr Val Leu Ala Thr Ile Leu Ser Ala Asn Val Ile Gly Leu
 325 330 335
 Leu Phe Ala Arg Ser Leu His Tyr Gln Phe Tyr Ala Tyr Leu Ala Trp
 340 345 350
 Ala Thr Pro Phe Leu Leu Trp Thr Ala Tyr Pro Asn Leu Leu Val Val
 355 360 365
 Val Pro Leu Trp Leu Ala Gln Glu Trp Ala Trp Asn Val Phe Pro Ser
 370 375 380
 Thr Pro Leu Ser Ser Ser Val Val Ile Ser Leu Val Pro Val Cys Leu
 385 390 395 400
 Leu Ser Pro Gln Leu Leu Val Ser His Asp Ile Tyr Asn Phe Ala Asn
 405 410 415
 Cys Ser Ala Ile Leu Arg Pro Arg Gly Ile Ala Phe Gly Gln Asp Ile
 420 425 430
 Ser Ala Thr Leu Asn Pro Asp Gly Val Ala Lys Pro Leu Gly Glu Leu
 435 440 445
 Glu Asn Asp Gly Leu Arg Val Trp His Leu Ala Ser Val Gln Val Val
 450 455 460
 Ser Phe Gly Leu His His Ala His Asn Glu Leu Gly Gly Leu Gln Phe
 465 470 475 480
 Gly Trp Trp Arg Glu Arg Phe Leu Arg Gly Gly Glu Asp Val Ala Leu
 485 490 495

Trp Phe Ala His Gly Gly Phe Glu Phe Arg Phe Phe Ser Glu Leu Leu
 500 505 510
 Val Arg Leu Ala Asp Thr Ser Asp Ile Lys Lys Ser Phe
 515 520 525

<210> 128
 <211> 419
 <212> PRT
 <213> 绿木霉

<400> 128
 Met Ala Ser Leu Ile Lys Phe Ala Ser Asp Val Ala Asn Gly Arg His
 1 5 10 15
 Ala Leu Ser Lys Phe Ile Pro Met Gly Leu Trp Leu Ala Asp Ala Val
 20 25 30
 Leu Cys Gly Leu Ile Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp
 35 40 45
 Val Ala Tyr Met Glu Gln Ile Thr Gln Phe Val His Gly Glu Arg Asp
 50 55 60
 Tyr Pro Lys Met Glu Gly Thr Gly Pro Leu Val Tyr Pro Ala Ala
 65 70 75 80
 His Val Tyr Ile Tyr Thr Gly Leu Tyr Tyr Leu Thr Asn Lys Gly Thr
 85 90 95
 Asp Ile Leu Leu Ala Gln Gln Leu Phe Ala Val Leu Tyr Met Ala Thr
 100 105 110
 Leu Gly Val Val Met Leu Cys Tyr Trp Lys Ala Lys Val Pro Pro Tyr
 115 120 125
 Ile Phe Pro Leu Leu Ile Leu Ser Lys Arg Leu His Ser Val Phe Val
 130 135 140
 Leu Arg Cys Phe Asn Asp Cys Phe Ala Ala Phe Phe Leu Trp Leu Ser
 145 150 155 160
 Ile Phe Phe Phe Gln Arg Arg Val Trp Thr Leu Gly Ala Ile Ala Tyr
 165 170 175
 Thr Ile Gly Leu Gly Val Lys Met Ser Leu Leu Leu Val Leu Pro Ala
 180 185 190
 Val Val Ile Val Leu Phe Leu Gly Arg Gly Phe Lys Gly Ala Leu Arg
 195 200 205
 Leu Leu Trp Leu Met Val Gln Val Gln Leu Leu Ala Ile Pro Phe
 210 215 220
 Ile Thr Thr Asn Trp Lys Gly Tyr Leu Gly Arg Ala Phe Glu Leu Ser
 225 230 235 240
 Arg Gln Phe Lys Phe Glu Trp Thr Val Asn Trp Arg Met Leu Gly Glu
 245 250 255
 Glu Leu Phe Leu Ser Arg Gly Phe Ser Ile Thr Leu Leu Ala Phe His
 260 265 270
 Ala Leu Phe Leu Leu Ile Phe Ile Leu Gly Arg Trp Leu Arg Ile Lys
 275 280 285
 Glu Arg Ser Phe Leu Gly Met Ile Pro Tyr Val Leu Arg Phe Thr Ser
 290 295 300
 Pro Phe Thr Glu His Glu Glu Ala Ser Ile Ser His Arg Val Val Thr
 305 310 315 320
 Pro Glu Tyr Ile Met Ser Ala Met Leu Ser Ala Asn Val Val Gly Leu
 325 330 335
 Leu Phe Ala Arg Ser Leu His Tyr Gln Phe Tyr Ala Tyr Leu Ala Trp
 340 345 350
 Ala Thr Pro Phe Leu Leu Trp Thr Ala Ser Pro Asn Leu Leu Val Val
 355 360 365
 Val Pro Leu Trp Ala Ala Gln Glu Trp Ala Trp Asn Val Phe Pro Ser
 370 375 380
 Thr Pro Leu Ser Ser Asn Val Val Val Ser Val Leu Ala Val Thr Val
 385 390 395 400
 Ala Met Ala Phe Val Gly Ser Asn Pro Gln Arg Gly Ala Pro Lys Pro
 405 410 415
 Lys Gln Leu

[0067]

<210> 129
 <211> 434
 <212> PRT
 <213> 尖孢镰刀菌

<400> 129
 Met Pro Glu Ser Ala Ser Gly Thr Leu Ser Gln Gly Val Arg Phe Leu
 1 5 10 15

[0068]

```

Arg Asn Val Leu Asn Gly Arg His Ala Leu Ser Lys Leu Ile Pro Ile
      20      25      30
Ala Leu Trp Leu Val Asp Ala Leu Gly Cys Gly Leu Ile Trp Lys
      35      40      45
Ile Pro Tyr Thr Glu Ile Asp Trp Val Ala Tyr Met Gln Gln Ile Ser
      50      55      60
Gln Phe Val Ser Gly Glu Arg Asp Tyr Thr Lys Met Glu Gly Asp Thr
      65      70      75      80
Gly Pro Leu Val Tyr Pro Ala Ala His Val Tyr Thr Tyr Thr Gly Leu
      85      90      95
Tyr Tyr Ile Thr Asp Lys Gly Thr Asn Ile Leu Leu Ala Gln Gln Ile
      100     105     110
Phe Ala Val Leu Tyr Met Ala Thr Leu Ala Val Val Met Leu Cys Tyr
      115     120     125
Trp Lys Ala Lys Val Pro Pro Tyr Met Phe Ile Phe Leu Ile Ala Ser
      130     135     140
Lys Arg Leu His Ser Leu Phe Val Leu Arg Cys Phe Asn Asp Cys Phe
      145     150     155     160
Ala Val Phe Phe Leu Trp Leu Thr Ile Phe Leu Phe Gln Arg Arg Gln
      165     170     175
Trp Thr Val Gly Ser Leu Val Tyr Ser Trp Gly Leu Gly Ile Lys Met
      180     185     190
Ser Leu Leu Leu Val Leu Pro Ala Ile Gly Val Ile Leu Phe Leu Gly
      195     200     205
Arg Gly Leu Trp Pro Ser Leu Arg Leu Ala Trp Leu Met Ala Gln Ile
      210     215     220
Gln Phe Ala Ile Gly Leu Pro Phe Ile Thr Lys Asn Pro Arg Gly Tyr
      225     230     235     240
Ala Ala Arg Ala Phe Glu Leu Ser Arg Gln Phe Gln Phe Lys Trp Thr
      245     250     255
Val Asn Trp Arg Met Leu Gly Glu Glu Val Phe Leu Ser Lys Tyr Phe
      260     265     270
Ala Leu Ser Leu Leu Ala Cys His Ile Leu Val Leu Leu Ile Phe Ile
      275     280     285
Ser Lys Arg Trp Ile Gln Pro Thr Gly Arg Ser Leu Tyr Asp Leu Ile
      290     295     300
Pro Ser Phe Leu Arg Leu Lys Ser Pro Phe Thr Met Gln Glu Gln Leu
      305     310     315     320
Arg Ile Ser His Tyr Val Thr Pro Glu Tyr Ala Met Thr Thr Met Leu
      325     330     335
Thr Ala Asn Leu Ile Gly Leu Leu Phe Ala Arg Ser Leu His Tyr Gln
      340     345     350
Phe Tyr Ala Tyr Leu Ala Trp Ala Thr Pro Tyr Leu Leu Trp Arg Ala
      355     360     365
Thr Glu Asp Pro Val Ile Val Ala Ile Ile Trp Ala Ala Gln Glu Trp
      370     375     380
Ala Trp Asn Val Tyr Pro Ser Thr Asp Leu Ser Ser Thr Ile Ala Val
      385     390     395     400
Asn Thr Met Leu Ala Thr Val Val Leu Val Tyr Leu Gly Thr Ala Arg
      405     410     415
Arg Ala Val Pro Ala Pro Ala Ala Gln Val Gly Asn Val Asp Asp Lys
      420     425     430
Asn Lys

```

<210> 130
 <211> 432
 <212> PRT
 <213> 玉蜀黍赤霉

```

<400> 130
Met Ala Asp Pro Ala Pro Gly Ala Leu Ala Arg Gly Thr Arg Phe Val
  1      5      10      15
Arg Asn Val Leu Thr Gly Gln His Ala Leu Ser Lys Leu Ile Pro Val
      20      25      30
Ala Leu Trp Leu Ala Asp Ala Val Gly Thr Ser Leu Ile Ile Trp Lys
      35      40      45
Val Pro Tyr Thr Glu Ile Asp Trp Glu Ala Tyr Met Gln Gln Val Ser
      50      55      60
Gln Phe Ile Ser Gly Glu Arg Asp Tyr Thr Lys Ile Glu Gly Gly Thr
      65      70      75      80
Gly Pro Leu Val Tyr Pro Ala Ala His Val Tyr Thr Phe Thr Gly Leu
      85      90      95
Tyr His Ile Thr Asn Glu Gly Glu Asn Ile Phe Leu Ala Gln Gln Ile
      100     105     110

```

Phe Gly Val Leu Tyr Met Ala Thr Leu Ala Val Val Met Leu Cys Tyr
 115 120 125
 Trp Lys Ala Lys Val Pro Pro Tyr Met Phe Val Phe Leu Ile Ala Ser
 130 135 140
 Lys Arg Leu His Ser Leu Phe Val Leu Arg Cys Phe Asn Asp Cys Phe
 145 150 155 160
 Ala Val Phe Phe Leu Trp Leu Ser Ile Tyr Phe Phe Gln Arg Arg Asn
 165 170 175
 Trp Thr Phe Gly Ser Leu Ala Tyr Thr Trp Gly Leu Gly Ile Lys Met
 180 185 190
 Ser Leu Leu Val Leu Pro Ala Ile Gly Val Ile Leu Leu Gly
 195 200 205
 Arg Gly Phe Trp Pro Gly Leu Arg Leu Ala Trp Leu Met Ala Gln Val
 210 215 220
 Gln Phe Ala Ile Gly Ile Pro Phe Ile Met Lys Asn Ser Arg Gly Tyr
 225 230 235 240
 Ala Ala Arg Ala Phe Glu Leu Ser Arg Glu Phe Lys Phe Glu Trp Thr
 245 250 255
 Val Asn Trp Arg Met Leu Gly Glu Glu Val Phe Leu Ser Lys Ser Phe
 260 265 270
 Ala Ile Phe Leu Leu Ala Cys His Val Thr Ala Leu Leu Val Phe Ile
 275 280 285
 Ser Gln Arg Trp Leu Gln Pro Thr Gly Arg Pro Leu Ser Ala Met Ile
 290 295 300
 Pro Ser Phe Leu Gln Leu Lys Ser Pro Phe Thr Leu Gln Glu Gln Leu
 305 310 315 320
 Arg Ile Ser His Tyr Val Thr Pro Glu Tyr Val Met Thr Thr Met Leu
 325 330 335
 Ser Ala Asn Val Ile Gly Leu Leu Phe Ala Arg Ser Leu His Tyr Gln
 340 345 350
 Phe Tyr Ala Tyr Leu Ala Trp Ala Ser Pro Tyr Leu Ile Trp Arg Ala
 355 360 365
 Thr Glu Asp Pro Phe Ile Val Leu Leu Ile Trp Ala Ala Gln Glu Trp
 370 375 380
 Ala Trp Asn Val Phe Pro Ser Thr Asp Leu Ser Ser Arg Val Thr Val
 385 390 395 400
 Gly Ala Met Leu Ala Thr Val Val Leu Ala Tyr Arg Gly Thr Ala Arg
 405 410 415
 [0069] Leu Ala Val Pro Ser Gln Ala Arg Lys Ile Glu Ala Lys Asn Lys
 420 425 430

<210> 131
 <211> 446
 <212> PRT
 <213> 嗜热毁丝菌

<400> 131
 Met Thr Arg Met Arg Ser Ser Pro Lys Thr Pro Thr Ala Thr Met Ala
 1 5 10 15
 Asp Gln Asn Arg Pro Ile His Ile Arg Ala Thr Arg Leu Val Phe Asp
 20 25 30
 Ile Leu Asn Gly Arg His Val Leu Ser Lys Leu Ile Pro Pro Leu Val
 35 40 45
 Phe Leu Ala Asp Ala Leu Leu Cys Ala Leu Ile Ile Trp Lys Val Pro
 50 55 60
 Tyr Thr Glu Ile Asp Trp Asn Ala Tyr Met Glu Gln Val Ala Gln Ile
 65 70 75 80
 Leu Ser Gly Glu Arg Asp Tyr Thr Lys Ile Arg Gly Asn Thr Gly Pro
 85 90 95
 Leu Val Tyr Pro Ala Ala His Val Tyr Ile Tyr Thr Gly Leu Tyr His
 100 105 110
 Leu Thr Asp Glu Gly Arg Asn Ile Leu Thr Ala Gln Lys Leu Phe Gly
 115 120 125
 Phe Leu Tyr Met Val Thr Leu Ala Val Val Met Ala Cys Tyr Trp Gln
 130 135 140
 Ala Lys Val Pro Pro Tyr Val Phe Pro Leu Leu Ile Leu Ser Lys Arg
 145 150 155 160
 Leu His Ser Ile Phe Val Leu Arg Cys Phe Asn Asp Cys Phe Ala Thr
 165 170 175
 Leu Phe Leu Trp Leu Ala Ile Phe Ala Leu Gln Arg Arg Ala Trp Arg
 180 185 190
 Thr Gly Ala Leu Met Tyr Thr Leu Gly Leu Gly Val Lys Met Ser Leu
 195 200 205
 Leu Leu Val Leu Pro Ala Val Gly Val Val Leu Leu Leu Gly Ala Gly
 210 215 220

Phe Ala Thr Ser Leu Arg Leu Ala Ala Val Ile Gly Leu Val Gln Val
 225 230 235 240
 Leu Ile Ala Val Pro Phe Leu Ser Asn Asn Pro Trp Gly Tyr Leu Gly
 245 250 255
 Arg Ala Phe Glu Leu Ser Arg Gln Phe Phe Phe Lys Trp Thr Val Asn
 260 265 270
 Trp Arg Phe Val Gly Glu Glu Val Phe Leu Ser Lys Glu Phe Ser Leu
 275 280 285
 Ala Leu Leu Gly Leu His Val Ala Val Leu Ala Ile Phe Val Thr Thr
 290 295 300
 Arg Trp Leu Lys Pro Ala Arg Lys Pro Val Ser Gln Leu Ile Val Pro
 305 310 315 320
 Ile Leu Leu Gly Lys Ser Pro Phe Thr Glu Glu Glu Gln Arg Ala Val
 325 330 335
 Ser Arg Asp Val Thr Pro Arg Phe Ile Leu Thr Ser Ile Leu Ser Ala
 340 345 350
 Asn Val Val Gly Leu Leu Phe Ala Arg Ser Leu His Tyr Gln Phe Tyr
 355 360 365
 Ser Tyr Leu Ala Trp Met Thr Pro Tyr Leu Leu Trp Arg Ser Gly Val
 370 375 380
 His Pro Ile Leu Gln Tyr Ala Ile Trp Thr Ala Gln Glu Trp Ala Trp
 385 390 395 400
 Asn Val Tyr Pro Ser Thr Pro Ile Ser Ser Gly Val Val Val Gly Val
 405 410 415
 Leu Ala Leu Thr Ala Ala Leu Val Trp Leu Gly Ala Arg Glu Asp Trp
 420 425 430
 Glu Pro Arg Arg Val Leu Leu Lys Gly Glu Ala Ala Lys Arg
 435 440 445

<210> 132
 <211> 442
 <212> PRT
 <213> 粗糙脉孢菌

[0070]

<400> 132
 Met Ala Ala Pro Ser Ser Arg Pro Glu Ser Asn Pro Pro Leu Tyr Lys
 1 5 10 15
 Gln Ala Leu Asp Phe Ala Leu Asp Val Ala Asn Gly Arg His Ala Leu
 20 25 30
 Ser Lys Leu Ile Pro Pro Ala Leu Phe Leu Val Asp Ala Leu Leu Cys
 35 40 45
 Gly Leu Ile Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp Ala Ala
 50 55 60
 Tyr Met Glu Gln Val Ser Gln Ile Leu Ser Gly Glu Arg Asp Tyr Thr
 65 70 75 80
 Lys Val Arg Gly Gly Thr Gly Pro Leu Val Tyr Pro Ala Ala His Val
 85 90 95
 Tyr Ile Tyr Thr Gly Leu Tyr His Leu Thr Asp Glu Gly Arg Asn Ile
 100 105 110
 Leu Leu Ala Gln Gln Leu Phe Ala Gly Leu Tyr Met Val Thr Leu Ala
 115 120 125
 Val Val Met Gly Cys Tyr Trp Gln Ala Lys Ala Pro Pro Tyr Leu Phe
 130 135 140
 Pro Leu Leu Thr Leu Ser Lys Arg Leu His Ser Ile Phe Val Leu Arg
 145 150 155 160
 Cys Phe Asn Asp Cys Phe Ala Val Leu Phe Leu Trp Leu Ala Ile Phe
 165 170 175
 Phe Phe Gln Arg Arg Asn Trp Gln Ala Gly Ala Leu Leu Tyr Thr Leu
 180 185 190
 Gly Leu Gly Val Lys Met Thr Leu Leu Leu Ser Leu Pro Ala Val Gly
 195 200 205
 Ile Val Leu Phe Leu Gly Ser Gly Ser Phe Val Thr Leu Gln Leu
 210 215 220
 Val Ala Thr Met Gly Leu Val Gln Ile Leu Ile Gly Val Pro Phe Leu
 225 230 235 240
 Ala His Tyr Pro Thr Glu Tyr Leu Ser Arg Ala Phe Glu Leu Ser Arg
 245 250 255
 Gln Phe Phe Phe Lys Trp Thr Val Asn Trp Arg Phe Val Gly Glu Glu
 260 265 270
 Ile Phe Leu Ser Lys Gly Phe Ala Leu Thr Leu Leu Ala Leu His Val
 275 280 285
 Leu Val Leu Gly Ile Phe Ile Thr Thr Arg Trp Ile Lys Pro Ala Arg
 290 295 300
 Lys Ser Leu Val Gln Leu Ile Ser Pro Val Leu Leu Ala Gly Lys Pro
 305 310 315 320

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

<210> 133
 <211> 413
 <212> PRT
 <213> 米曲霉

<400> 133
 Met Glu Leu Lys His Phe Ile His Glu Leu Cys Leu Asn Pro Arg His
 1 5 10 15
 Thr Lys Trp Ile Ala Pro Leu Leu Val Ile Gly Asp Ala Phe Leu Cys
 20 25 30
 Ala Leu Ile Ile Trp Lys Ile Pro Tyr Thr Glu Ile Asp Trp Thr Thr
 35 40 45
 Tyr Met Gln Gln Ile Ala Leu Tyr Ile Ser Gly Glu Arg Asp Tyr Thr
 50 55 60
 Leu Ile Lys Gly Ser Thr Gly Pro Leu Val Tyr Pro Ala Ala His Val
 65 70 75 80
 Tyr Ser Tyr Met Ala Leu Tyr His Leu Thr Asp Glu Gly Arg Asp Ile
 85 90 95
 Leu Phe Gly Gln Ile Leu Phe Ala Val Leu Tyr Leu Val Thr Leu Ala
 100 105 110
 Val Val Met Val Cys Tyr Arg Gln Ser Gly Ala Pro Pro Tyr Leu Phe
 115 120 125
 Pro Leu Leu Val Leu Ser Lys Arg Leu His Ser Val Phe Val Leu Arg
 130 135 140
 Leu Phe Asn Asp Gly Leu Ala Val Cys Ala Met Trp Ile Ala Ile Leu
 145 150 155 160
 Leu Phe Gln Asn Lys Lys Trp Thr Ala Gly Val Thr Ala Trp Thr Val
 165 170 175
 Gly Val Gly Ile Lys Met Thr Leu Leu Leu Leu Ala Pro Ala Ile Ala
 180 185 190
 Val Val Thr Val Leu Ser Leu Ser Leu Val Pro Ser Ile Arg Leu Gly
 195 200 205
 Ile Leu Ala Leu Leu Ile Gln Val Leu Leu Ala Ile Pro Phe Leu Gln
 210 215 220
 Gly Asn Pro Ile Gly Tyr Val Ala Arg Ala Phe Glu Leu Thr Arg Gln
 225 230 235 240
 Phe Met Phe Lys Trp Thr Val Asn Trp Arg Phe Val Gly Glu Asp Leu
 245 250 255
 Phe Leu Ser Lys Gln Phe Ser Leu Ala Leu Leu Gly Leu His Ile Phe
 260 265 270
 Leu Leu Gly Leu Phe Val Thr Thr Gly Trp Leu Arg Pro Ser Gly Ser
 275 280 285
 Asn Val Pro Asp Phe Leu Arg Ser Leu Leu Gln Gly Arg Gln Arg Thr
 290 295 300
 Val Val Leu Ser Lys Ser Phe Ile Met Thr Val Met Leu Thr Ser Leu
 305 310 315 320
 Ala Ile Gly Leu Leu Cys Ala Arg Ser Leu His Tyr Gln Phe Phe Ala
 325 330 335
 Tyr Leu Ser Trp Ala Thr Pro Cys Leu Leu Trp Arg Ala Arg Leu His
 340 345 350
 Pro Ile Leu Ile Tyr Ala Ile Trp Ala Leu Gln Glu Trp Ala Trp Asn
 355 360 365
 Val Tyr Pro Ser Thr Asn Ala Ser Ser Val Val Val Phe Ser Leu
 370 375 380
 Ala Val Gln Val Phe Gly Val Leu Leu Asn Ser Arg Asn Ala Leu Ser
 385 390 395 400
 Asp Ala Pro Pro Arg Arg Lys Gly Lys Glu His Ile Gln
 405 410

[0071]

<210> 134
 <211> 411
 <212> PRT
 <213> 费氏新萨托菌

<400> 134
 Met Asp Leu Lys His Thr Leu Arg Asp Leu Cys Met Asn Pro Arg His
 1 5 10 15
 Thr Arg Trp Val Ala Pro Leu Leu Ile Leu Gly Asp Ala Val Leu Cys
 20 25 30
 Ala Leu Ile Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp Thr Thr
 35 40 45
 Tyr Met Gln Gln Ile Ser Leu Tyr Ile Ser Gly Glu Arg Asp Tyr Thr
 50 55 60
 Leu Ile Lys Gly Ser Thr Gly Pro Leu Val Tyr Pro Ala Ala His Val
 65 70 75 80
 Tyr Ile Phe Asn Ile Leu Tyr His Leu Thr Asp Glu Gly Arg Asp Ile
 85 90 95
 Phe Leu Gly Gln Ile Leu Phe Ala Ile Leu Tyr Leu Ala Thr Leu Thr
 100 105 110
 Val Ala Met Thr Cys Tyr Arg Gln Ala Gly Ala Pro Pro Tyr Leu Leu
 115 120 125
 Val Pro Leu Val Leu Ser Lys Arg Leu His Ser Val Phe Met Leu Arg
 130 135 140
 Leu Phe Asn Asp Gly Phe Ala Ala Tyr Ala Met Trp Val Ser Ile Leu
 145 150 155 160
 Leu Phe Met Asn Lys Lys Trp Thr Ala Gly Ala Ile Val Trp Ser Thr
 165 170 175
 Gly Val Gly Ile Lys Met Thr Leu Leu Leu Leu Ala Pro Ala Ile Ala
 180 185 190
 Val Val Leu Val Leu Ser Leu Ser Leu Gly Pro Ser Met Gln Leu Gly
 195 200 205
 Phe Leu Ala Val Leu Ile Gln Val Leu Phe Gly Ile Pro Phe Leu Gln
 210 215 220
 Asn Asn Pro Ala Gly Tyr Val Ser Arg Ala Phe Glu Leu Thr Arg Gln
 225 230 235 240
 Phe Met Phe Lys Trp Thr Val Asn Trp Arg Phe Val Gly Glu Glu Leu
 245 250 255
 Phe Leu Ser Arg Lys Phe Ser Leu Ala Leu Leu Ala Leu His Ile Leu
 260 265 270
 Leu Leu Gly Leu Phe Val Ala Thr Val Trp Leu Lys Pro Ser Gly Ser
 275 280 285
 Asp Leu Pro Ser Phe Leu Gln Arg Leu Ile Gln Arg Arg Tyr Arg Thr
 290 295 300
 Ala Ser Leu Ser Lys Ser Phe Ile Met Thr Ala Met Leu Ser Ser Leu
 305 310 315 320
 Ala Ile Gly Leu Leu Cys Ala Arg Ser Leu His Tyr Gln Phe Phe Ala
 325 330 335
 Tyr Leu Ala Cys Ala Thr Pro Phe Leu Leu Trp Gln Ala Gly Phe His
 340 345 350
 Pro Ile Leu Val Tyr Val Val Trp Val Ala Gln Glu Trp Ala Trp Asn
 355 360 365
 Thr Tyr Pro Ser Thr Asn Ala Ser Ser Leu Val Val Ile Leu Ser Leu
 370 375 380
 Ala Ala Gln Val Phe Gly Val Leu Gly Asn Ser Phe Ser Arg Lys His
 385 390 395 400
 Leu Asp Gln Ser Ser Gln Lys Glu His Leu Gln
 405 410

[0072]

<210> 135
 <211> 413
 <212> PRT
 <213> 黑曲霉

<400> 135
 Met Asp Trp Met Arg Leu Ile Arg Asp Leu Cys Phe Asn Pro Arg His
 1 5 10 15
 Thr Lys Trp Met Ala Pro Leu Leu Val Leu Gly Asp Ala Phe Leu Cys
 20 25 30
 Ala Leu Ile Ile Trp Lys Val Pro Tyr Thr Glu Ile Asp Trp Ala Thr
 35 40 45
 Tyr Met Gln Gln Ile Ser Leu Tyr Leu Ser Gly Glu Arg Asp Tyr Thr
 50 55 60

[0073]

Leu Ile Arg Gly Ser Thr Gly Pro Leu Val Tyr Pro Ala Ala His Val
 65 70 75 80
 Tyr Ser Tyr Thr Ala Leu Tyr His Leu Thr Asp Glu Gly Arg Asp Ile
 85 90 95
 Phe Phe Gly Gln Ile Leu Phe Ala Val Leu Tyr Leu Ile Thr Leu Val
 100 105 110
 Val Val Leu Cys Cys Tyr Arg Gln Ser Gly Ala Pro Pro Tyr Leu Leu
 115 120 125
 Pro Leu Leu Val Leu Ser Lys Arg Leu His Ser Val Tyr Val Leu Arg
 130 135 140
 Leu Phe Asn Asp Gly Leu Ala Ala Leu Ala Met Trp Val Ala Ile Leu
 145 150 155 160
 Leu Phe Met Asn Arg Lys Trp Thr Ala Ala Val Ala Val Trp Ser Thr
 165 170 175
 Gly Val Ala Ile Lys Met Thr Leu Leu Leu Leu Ala Pro Ala Ile Ala
 180 185 190
 Val Val Thr Val Leu Ser Leu Ser Leu Gly Pro Ser Val Gly Leu Gly
 195 200 205
 Val Leu Ala Val Leu Val Gln Val Leu Leu Ala Ile Pro Phe Leu Gln
 210 215 220
 Asn Asn Pro Ala Gly Tyr Leu Ser Arg Ala Phe Glu Leu Thr Arg Gln
 225 230 235 240
 Phe Met Phe Lys Trp Thr Val Asn Trp Arg Phe Val Gly Glu Glu Val
 245 250 255
 Phe Leu Ser Lys Ser Phe Ser Leu Ala Leu Leu Ala Val His Ile Val
 260 265 270
 Leu Leu Gly Ala Phe Ala Val Thr Gly Trp Leu Arg Tyr Ser Arg Ser
 275 280 285
 Ser Leu Pro Ala Phe Ile Arg Asn Leu Leu Ala Gly Arg His Arg Thr
 290 295 300
 Val Ser Leu Pro Lys Pro Tyr Ile Met Ser Val Met Leu Ser Ser Leu
 305 310 315 320
 Thr Val Gly Leu Leu Cys Ala Arg Ser Leu His Tyr Gln Phe Phe Ala
 325 330 335
 Tyr Leu Ser Trp Ala Thr Pro Phe Leu Leu Trp Arg Ala Gly Phe His
 340 345 350
 Pro Ile Leu Leu Tyr Leu Ile Trp Ala Met Gln Glu Trp Ala Trp Asn
 355 360 365
 Thr Phe Pro Ser Thr Asn Leu Ser Ser Ile Ile Val Val Leu Ser Leu
 370 375 380
 Ala Thr Gln Ser Phe Gly Val Leu Ala Asn Ser Ala Ser Ala Phe Tyr
 385 390 395 400
 Thr Met Arg Ser Asn Pro Ser Gly Lys Glu His Asn Gln
 405 410

<210> 136
 <211> 357
 <212> PRT
 <213> 稻瘟病菌

<400> 136
 Met Ala Ala Glu Arg Pro Ser Thr Leu Gly Lys Pro Val Gln Phe Val
 1 5 10 15
 Phe Asp Val Ala Asn Gly Arg His Pro Leu Ser Arg Ala Ile Pro Pro
 20 25 30
 Met Leu Leu Ala Phe Asp Gly Leu Leu Cys Gly Leu Ile Ile Lys Lys
 35 40 45
 Val Pro Ser Cys Tyr Arg Lys Ala Lys Val Pro Pro Tyr Val Leu Pro
 50 55 60
 Leu Leu Val Leu Ser Lys Arg Leu His Ser Ile Phe Val Leu Arg Cys
 65 70 75 80
 Phe Asn Asp Cys Phe Ala Val Leu Phe Phe Trp Leu Ala Ile Tyr Cys
 85 90 95
 Phe Gln Arg Arg Ala Trp Ser Leu Gly Gly Val Phe Tyr Ser Phe Gly
 100 105 110
 Leu Gly Ile Lys Met Thr Val Leu Leu Ser Leu Pro Ala Val Gly Val
 115 120 125
 Ile Leu Leu Leu Gly Arg Gly Phe Gly Gly Ala Leu Asn Val Ala Ser
 130 135 140
 Ile Met Gly Gln Leu Gln Val Ala Ile Gly Leu Pro Phe Leu Ser Lys
 145 150 155 160
 Asn Ala Trp Gly Tyr Leu Ser Arg Ala Phe Glu Leu Ser Arg Gln Phe
 165 170 175
 Met Phe Lys Trp Thr Val Asn Trp Arg Phe Val Gly Glu Glu Thr Phe
 180 185 190

Leu Ser Lys Pro Phe Ala Ile Thr Leu Leu Ala Leu His Ala Ser Val
 195 200 205
 Leu Leu Ala Phe Val Thr Lys Arg Trp Leu Lys Pro Ala Ser Lys Ser
 210 215 220
 Ile Gly Gly Leu Ile Ala Pro Leu Leu Scr Gly Arg Pro Ile Phe Thr
 225 230 235 240
 Ala Glu Glu Ala Gln Thr Ala Ala Arg Ala Val Thr Pro Glu Tyr Val
 245 250 255
 Met Thr Thr Met Leu Thr Ala Asn Ile Val Gly Met Leu Phe Ala Arg
 260 265 270
 Ser Leu His Tyr Gln Phe Tyr Ala Tyr Leu Ala Trp Ser Thr Pro Tyr
 275 280 285
 Leu Leu Trp Arg Ser Gly Ile His Pro Leu Leu Gln Trp Gly Leu Trp
 290 295 300
 Ala Leu Gln Glu Trp Ala Trp Asn Val Tyr Pro Ser Thr Pro Val Ser
 305 310 315 320
 Ser Gly Val Val Val Gly Val Met Ala Ile Thr Val Gly Ala Val Met
 325 330 335
 Val Gly Ala Lys Ala Glu Phe Arg Pro Gln Val Pro Val Ala Lys Lys
 340 345 350
 Val Glu Ala Lys Arg
 355

<210> 137

<211> 406

<212> PRT

<213> 粟酒裂殖酵母

<400> 137

Met Ser Ser Val Glu Thr Arg Asn Ser Phe Asn Pro Phe Arg Val Leu
 1 5 10 15
 Phe Asp Leu Gly Ser Tyr Gly Trp Leu His Pro Ser Arg Leu Leu Leu
 20 25 30
 Leu Glu Ile Pro Phe Val Phe Ala Ile Ile Ser Lys Val Pro Tyr Thr
 35 40 45
 Glu Ile Asp Trp Ile Ala Tyr Met Glu Gln Val Asn Ser Phe Leu Leu
 50 55 60
 Gly Glu Arg Asp Tyr Lys Ser Leu Val Gly Cys Thr Gly Pro Leu Val
 65 70 75 80
 Tyr Pro Gly Gly His Val Phe Leu Tyr Thr Leu Leu Tyr Tyr Leu Thr
 85 90 95
 Asp Gly Gly Thr Asn Ile Val Arg Ala Gln Tyr Ile Phe Ala Phe Val
 100 105 110
 Tyr Trp Ile Thr Thr Ala Ile Val Gly Tyr Leu Phe Lys Ile Val Arg
 115 120 125
 Ala Pro Phe Tyr Ile Tyr Val Leu Leu Ile Leu Ser Lys Arg Leu His
 130 135 140
 Ser Ile Phe Ile Leu Arg Leu Phe Asn Asp Gly Phe Asn Ser Leu Phe
 145 150 155 160
 Ser Ser Leu Phe Ile Leu Ser Ser Cys Lys Lys Trp Val Arg Ala
 165 170 175
 Ser Ile Leu Leu Ser Val Ala Cys Ser Val Lys Met Ser Ser Leu Leu
 180 185 190
 Tyr Val Pro Ala Tyr Leu Val Leu Leu Gln Ile Leu Gly Pro Lys
 195 200 205
 Lys Thr Trp Met His Ile Phe Val Ile Ile Ile Val Gln Ile Leu Phe
 210 215 220
 Ser Ile Pro Phe Leu Ala Tyr Phe Trp Ser Tyr Trp Thr Gln Ala Phe
 225 230 235 240
 Asp Phe Gly Arg Ala Phe Asp Tyr Lys Trp Thr Val Asn Trp Arg Phe
 245 250 255
 Ile Pro Arg Ser Ile Phe Glu Ser Thr Ser Phe Ser Thr Ser Ile Leu
 260 265 270
 Phe Leu His Val Ala Leu Leu Val Ala Phe Thr Cys Lys His Trp Asn
 275 280 285
 Lys Leu Ser Arg Ala Thr Pro Phe Ala Met Val Asn Ser Met Leu Thr
 290 295 300
 Leu Lys Pro Leu Pro Lys Leu Gln Leu Ala Thr Pro Asn Phe Ile Phe
 305 310 315 320
 Thr Ala Leu Ala Thr Ser Asn Leu Ile Gly Ile Leu Cys Ala Arg Ser
 325 330 335
 Leu His Tyr Gln Phe Tyr Ala Trp Phe Ala Trp Tyr Ser Pro Tyr Leu
 340 345 350
 Cys Tyr Gln Ala Ser Phe Pro Ala Pro Ile Val Ile Gly Leu Trp Met
 355 360 365

[0074]

Leu Gln Glu Tyr Ala Trp Asn Val Phe Pro Ser Thr Lys Leu Ser Ser
 370 375 380
 Leu Ile Ala Val Cys Val Pro Leu Ile Thr Ile Leu Lys Leu Tyr Thr
 385 390 395 400
 Ser Asp Tyr Arg Lys Pro
 405

<210> 138
 <211> 30
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 138
 ggaggtgggg gcagtgaggg tggcggcagl 30

<210> 139
 <211> 2460
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成构建体

[0075]

<400> 139
 atgcgcttcc gaatctacaa gcggaaggte ctcattctga cccctgtcgt ggccgcttgc 60
 ggctttgttc tctggteccag caacggctgc cagcgtaaga acgaggccct ggccgctccc 120
 ctcttggacg ccgaaccggc cagaggcgca ggtggcaggg gaggggatca cccctcggtc 180
 gctgtcggca tccgccgcgt cagcaatgtg tccgccgct ctctggtecc ggccggttccg 240
 cagcctgagg cagacaacct cagctgcgc taccgatcac tegtgtatca acttaacttc 300
 gaccagactc tgcggaacgt cgacaaggcc ggaacctggg ctccgcgtga gttggltccic 360
 gtctgtcagg tgcacaacag gcccgagtac ctccgcctcc tgcctggattc gcttcgaaag 420
 gcccgaggca tgcacaacgt cctgggtgatt ttcagccatg acttttggte cacagagatc 480
 aatcagctca ttgcgggtgt caacttttgc cccgtcttgc aagttttctt ccttttctct 540
 atccaactct accccaacga gtccccgggc agtgaccccc gcgactgtcc tccggatctg 600
 ccaaaaaacg ccgctctcaa gclgggctgc atcaacgccg aataccccga cagctttggc 660
 cactatcgcg agggcaagtt ctccagacg aagcaccact ggtgtgtgaa gctccatttt 720
 gtctgggagc gagtgaagat ccttcgtgat taccgaggac tcattctgtt cttggaagag 780
 gaccactacc tggccccgga ctctaccac gtctttaaga agatgtggaa gctcaagcag 840
 caggaaatgcc ccgagtgcga cgttctgtcc ctggcaccct atagcgctgc ccgctcgttc 900
 tacggtatgg ctgacaaggc cgtgtgaaa acctggaagt caactgagca caatatgggc 960
 ctccgcttga cgaggaaacgc ctaccagaaa ctcatcgagt gtaccgacac cttctgcacg 1020
 tacgacgact ataactggga ttggacactg cagtacttga ctgtcagctg cctccctaag 1080
 ttttggaaag tccttgttcc ccagatcccc agaattttcc atgctggcga ctgcgggatg 1140
 caccacaaga aaacctgtcg cccatccacg cagctctgcc aaatcgagtc gctccgaaac 1200
 aacaacaagc agtacatgtt ccccgagaca ctgaccatta gcgagaagtt tacggtcgtg 1260
 gcgacttccc gcctcgaaga gaalggcggc tggggtgaca tccgcgatca cgagctgtgc 1320
 aagttttacc gccggctcca gggagggtgg ggcagtgag gtggcgagcag tgggagggtg 1380
 cccaccgcgc cccctccgcg ccagccgcgt gtgcctgtga ccccgcgcc ggccggtgatt 1440
 cccatcctgg tcatcgctg tgaccgcagc actgttcggc gctgccttga caagctgctg 1500
 cattatcggc cctcggttga gctcttcccc atcatcgta gccaggactg cgggcacgag 1560
 gagacggccc aggcctatgc ctctacggc agcgcggtca cgcacatccg gcagcccgac 1620
 ctgagcagca ttgcgggtgc gccggaccac cgeaagttcc agggctacta caagatcgcg 1680
 cgccactacc gctgggcgct gggccaggtc ttccggcagt ttgccttccc cggcgccgtg 1740
 gtgggtggag atgacctgga ggtggccccc gacttcttcg agtactttcg ggccacctat 1800
 ccgctgctga aggcggaccc ctccctgtgg tgcgtctcgg cctggaatga caacggcaag 1860
 gagcagatgg tggacgccag caggccctgag ctgctctacc gcaccgactt ttcccttggc 1920
 ctgggctggc tgcgttggc cgagctctgg gctgagctgg agcccaagtg gccaaaggcc 1980
 ttctgggacg actggatgcg gcggccggag cagcggcagg ggcgggcctg catccgccct 2040
 gagatctcaa gaacgatgac ctttggccgc aagggtgtga gccacgggca gttctttgac 2100
 cagcaactca agttcatcaa gctgaaccag cagtttgtgc acttcaccca gctggacctg 2160
 tcttacctgc agcgggaggc ctatgaccga gatttctcgc cccgcgtcta cgggtctccc 2220
 cagctgcagg tggagaaagt gaggaccaat gaccggaagg agctggggga ggtgcgggtg 2280
 cagtacacgg gcagggacag cttcaaggct ttccccaagg ctctgggtgt catggatgac 2340
 ctcaagtcgg gggttccgag agctggctac cggggcattg tcaccttcca gttccggggc 2400
 cgccgtgtcc acctggcgcc cccaccgacg tgggagggtc atgatccag ctggaattag 2460

<210> 140
 <211> 45
 <212> DNA
 <213> 人工序列

<220>
<223> 合成构建体

<400> 140
ggagggtgggg gcagtgagg tgccggcagtg ggcggcgggtg gaagt 45

<210> 141
<211> 2475
<212> DNA
<213> 人工序列

<220>
<223> 合成构建体

<400> 141
atgcgcttcc gaattctacaa gcggaaggct ctcattctga ccttctgtct ggccgcttgc 60
ggcctttgttc tctgggtccag caacgggtcgc cagcgtaaga acgaggccct ggccgctccc 120
ctcttggcag ccgaaccggc cagaggcgca ggtggcaggg gaggggatca cccctcggtc 180
gtctgtcgca tccgccgctg cagcaatgtg tccgccgctt ctcgtgtccc ggccggttccg 240
cagcctgagg cagacaacct cagcgtgcgc taccgatcac tctgttatca acttaacttc 300
gaccagactc tgcggaacct cgacaaggcc ggaacctggg ctccgcgtga gttggtcctc 360
gtcgttcagg tgcacaacag gcccgagtac ctccgcctcc tgcgtgattc gcttcgaaag 420
gcccagggca tcgacaacct cctggtgatt ttcagccatg acttttggtc cacagagatc 480
aatcagctca ttgcgggtgt caacttttgc cccgtcttgc aagttttctt ccttttctct 540
atccaactct accccaacga gttcccgggc agtgaccccc gcgactgttc tcgggatctg 600
ccaaaaaacg ccgctctcaa gctgggctgc atcaacgccg aataccccga cagctttggc 660
cactatcgcg aggccaaagt ctccgagacg aagcaccact ggtgtgtgaa gctccatttt 720
gtctgggagc gagtgaagat ccttcgtgat tacgcaggac tcatctgtt cttggaagag 780
gaccactacc tggccccgga cttctaccac gtctttaaga agatgtggaa gctcaagcag 840
caggaatgcc ccgagtgcga cgttctgtcc ctggcacct atagcgcgtc ccgctcgttc 900
tacggtatgg ctgacaaggt cgatgtgaaa acctggaagt caactgagca caatatgggc 960
ctgcctctga cgaggaaacgc ctaccagaaa ctcatcgagt gtaccgacac ctctgcacg 1020
lacgacgact ataatggga ttggacactg cagtacttga ctgtcagctg cctccctaag 1080
ttttggaagg tcttgttcc ccagatcccg agaattttcc atgtgggga ctgcgggatg 1140
caccacaaga aaacctgtcg cccatccacg cagctctccc aaatcgagtc gctcctgaac 1200
aacaacaagc agtacaatgt ccccgagaca ctgaccatta gcgagaagtt tacggtcgtg 1260
gcgatctccc cgcctcgaaa gaattggcggc tggggtgaca tccgcgatca cgagctgtgc 1320
aagtcttacc gccggctcca gggagggtggc ggcagtgagg gtggcgccag tggaggtggc 1380
ggcagtgagg ggggtgccac cgcgcgccct cccgcccagc cgcgtgtgcc tgtgacccc 1440
gcgcggcggt tgattcccat cctggtcatc gcctgtgacc gcagcactgt tcggcgctgc 1500
ctggacaagc tctgcatta tcggccctcg gctgagctct tcccctcat cgtcagccag 1560
gactgcgggc acgaggagac ggcccaggcc atcgcttctt acggcagcgc ggtcacgcac 1620
atccggcagc ccgacctgag cagcattgct gtgcccggc accaccgcaa gttccagggc 1680
tactacaaga tcgcgcgcca ctaccgctgg gcgctgggca aggtcttccg gcagtttcgc 1740
ttcccccgcc ccgtgggtgt ggaggatgac ctggagggtg ccccggaact cttcgagtac 1800
tttcgggcca cctatecgt gctgaaggcc gacccctccc tgtgtgtcgt ctccgcttgg 1860
aatgacaacg gcaaggagca gatggtggac gccagcaggc ctgagctgct ctaccgcacc 1920
gactttttcc ctggcctggg ctggctgctg ttggccgagc tctgggctga gctggagccc 1980
aagtggccaa aggccttctg ggacgactgg atgcggcgcc cggagcagcg gcaggggcgg 2040
gctgtcatcc gccctgagat ctcaagaacg atgacctttg gccgcaaggg tgtgagccac 2100
gggcagttct ttgaccagca cctcaagttc atcaagctga accagcagtt tgtgacttcc 2160
accagctgg acctgtctta cctgcagcgg gaggcctatg accgagattt cctcggcccgc 2220
gtctacggtg ctccccagct gcagggtggag aaagtgagga ccaatgaccg gaaggagctg 2280
ggggaggtgc ggggtgcagta cacgggcagg gacagcttca aggttttgc caaggctctg 2340
gggtgtcatg atgacctcaa gtccgggggt cccagagctg gctaccgggg cattgtcacc 2400
ttccagttcc ggggcggccg tgtccacctg gcgccccac cgacgtggga gggctatgat 2460
ccagctgga attag 2475

[0076]

<210> 142
<211> 57
<212> DNA
<213> 里氏木霉

<400> 142
agcaccggca accctagcgg cggcaacctt cccggcggaa acccgcttgg cagcacc 57

<210> 143
<211> 2487
<212> DNA
<213> 人工序列

<220>
<223> 合成构建体

<400> 143
atgcgcttcc gaattctacaa gcggaaggct ctcattctga ccttctgtct ggccgcttgc 60

```

ggctttgttc tctggtccag caacggtcgc cagcgtaaga acgaggccct ggccctccc 120
ctcttgagc ccgaaccggc cagaggcgca ggtggcaggg gaggggatca cccctcggtc 180
gctgtcggca tccgcccggt cagcaatgtg tccgcccgct ctctggtccc ggccggtccc 240
cagcctgagg cagacaacct cagcgtgcgc taccgatcac tctgtatca acttaacttc 300
gaccagactc tgcggaacgt cgacaaggcc ggaacctggg ctccgcgtga gttggtcctc 360
gtcgttcagg tgcacaacag gcccgagtag ctccgcctcc tgcgtgattc gcttcgaaag 420
gcccagggca tgcacaacgt cctggtgatt ttcagccatg acttttggtc cacagagatc 480
aatcagctca ttgcgggtgt caacttttgc cccgtcttgc aagttttctt ccttttctct 540
atccaactct accccaacga gtccccgggc agtgaccccc gcactgtcc tgggactctg 600
ccaaaaaacg ccgctctcaa gctgggctgc atcaacgccg aataccccga cagcttttgc 660
cactatcgcg aggccaaagt ctccgagacg aagcaccact ggtggtggaa gctccatttt 720
gtctgggagc gagtgaagat ccttcgtgat tacgcaggac tcattctgtt cttggaagag 780
gaccactacc tggccccgga cttctaccac gtctttaaga agatlgggaa gctcaagcag 840
caggaatgcc ccgagtgcga cgttctgtcc ctggcacct atagcgcgtc ccgctcgttc 900
tacggtatgg ctgacaaggt cgaatgtaaa acctggaagt caactgagca caatatgggc 960
ctgcgcctga caggaacgc ctaccagaaa ctcatcaggt gtaccgacac cttctgcacg 1020
tacgacgact ataactggga ttggacactg cagtacttga ctgtcagctg cctccctaag 1080
ttttggaagg tcttgttcc ccagatcccc agaattttcc atgctggcga ctgcgggatg 1140
caccacaaga aaacctgtcg cccatccacg cagtctgccc aaatcgagtc gctcctgaac 1200
aacaacaagc agtactlgt ccccgagaca ctgaccatta gcgagaagtt tacggtcgtg 1260
gctcgtctcc cgctcgaaa gaatggcggc tggggtgaca tccgcgatca cagactgtgc 1320
aagtcttacc gccggtccca gacacccggc aacctagcg gcggcaacct tcccggcgga 1380
aaccgcctg gcagacccgg gagggtgccc accgcggccc ctcccgccca gccgcgtgtg 1440
cctgtgacc ccgcgcgggc ggtgattccc atcttggtca tgcctgtga ccgcagcact 1500
gttcgcgcgt gccctggaca gctgctgcat tatcgccct cggtcagct cttccctc 1560
atcgtcagcc aggaactgccc gcacgaggag acggcccgag ccacgcctc ctacggcagc 1620
gcggtcacgc acatccggca gcccgacctg agcagcattg cgggtccgcc ggaccaccgc 1680
aagttccagg gctactacaa gatcgcgccc cactaccgct ggccgctggg ccaggtcttc 1740
cggcagtttc gcttccccgc cgcctgtgtg gtggaggatg acctggaggt ggccccggac 1800
ttcttcgagt actttcgggc caactatccc ctgctgaagg ccgaccctc cctgtgggtc 1860
gtctcggcct ggaatgacaa cggcaaggag cagatggtgg acccgacag ccctgagctg 1920
ctctaccgca ccgactttt ccttgccctg ggtggtgctc tgttggccga gctctgggtc 1980
gagctggagc ccaagtggcc aaaggccttc tgggacgact ggatgcggcg gccggagcag 2040
ggcagggggc gggcctgcat ccgcctgag atctcaagaa cgaatgacct tggccgcaag 2100
ggttggagcc accggcagtt cttgaccag cacctcaagt tcataagct gaaccagcag 2160
tttgtgact tcaccagct ggacctgtct tacctgcagc gggaggccta tgaccgagat 2220
ttcttcgccc gcgtlclagg tgcctcccag ctgcaggtgg agaaagttag gaccaatgac 2280
cggaaggagc tggggagggt gcgggtgcag tacacgggca gggacagctt caaggcttcc 2340
gccaaggctc tgggtgtcal ggaatgacct aagtcggggg ttccgagagc tggctaccgg 2400
ggcattgtca ccttccagtt ccggggccgc cgtgtccacc tggcgcccc accgacgtgg 2460
gagggtatg atccacgtg gaattag 2487

```

[0077]

<210> 144
 <211> 65
 <212> DNA
 <213> 里氏木霉

<400> 144
 agctccgccg cgacggccac cgccagcgcc actgttctctg gagcggttag cggccccacc 60
 agcgg 65

<210> 145
 <211> 2493
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 145
 atgcgcttcc gaattctaaa ggggaaggtc ctcattctga ccttgtcgt ggccgcttgc 60
 ggctttgttc tctggtccag caacggtcgc cagcgtaaga acgaggccct ggccctccc 120
 ctcttgagc ccgaaccggc cagaggcgca ggtggcaggg gaggggatca cccctcggtc 180
 gctgtcggca tccgcccggt cagcaatgtg tccgcccgct ctctggtccc ggccggtccc 240
 cagcctgagg cagacaacct cagcgtgcgc taccgatcac tctgtatca acttaacttc 300
 gaccagactc tgcggaacgt cgacaaggcc ggaacctggg ctccgcgtga gttggtcctc 360
 gtcgllcagc tgcacaacag gcccgagtag ctccgcctcc tgcgtgattc gcttcgaaag 420
 gcccagggca tgcacaacgt cctggtgatt ttcagccatg acttttggtc cacagagate 480
 aatcagctca ttgcgggtgt caacttttgc cccgtcttgc aagttttctt ccttttctct 540
 atccaactct accccaacga gtccccgggc agtgaccccc gcactgtcc tgggactctg 600
 ccaaaaaaac ccgctctcaa gctgggctgc atcaacgccg aataccccga cagcttttgc 660
 cactatcgcg aggccaaagt ctccgagacg aagcaccact ggtggtggaa gctccatttt 720
 gtctgggagc gagtgaagat ccttcgtgat tacgcaggac tcattctgtt cttggaagag 780
 gaccactacc tggccccgga cttctaccac gtctttaaga agatgtggaa gctcaagcag 840
 caggaatgcc ccgagtgcga cgttctgtcc ctggcacct atagcgcgtc ccgctcgttc 900
 tacggtatgg ctgacaaggt cgaatgtaaa acctggaagt caactgagca caatatgggc 960

ctcgccctga cgaggaacgc ctaccagaaa ctcatcgagt gtaccgacac cttctgcacg 1020
 tacgacgact ataactggga ttggacactg cagtacttga ctgtcagctg cctccctaag 1080
 ttttgggaagg tccttgttcc ccagatcccg agaattttcc atgctggcga ctgctgggatg 1140
 caccacaaga aaacctgtcg cccatccacg cagtcgtccc aaatcgagtc gctcctgaac 1200
 aacaacaagc agtacatgtt ccccgagaca ctgaccatta gcgagaagtt tacggctcgtg 1260
 gcgatctccc cgcctcgaaa gaattggcggc tggggtgaca tccgcgatca cgagctgtgc 1320
 aagtccttacc gccggctcca gagctcccgc gcgacggcca ccgccagcgc cactgttcct 1380
 ggaggcggta cgggcccgcac cagcggggagg gtgcccaccg ccgccctcc cggccagccg 1440
 cgtgtgcctg tgacccccgc gccggcgggtg attcccaccc tggtcacgc ctgtgaccgc 1500
 agcactgttc ggcgctgcct ggacaagctg ctgcattatc ggccctcggc tgagctcttc 1560
 cccatcatcg tcagccagga ctgcgggcac gaggagacgg cccaggccat cgcctcctac 1620
 ggagcgcggg tcacgcacat ccggcagccc gacctgagca gcattgcggg gccgccggac 1680
 caccgcaagt tccagggtca ctacaagatc gcgcgccact accgctgggc gctgggccag 1740
 gtcttcggc agtttcgctt ccccgccgcc gtgggtgtgg aggatgacct ggagggtggc 1800
 ccggacttct tcgagtactt tcgggccacc tatccgtgc tgaaggccga ccctccctg 1860
 tgggtgcctc cggcctggaa tgacaacggc aaggagcaga tgggtggacgc cagcaggcct 1920
 gagctgctct accgcaccga cttttccct ggcttgggt ggctgctgtt ggccgagctc 1980
 tgggctgagc tggagcccaa gtggccaaag gccttctggg acgactggat gcggcggccg 2040
 gagcagcggc aggggcgggc ctgcatccgc cctgagatct caagaacgat gacctttggc 2100
 cgcaagggtg tgagccacgg gcagttcttt gaccagcacc tcaagtcat caagctgaac 2160
 cagcagtttg tgcacttcac ccagctggac ctgtcttacc tgcagcggga ggcttatgac 2220
 cgagatttcc tcgcccgcgt ctacgggtgt ccccagctgc aggtggagaa agtgaggacc 2280
 aatgaccgga aggagctggg ggaggtgcgg gtgcagtaca cgggcaggga cagcttcaag 2340
 gctttcgcca aggtcttggg tgtcatggat gacctcaagt cgggggttcc gagagctggc 2400
 taccggggca ttgtcacctt ccagttccgg gcccgccgtg tccacctggc gccccaccg 2460
 acgtgggagg gctatgatcc cagctggaat tag 2493

[0078]

<210> 146
 <211> 45
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 146
 ggtaccgggc ccactgcgca tcatgcgctt ccgaatctac aagcg 45

<210> 147
 <211> 38
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 147
 ggcgcgccac tagtctaatt ccagctggga tcatagcc 38

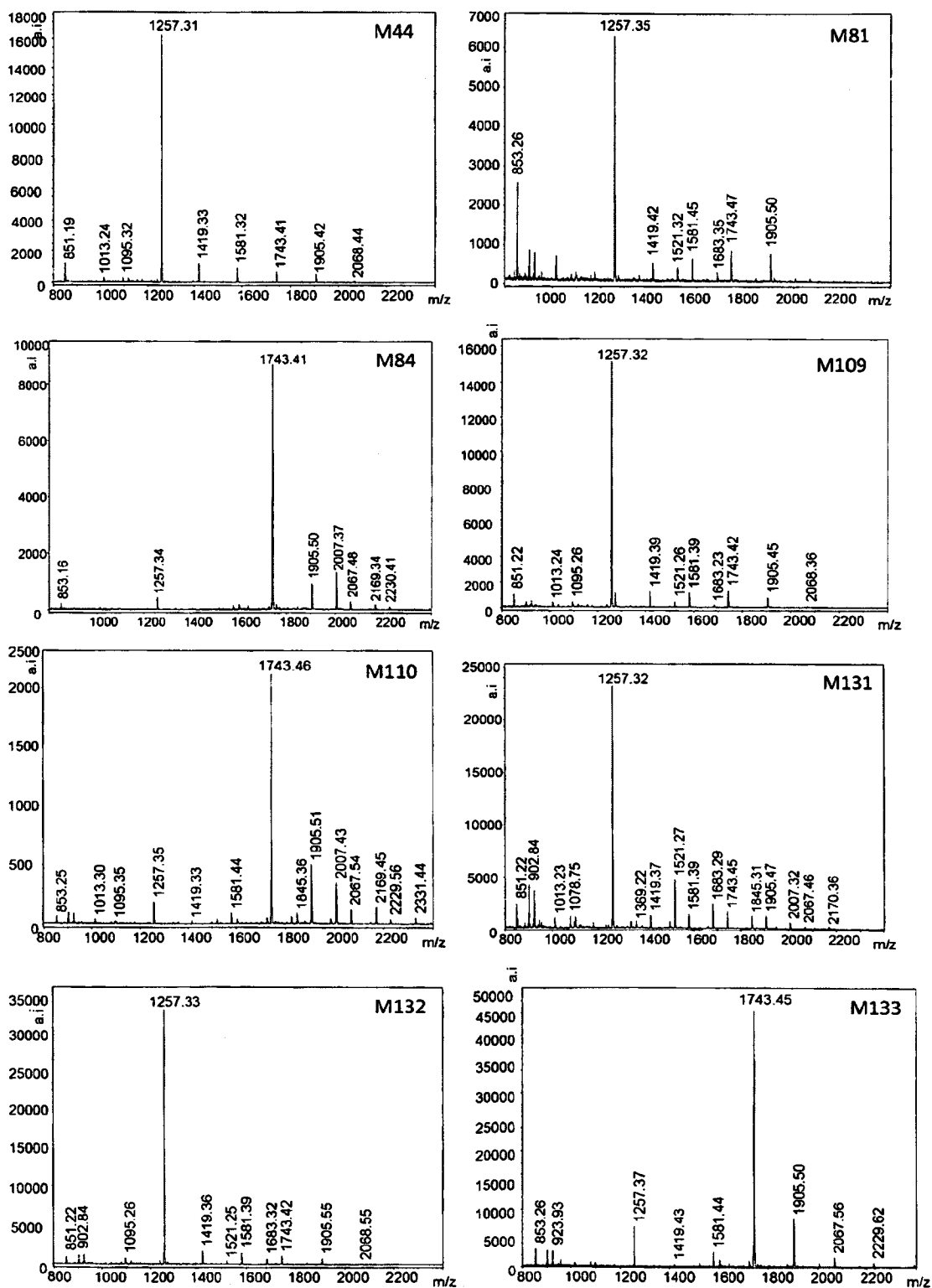


图1

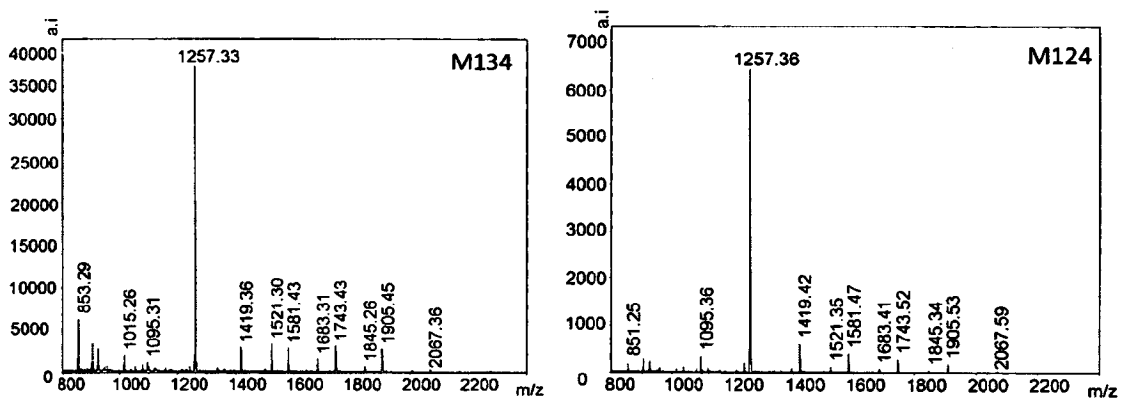


图1(续)

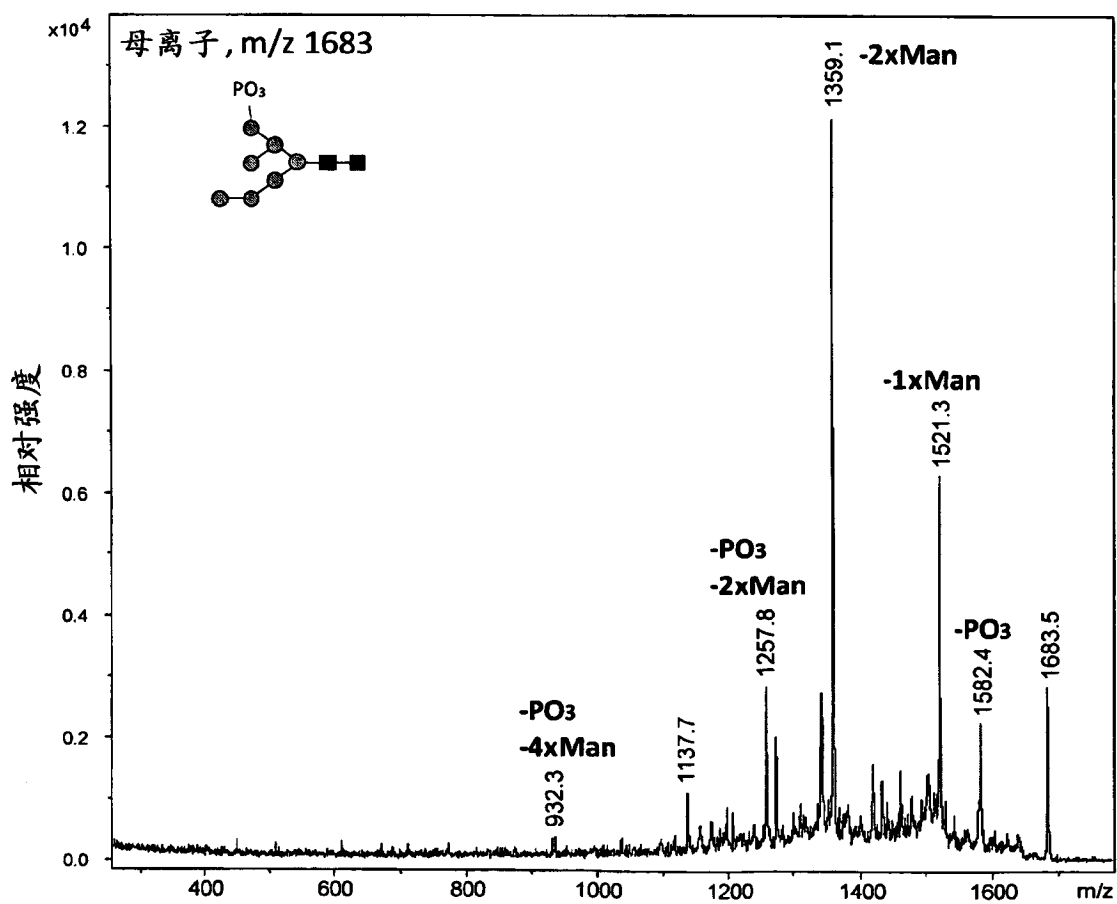


图2

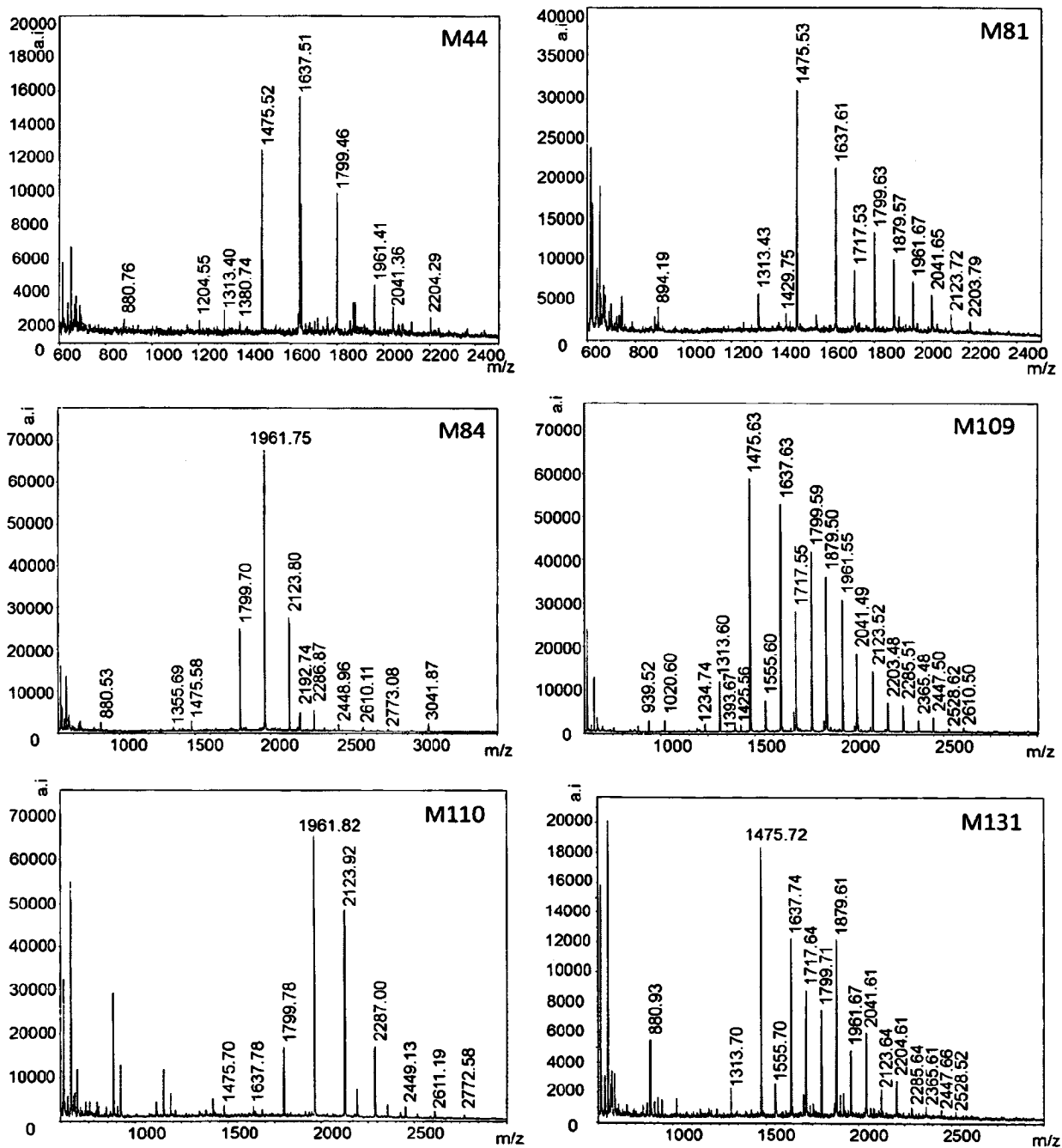


图3

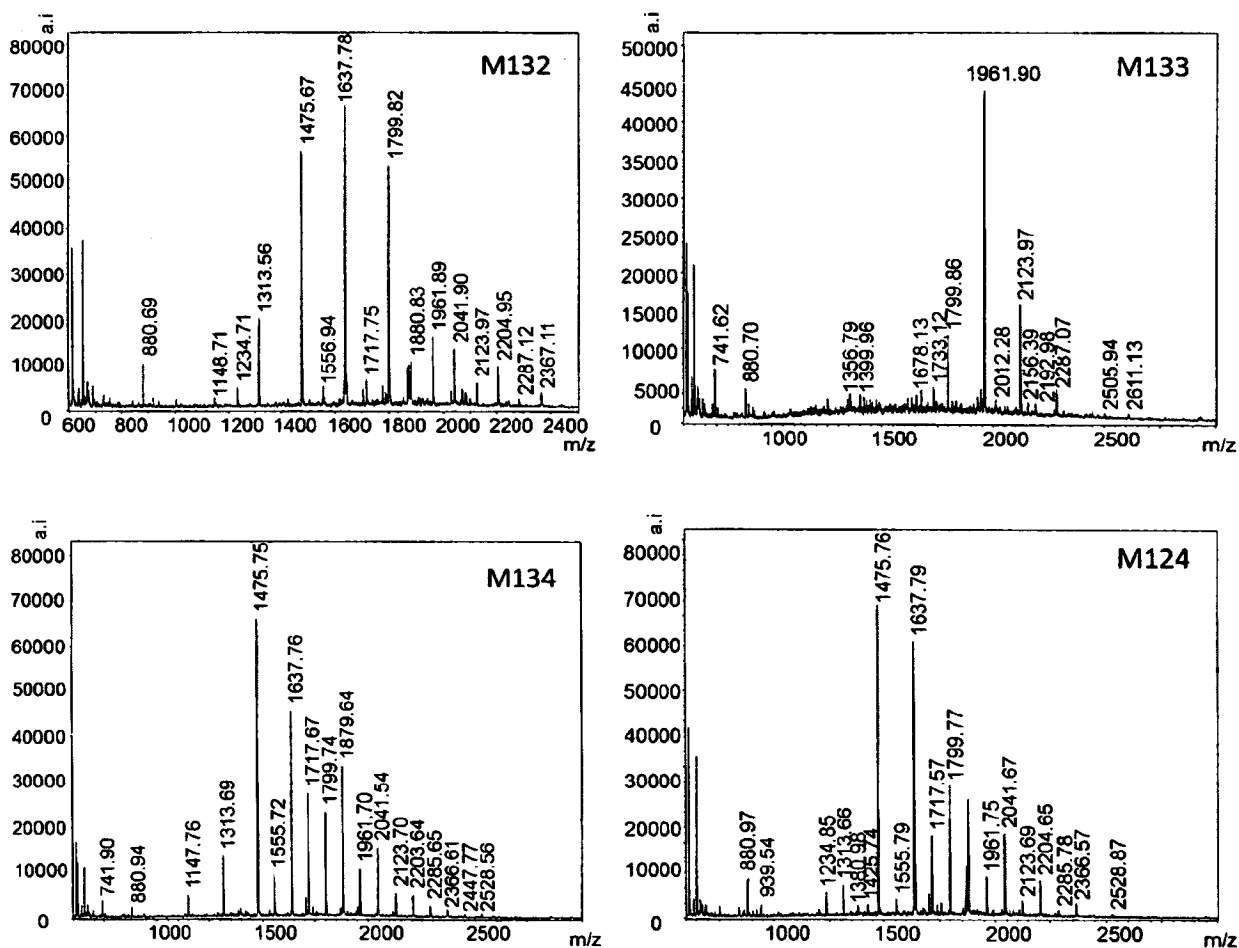
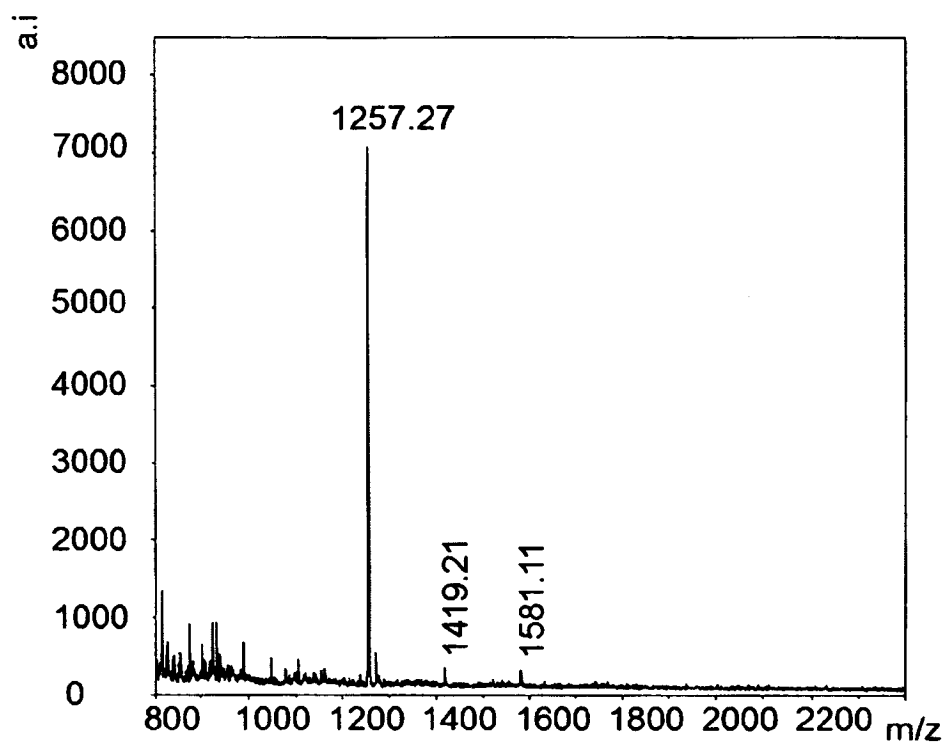


图3(续)

A.



B.

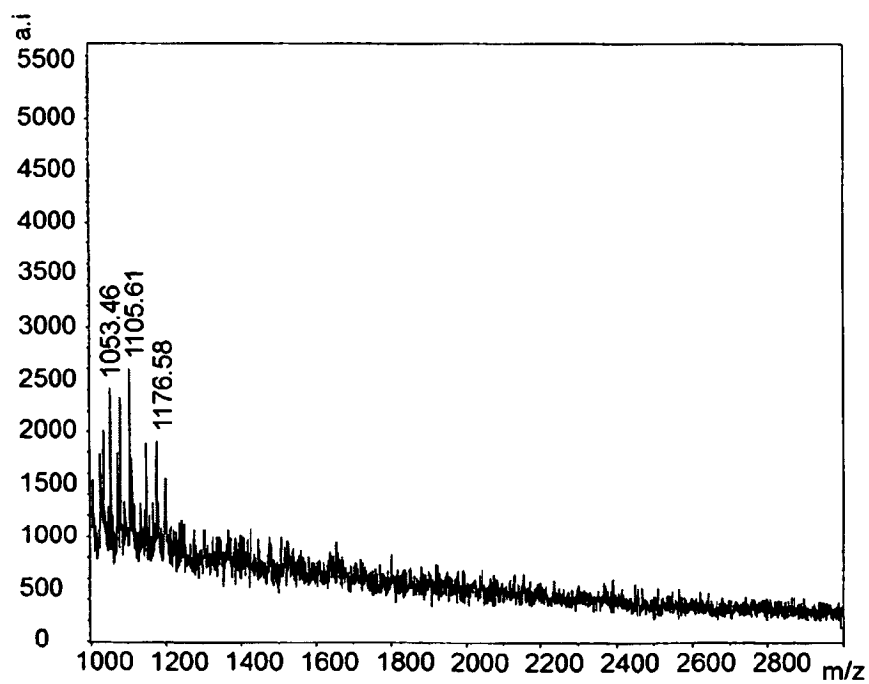
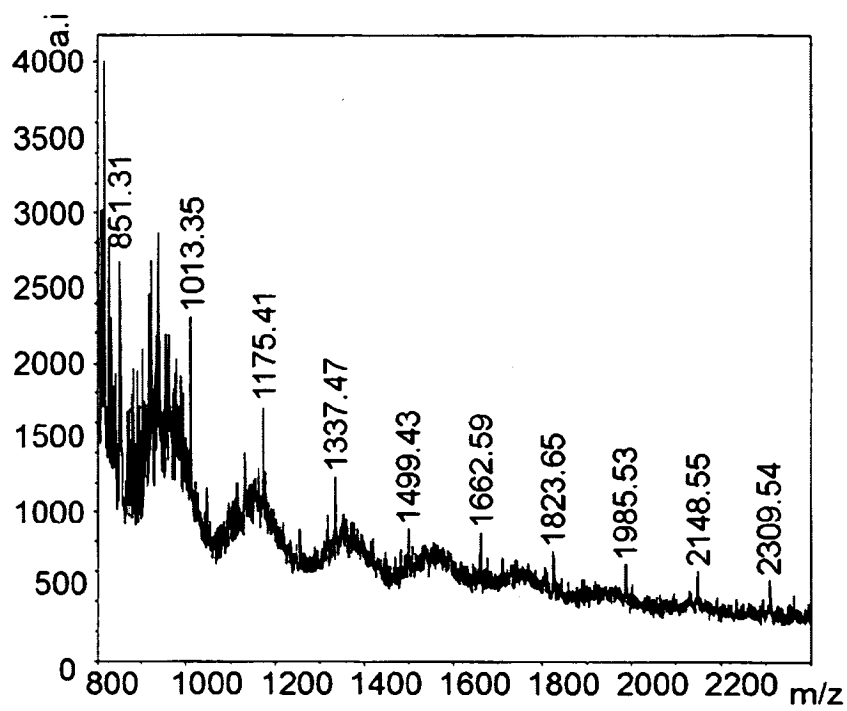


图4

A.



B.

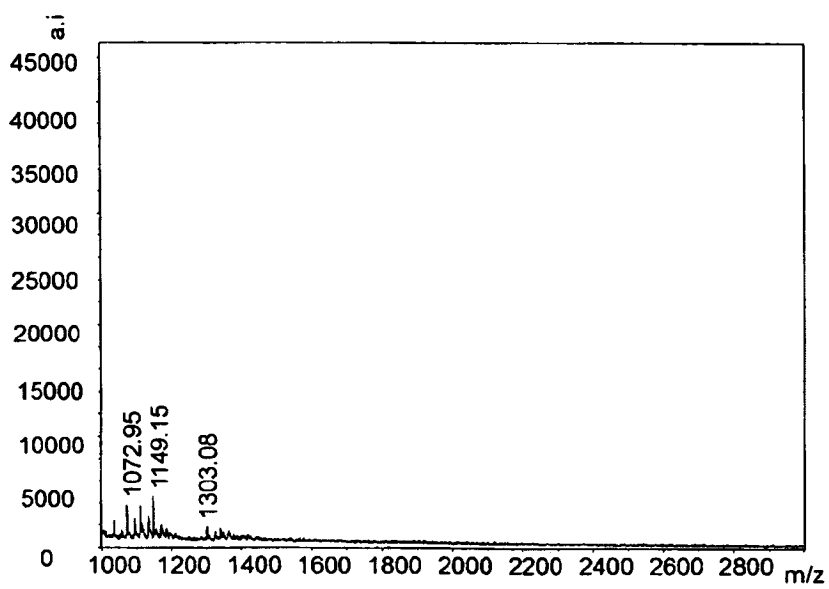


图5

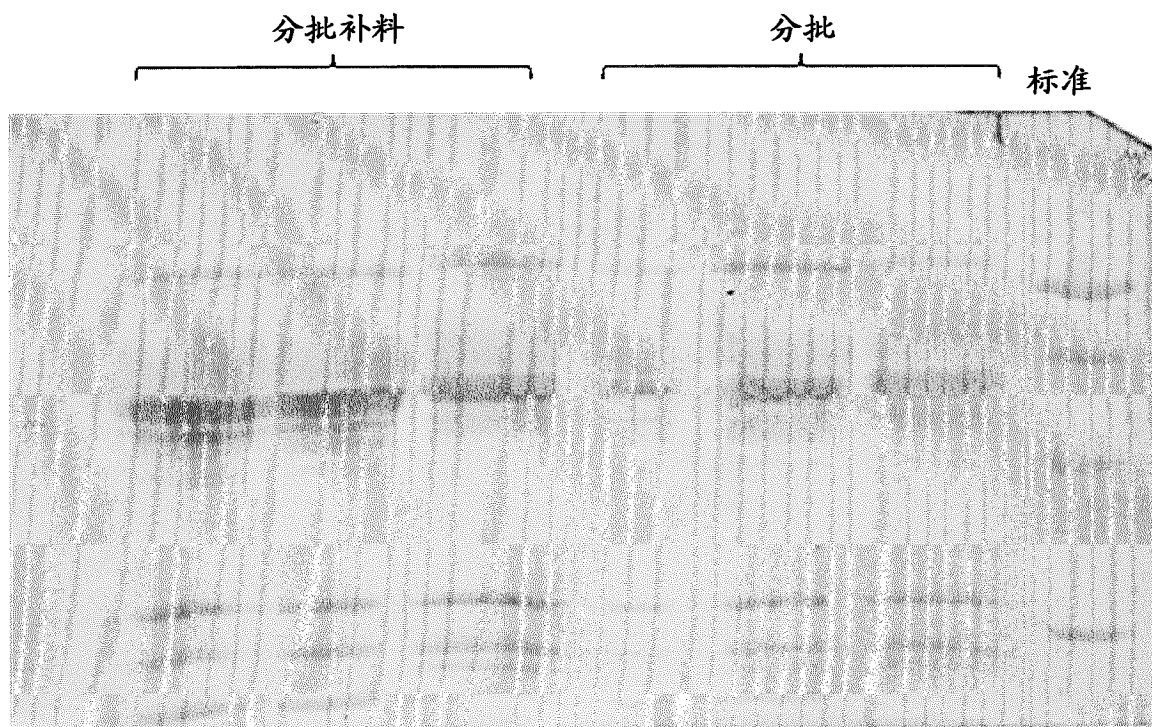


图6

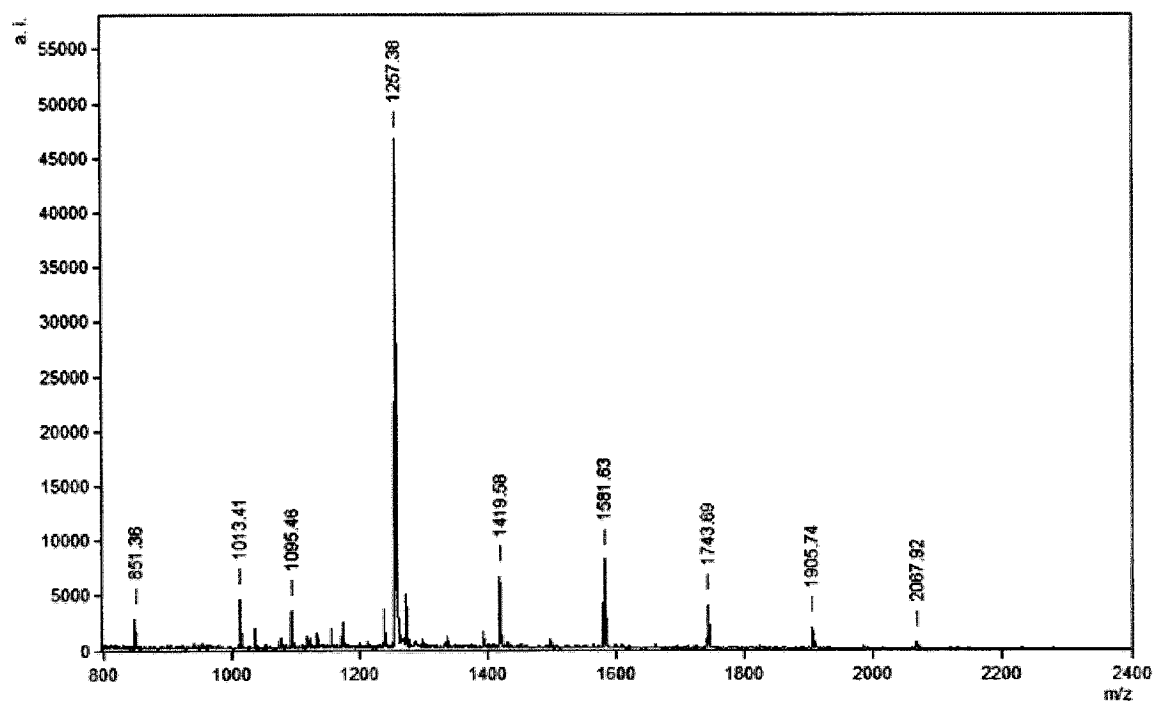


图7

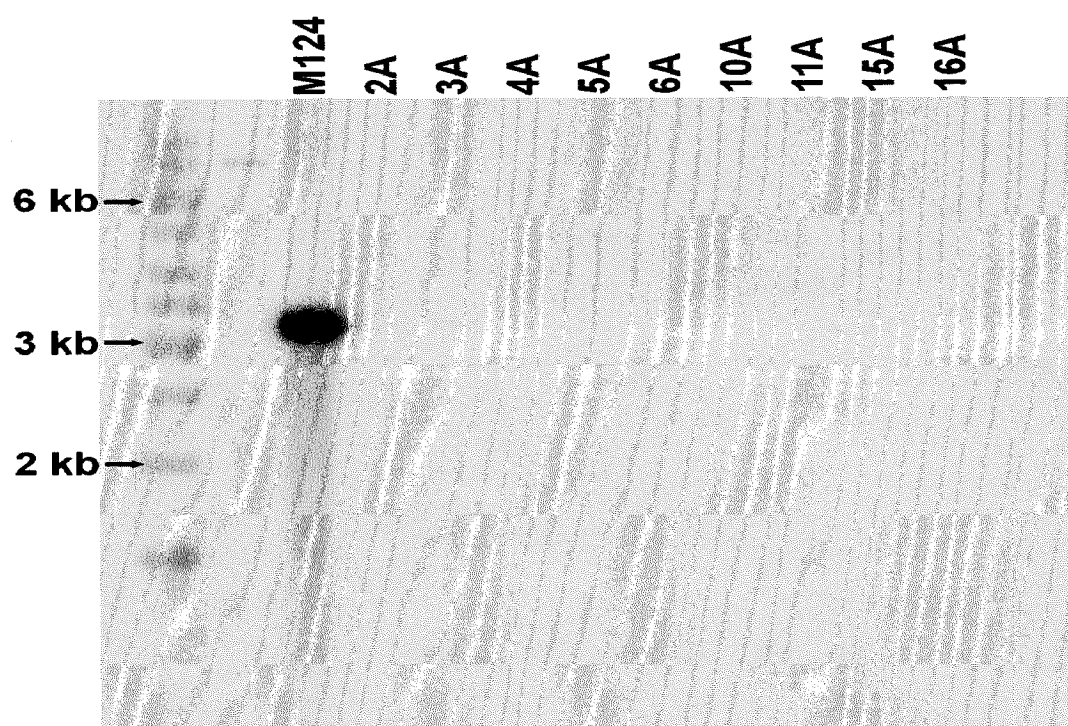
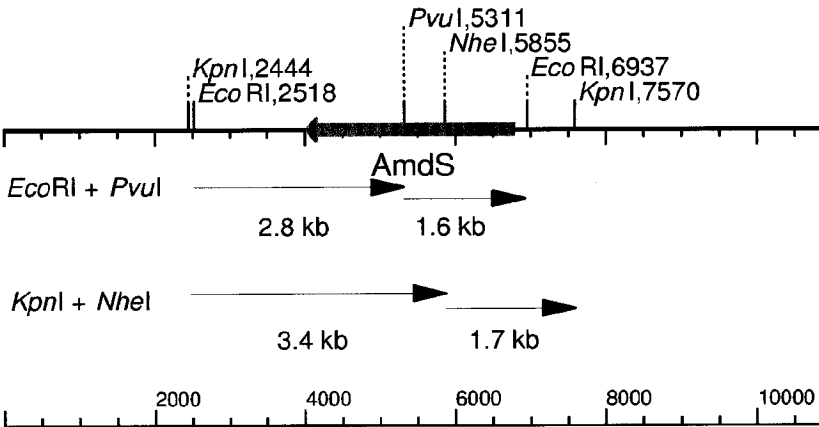


图8

A.



B.

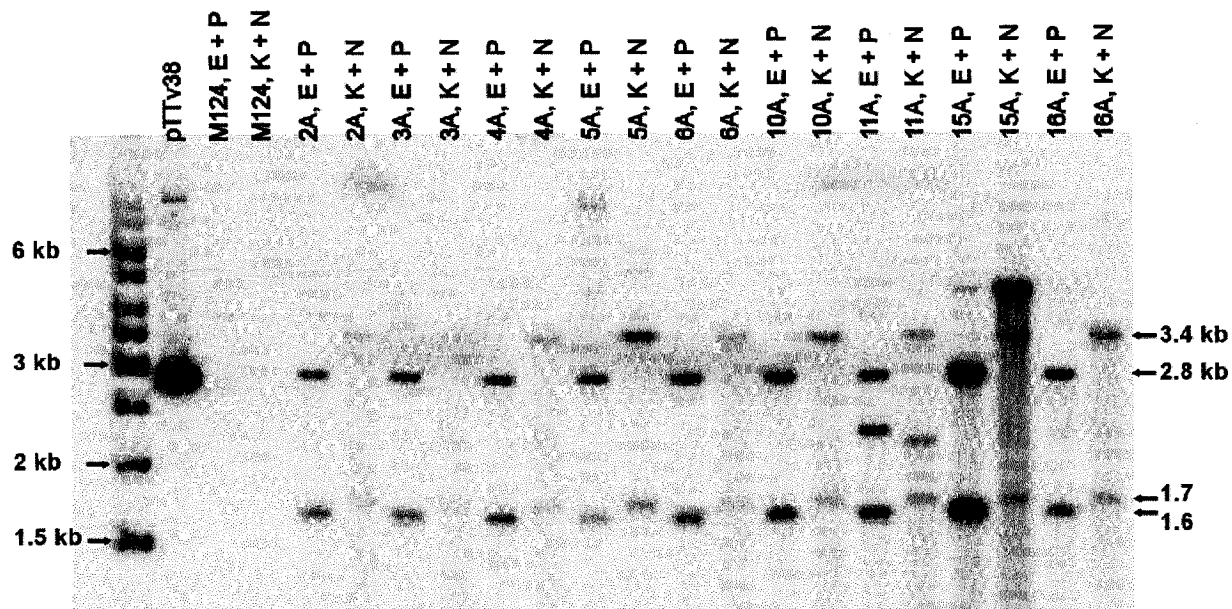


图9

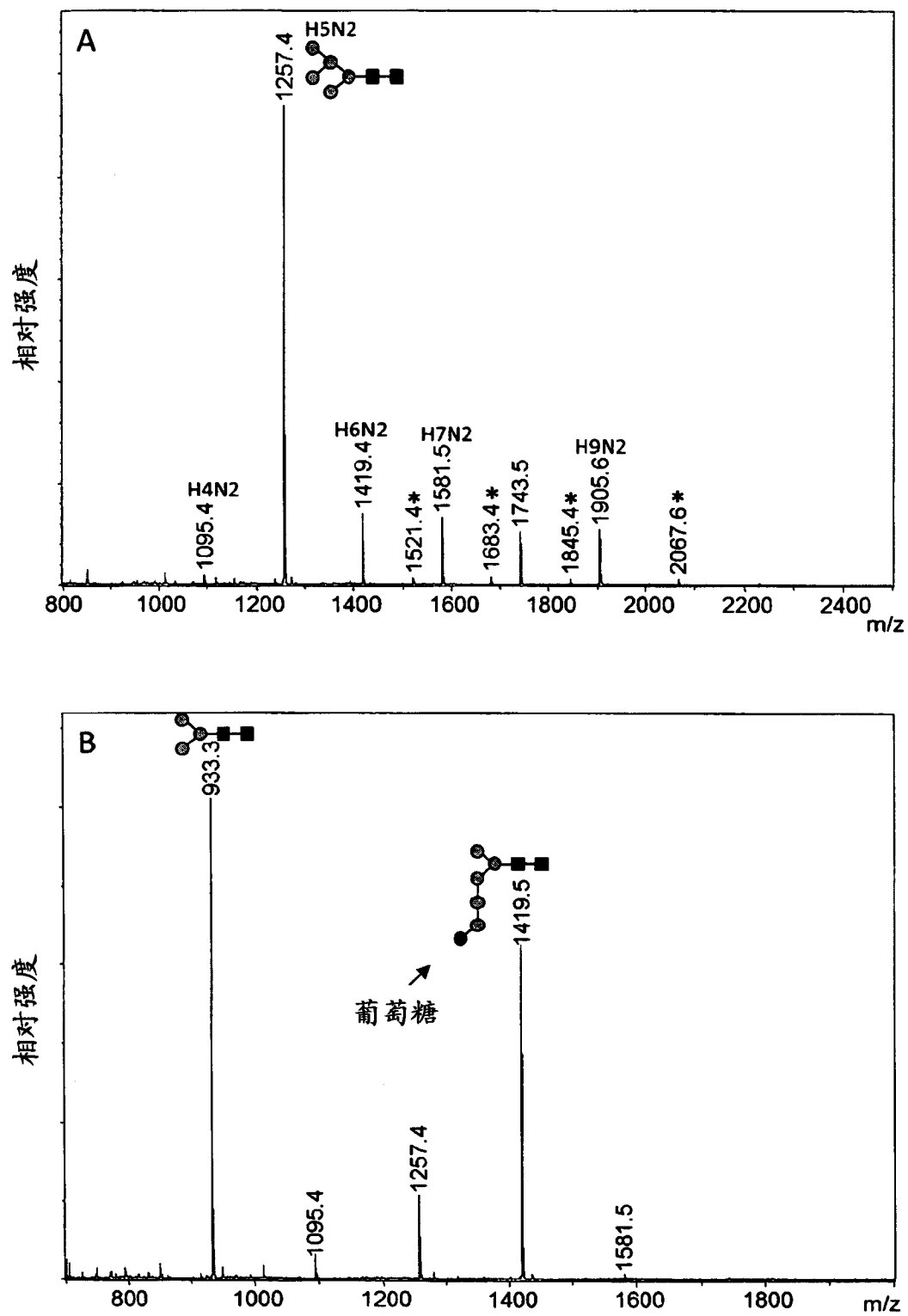


图10

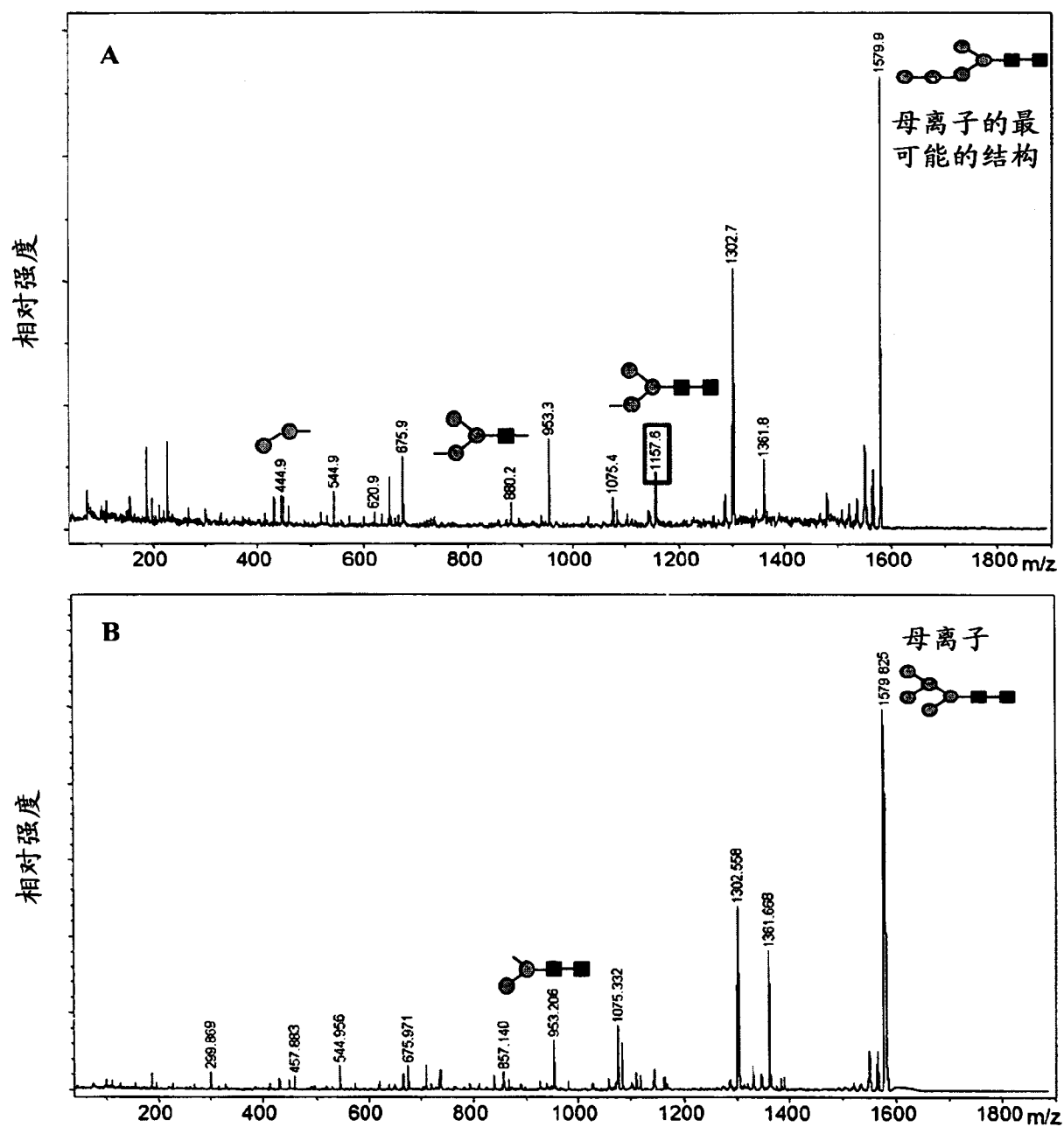


图12

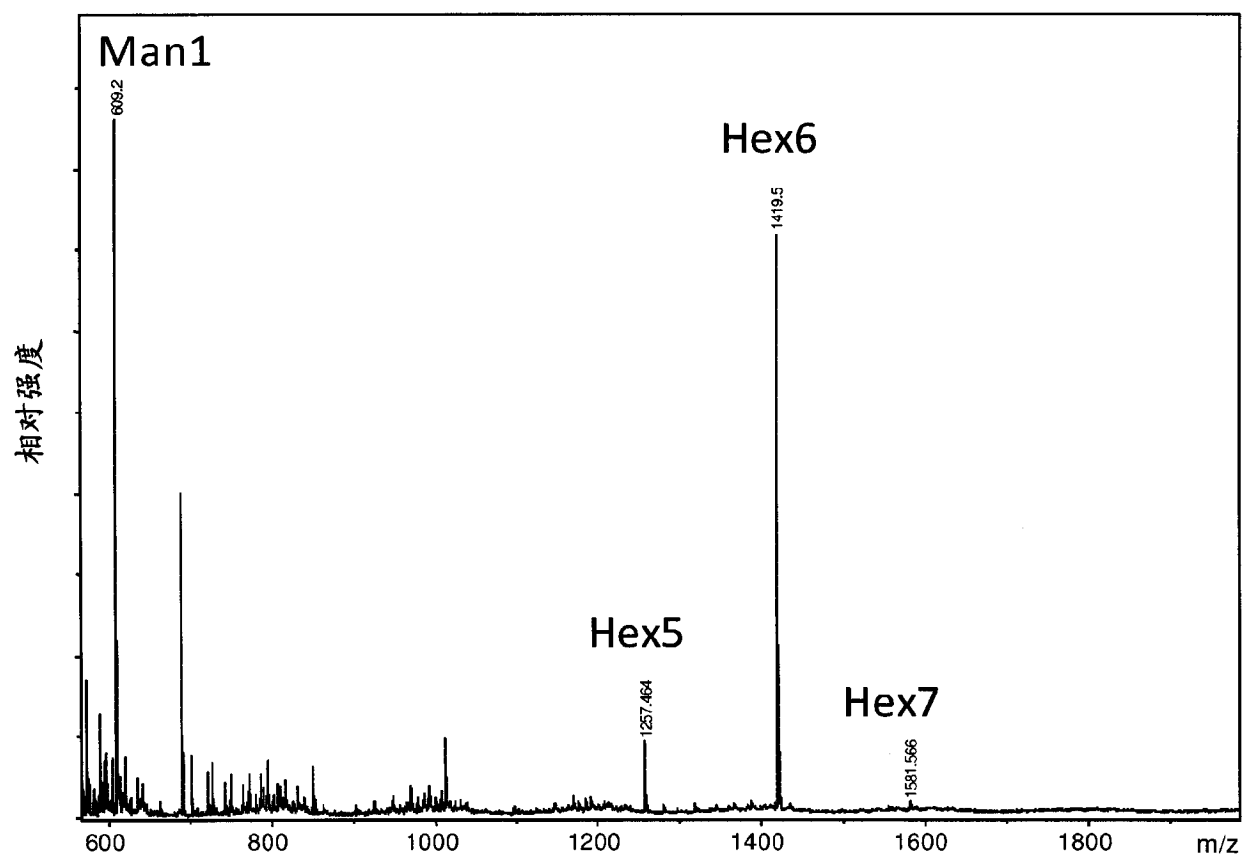


图13

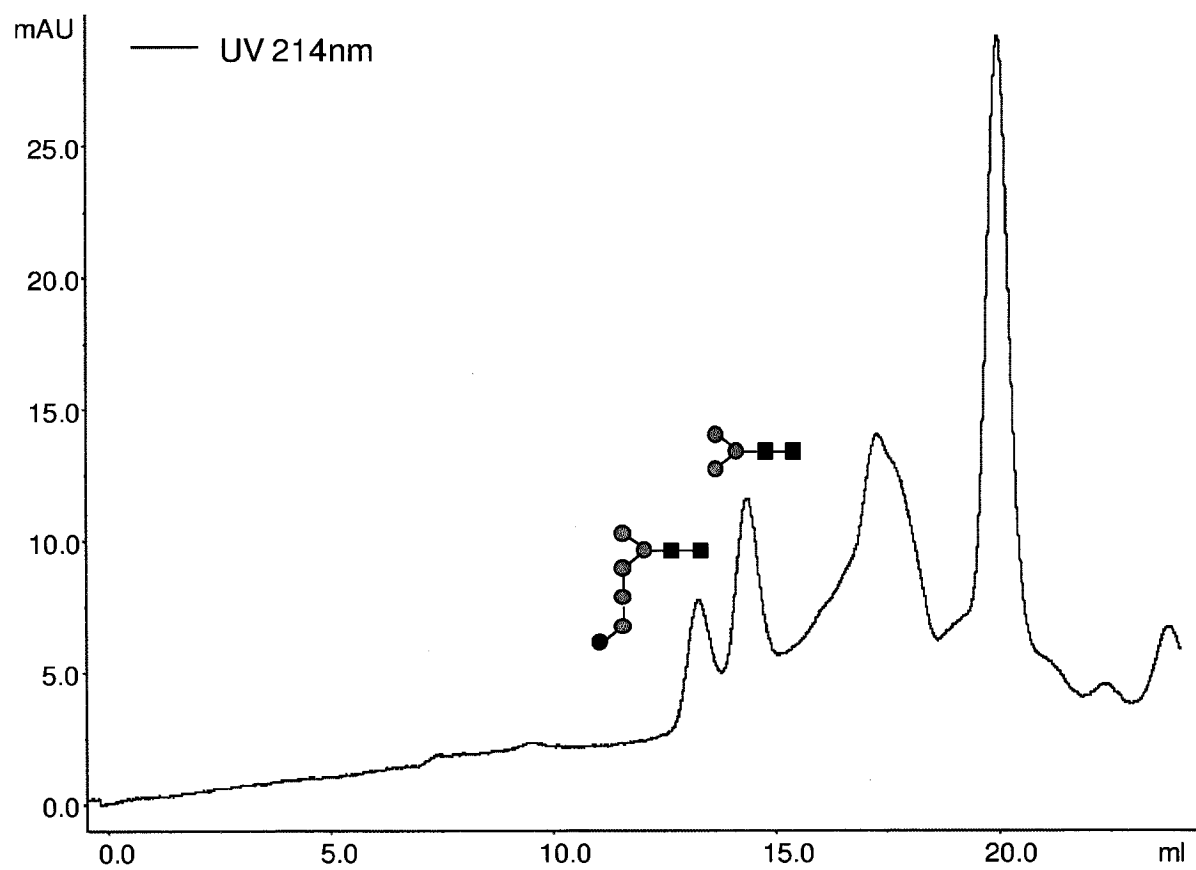


图14

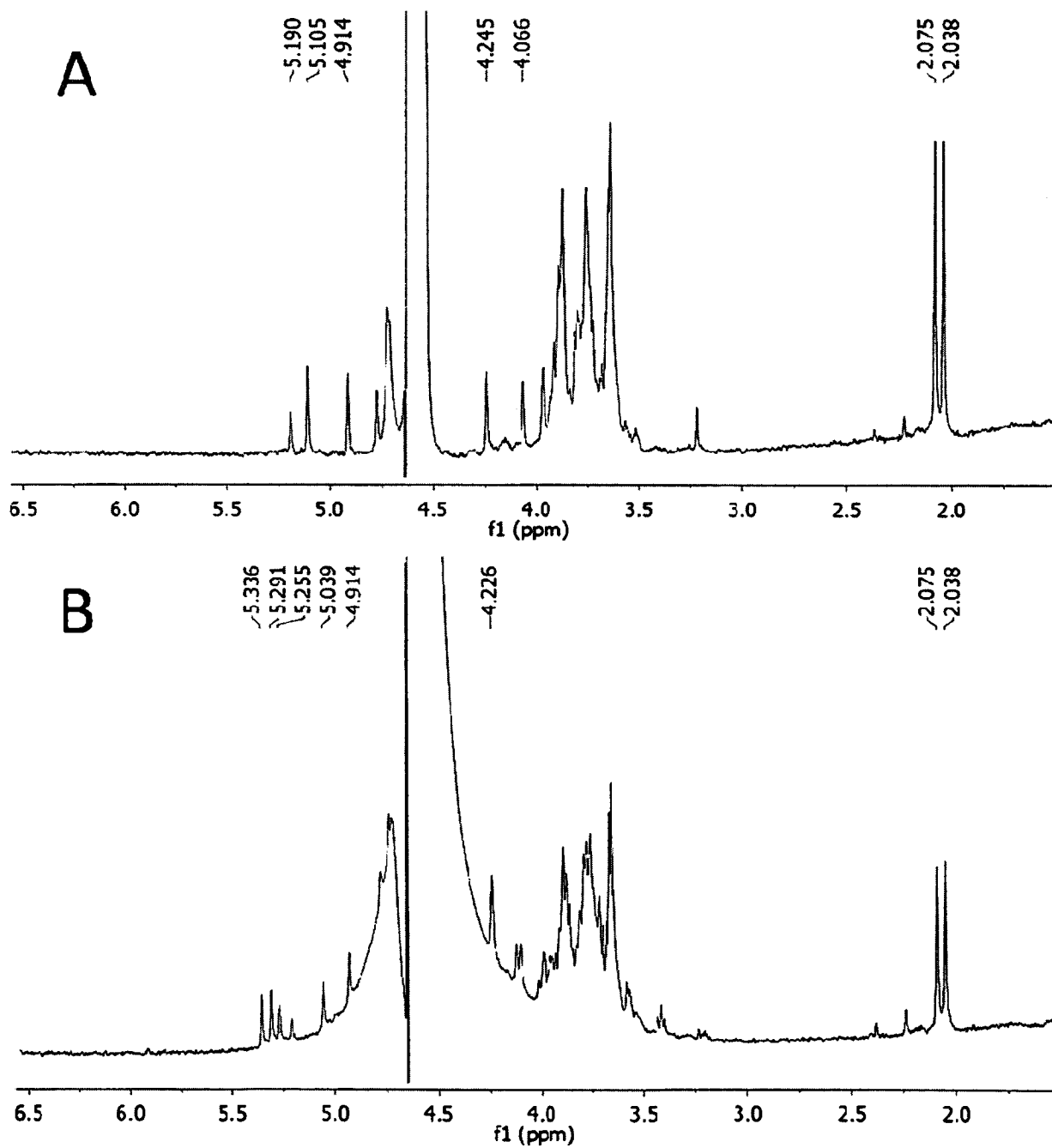


图15

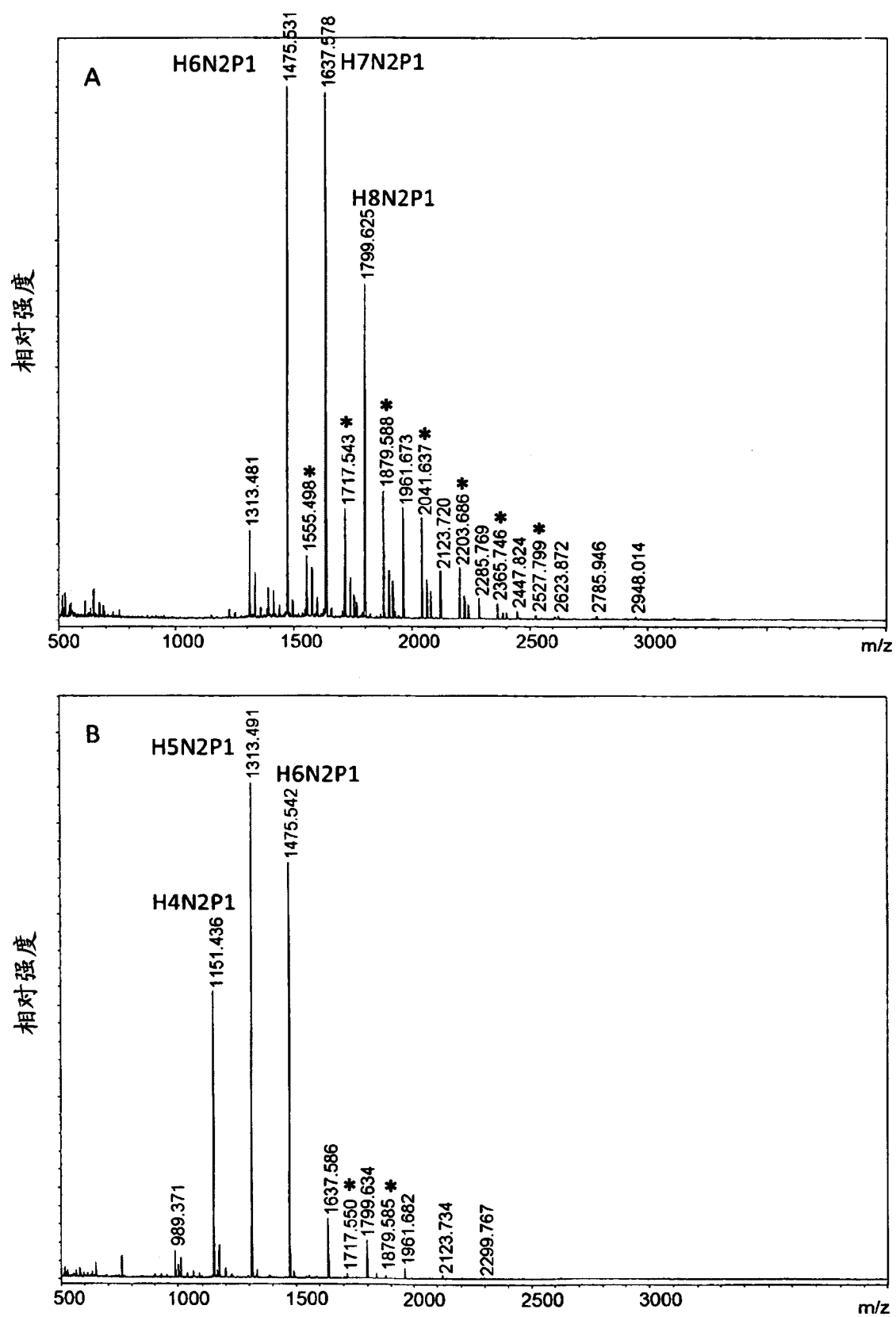


图16

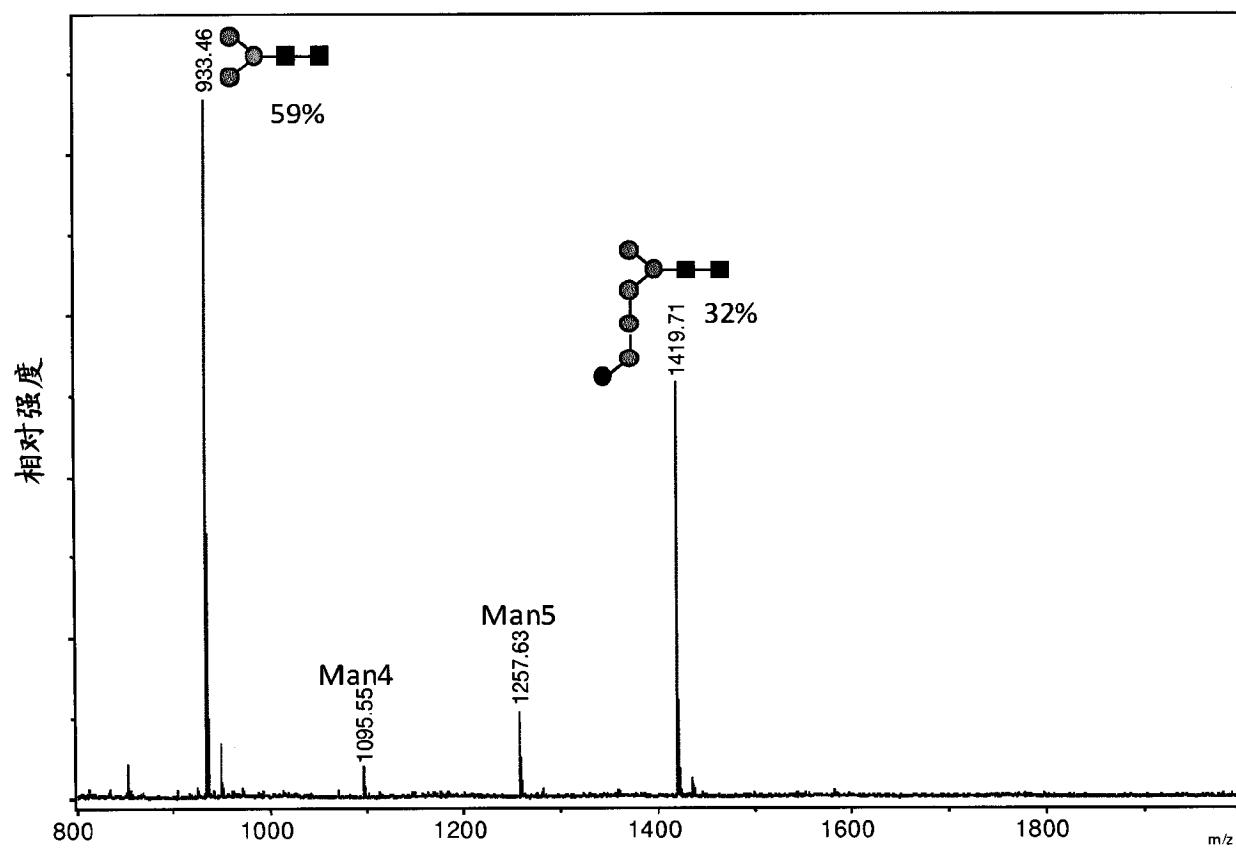


图17

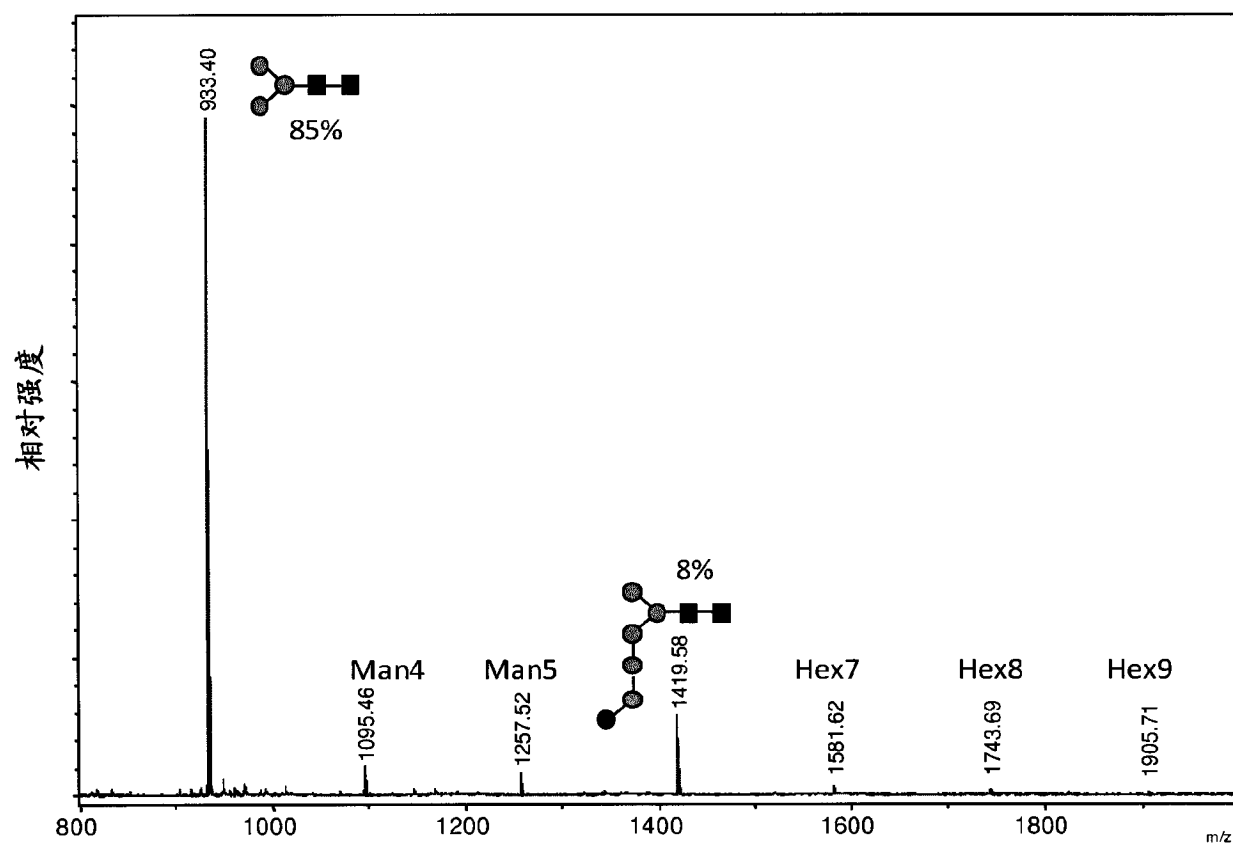


图18

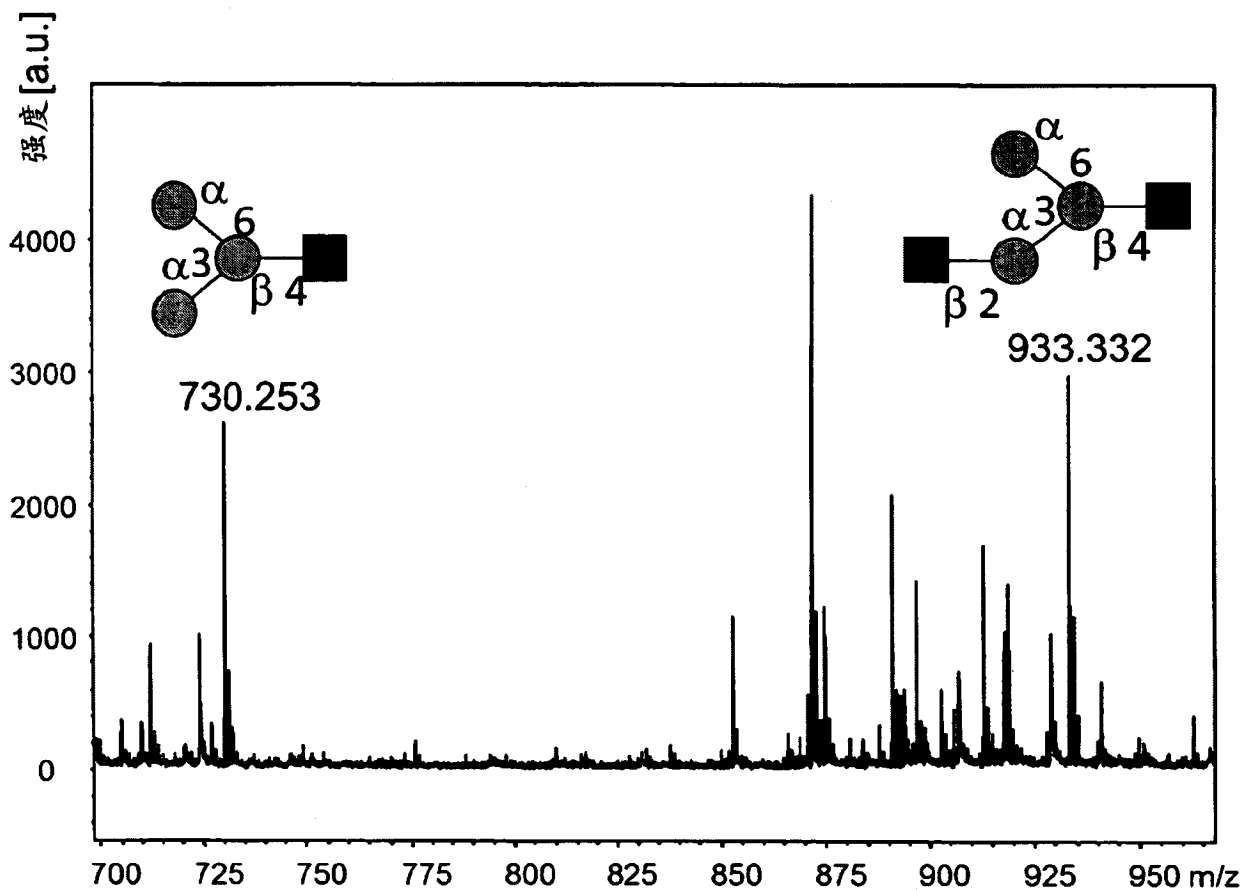


图19

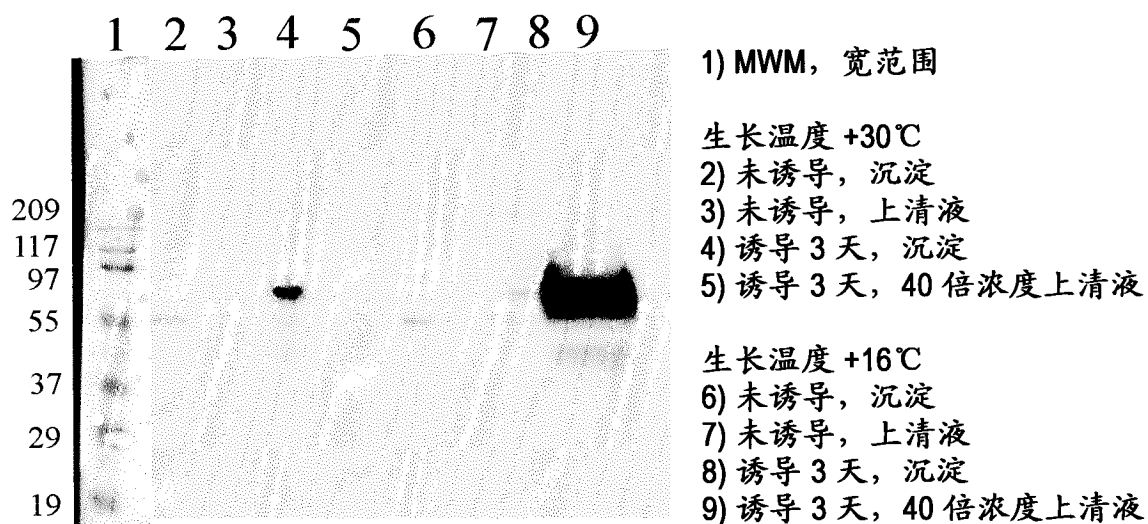


图20

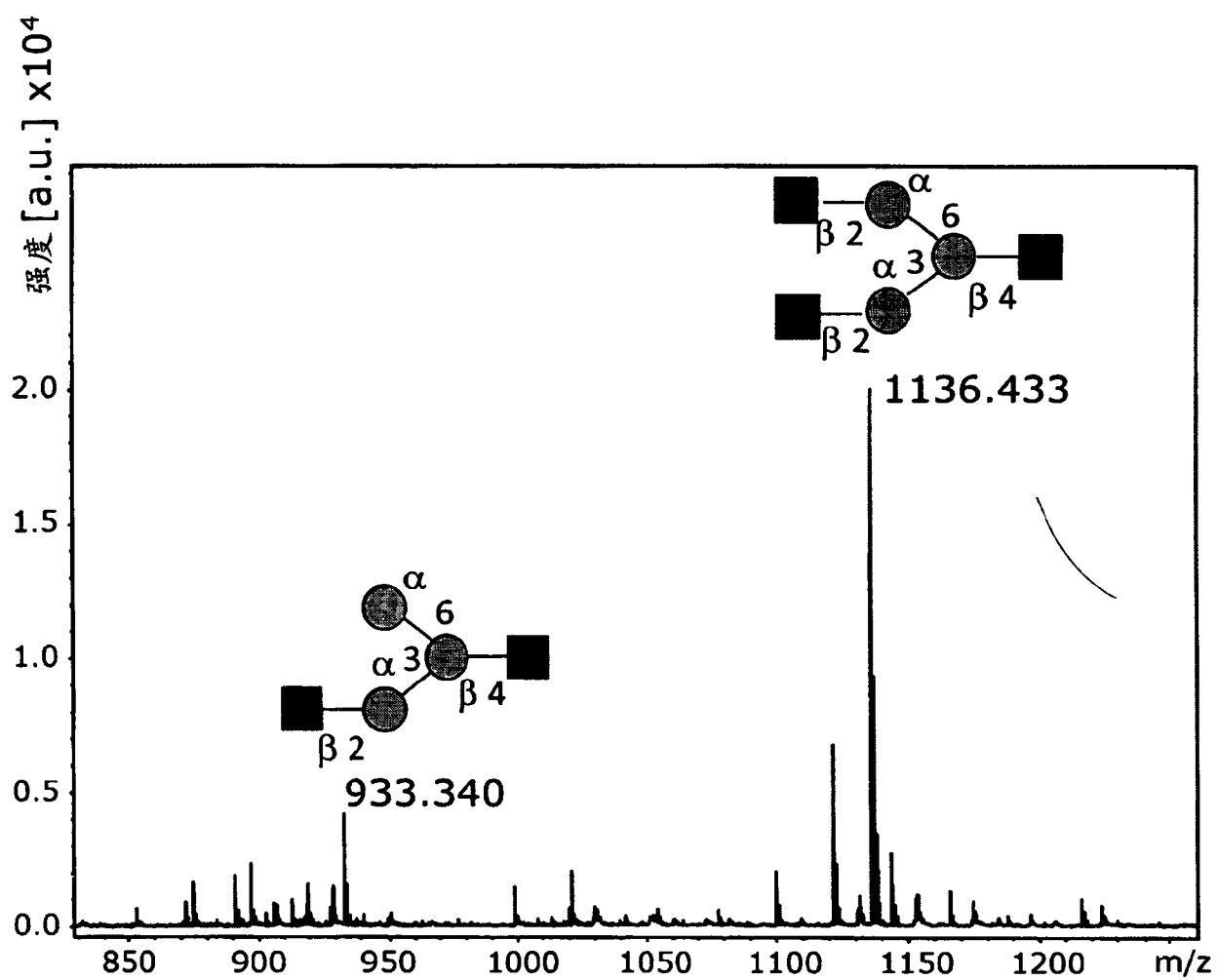


图21

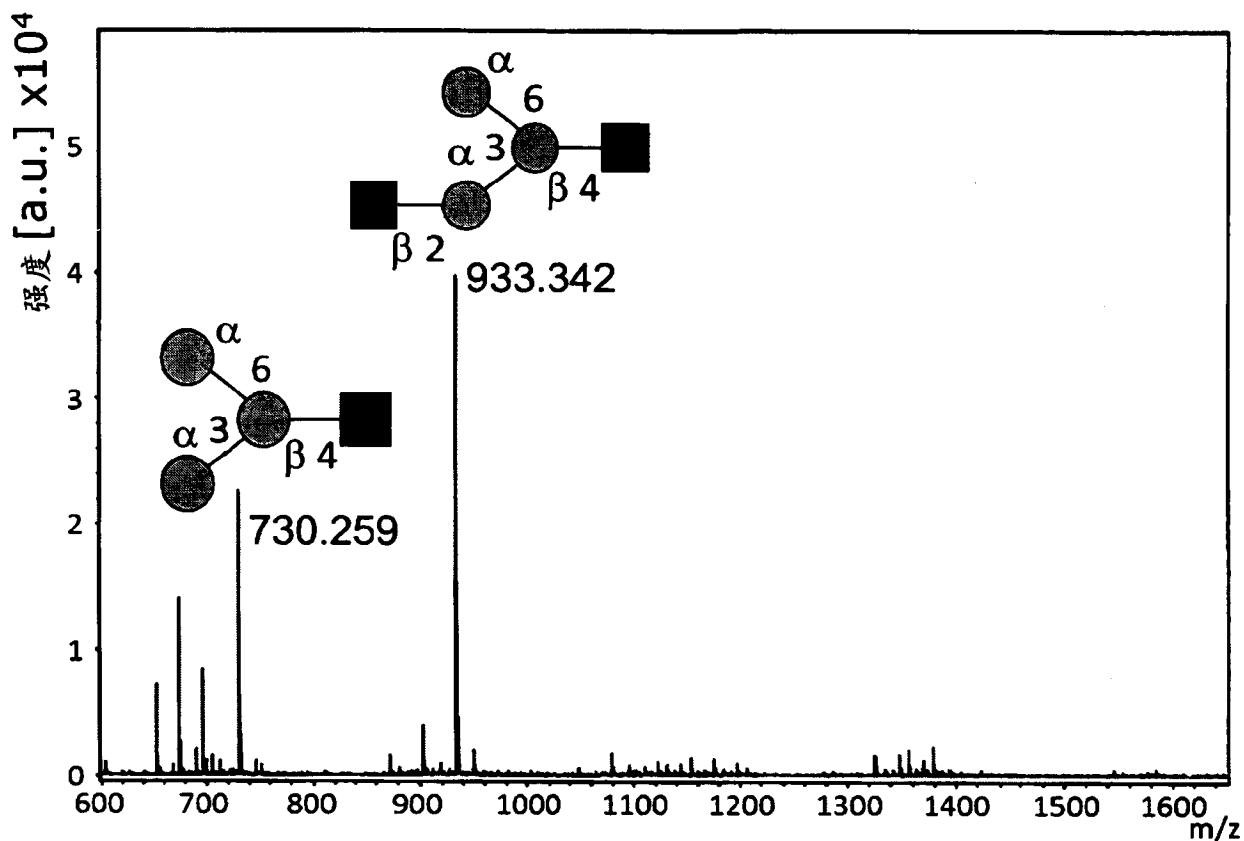


图22

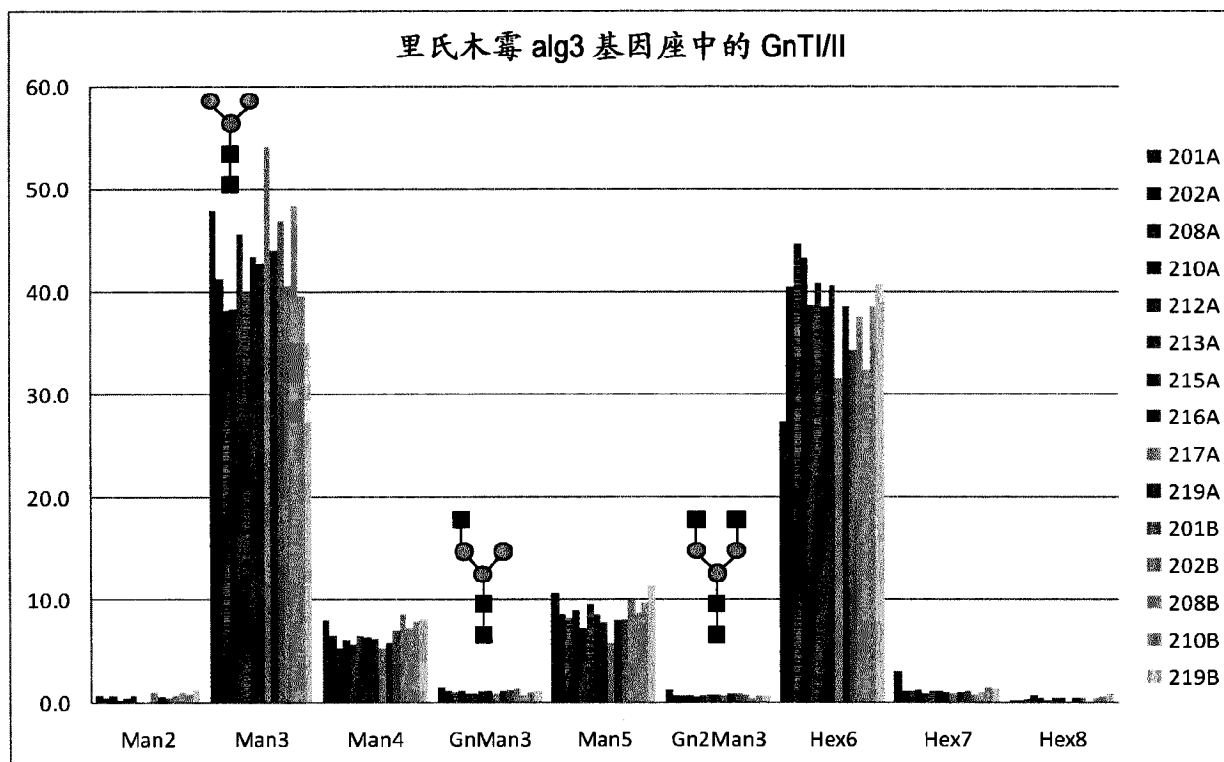


图23

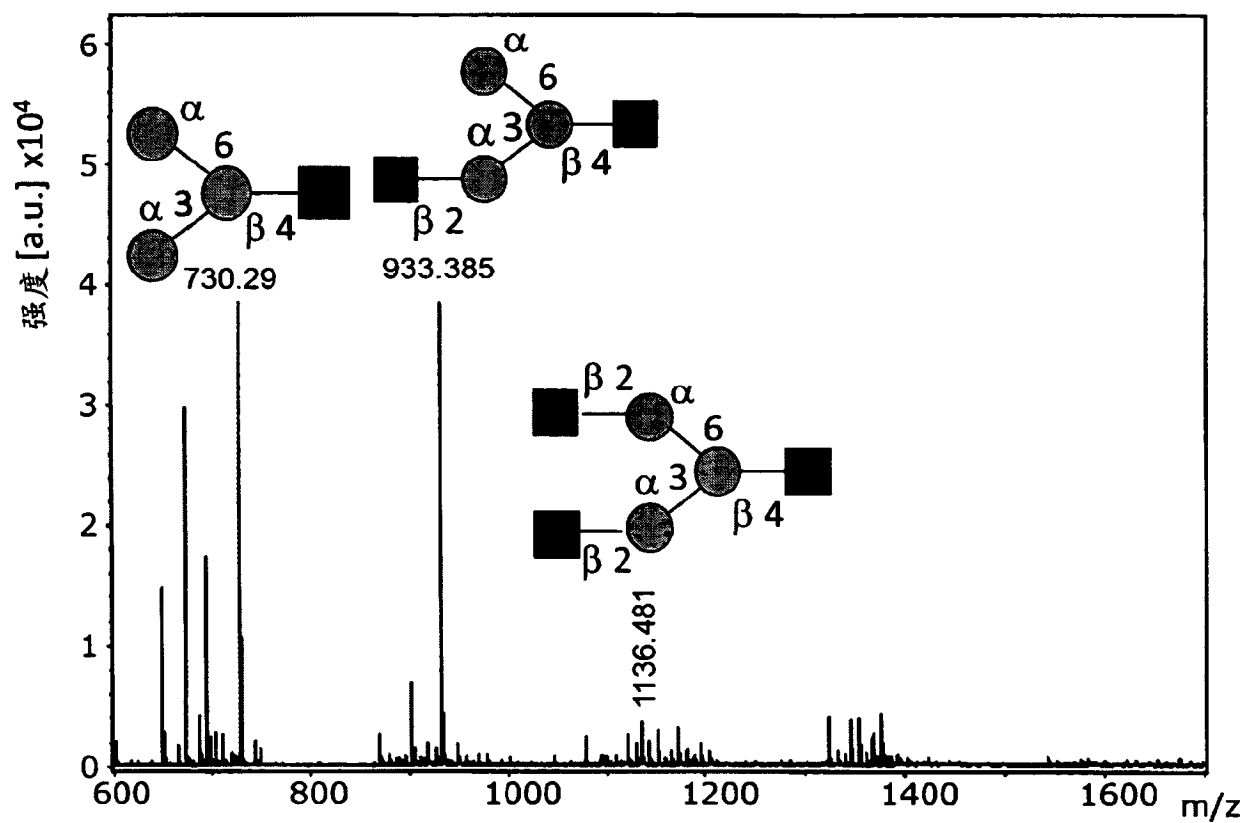


图24

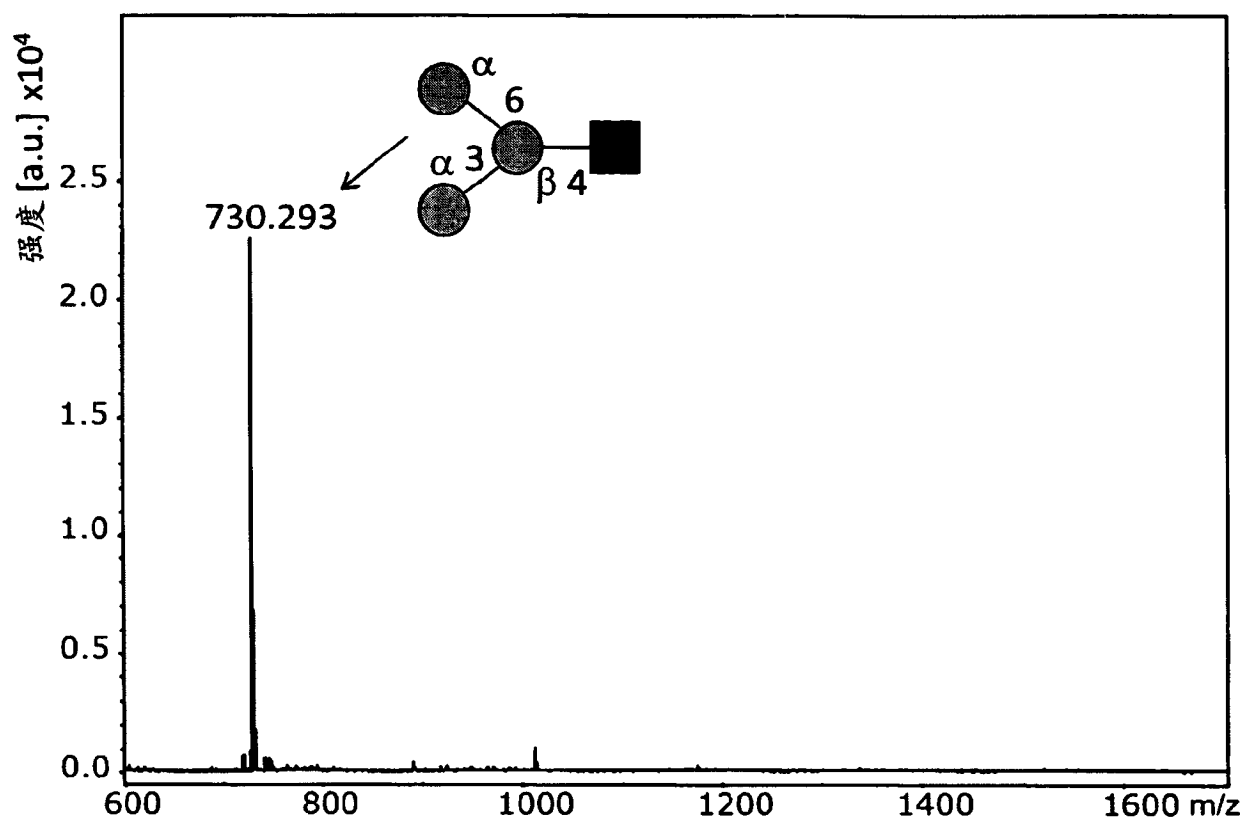


图25

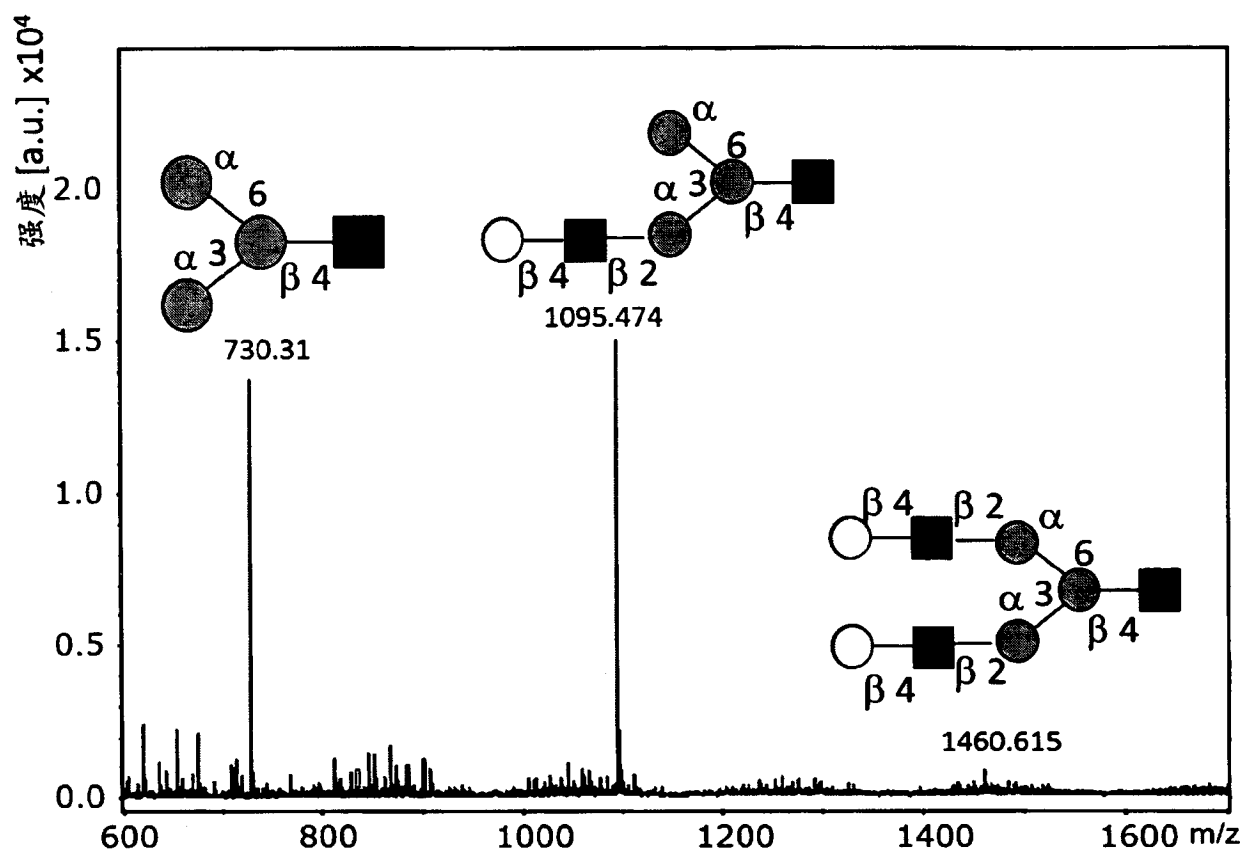
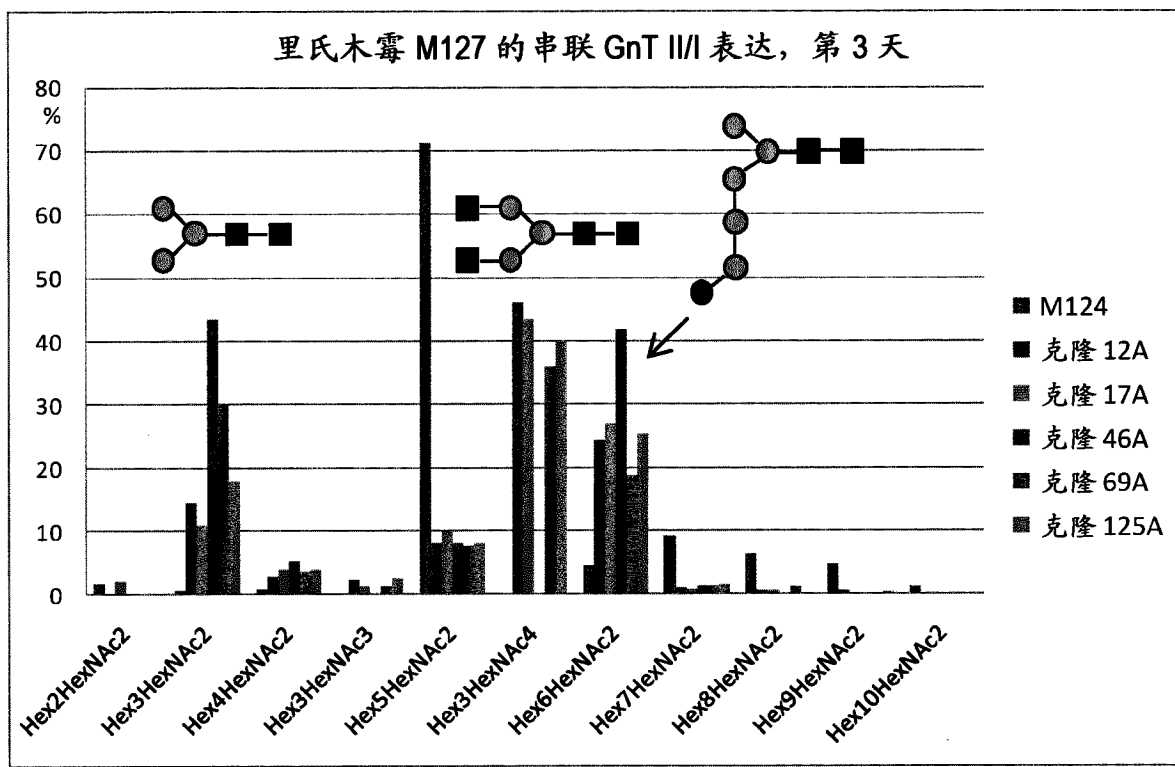


图26

A.



B.

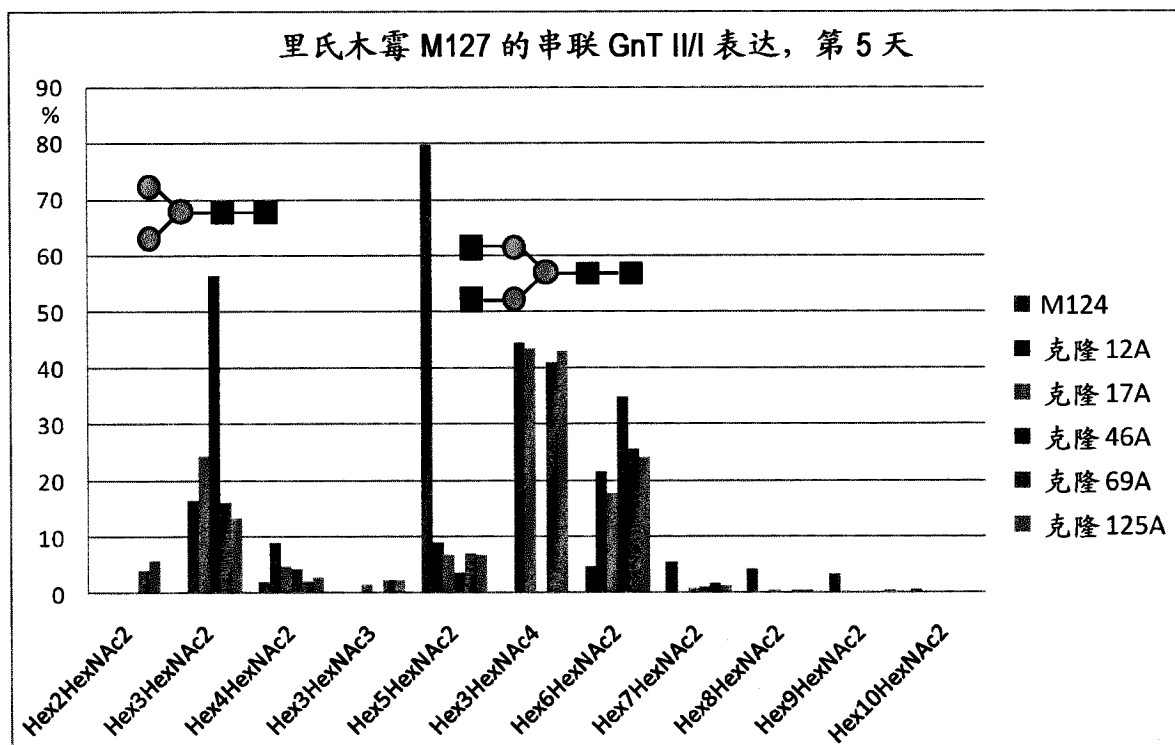
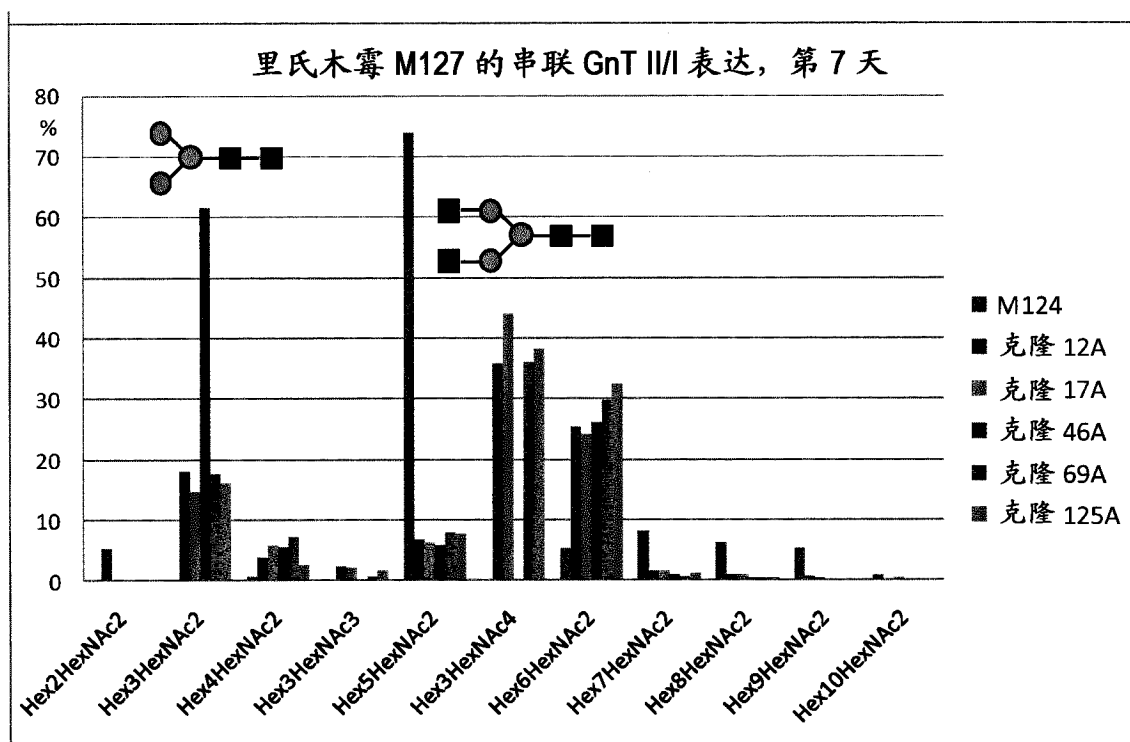


图27

C.



D.

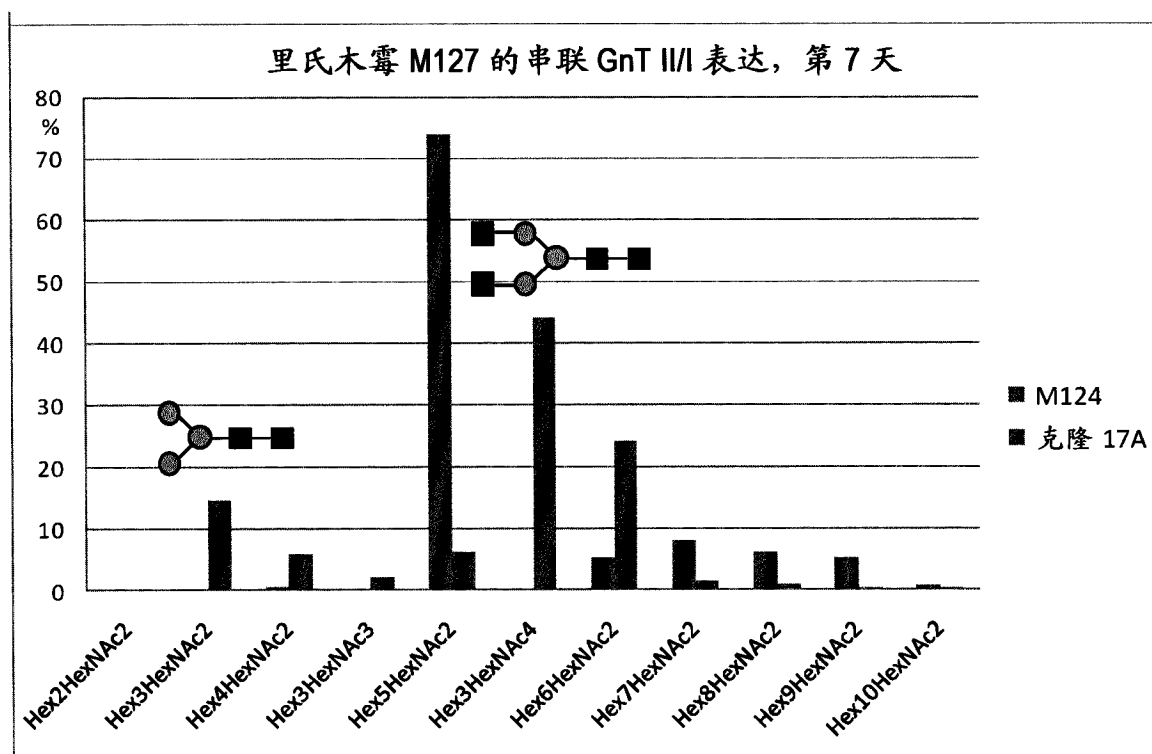


图27(续)

E.

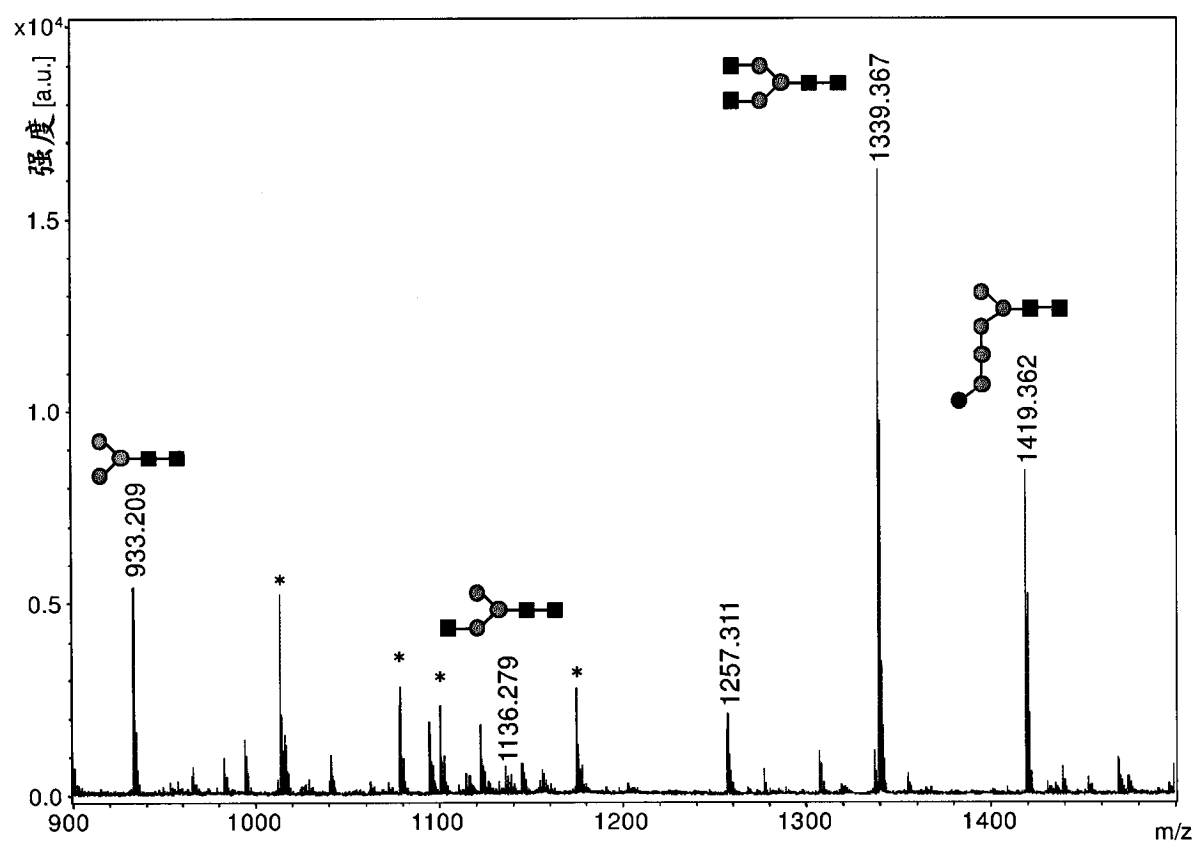
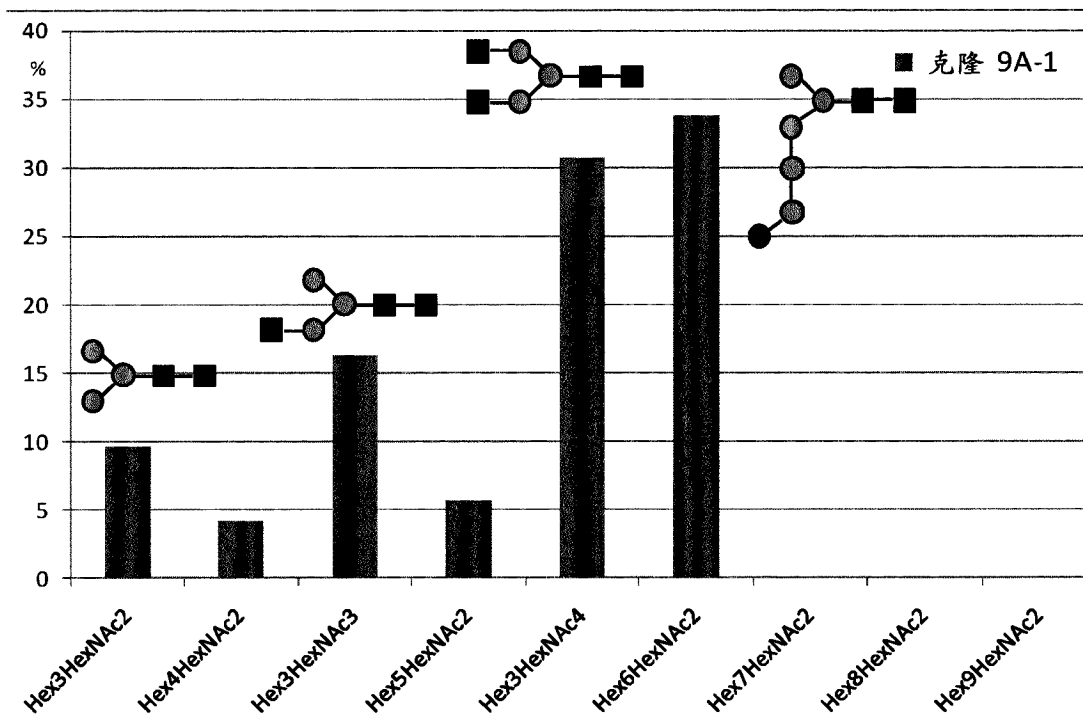


图27(续)

A.



B.

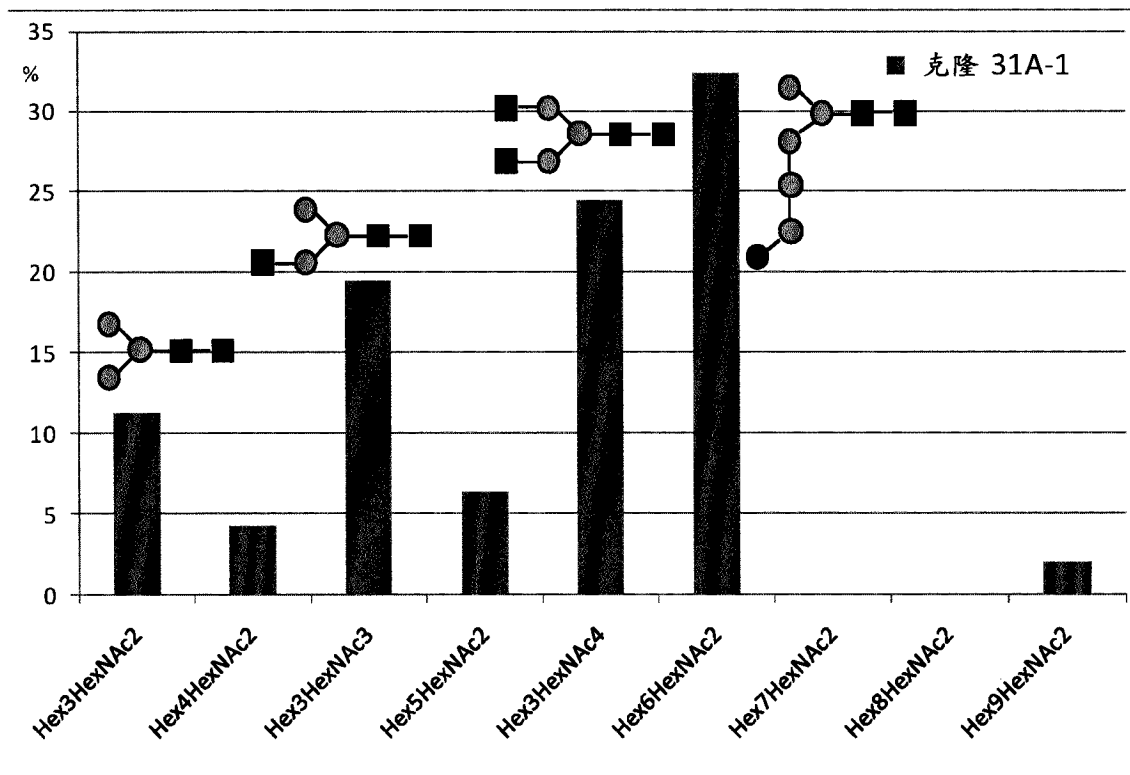


图28

C.

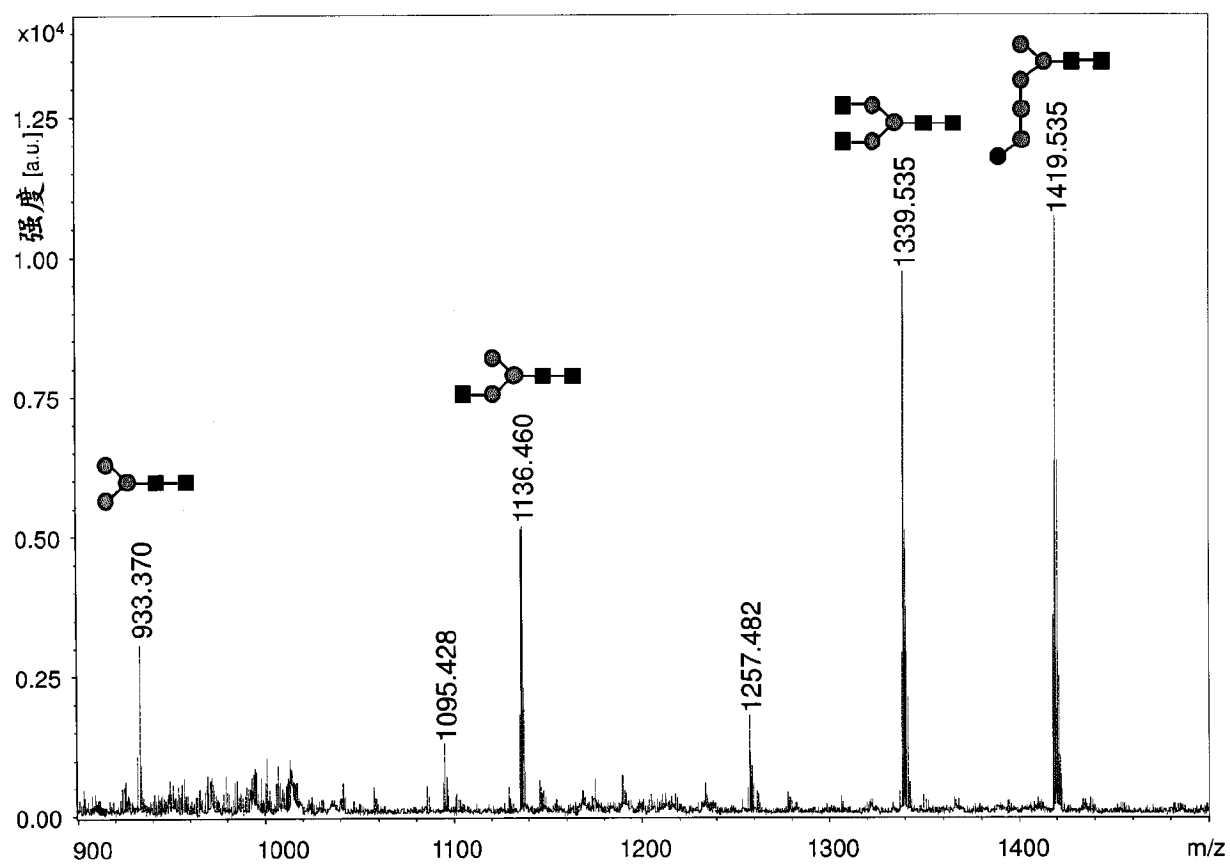


图28(续)

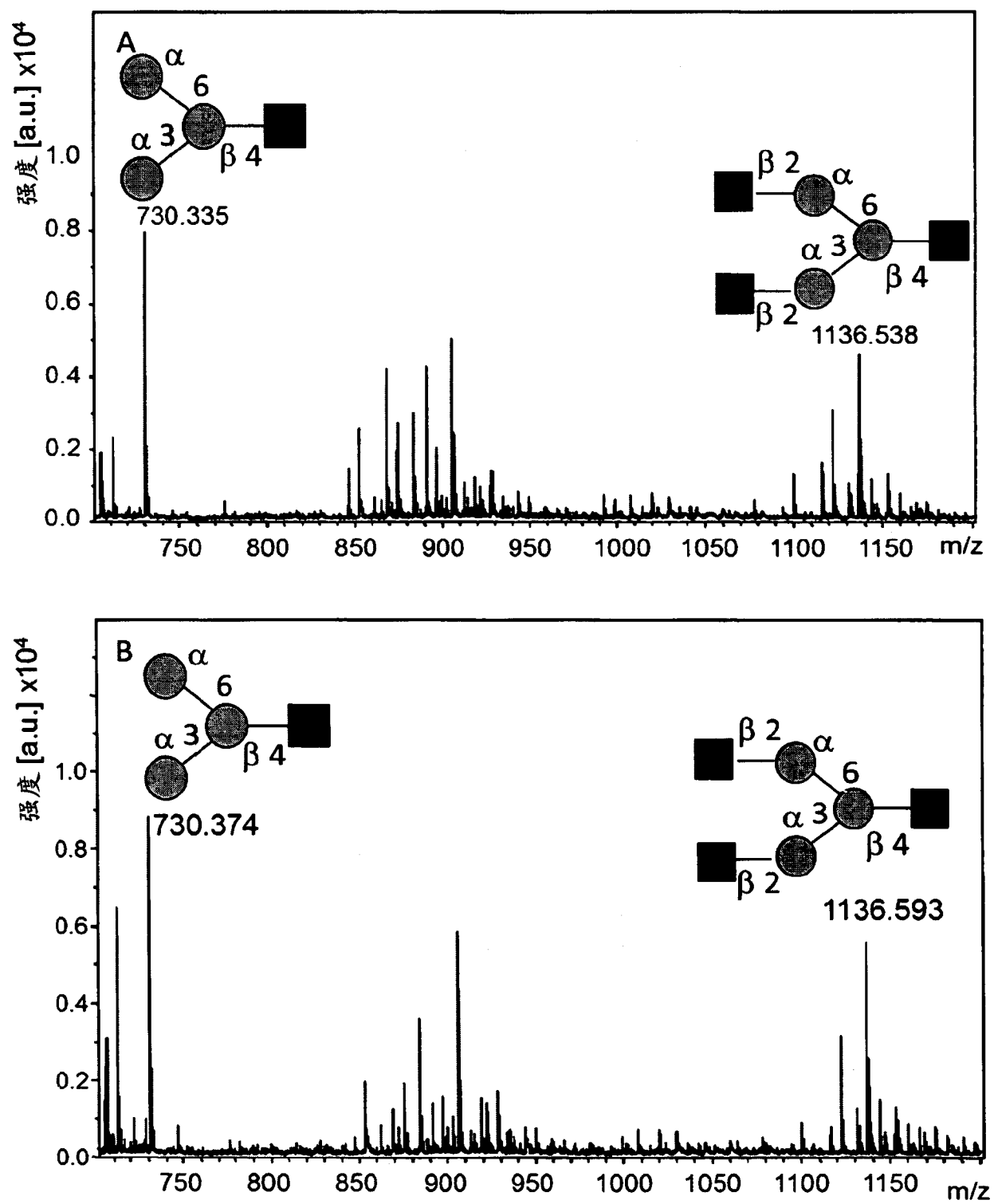


图29

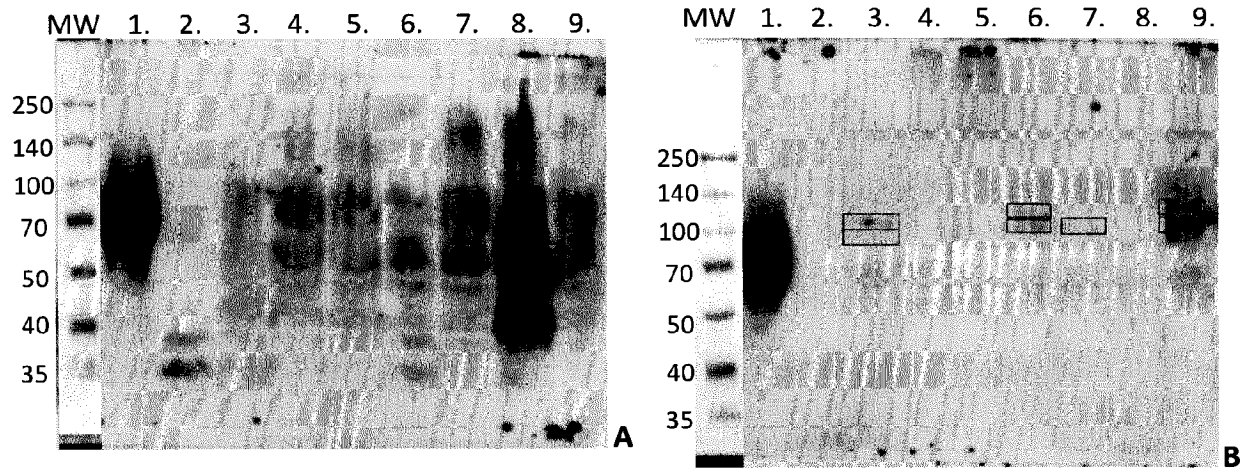
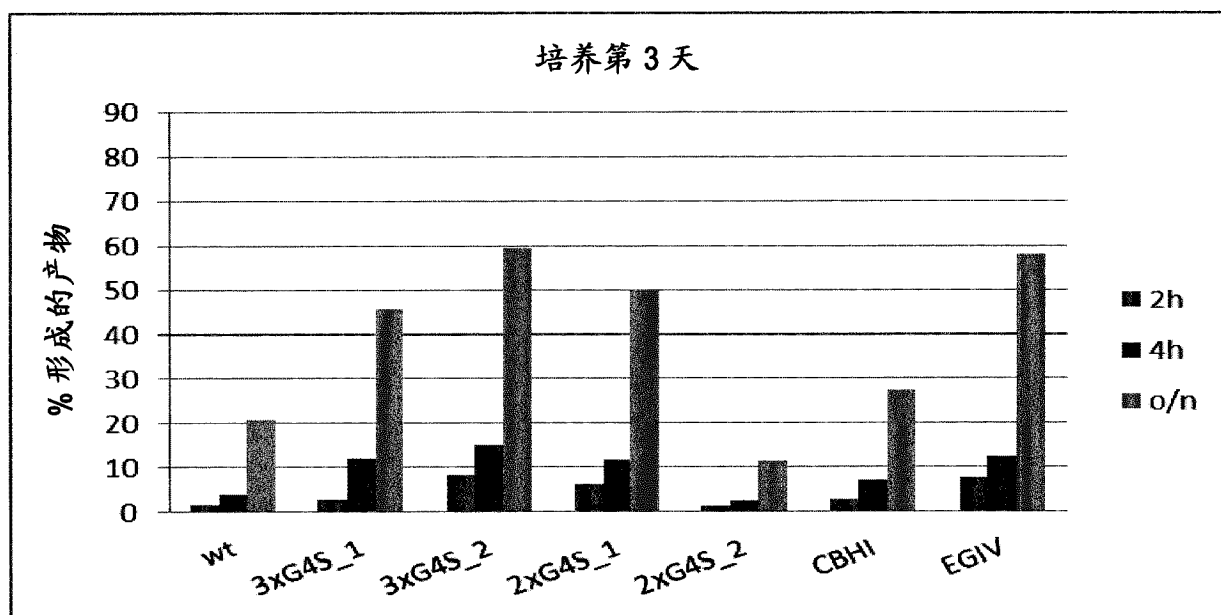


图30

A.



B.

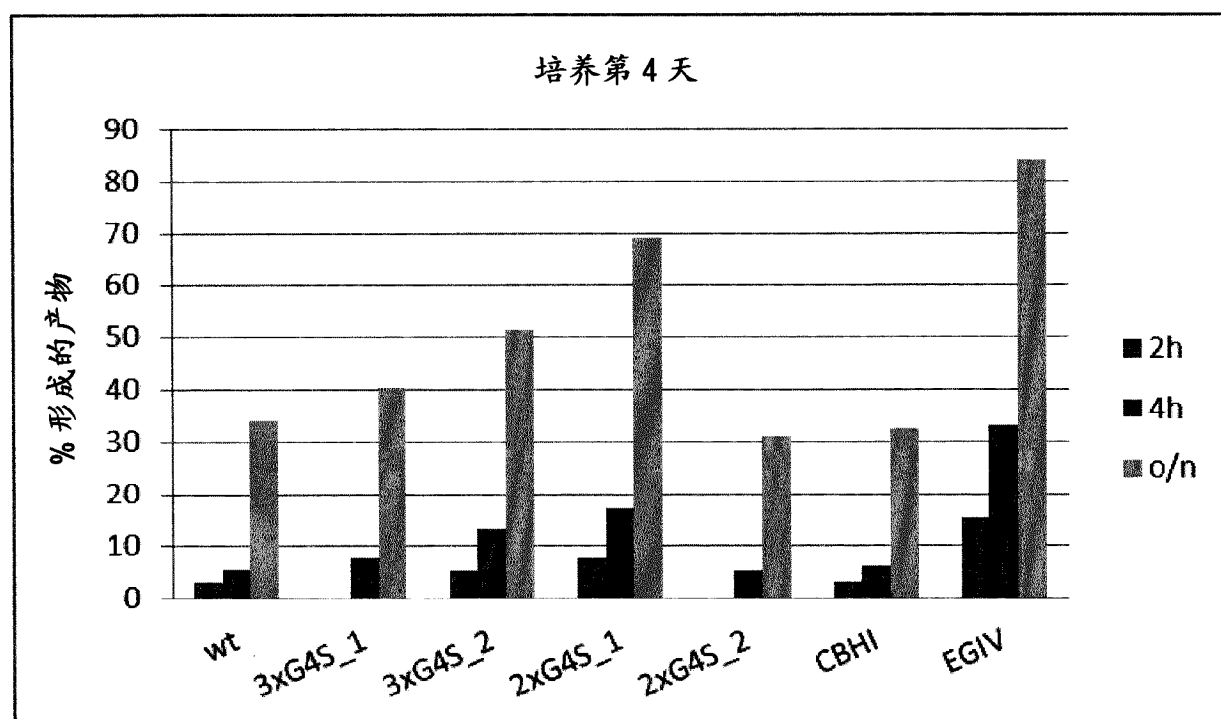


图31

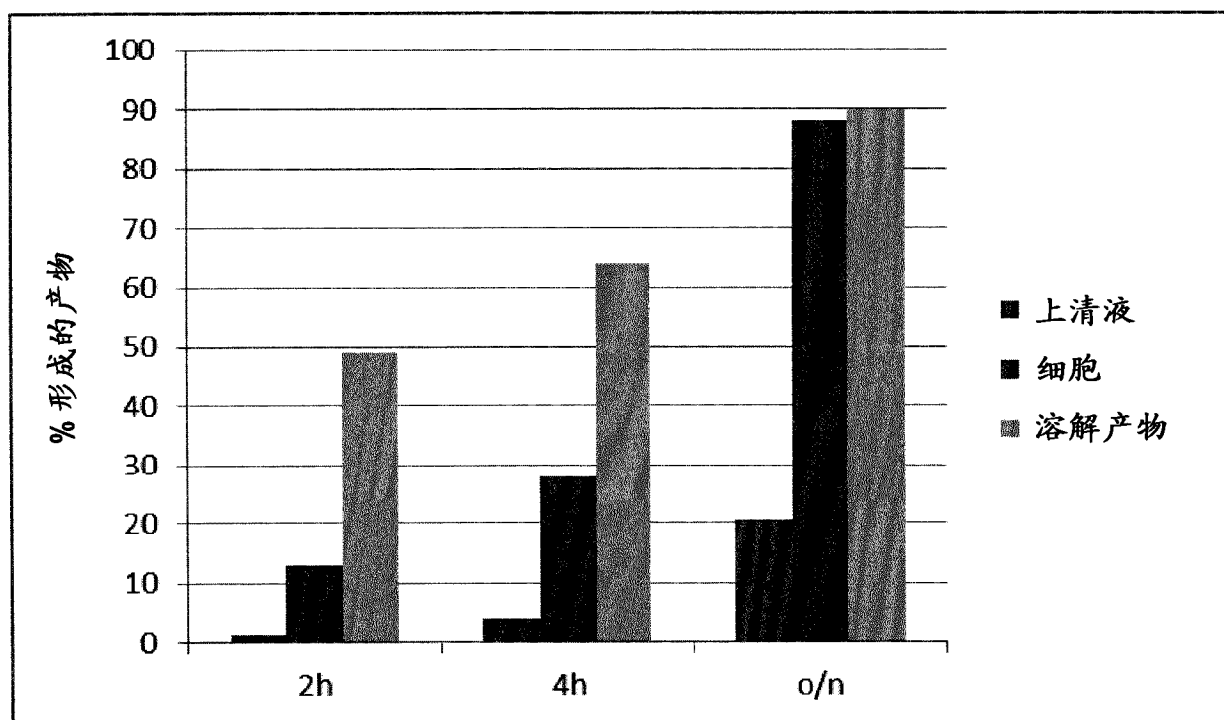


图32

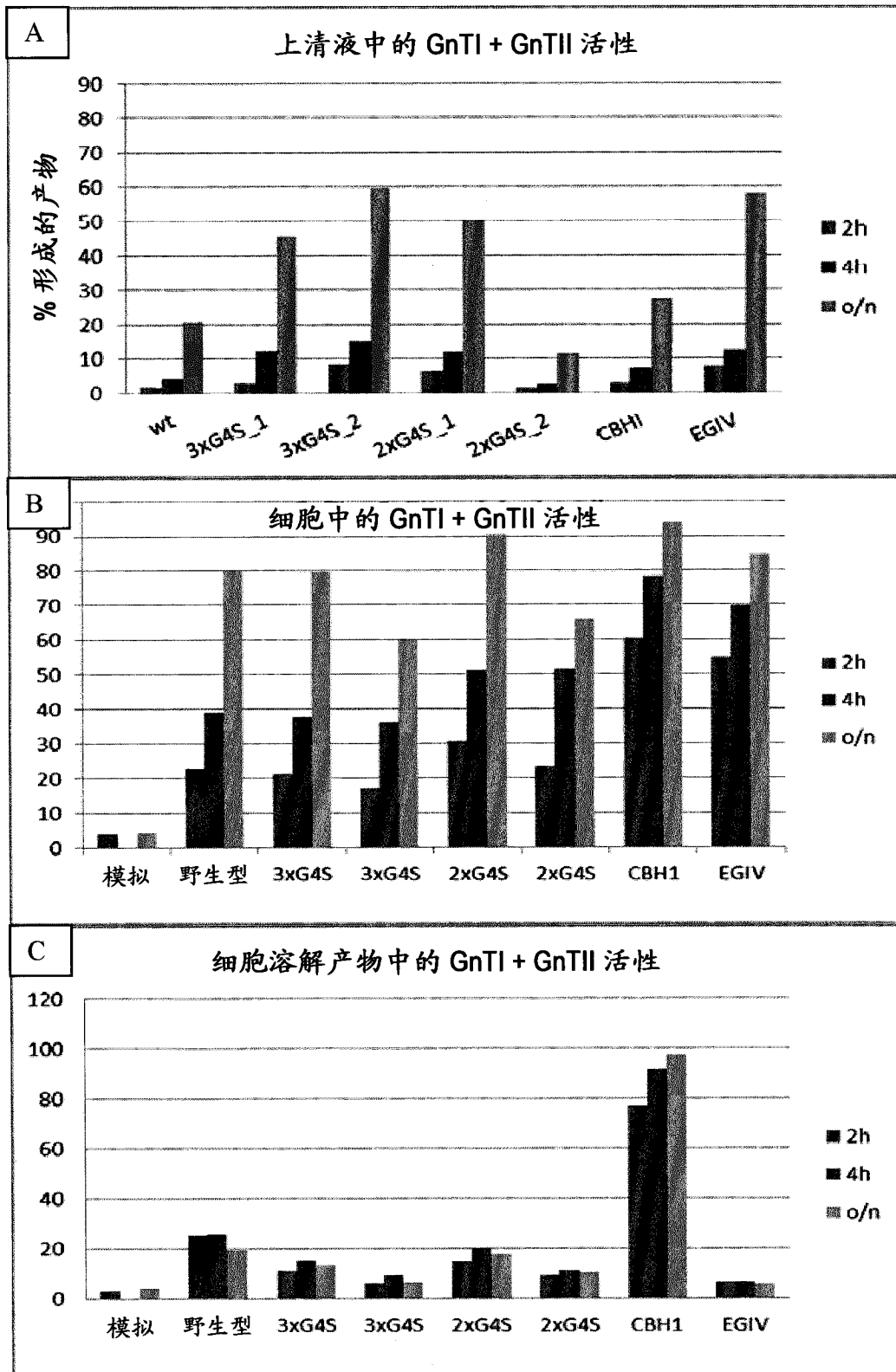


图33

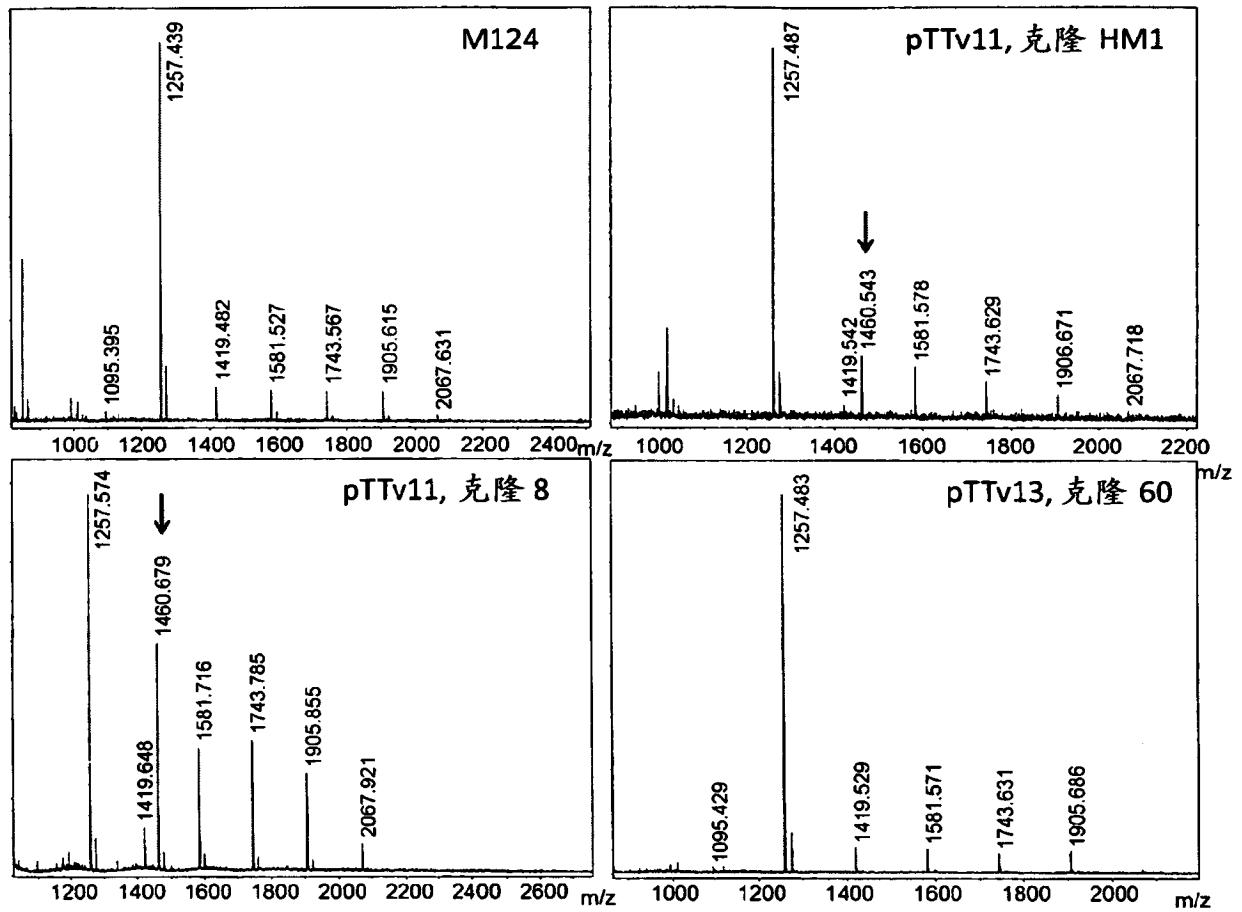


图34

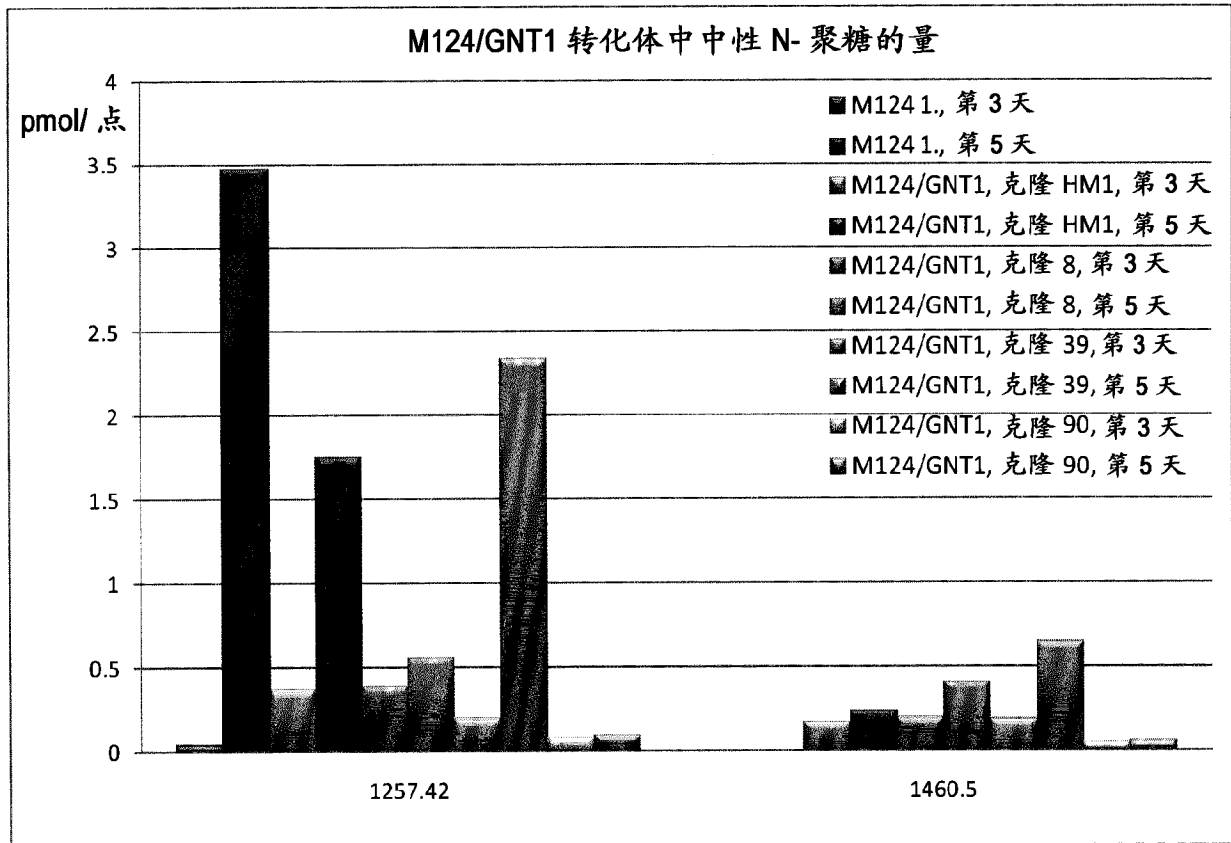


图35

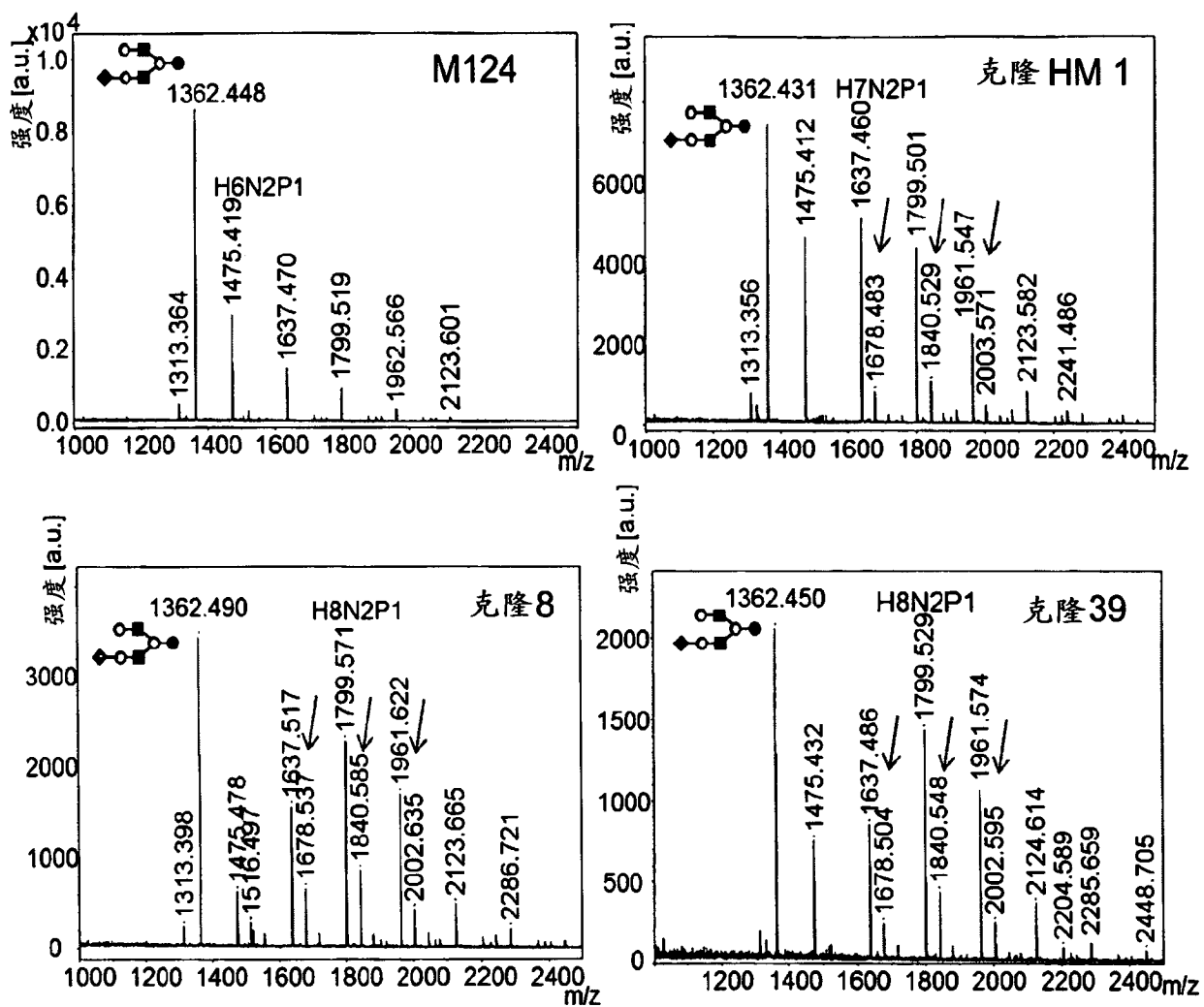


图36

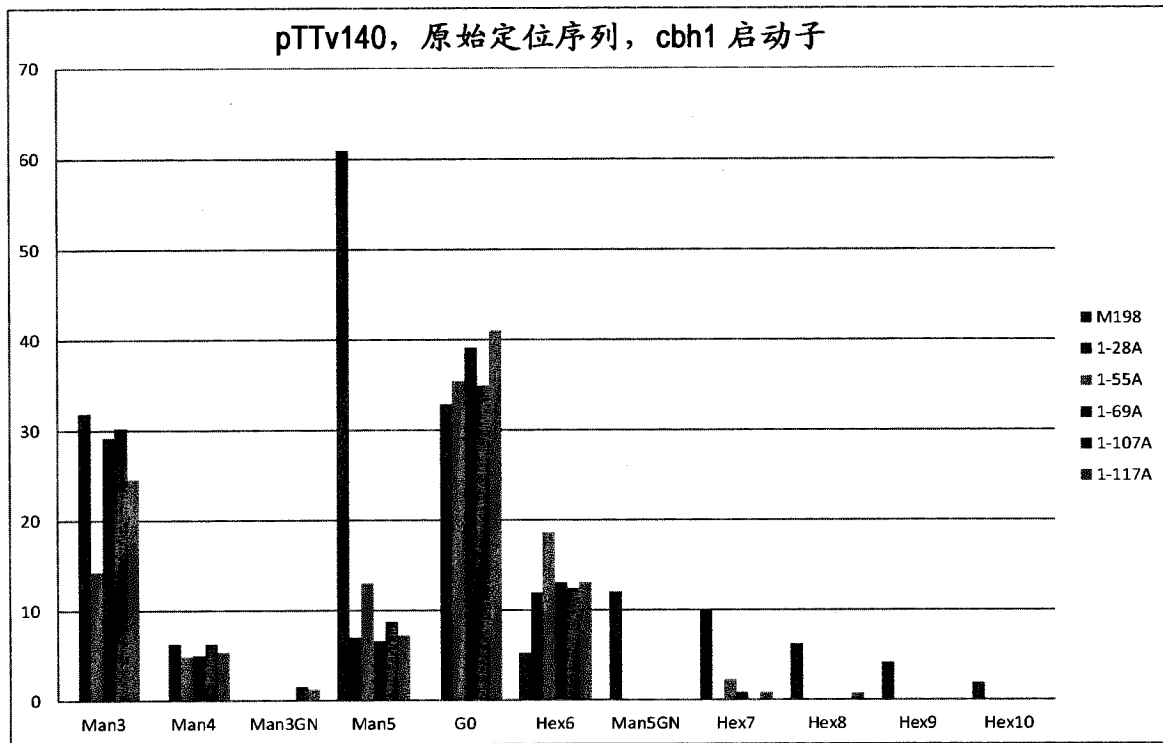
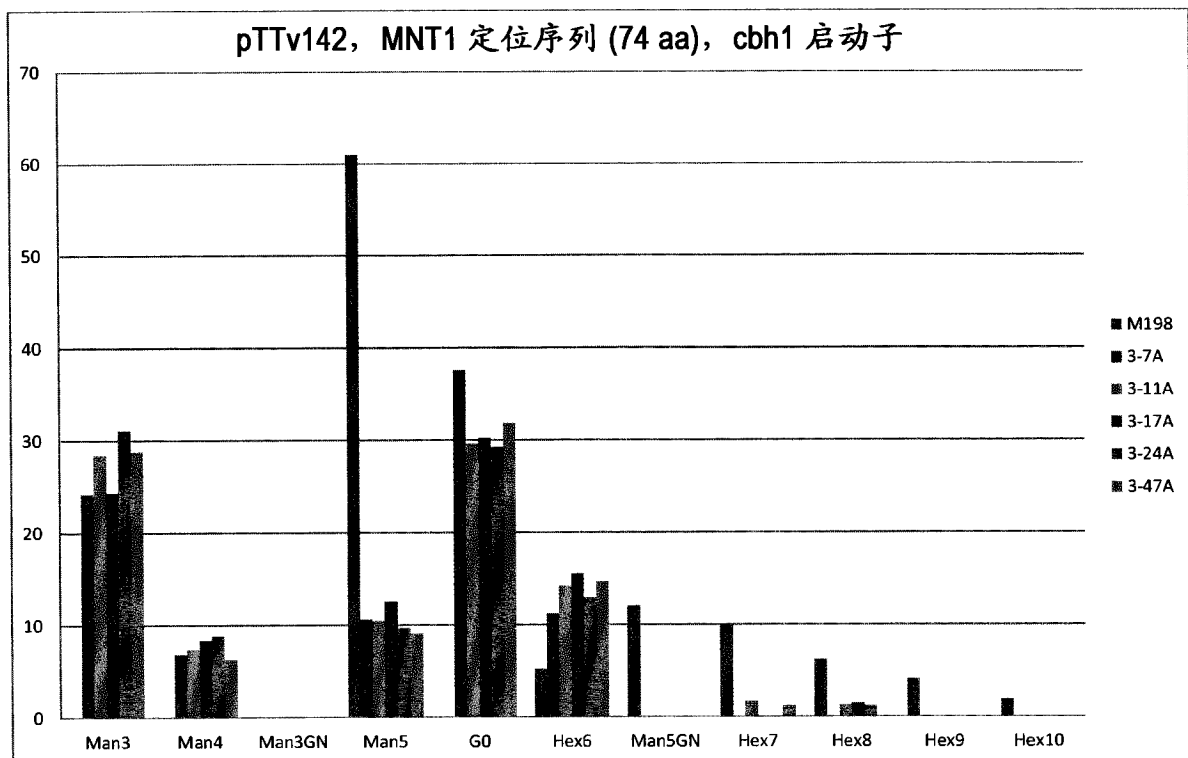
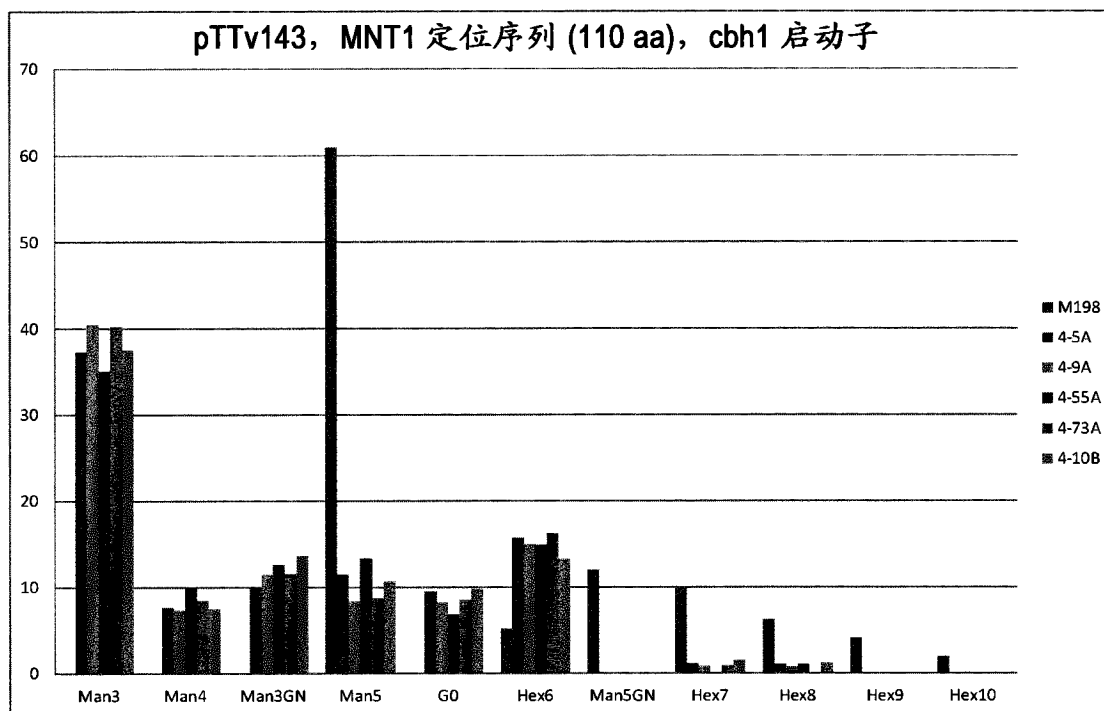
A.**B.**

图37

C.



D.

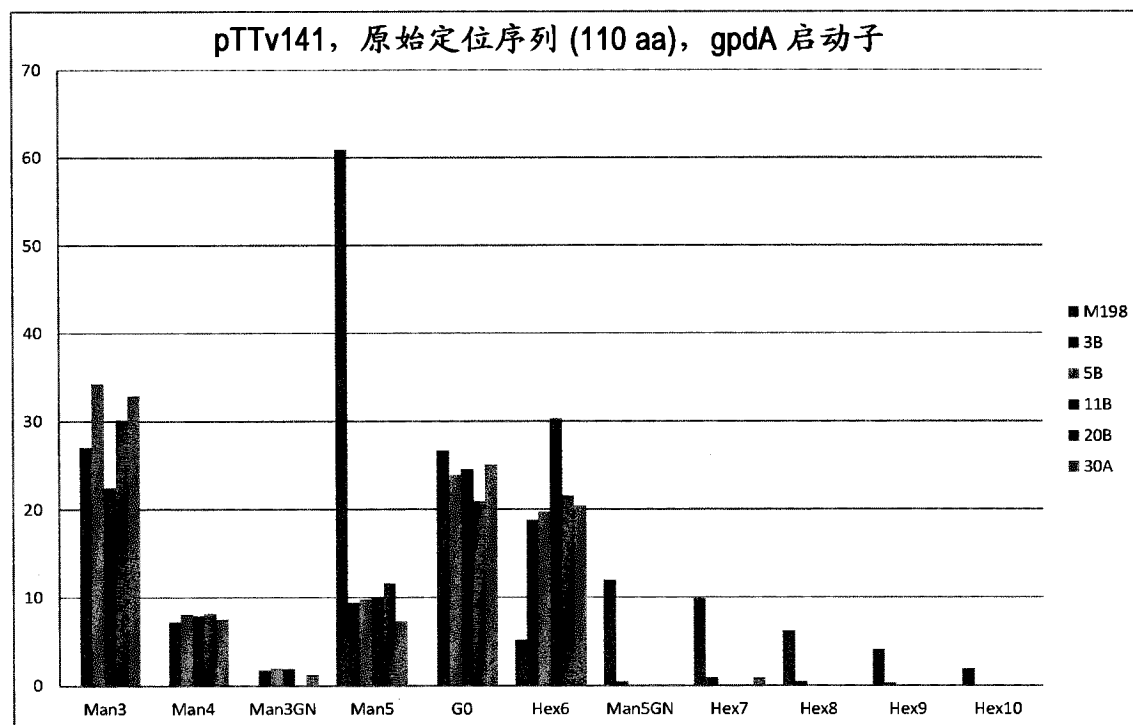


图37(续)

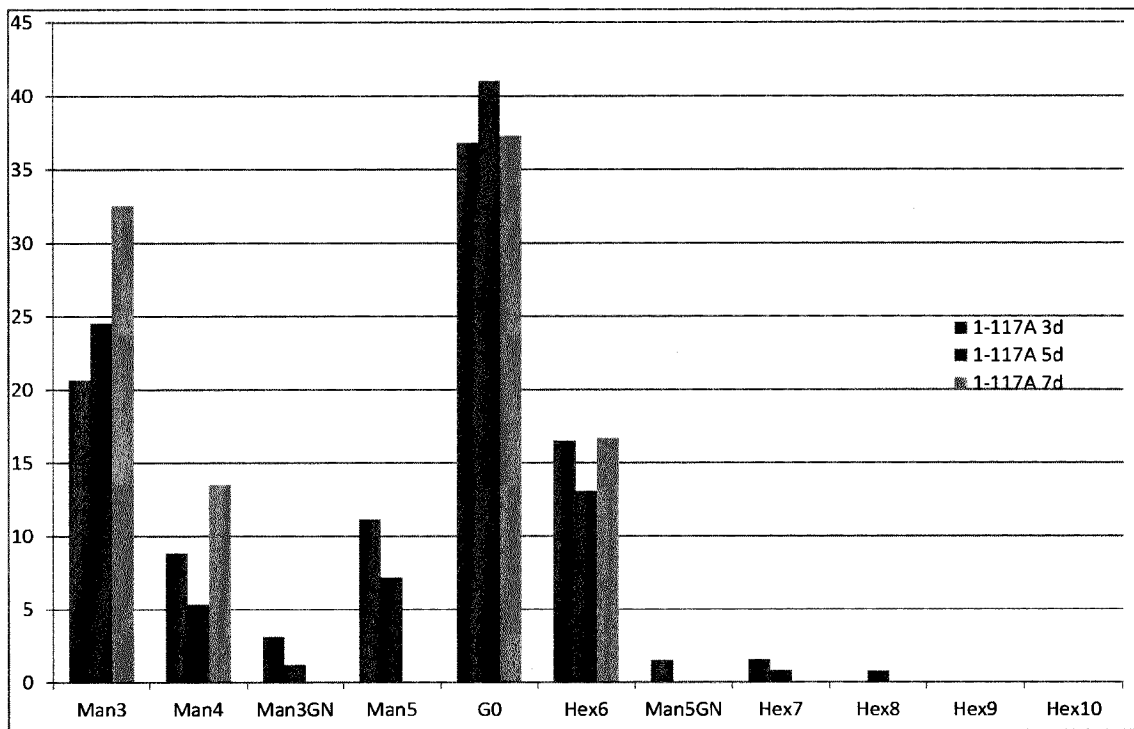
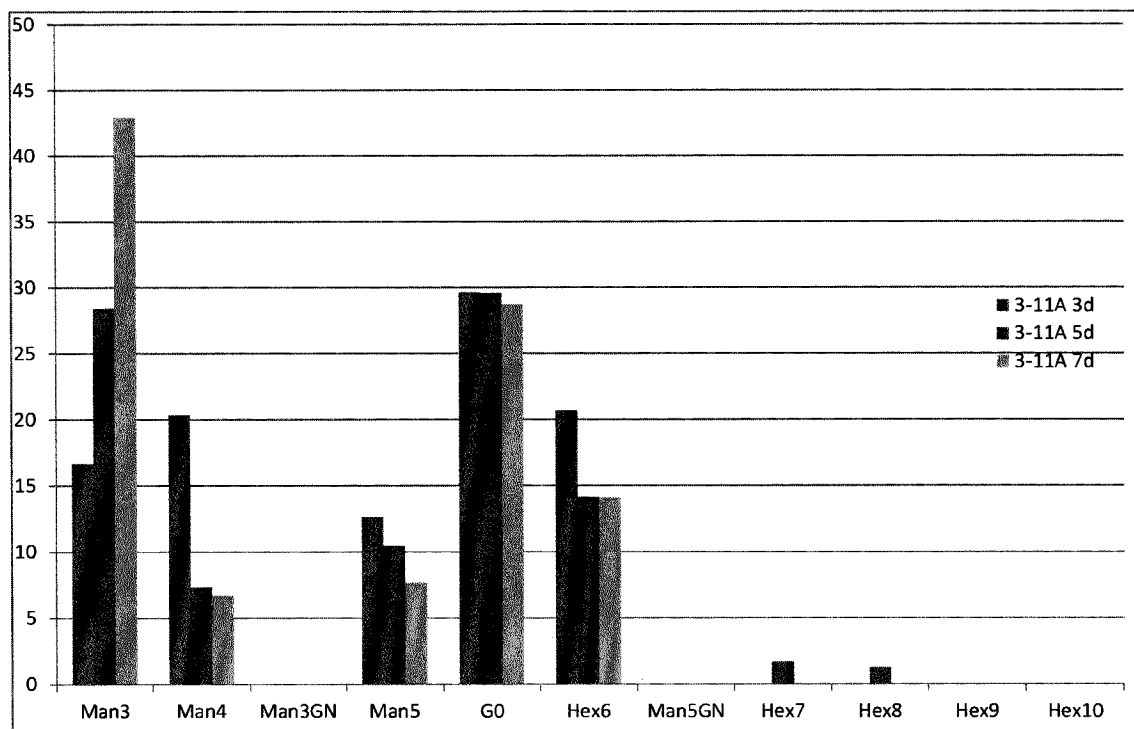
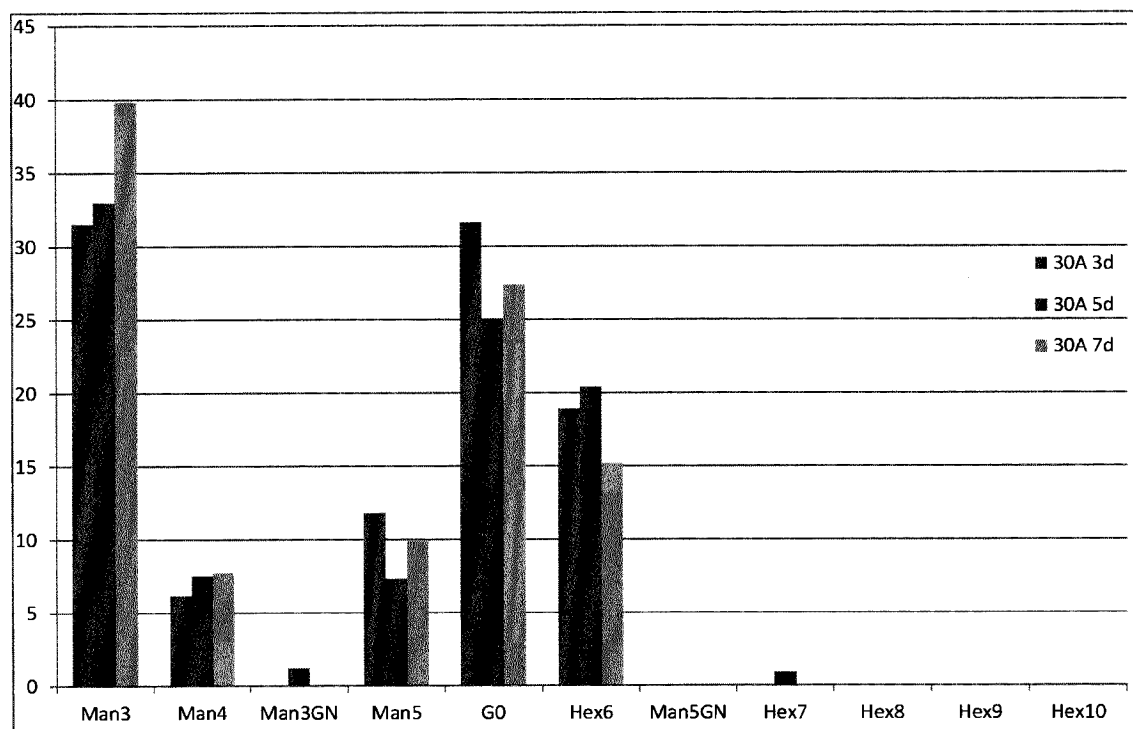
A.**B.**

图38

C.



D.

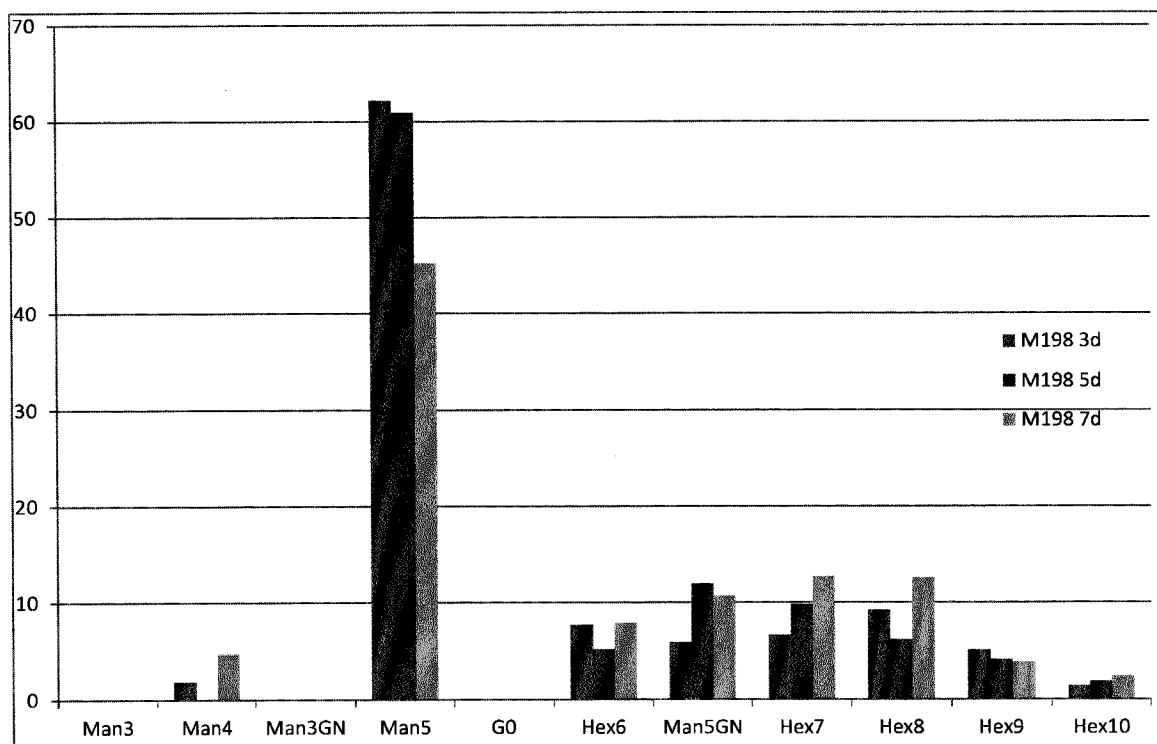


图38(续)

BSA

3-17AM198

标准

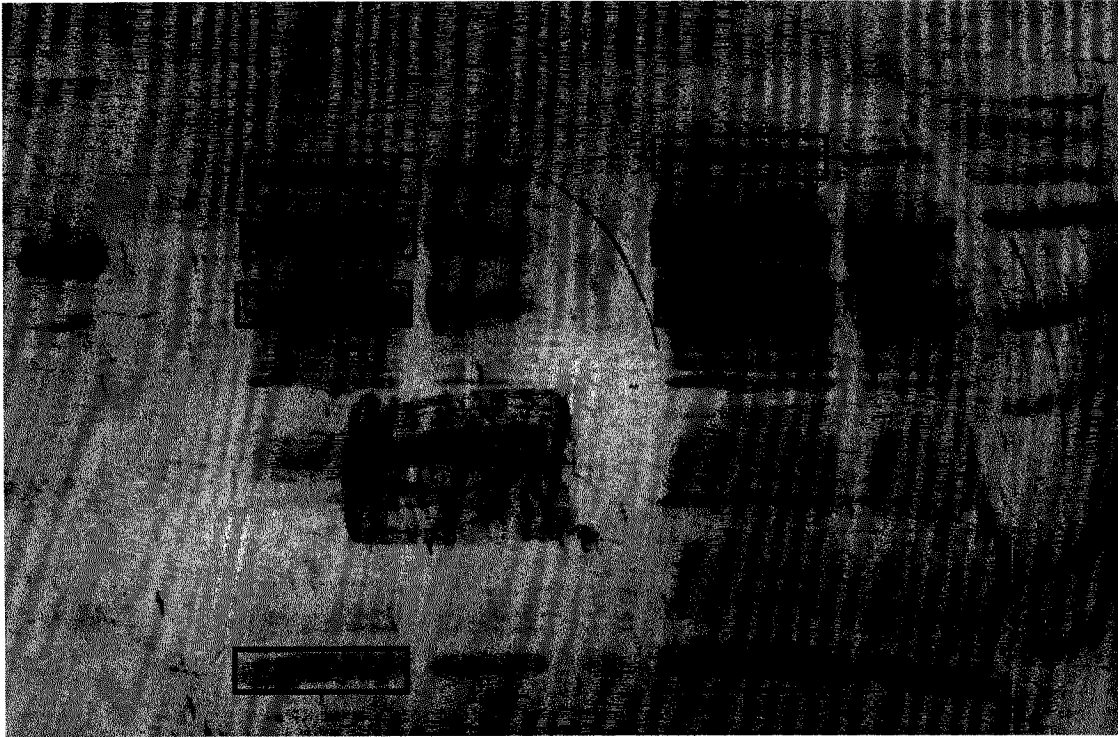


图39

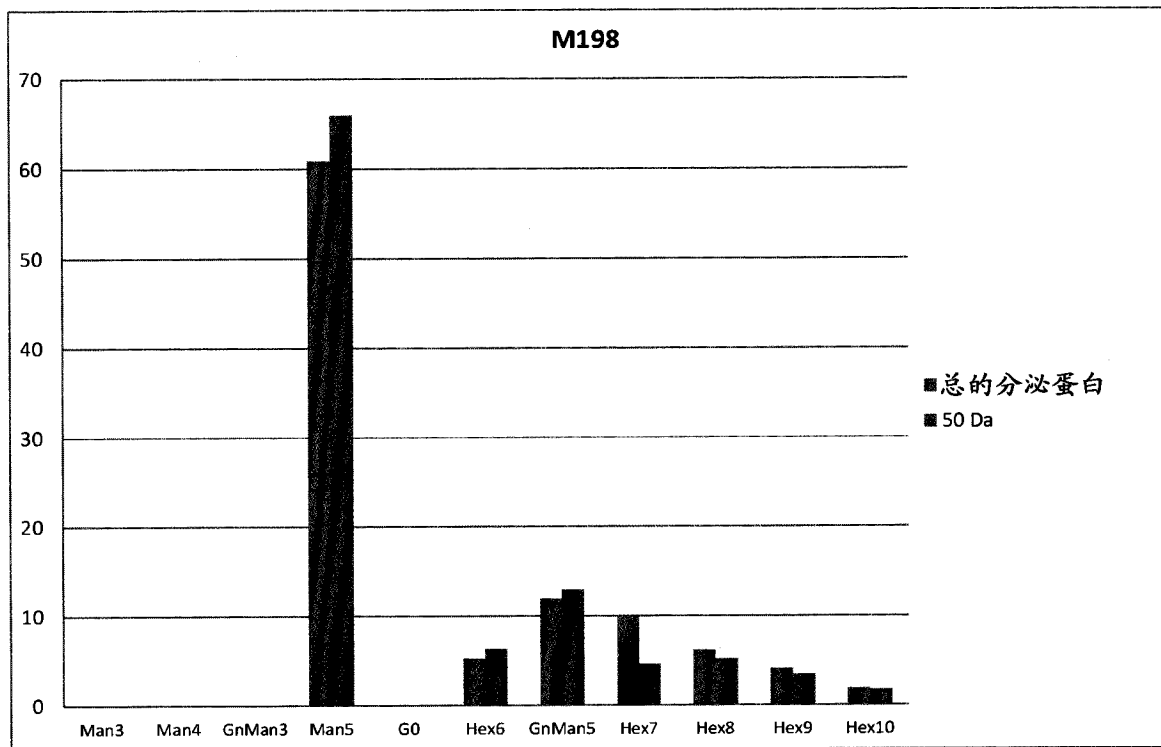
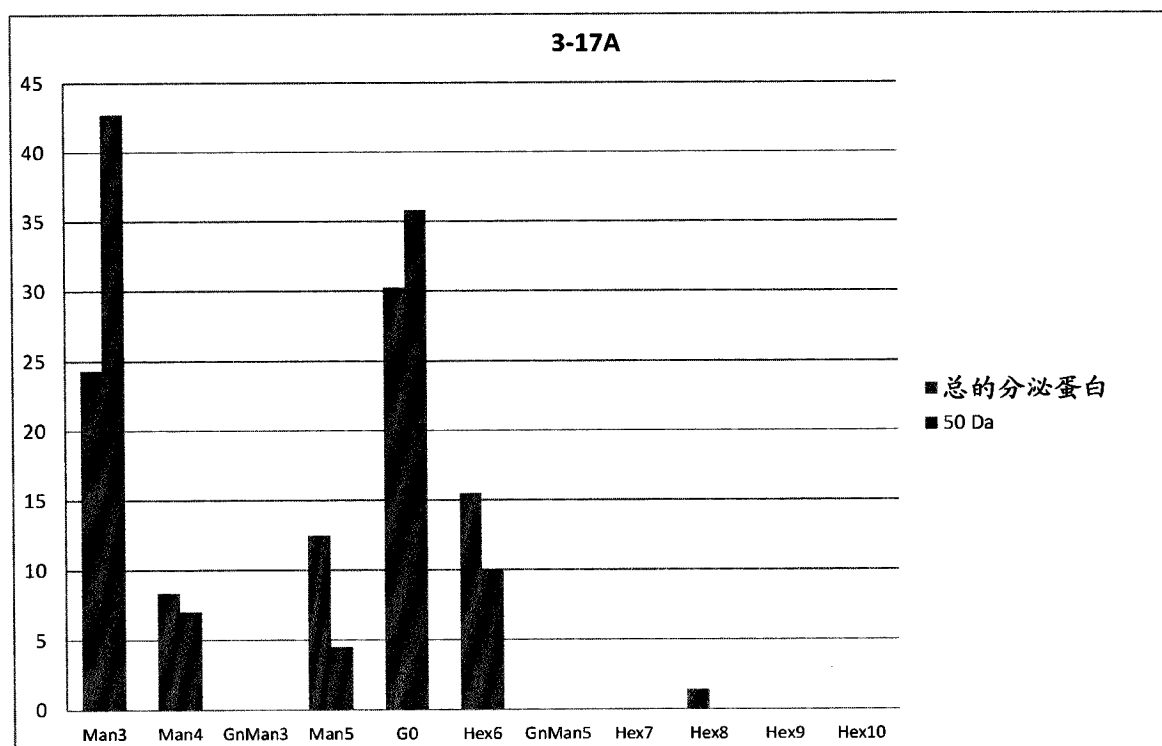
A.**B.**

图40

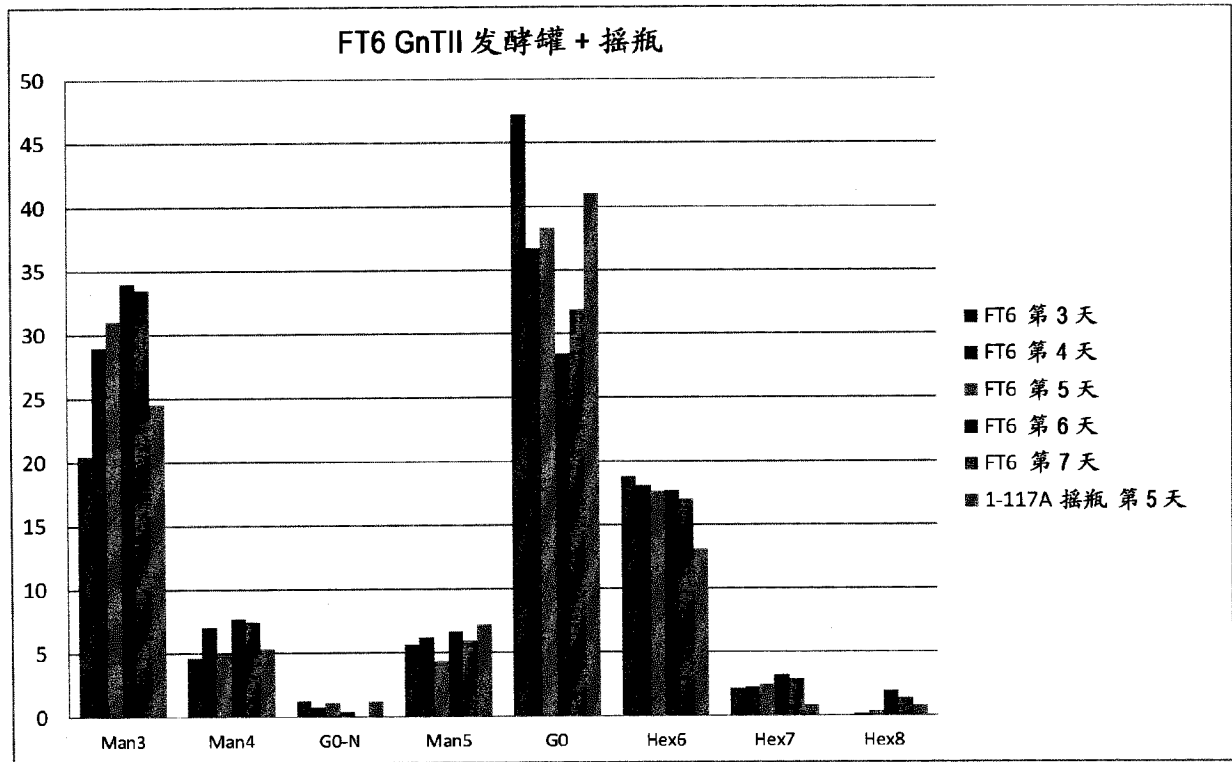


图41

图42

图42(续)

Aoryzae_gi 169779159 ref XP_00	AWNVPSTNASSSVVFS LAVQVFG-----V	391
Aniger_gi 317037568 ref XP_001	AWNTFPSTNLSSTIIVVLSLATQSFG-----V	391
Nfischeri_gi 119478936 ref XP_	AWNTPSTNASSLVVILSLAAQVFG-----V	391
Treesei_jgi Trire2 104121 fgen	AWNVPSTPLSSSVVSVLAVTVA-----MA	403
Tvirens_jgi TriviGv29 8 2 1944	AWNVPSTPLSSNVVSVLAVTVA-----MA	403
Tatroviride_jgi Triat2 270085	AWNVPSTPLSSSVVISLVPVCLLSPQLLVSHDIYNFANC SAILRPRGIA	427
Foxysporum_gi 342880904 gb EGU	AWNVPSTDLSSSTIAVNTMLATVV-----LV	410
Gzea_gi 46136275 ref XP_389829	AWNVPSTDLSSRVTVGAMLATVV-----LA	410
Mthermophila_gi 347013322 gb A	AWNVPSTPISSGVVVGVLALTAA-----LV	424
Ncrassa_gi 85106839 ref XP_962	AWNVPSTPASSAVVGV LGVTVA-----GV	418
Moryzae_gi 145611997 ref XP_36	AWNVPSTPVSSGVVGVMAITVG-----AV	335
Spombe_gi 19114765 ref NP_5938	AWNVPSTKLSSLI AVCVPLITILK-----L	398
	.: ** :.:	
Aoryzae_gi 169779159	LLNSRNALSDAPRRKGKEHIQ-----	413 (SEQ ID NO: 133)
Aniger_gi 317037568	LANSAFAFYTMRSNPSGKEHNQ-----	413 (SEQ ID NO: 135)
Nfischeri_gi 1194789	LGNSFSRKHL DQS--SQKEHLQ-----	411 (SEQ ID NO: 134)
Treesei_jgi Trire2 1	F---AGSNPQPRETSKPKQH-----	420 (SEQ ID NO: 126)
Tvirens_jgi TriviGv2	F---VGSNPQ--RGAPKPKQL-----	419 (SEQ ID NO: 128)
Tatrovirice_jgi Tria	FGQDISATINPDGVAKPLGELENDGLRVWHLASVQVVSFGLHFAHNELGG	477 (SEQ ID NO: 127)
Foxysporum_gi 342880	YLG TARRAVPAPAAQVGNVDDKNK-----	434 (SEQ ID NO: 129)
Gzea_gi 46136275 ref	YRG TARLAVFP--SQARKIEAKNK-----	432 (SEQ ID NO: 130)
Mthermophila gi 3470	WLGAREDWEP--RRVLLKGEAAKR-----	446 (SEQ ID NO: 131)
Ncrassa_gi 85106839	WFGAREEWEPGMKSSSKKEEAAMR-----	442 (SEQ ID NO: 132)
Moryzae_gi 145611997	MVGAKAEFRP--QVPVAKKVEAKR-----	357 (SEQ ID NO: 136)
Spombe_gi 19114765 r	YTS DYRKP-----	406 (SEQ ID NO: 137)

图42(续)