

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200610129096.7

[51] Int. Cl.

C12N 15/09 (2006.01)

C12N 15/31 (2006.01)

C12N 1/21 (2006.01)

C12P 21/02 (2006.01)

C12N 15/70 (2006.01)

C07K 14/195 (2006.01)

[43] 公开日 2007 年 8 月 8 日

[11] 公开号 CN 101012458A

[51] Int. Cl. (续)

C07K 19/00 (2006.01)

C12N 15/62 (2006.01)

A61K 38/16 (2006.01)

A61K 39/02 (2006.01)

A61P 31/04 (2006.01)

[22] 申请日 2001.7.6

[21] 申请号 200610129096.7

分案原申请号 01812409.7

[30] 优先权

[32] 2000. 7. 8 [33] GB [31] 0016702.3

[71] 申请人 英国国防部

地址 英国威尔特郡

[72] 发明人 E·D·威廉森 J·米勒

N·J·瓦尔克 L·W·J·拜利

P·T·霍尔登

H·C·弗利克 - 史密斯

H·L·布利芬特

R·W·蒂特巴尔 A·W·托平

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 梁 谋

权利要求书 4 页 说明书 41 页 附图 10 页

[54] 发明名称

表达系统

[57] 摘要

表达系统。一种引起抗炭疽芽孢杆菌的保护性免疫应答的免疫原性试剂，所述试剂包含结合在一起代表炭疽芽孢杆菌全长保护性抗原(PA)的至多3个结构域或这些结构域的变异体的一种或多种多肽，并且所述结构域中的至少一个包含PA的结构域1或结构域4或其变异体。通过从大肠杆菌表达，生产所述免疫原性试剂的多肽以及全长PA。用该方法获得高产量的多肽。也描述并要求保护了在该方法中使用的细胞、载体和核酸。

1. 一种用于生产引起抗炭疽芽孢杆菌(*Bacillus anthracis*)的保护性免疫应答的免疫原性多肽的方法, 所述方法包括用编码炭疽芽孢杆菌保护性抗原(PA)或其变异体的核酸转化大肠杆菌宿主, 所述变异体是通过删除和/或取代氨基酸序列中一个或多个氨基酸产生的并且其保护性免疫原性活性没有因此受损; 培养所述转化宿主, 并且从中回收所述多肽, 其中核酸中至少部分野生型密码子被大肠杆菌偏好的密码子置换。

2. 一种权利要求 1 的方法, 其中所述核酸编码炭疽芽孢杆菌保护性抗原(PA)。

3. 一种权利要求 1 或 2 的方法, 其中所述核酸中胍和胞嘧啶残基的百分比超过 35%。

4. 一种权利要求 3 的方法, 其中所述核酸中胍和胞嘧啶残基的百分比超过 45%。

5. 一种权利要求 4 的方法, 其中所述核酸中胍和胞嘧啶残基的百分比为 50-52%。

6. 一种重组大肠杆菌(*Escherichia coli*)细胞, 所述细胞已经用编码炭疽芽孢杆菌保护性抗原(PA)或其变异体的核酸转化, 所述变异体是通过删除和/或取代氨基酸序列中一个或多个氨基酸产生的并且其保护性免疫原性活性没有因此受损, 以及其中核酸中至少部分野生型密码子被大肠杆菌偏好的密码子置换。

7. 一种权利要求 6 的重组大肠杆菌细胞, 其中所述核酸中胍和胞嘧啶残基的百分比超过 35%。

8. 一种权利要求 7 的重组大肠杆菌细胞, 其中所述核酸中胍和胞嘧啶残基的百分比超过 45%。

9. 一种权利要求 8 的重组大肠杆菌细胞, 其中所述核酸中胍和胞嘧啶残基的百分比为 50-52%。

10. 一种权利要求 6 的重组大肠杆菌细胞, 其中所述核酸是图 2 显示的 SEQ ID NO 1 或其修饰形式。

11. 一种权利要求 10 的重组大肠杆菌细胞, 其中所述核酸是 SEQ ID NO 1。

12. 一种用于生产引起抗炭疽芽孢杆菌的保护性免疫应答的多肽的方法, 所述方法包括培养权利要求 6-11 中任一项的细胞, 并且从培养物中回收所述保护性多肽。

13. 一种大肠杆菌转化载体, 所述载体包含编码炭疽芽孢杆菌保护性抗原(PA)或其变异体的核酸, 所述变异体是通过删除和/或取代氨基酸序列中一个或多个氨基酸产生的并且其保护性免疫原性活性没有因此受损, 以及其中核酸中至少部分野生型密码子被大肠杆菌偏好的密码子置换。

14. 一种权利要求 13 的大肠杆菌转化载体, 其中所述核酸中脯氨酸和胞嘧啶残基的百分比超过 35%。

15. 一种 SEQ ID NO 1 的核酸或其修饰形式, 所述核酸或其修饰形式编码引起保护性免疫应答的 PA 或其变异体, 以及其中核酸中至少部分野生型密码子被大肠杆菌偏好的密码子置换。

16. 一种权利要求 15 的核酸, 其具有至少 35%GC 含量。

17. 一种权利要求 15 的核酸, 其为 SEQ ID NO 1。

18. 一种权利要求 15 的核酸, 其与 SEQ ID NO 1 至少 90%相同。

19. 一种权利要求 15 的核酸, 其包含 SEQ ID NO 1。

20. 一种引起抗炭疽芽孢杆菌(*Bacillus anthracis*)的保护性免疫应答的免疫原性试剂, 所述试剂包含一种或多种多肽, 所述多肽结合在一起代表炭疽芽孢杆菌全长保护性抗原(PA)的至多 3 个结构域或这些结构域的变异体, 并且所述结构域中的至少一个包含 PA 的结构域 1 或结构域 4 或其变异体。

21. 一种权利要求 20 的免疫原性试剂, 所述试剂包含炭疽芽孢杆菌 PA 的结构域 4。

22. 一种权利要求 20 或权利要求 21 的免疫原性试剂, 所述试剂包含结构域 1 和 4 或它们的保护性区的组合。

23. 一种权利要求 22 的免疫原性试剂, 所述试剂包含野生型 PA 的结构域 1 和结构域 4 的序列。

24. 一种权利要求 23 的免疫原性试剂, 其中所述结构域以融合多肽的形式存在。

25. 一种权利要求 24 的免疫原性试剂, 所述试剂包含与所述 PA 序列的结构域 2 融合的结构域 1。

26. 一种权利要求 25 的免疫原性试剂, 所述试剂与所述 PA 序列的结构域 3 融合。

27. 一种权利要求 23 的免疫原性试剂, 所述试剂包含多种多肽的混合物, 其中一种多肽包含所述 PA 序列的结构域 1, 而其中一种多肽包含所述 PA 序列的结构域 4。

28. 一种权利要求 20-27 中任一项的免疫原性试剂, 其中一种多肽与另一种多肽融合。

29. 一种权利要求 28 的免疫原性试剂, 其中所述另一种肽是谷胱甘肽-S-转移酶(GST)。

30. 一种核酸, 所述核酸编码权利要求 20-27 中任一项的免疫原性试剂的多肽。

31. 一种表达载体, 所述表达载体包含权利要求 30 的核酸。

32. 一种用权利要求 31 的载体转化的细胞。

33. 一种用于生产引起抗炭疽芽孢杆菌的保护性免疫应答的免疫原性多肽的方法, 所述方法包括用编码炭疽芽孢杆菌保护性抗原(PA)的保护性结构域或其引起保护性免疫应答的变异体的核酸转化大肠杆菌宿主, 培养所述转化宿主, 并且从中回收所述多肽。

34. 一种权利要求 33 的方法, 其中核酸中至少部分野生型密码子被大肠杆菌偏好的密码子置换。

35. 一种权利要求 34 的方法, 其中所述核酸中胍和胞嘧啶残基

的百分比超过 45%。

36. 一种权利要求 35 的方法，其中所述核酸中胍和胞嘧啶残基的百分比为 50-52%。

37. 一种权利要求 33 的方法，其中所述结构域是炭疽芽孢杆菌 PA 的结构域 1 和/或结构域 4。

38. 一种重组大肠杆菌细胞，所述细胞已经用编码炭疽芽孢杆菌保护性抗原(PA)的保护性结构域或其引起保护性免疫应答的变异体的核酸转化。

39. 一种权利要求 38 的重组大肠杆菌细胞，其中所述核酸编码炭疽芽孢杆菌 PA 的结构域 1 或结构域 4。

40. 一种用于生产引起抗炭疽芽孢杆菌的保护性免疫应答的多肽的方法，所述方法包括培养权利要求 38 或 39 的细胞，并且从所述培养物中回收所述保护性多肽。

41. 一种大肠杆菌转化载体，所述载体包含编码炭疽芽孢杆菌保护性抗原(PA)的保护性结构域或其可引起保护性免疫应答的变异体的核酸。

42. 一种预防或治疗炭疽芽孢杆菌感染的方法，所述方法包括给予需要预防或治疗的哺乳动物足量的权利要求 20-29 中任一项的免疫原性试剂。

43. 权利要求 20-29 中任一项的免疫原性试剂在制备用于预防或治疗炭疽芽孢杆菌感染的药物中的用途。

表达系统

本申请是专利申请号为 01812409.7、国际申请日为 2001 年 7 月 6 日（国际申请号为 PCT/GB01/03065）、发明名称为“表达系统”的发明专利申请的分案申请。

发明领域

本发明涉及引起抵抗炭疽芽孢杆菌(*Bacillus anthracis*)感染的保护性免疫应答的多肽，涉及生产这些多肽的方法、可用于所述方法的重组大肠杆菌(*Escherichia coli*)细胞以及采用的核酸和转化载体。

背景技术

用于表达供疫苗系统用的保护性抗原(PA)的本发明系统，用蛋白酶缺陷型枯草杆菌(*Bacillus subtilis*)作为表达宿主。尽管这类系统就产物的量和纯度而论是可接受的，但仍有显著缺陷。首先，管理机构一般不熟悉这种宿主，批准决定可能因此而延迟。更重要的是，目前使用的枯草杆菌菌株产生热稳定孢子，因而需要使用专用的生产工厂。

WO00/02522 具体描述了表达 PA 或某些免疫原性片段的 VEE 病毒复制子。

众所周知大肠杆菌是多种人用疫苗的表达系统。虽然容易使大肠杆菌发酵至非常高的细胞密度的能力，使得这种细菌成为表达许多蛋白质的理想宿主，但先前为用表达和从大肠杆菌细胞溶胶中纯化重组 PA 所作的尝试由于蛋白质收率低和蛋白酶降解而受到阻碍(Singh 等, J. Biol. Chem. (1989) 264; 11099-11102, Vodkin 等, Cell (1993) 34; 693-697 以及 Sharma 等, Protein Expr. Purif. (1996), 7, 33-38)。

在大肠杆菌细胞溶胶中过量表达作为稳定的可溶性蛋白的 PA 的策略最近已有描述(Willhite 等, Protein and Peptide Letters, (1998), 5; 273-278)。所用的策略是将亲和标志序列加至 PA 的 N 末端的一种策略, 这提供了一种简单的纯化系统。

这种系统的一个问题是需要另一个下游的加工步骤, 以便在可以使用 PA 之前除去所述标志。

密码子优化是现在众所周知并且用于合成基因设计的一种技术。遗传密码有一定程度的丰余性, 因此大多数氨基酸由不止一个密码子序列编码。不同的生物优先利用这些不同密码子中的某一个。一般可预期, 通过优化密码子, 将提高特定蛋白的表达水平。

这一般是理想的, 但 PA 除外, 就 PA 而论, 较高的表达水平会导致蛋白酶降解和/或细胞毒性。在这种情况下, 提高表达水平可能是逆生产性的(counter-productive), 并且引起显著的细胞毒性。

然而意外的是, 本申请人已经发现, 在大肠杆菌中情况不是这样, 在该系统中, 无论在该菌株中是否存在蛋白酶, 密码子优化均导致意想不到的高水平重组 PA 的表达。

此外, 似乎 PA 保护性结构域的表达并不抑制大肠杆菌中的表达。

天然 PA 的晶体结构已经阐明(Petosa C.等 Nature 385:833-838, 1997), 表明 PA 由 4 个不同的功能独立的结构域组成: 结构域 1, 分为 1a (1-167 号氨基酸)和 1b (168-258 号氨基酸); 结构域 2, 259-487 号氨基酸; 结构域 3, 488-595 号氨基酸; 和结构域 4, 596-735 号氨基酸。

发明内容

本申请人已经确定, 某些结构域当以分离形式、融合蛋白形式或者相互组合形式使用时, 看来产生意想不到的良好保护性效应。

按照本发明, 提供一种引起抗炭疽芽孢杆菌的保护性免疫应答的免疫原性试剂, 所述试剂包含结合在一起代表炭疽芽孢杆菌全长保护性抗原(PA)的至多 3 个结构域或这些结构域的变异体的一种或多种多肽, 并且所述结构域中的至少一个包含 PA 的结构域 1 或结构域 4 或其变异体。

具体地说, 所述试剂将包含多肽的混合物或融合肽, 其中所述各个多肽包含 PA 多个个别结构域中的一个。

具体地说, 所述试剂包括非全长 PA 形式的包含 PA 结构域 1 或结构域 4 或其变异体的多肽。当存在时, 结构域适为完整的, 特别是结构域 1 全部存在。

本文使用的术语“多肽”包括蛋白质和肽。

用于本文时, 表述“变异体”是指由于基础序列中一个或多个氨基酸缺失或取代为其它氨基酸而不同于所述序列、但仍引起抗炭疽芽孢杆菌的保护性免疫应答的氨基酸序列。当一种氨基酸被一种不同的、其性质大致相似的氨基酸取代时, 所述氨基酸取代可以被称为“保守的”。当氨基酸被不同类型的氨基酸取代时, 为非保守取代。广义上讲, 较少的非保守取代在不影响所述多肽的生物活性的情况下是可行的。合适的变异体与所述 PA 序列将有至少 60%相同, 优选至少 75%相同, 更优选至少 90%相同。

具体地说, 特定变异体序列与所述 PA 序列的同一性可以采用 Lipman 和 Pearson 描述的多重比对法(Lipman D.J.和 Pearson, W.R. (1985) Rapid and Sensitive Protein Similarity Searches, Science, vol 227, pp1435-1441)来评价。“优化”百分比分值可以用 Lipman-Pearson 算法的以下参数来计算: ktup=1, gap penalty (空位罚分)=4, 和 gap penalty length (空位罚分长度)=12。待评价相似性的序列应该用作“试验序列”, 这意指用于比较的碱基序列(SEQ ID NO 1)应该首先进入该算法。

最好是, 本发明的试剂包括具有野生型 PA 的结构域 1 序列和/或

结构域 4 序列的多肽。

本发明的一个特别优选的实施方案包括炭疽芽孢杆菌 PA 的结构域 4。

这些结构域包含以下表 1 中所示的以下序列。

表 1

结构域	全长 PA 的氨基酸*
4	596-735
1	1-258

这些氨基酸编号是指 *在 Welkos 等 Gene 69 (1988) 287-300 中所示的序列，在下文中分别描述为 SEQ ID NO 15 (图 4) 和 SEQ ID NO 3 (图 3)。*

结构域 1 包括两个区，名为 1a 和 1b。区 1a 包括氨基酸 1-167，而区 1b 从氨基酸 168 至氨基酸 258。看来区 1a 对于良好的保护性应答的产生是重要的，可能最好是全长的结构域。

在一个特别优选的实施方案中，用结构域 1 和结构域 4 或其保护性区的组合作为免疫原性试剂，该试剂引起抗炭疽芽孢杆菌的保护性免疫应答。这种组合例如可作为融合肽，可以用下文概述的本发明的表达系统来表达。

当使用结构域 1 时，适宜将其与所述 PA 序列的结构域 2 融合，可能最好将其与结构域 2 和结构域 3 融合。

这类组合及其在预防或治疗中的应用，构成了本发明的另一方面。

上述结构域适宜作为融合蛋白的部分，最好具有 N 末端谷胱甘肽-s-转移酶蛋白(GST)。GST 不仅有助于所述蛋白的纯化，它也可以提供佐剂效应，可能是由于大小增加了。

本发明的多肽适宜用常规方法制备。例如，可以合成它们，或者

可以用重组 DNA 技术制备它们。特别是，将编码所述结构域的核酸包括在表达载体中，用所述载体转化宿主细胞。然后可以采用常规方法进行宿主细胞的培养以及随后所需多肽的分离。这些方法中所用的核酸、载体和转化细胞构成了本发明的再一方面。

一般而言，所用的宿主细胞是常规用来制备 PA 的宿主细胞，例如枯草杆菌。

本申请人意外的发现，或者分离的或者组合的所述结构域可以在一定条件下在大肠杆菌中成功地表达。

因此，本发明还提供一种用于生产引起抗炭疽芽孢杆菌的保护性免疫应答的免疫原性多肽的方法，所述方法包括用或者(a)编码炭疽芽孢杆菌保护性抗原(PA)或可以引起保护性免疫应答的其变异体的核酸，或者(b)编码如上所述的炭疽芽孢杆菌保护性抗原(PA)的至少一个保护性结构域或可以引起保护性免疫应答的其变异体的核酸，转化大肠杆菌宿主；培养所述转化宿主，并且从中回收所述多肽，前提是：当所述多肽是炭疽芽孢杆菌保护性抗原(PA)或可以引起保护性免疫应答的其变异体时，所述核酸中胍和胞嘧啶残基的百分比超过 35%。

应用这些选项，可以用优选的表达宿主获得高收率的产物。

在图 1 中显示了一个表，该表显示了在大肠杆菌和炭疽芽孢杆菌基因组中出现的密码子及其频率。显然，胍和胞嘧啶在大肠杆菌中出现的频率比在炭疽芽孢杆菌中出现的频率高得多。密码子选择含量的分析显示出以下结果：

物种	密码子第一个字母为 GC	密码子第二个字母为 GC	密码子第三个字母为 GC	总 GC 含量
大肠杆菌	58.50%	40.70%	54.90%	51.37%
炭疽芽孢杆菌	44.51%	31.07%	25.20%	33.59%

因此，看来大肠杆菌偏好的密码子是可能时包括胍或胞嘧啶的那些密码子。

通过在用来编码所述免疫原性蛋白的序列中将胍和胞嘧啶核苷酸的百分比提高至野生型炭疽芽孢杆菌基因中正常存在的百分比，密码子选择将使得在大肠杆菌中的表达得以改善。

在本发明中使用的编码核酸中胍和胞嘧啶残基的百分比至少在所述多肽为炭疽芽孢杆菌保护性抗原(PA)或可以引起保护性免疫应答的其变异体时，适宜超过 40%，优选超过 45%，最优选为 50-52%。

保护性结构域的高水平表达可以采用编码这些单位的野生型炭疽芽孢杆菌序列来达到。然而，如上所述，通过提高所述核酸的 GC%，可以进一步提高收率。

在一个具体实施方案中，所述方法包括表达炭疽芽孢杆菌的 PA。

此外，按照本发明，提供一种已经用编码炭疽芽孢杆菌保护性抗原(PA)或可以引起保护性免疫应答的其变异体的核酸转化的重组大肠杆菌细胞，其中所述核酸中胍和胞嘧啶残基的百分比超过 35%。

如上所述，所述编码核酸中胍和胞嘧啶残基的百分比适宜超过 40%，优选超过 45%，最优选为 50-52%。

本发明的用来转化大肠杆菌细胞的核酸适宜为合成基因。特别是，所述核酸为图 2 中所示的 SEQ ID NO 1 的核酸或其修饰形式。

表述“修饰形式”是指编码 PA 或引起保护性免疫应答但利用某些不同密码子的片段或其变异体的其它核酸序列，条件是符合对按照本发明的 GC 百分含量的要求。合适的修饰形式是与 SEQ ID NO 1 至少 80%相似，优选 90%相似，最优选至少 95%相似。特别是，所述核酸包含 SEQ ID NO 1。

在一个替代实施方案中，本发明提供一种已经用编码炭疽芽孢杆菌保护性抗原(PA)的保护性结构域或可以引起保护性免疫应答的其变异体的核酸转化的重组大肠杆菌细胞。

所述核酸最好编码炭疽芽孢杆菌的结构域 1 或结构域 4。

按照本发明，再提供一种生产免疫原性多肽的方法，所述多肽引起抗炭疽芽孢杆菌的保护性免疫应答，所述方法包括培养如上所述的

细胞，并且从培养物中回收所需的多肽。这类方法是本领域众所周知的。

在再一方面，本发明提供一种大肠杆菌转化载体，所述载体包含编码炭疽芽孢杆菌保护性抗原(PA)或可以引起保护性免疫应答的其变异体的核酸，其中所述核酸中的胍和胞嘧啶残基的百分比超过 35%。

本发明的又一方面包括一种大肠杆菌转化载体，所述载体包含编码炭疽芽孢杆菌保护性抗原(PA)的保护性结构域或可以引起保护性免疫应答的其变异体的核酸。

可供用于大肠杆菌转化的合适载体是本领域众所周知的。例如，T7 表达系统提供良好的表达水平。然而，一种特别优选的载体包括可得自 Avecia (UK)的 pAG163。

SEQ ID NO 1 的核酸或者编码 PA 并且 GC 含量至少为 35%、优选至少 40%、更优选至少 45%、最优选为 50-52%的其变异体，构成了本发明的另一方面。

如有必要，所述变异体的 PA 或结构域可以作为与另一蛋白的融合体表达，所述另一蛋白例如为提供一种不同免疫的蛋白、有助于所述产物纯化的蛋白或者用以确保翻译的良好起始的高表达蛋白(例如硫氧还蛋白，GST)。

可任选地加入其它系统例如 T7 溶菌酶至所述表达系统中，以改善对所述系统的阻抑，尽管就本发明而论，尚未记录到与细胞毒性相关的问题。

任何合适的大肠杆菌菌株都可以用于本发明的方法中。可得到许多蛋白酶缺陷型的菌株(例如 *lon⁻*、*ompT⁻*)，预期所述菌株会将蛋白水解减少至最低。然而，本发明人已经发现，不需要应用这类菌株来达到产物的良好收率，其它已知菌株例如 K12 可产生意想不到的高产物收率。

所述菌株的发酵一般在本领域知道的常规条件下进行。例如，发酵可以作为分批培养物进行，优选在大摇瓶中进行，使用含有抗生素

以保持质粒并添加 IPTG 进行诱导的复合培养基。

适当地收获培养物，将细胞贮存于-20℃直至需要用于纯化。

对于大肠杆菌 PA (或变异体或结构域)表达的合适的纯化方案可以根据在枯草杆菌表达中所用的方法进行改进。待使用的各个纯化步骤将取决于重组 PA 的物理特性。通常，在允许最大差别结合于所述离子交换柱条件下进行离子交换色谱分离，然后从窄梯度收集流分。在某些情况下，单个色谱分离步骤可能足以获得所需规格的产物。

可以根据需要用 SDS PAGE 或蛋白质印迹，分析流分中所述产物的存在。

正如下文说明的，已经达到了代表 rPA 的完整结构域或部分结构域的一组融合蛋白的成功克隆和表达。在 A/J 小鼠模型中评价了这些融合蛋白的免疫原性和抗 STI 孢子攻击的保护性功效。

所有所述 rPA 结构域蛋白在 A/J 小鼠中都有免疫原性，并且与 GST 对照免疫小鼠相比，至少赋予抗攻击的部分保护。连接于所述结构域蛋白 N 末端的载体蛋白 GST，通过在免疫动物体内所激发的抗体应答表明，并不损害所述融合蛋白的体内免疫原性，或者不损害所述融合蛋白的体外免疫原性，因为用抗 rPA 抗血清在蛋白质印迹后可以检测到所述融合蛋白，这表明 GST 标志并不干扰 rPA 表位的识别。用较大的融合蛋白免疫，产生最高效价。具体地说，用全长 GST 1-4 融合蛋白免疫的小鼠，产生的平均血清抗 rPA 浓度大约是 rPA 免疫组的 8 倍 (图 5)。用切除所述 GST 的 rPA 结构域 1-4 免疫小鼠，产生的效价大约是用所述融合蛋白免疫所产生的效价的一半。这种融合蛋白的免疫原性高得多的原因尚不清楚。有可能这种蛋白大小的增加可能对免疫效应细胞有佐剂效应。它并未刺激这种应答达到在所述其它融合蛋白中的同一程度，并且 GST 标志的任何佐剂效应并不增强抗攻击的保护，因为经切割的蛋白的保护性与其融合蛋白对应物相似。

尽管用 GST1、经切割的 1、GST1b-2、GST1b-3 和 GST1-3 免疫的组的抗 rPA 效价高，但在这些组中在 10^2 MLD 的较低攻击水平下发

生保护有一些被突破,并且与 GST 对照免疫小鼠相比,用这些蛋白质免疫并未延长那些小鼠的存活时间,这些小鼠的确死于攻击。这提示这些蛋白尚未恰当地引发所述免疫应答,以达到完全抗所述感染。正如小鼠和豚鼠的其它研究(Little S.F.等 1986. *Infect. Immun.* 52: 509-512, Turnbull P.C.B.等, 1986. *Infect. Immun.* 52: 356-363)所表明的,在抗 PA 的抗体效价与抗攻击的保护之间没有明确的关联。然而,保护需要一定阈值的抗体(Cohen S 等, 2000 *Infect. Immun.* 68: 4549-4558),提示也需要免疫应答的细胞介导组分受刺激以提供保护(Williamson 1989)。

SDS Page 和蛋白质印迹结果表明, GST1、GST1b-2 和 GST1-2 是所产生的稳定性最低的融合蛋白,可能是由于所述蛋白在无结构域 3 的情况下对降解更为敏感,这种不稳定性可能导致了保护性表位的丧失。

所述蛋白质的结构构象对于刺激保护性免疫应答也可能是重要的。与其完整的对应物 GST1-2 和 GST1-3 相比时,从所述融合蛋白去除结构域 1a,导致抗体效价降低和抵抗攻击的保护减弱。同样,仅用 GST 1 免疫的小鼠受到抵抗攻击的部分保护,但与结构域 2 联合时,如 GST1-2 融合蛋白,在 10^2 MLD 攻击水平下观测到完全的保护。然而,通过用 GST1-2 融合蛋白免疫激发的免疫应答不足以提供抗较高的 10^3 MLD 攻击水平的完全保护,这又可能是由于该蛋白降解所致的保护性表位损失引起的。

用含结构域 4 的截短物免疫的所有组都受到完全的保护,抵抗 10^3 MLD 的 STI 孢子的攻击(表 1),所述含结构域 4 的截短物包括单独的 GST 4、单独的经切割的 4、以及两种单独表达的结构域 GST 1 和 GST 4 的混合物。Brossier 等指出,用表达无结构域 4 的 PA 的炭疽芽孢杆菌突变型菌株免疫的小鼠的保护减弱(Brossier F.等 2000. *Infect. Immun.* 68: 1781-1786),并且这一点在该项研究中得以证实,在该项研究中,用 GST 1-3 免疫导致保护被突破,尽管有良好的抗体效价。这些数据表明,结构域 4 是 PA 的免疫显性亚单位。结构域 4 代表所述 PA 多肽

羧基末端的 139 个氨基酸。它含有宿主细胞受体结合区(Little S.F.等, 1996 Microbiology 142: 707-715), 被鉴定为在位于氨基酸残基 679-693 之间的小环中或小环附近(Varughese M.等 1999 Infect. Immun. 67: 1860-1865)。因此, 它对于宿主细胞中毒是必不可少的, 因为已经证明, 已表达的在结构域 4 区中含突变(Varughese 1999 出处同上)或缺失(Brossier 1999 出处同上)的 PA 形式是无毒的。PA 的晶体结构表明, 结构域 4、特别是该结构域的 19 个氨基酸的环(703-722)比相互之间紧密连接的其它 3 个结构域为暴露(Petosa 1997 出处同上)。这一结构安排可能使得结构域 4 是对于免疫效应细胞的识别最为突出的表位, 因此含结构域 4 的融合蛋白能够引发保护性最强的免疫应答。

这一研究进一步阐明了 PA 在保护性免疫应答激发中的作用, 证明抗炭疽感染的保护作用可能归因于 PA 的单个结构域。

附图说明

现在将通过实施例并参考附图, 具体描述本发明, 在附图中:

图 1 是在大肠杆菌和炭疽芽孢杆菌中发现的密码子频率的表;

图 2 显示了按照本发明的核酸的序列, 该核酸编码炭疽芽孢杆菌的 PA, 如 Welkos 等(出处同上)所发表的; 和

图 3 显示了 SEQ ID NO 3-14, 它们是用来编码 PA 的各个结构域或结构域组合的氨基酸序列和 DNA 序列, 如下文详述;

图 4 显示了 SEQ ID NO 15-16, 它们分别是 PA 结构域 4 的氨基酸序列和 DNA 序列; 和

图 5 是一个表, 显示了在初次免疫后 37 天, 得自在第 1 天和第 28 天用 10 μ g 包含 PA 片段的融合蛋白肌肉免疫的 A/J 小鼠的抗 rPA IgG 浓度; 所示结果为取自每个治疗组 5 只小鼠的样品的平均数 $\pm$ sem。

具体实施方式

实施例 1

在大肠杆菌中表达的研究

已经对 rPA 表达质粒 pAG163::rPA 进行了修饰, 以用 Km^R 标记取代原始的 Tc^R 基因。将该质粒转化到表达宿主大肠杆菌 BLR (DE3) 中, 并且评价了表达水平和溶解性。该菌株的胞内蛋白酶 La (Ion 基因产物) 和外膜蛋白酶 OmpT 有缺陷。

然而, 表达研究并未表明与 Ion+ K12 宿主菌株相比在该菌株内可溶性蛋白的积累有任何改善(即由于过度的蛋白水解, 积累受到阻碍)。因此推断, rPA 的任何胞内蛋白水解都不是 La 蛋白酶的作用所致。

实施例 2

发酵分析

使用 K12 菌株 UT5600 (DE3) pAG163::rPA 进行的发酵的进一步分析。

发现在该培养物中的 rPA 被分为可溶性部分不可溶性部分(估计有 350 mg/L 不溶性部分, 650 mg/L 全长可溶性部分)。所用的条件(37 °C, 1mM IPTG 用于诱导)在摇瓶培养物中未产生任何可检测的可溶性 rPA, 也未得到以上实施例 1 中所述的结果, 而存在大量的可溶性 rPA 是出乎意料的。然而, 看来发酵的操作、诱导和收获点可能提供在大肠杆菌 K12 表达菌株中 rPA 的稳定积累。

实施例 3

用最初作为不溶性内含体从 UT5600 (DE3) pAG163::rPA 发酵分离出的物质, 制备 rPA 样品。内含体用 25mM Tris-HCl pH 8 洗涤 2 次, 用同一缓冲液+2M 尿素洗涤 1 次。然后将它们溶于缓冲液+8M 尿素中, 沉淀出碎片。通过在 25mM Tris-HCl pH 8 中稀释并于 4°C 静置温

育过夜，而去除尿素。将经稀释的样品上样至 Q sepharose 柱，用 NaCl 梯度洗脱蛋白质。合并含有最高纯度 rPA 的流分，将其分为等分样品并冻存于-70℃。用 4-12% MES-SDS NuPAGE 凝胶，与已知标准品对比测试该样品，表明其纯度高，并且内毒素污染少。

实施例 4

所述产物的进一步表征

该产物的 N 末端测序表明，N 末端序列的组成为

MEVKQENRLL (SEQ ID NO 2)

这证明该产物正如所预期的，且留有起始密码子甲硫氨酸。

发现该物质在蛋白质印迹中有反应；该样品的 MALDI -MS 表明，质量大约为 82 700 (与预期的质量为 82 915 相似)。已知分子量高并且与所用的质量标准(66KDa)有一定的距离，则认为这表明该物质没有显著地截短，而并未消除该样品内的微不均一性。

实施例 5

PA 各个结构域的测试

使用 Pharmacia pGEX-6P-3 表达系统, 以与载体蛋白谷胱甘肽-s 转移酶(GST)的融合蛋白在大肠杆菌中产生作为重组蛋白的 PA 的各个结构域。各个结构域的序列以及用来编码所述结构域的 DNA 序列如图 3 进行连接。相应的氨基酸序列和 DNA 序列在以下表 2 中提供。

用这些融合蛋白, 以总共 100 μ l 体积、吸附于 20% v/v 铝胶的 10 μ g 相应融合蛋白, 肌肉免疫 A/J 小鼠(Harlan Olac)。

动物经两次免疫, 通过用指定剂量水平的炭疽芽孢杆菌(STI 菌株)孢子攻击, 测定其保护性免疫的产生。下表显示了攻击后 14 天的存活者。

孢子的攻击水平/小鼠

结构域	氨基酸 SEQ ID NO	DNA SEQ ID NO	5x10 ⁴	9x10 ⁴	9x10 ⁵	1x10 ⁶	5x10 ⁶
GST-1	3	4	4/4	3/5			
GST-1+2	5	6	4/4; 5/5	4/5; 5/5			
GST-1b+2	7	8	2/5	1/5			
GST-1b+2+3	9	10	2/5	3/5			
GST-1+2+3	11	12	Nd	4/5	3/5		
GST-1+2+3+4	13	14	Nd	5/5	5/5		
1+2+3+4	13	14	Nd	Nd		5/5	5/5

所述数据表明, PA 所有 4 个结构域的组合无论是否以与 GST 的融合蛋白存在, 在直至高攻击水平下都是保护性的。去除结构域 4, 留下 1+2+3, 导致在所测试的最高攻击水平 9x10⁵下被突破。在 9x10⁴孢子下, 结构域 1+2 的保护性与结构域 1+2+3 的组合一样。然而, 去除结构域 1a 而留下 GST 与结构域 1b+2 的融合体, 导致在所测试最高

攻击水平(9×10^4)下保护被突破, 加入结构域 3 仅有略微改善。

该数据表明, PA 诱导的保护性免疫可以归因于单个结构域(完整的结构域 1 和结构域 4)或者得自所有 4 个结构域作为全变物(permutation)的结构域组合。

结构域 4 的氨基酸序列和 DNA 编码序列示于图 4, 分别为 SEQ ID NO 15 和 SEQ ID NO 16。

实施例 6

结构域作为疫苗的进一步测试

从炭疽芽孢杆菌 Sterne DNA, 通过 PCR 扩增编码所述 PA 结构域一氨基酸 1-259、168-488、1-488、168-596、1-596、260-735、489-735、597-735 和 1-735 (分别为截短物 GST1、GST1b-2、GST1-2、GST1b-3、GST1-3、GST2-4、GST3-4、GST4 和 GST1-4)的 DNA, 将其克隆到表达载体 pGEX-6-P3 (Amersham-Pharmacia)的 *XhoI/BamHI* 位点中, 置于 *lac* 启动子下游并且与所述启动子符合读框。使用该系统产生的蛋白以具有 N 末端谷胱甘肽-s-转移酶蛋白(GST)的融合蛋白表示。然后, 将带有编码所述 PA 结构域的 DNA 的重组质粒 DNA 转化到大肠杆菌 BL21 中, 以供蛋白质表达研究用。

在含有 50 μ g/ml 氨苄青霉素、30 μ g/ml 氯霉素和 1% w/v 葡萄糖的 L-肉汤中, 培养带有重组 pGEX-6-P3 质粒的大肠杆菌 BL21。将培养物于 30 $^{\circ}$ C 振荡(170 rev/min)培养至 A_{600nm} 为 0.4, 然后用 0.5mM IPTG 诱导。将培养物再孵育 4 小时, 然后通过以 10 000 rpm 离心 15 分钟收获培养物。

所述 PA 截短物-融合蛋白的初步提取表明, 它们以内含体产生。将细胞沉淀重悬于磷酸缓冲盐溶液(PBS)中, 在冰水浴中超声处理 4 $\times$ 20 秒。将悬浮液以 15 000 rpm 离心 15 分钟, 然后通过悬浮于 8M 尿素中于室温搅拌 1 小时, 用尿素提取细胞沉淀。将悬浮液以 15000 rpm 离心 15 分钟, 上清液对含有 400mM L-精氨酸和 0.1mM EDTA 的 100mM

Tris pH 8 透析, 然后在 PBS 中透析。

所述 PA 截短物-融合蛋白的成功重折叠, 使其可以在谷胱甘肽 Sepharose CL-4B 亲和柱上纯化。将所有提取物(截短物 GST1b-2-氨基酸残基 168-487 除外)上样到先前用 PBS 平衡的 15 ml 谷胱甘肽 Sepharose CL-4B 柱(Amersham-Parmacia), 于 4℃ 旋转温育过夜。柱子用 PBS 洗涤, 用含 150mM NaCl、1mM EDTA 和 20mM 还原型谷胱甘肽的 50mM Tris pH 7 洗脱所述融合蛋白。合并通过 SDS-PAGE 分析鉴定的含 PA 截短物的流分, 然后对 PBS 透析。用 BCA (Perbio)测定蛋白浓度。

然而, 用还原型谷胱甘肽未能从谷胱甘肽 sepharose CL-4B 亲和柱洗脱下载短物 GST1b-2, 因此用离子交换色谱纯化。具体地说, 将截短物 GST1b-2 对 20mM Tris pH 8 透析, 然后上样至用同一缓冲液平衡的 HiTrap Q 柱(Amersham-Parmacia)。用 20mM Tris pH 8 中的 0-1M 增加的 NaCl 梯度洗脱融合蛋白。合并含 GST-蛋白的流分, 浓缩, 然后上样至先前用 PBS 平衡的 HiLoad 26/60 Superdex 200 凝胶过滤柱(Amersham-Parmacia)。合并含融合蛋白的流分, 通过 BCA (Perbio)测定蛋白浓度。产量在每升培养物 1 mg 和 43 mg 之间。

所述片段的分子量及其被抗 PA 抗体的识别用 SDS PAGE 和蛋白质印迹法来证实。通过 SDS Page 和蛋白质印迹法分析所述 rPA 截短物, 显示出预期大小的蛋白条带。在所调查的所有所述 rPA 截短物中显然有一些降解, 显示出与在枯草杆菌中表达的重组 PA 有相似性。rPA 截短物 GST1、GST1b-2 和 GST1-2 在无结构域 3 的情况下对降解特别敏感。同样报道了结构域 3 中含有突变的 rPA 构建体不能从炭疽芽孢杆菌培养上清液中纯化(Brossier 1999), 表明结构域 3 可能使结构域 1 和结构域 2 稳定化。

在该项研究中使用雌性无特定病原体的 A/J 小鼠(Harlan UK), 因为它们炭疽感染的稳定模型(Welkos 1986)。小鼠在年龄上相当, 在该项研究开始时为 7 周龄。

在该项研究的第 1 天和第 28 天, 用 100 μ l 总体积 PBS 中吸附于 20% 1.3% v/v 铝胶(HCI Biosector, Denmark)的 10 μ g 融合蛋白为 A/J 小鼠免疫。也包括用得自枯草杆菌的 rPA (Miller 1998)、重组 GST 对照蛋白或去除了 GST 标志的编码结构域 1、4 以及 1-4 的融合蛋白免疫的组。在后腿两个部位肌内给予免疫剂量。初次免疫后 37 天, 取小鼠的血样, 以供通过酶联免疫吸附测定(ELISA)进行血清抗体分析。

微量滴定板(Immulon 2, Dynex Technologies)用 PBS 中从枯草杆菌表达的(Miller 1998) 5 μ g/ml rPA 于 4 $^{\circ}$ C 包被过夜, 每块板的两列除外, 这两列用 5 μ g/ml 抗小鼠 Fab (Sigma, Poole, Dorset)包被。用含 1% v/v Tween 20 的 PBS (PBS-T)洗涤板, 5% w/v 脱脂奶粉的 PBS 溶液(blotto)于 37 $^{\circ}$ C 封闭 2 小时。加入在 1% blotto 中两倍稀释的血清至所述 rPA 包被孔中, 与加入至抗 fab 包被孔的鼠 IgG 标准(Sigma)一起, 以双份重复进行分析, 于 4 $^{\circ}$ C 温育过夜。洗涤后, 向所有孔中加入在 PBS 以 1:2000 稀释的辣根过氧化物酶缀合的山羊抗小鼠 IgG (Southern Biotechnology Associates Inc.), 于 37 $^{\circ}$ C 温育 1 小时。再次洗涤板, 然后加入底物 2,2'-Azinobis (3-乙基苯并噻唑啉磺酸) (1.09mM ABTS, Sigma)。于室温温育 20 分钟后, 测定 414 nm 下各孔的吸光度(Titertek Multiscan, ICN Flow)。用 Tittersoft version 3.1c 软件计算标准曲线。效价以每 ml 血清 μ g IgG 表示, 计算组平均数 $\pm$ 平均标准误差(sem)。结果示于图 5。

所产生的所有 rPA 截短物都是免疫原性的, 在 A/J 小鼠体内激发的平均血清抗 rPA IgG 浓度范围从 GST1b-2 截短物免疫组的 6 μ g/ml 至 GST 1-4 截短物免疫组的 1488 μ g/ml (图 5)。GST 对照免疫小鼠没有可检测的抗 rPA 抗体。

在免疫方案的第 70 天, 用炭疽芽孢杆菌 STI 孢子攻击小鼠。从原液取出足够的 STI 孢子用于攻击, 在无菌蒸馏水中洗涤, 然后重悬于 PBS 中至 1×10^7 孢子/ml 和 1×10^6 孢子/ml 的浓度。分别以每只小鼠含 1×10^6 孢子和 1×10^5 孢子的 0.1 ml 体积腹膜内攻击小鼠, 在攻击后监测

14 天，以确定其保护状况。严格观测人道终点，使得表现出综合指示具有致死感染的大量临床体征的小鼠被剔除。攻击后存活 14 天的免疫小鼠数目示于表 3。

表 3
攻击水平 MLD

结构域	存活者/受攻击的数目(%)	
	10^2 MLD	10^3 MLD
GST 1	3/5 (60)	1/5 (20)
GST 1b-2	1/5 (20)	nd
GST 1-2	5/5 (100)	3/5 (60)
GST 1b-3	3/5 (60)	nd
GST 1-3	4/5 (80)	nd
GST 1-4	nd	5/5 (100)
GST 2-4	nd	5/5 (100)
GST 3-4	nd	5/5 (100)
GST 4	5/5 (100)	5/5 (100)
GST 1+ GST 4	nd	5/5 (100)
经切割的 1	1/5 (20)	2/5
经切割的 4	5/5 (100)	5/5
经切割的 1-4	nd	5/5
rPA	nd	4/4 (100)
对照	0/5 (0)	0/5 (0)

1 MLD = 约 1×10^3 STI 孢子

nd = 未进行

用 10^3 MLD 的 STI 孢子攻击的组都是完全受保护的，除了 GST1、GST1-2 和经切割的 1 免疫的组，在这些组中有一些保护被突破，仅用 GST 免疫的对照组全都死于感染，至死亡时的平均时间

(MTTD)为 2.4 ± 0.2 天。在 10^2 MLD 的较低攻击水平下, GST1-2、GST4 和经切割的 4 免疫的组都受到完全的保护, 但在其它组中有一些保护被突破。在这些组中死亡小鼠的 MTTD 为 4.5 ± 0.2 天, 这与 GST 对照免疫组没有显著差异, GST 对照免疫组全部死亡, MTTD 为 4 ± 0.4 天。

序 列 表

<110> 英国国防部(The Secretary of State for Defence)

Williamson, Ethel D

Miller, Julie

Walker, Nicola J

Baillie, Leslie WJ

Holden, Paula T

Flick-Smith, Helen C

Bullifent, Helen L

Titball, Richard W

<120> 表达系统

<130> CG/P/110/WOD

<140> PCT/GB01/03065

<141> 2001-07-06

<150> GB 0016702.3

<151> 2000-07-08

<160> 16

<170> PatentIn Ver. 2.1

<210> 1

<211> 2228

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 编码炭疽芽孢杆菌保护性抗原的核酸

<400> 1

```
aagcttcata tggaagtaaa gcaagagaac cgtctgctga acgaatctga atccagctct 60
cagggcctgc ttggttacta tttctctgac ctgaacttcc aagcacgat ggttgtaacc 120
agctctacca ctggcgatct gtccatcccg tctagtgaac ttgagaacat tccaagegag 180
aaccagtatt tccagtctgc aatctggtcc ggttttatca aagtcaagaa atctgatgaa 240
tacacgtttg ccacctctgc tgataaccac gtaacctgt gggttgacga tcaggaagtg 300
atcaacaaag cateccaatc caacaaaatt cgtctggaaa aaggcogtct gtatcagatc 360
```

```

aagattcagt accaacgcga gaacccgact gaaaaaggcc tggactttaa actgtattgg 420
actgattctc agaacaagaa agaagtgate agctctgaca atctgcaact gccggaattg 480
aaacagaaaa gctccaactc tcgtaagaaa cgttccacca gcgctggccc gaccgtacca 540
gatcgcgaca acgatgggtat tccggactct ctggaagttg aaggctacac ggttgatgta 600
aagaacaaac gtaccttcct tagtccgtgg atctccaata ttcacgagaa gaaaggtctg 660
accaaataca aatccagtcc ggaaaaatgg tccactgcat ctgatccgta ctctgacttt 720
gagaaagtga ccggtcgtat cgacaagaac gtctctccgg aagcacgcca tccactggtt 780
gctgcgtatc cgatcgtaca tgttgacatg gaaaacatca ttttgtccaa gaacgaagac 840
cagtccactc agaacactga ctctgaaact cgtaccatct ccaagaacac ctccacgtct 900
cgtactcaca ccagtgaagt acatggtaac gctgaagtac acgcctcttt ctttgacatc 960
ggcggctctg ttagcgtctg cttctccaac tctaattctt ctactgttgc cattgatcac 1020
tctctgagtc tggctggcga acgtacctgg gcagagacca tgggtcttaa cactgctgat 1080
accgcgcgtc tgaatgctaa cattcgtac gtcaacactg gtacggcacc gatctacaac 1140
gtactgccaa ccaccagcct ggttctgggt aagaaccaga ctcttgcgac catcaaagcc 1200
aaagagaacc aactgtctca gattctggca ccgaataact actatccttc caagaacctg 1260
gctccgatcg cactgaacgc acaggatgac ttctcttcca ctccgatcac catgaactac 1320
aaccagttcc tggaacttga gaagacaaa cagctgcgtc ttgacactga ccaagtgtac 1380
ggtaacatcg cgacctaaa ctttgagaac ggctcgcgtc gcgttgacac aggcctctaat 1440
tggctctgaag tactgcctca gattcaggaa accaccgctc gtatcatctt caacggtaaa 1500
gacctgaacc tgggtgaacg tcgtattgct gctgtgaacc cgtctgatcc attagagacc 1560
accaaaccgg atatgactct gaaagaagcc ctgaagatcg cctttggctt caacgagccg 1620
aacggtaatc ttcagtacca aggtaaagac atcactgaat ttgacttcaa ctttgatcag 1680
cagacctctc agaatatcaa gaaccaactg gctgagctga acgcgaccaa tatctatacg 1740
gtactcgaca agatcaaact gaacgcgaaa atgaacattc tgattcgcga caaacgtttc 1800
cactacgate gtaataacat cgctgttggc gctgatgaat ctgttgtgaa agaagcgcac 1860
cgcgaaagtca tcaactccag caccgaaggc ctgcttctga acatcgacaa agacattcgt 1920
aagatcctgt ctggttacat tgttgagatc gaagacaccg aaggcctgaa agaagtgatc 1980
aatgatcgtt acgacatgct gaacatcagc tctctgcgtc aagatggtaa gacgttcatt 2040
gacttcaaga aatacaacga caaacttcgc ctgtatatct ctaatccgaa ctacaaagtg 2100
aacgtttacg ctgttaccaa agagaacacc atcatcaate catctgagaa cggcgatacc 2160
tctaccaacg gtatcaagaa gattctgacg ttctccaaga aaggttacga gatcggttaa 2220
taggatcc
2228

```

<210> 2

<211> 10

<212> PRT

<213> 炭疽芽孢杆菌(Bacillus anthracis)

<400> 2

Met Glu Val Lys Gln Glu Asn Arg Leu Leu

1

5

10

<210> 3

<211> 258

<212> PRT

<213> 炭疽芽孢杆菌(Bacillus anthracis)

<400> 3

Glu Val Lys Gln Glu Asn Arg Leu Leu Asn Glu Ser Glu Ser Ser Ser
 1 5 10 15

Gln Gly Leu Leu Gly Tyr Tyr Phe Ser Asp Leu Asn Phe Gln Ala Pro
 20 25 30

Met Val Val Thr Ser Ser Thr Thr Gly Asp Leu Ser Ile Pro Ser Ser
 35 40 45

Glu Leu Glu Asn Ile Pro Ser Glu Asn Gln Tyr Phe Gln Ser Ala Ile
 50 55 60

Trp Ser Gly Phe Ile Lys Val Lys Lys Ser Asp Glu Tyr Thr Phe Ala
 65 70 75 80

Thr Ser Ala Asp Asn His Val Thr Met Trp Val Asp Asp Gln Glu Val
 85 90 95

Ile Asn Lys Ala Ser Asn Ser Asn Lys Ile Arg Leu Glu Lys Gly Arg
 100 105 110

Leu Tyr Gln Ile Lys Ile Gln Tyr Gln Arg Glu Asn Pro Thr Glu Lys
 115 120 125

Gly Leu Asp Phe Lys Leu Tyr Trp Thr Asp Ser Gln Asn Lys Lys Glu
 130 135 140

Val Ile Ser Ser Asp Asn Leu Gln Leu Pro Glu Leu Lys Gln Lys Ser
 145 150 155 160

Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro
 165 170 175

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

Thr Val Asp Val Lys Asn Lys Arg Thr Phe Leu Ser Pro Trp Ile Ser
195 200 205

Asn Ile His Glu Lys Lys Gly Leu Thr Lys Tyr Lys Ser Ser Pro Glu
210 215 220

Lys Trp Ser Thr Ala Ser Asp Pro Tyr Ser Asp Phe Glu Lys Val Thr
225 230 235 240

Gly Arg Ile Asp Lys Asn Val Ser Pro Glu Ala Arg His Pro Leu Val
245 250 255

Ala Ala

<210> 4

<211> 774

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：用来编码 SEQ ID NO: 3 的 DNA 序列

<400> 4

```
gaagttaaac aggagaaccg gttattaaat gaatcagaat caagttccca ggggttacta 60
ggatactatt ttagtgattt gaattttcaa gcacccatgg tggttacctc ttctactaca 120
ggggatttat ctattcctag ttctgagtta gaaaatattc catcggaata ccaatatatt 180
caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct 240
acttccgctg ataatacatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct 300
tctaattcta acaaaatcag attagaaaaa ggaagattat atcaaataaa aattcaatat 360
caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgatttctca 420
aataaaaaag aagtgatttc tagtgataac ttacaattgc cagaattaaa acaaaaatct 480
tcgaactcaa gaaaaaagcg aagtacaagt gctggacctc cggttcaga cegtgacaat 540
gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga 600
acttttcttt caccatggat ttctaattat catgaaaaga aaggattaac caaatataaa 660
tcattctctg aaaaatggag cacggttctt gatccgtaca gtgatttcga aaaggttaca 720
ggacggattg ataagaatgt atcaccagag gcaagacacc ccctgtggc agct 774
```

<210> 5

<211> 487

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：融合蛋白

<400> 5

Glu Val Lys Gln Glu Asn Arg Leu Leu Asn Glu Ser Glu Ser Ser Ser
1 5 10 15

Gln Gly Leu Leu Gly Tyr Tyr Phe Ser Asp Leu Asn Phe Gln Ala Pro
20 25 30

Met Val Val Thr Ser Ser Thr Thr Gly Asp Leu Ser Ile Pro Ser Ser
35 40 45

Glu Leu Glu Asn Ile Pro Ser Glu Asn Gln Tyr Phe Gln Ser Ala Ile
50 55 60

Trp Ser Gly Phe Ile Lys Val Lys Lys Ser Asp Glu Tyr Thr Phe Ala
65 70 75 80

Thr Ser Ala Asp Asn His Val Thr Met Trp Val Asp Asp Gln Glu Val
85 90 95

Ile Asn Lys Ala Ser Asn Ser Asn Lys Ile Arg Leu Glu Lys Gly Arg
100 105 110

Leu Tyr Gln Ile Lys Ile Gln Tyr Gln Arg Glu Asn Pro Thr Glu Lys
115 120 125

Gly Leu Asp Phe Lys Leu Tyr Trp Thr Asp Ser Gln Asn Lys Lys Glu
130 135 140

Val Ile Ser Ser Asp Asn Leu Gln Leu Pro Glu Leu Lys Gln Lys Ser
145 150 155 160

Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro
165 170 175

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

Thr Val Asp Val Lys Asn Lys Arg Thr Phe Leu Ser Pro Trp Ile Ser
195 200 205

Asn Ile His Glu Lys Lys Gly Leu Thr Lys Tyr Lys Ser Ser Pro Glu
210 215 220

Lys Trp Ser Thr Ala Ser Asp Pro Tyr Ser Asp Phe Glu Lys Val Thr
225 230 235 240

Gly Arg Ile Asp Lys Asn Val Ser Pro Glu Ala Arg His Pro Leu Val
245 250 255

Ala Ala Tyr Pro Ile Val His Val Asp Met Glu Asn Ile Ile Leu Ser
260 265 270

Lys Asn Glu Asp Gln Ser Thr Gln Asn Thr Asp Ser Glu Thr Arg Thr
275 280 285

Ile Ser Lys Asn Thr Ser Thr Ser Arg Thr His Thr Ser Glu Val His
290 295 300

Gly Asn Ala Glu Val His Ala Ser Phe Phe Asp Ile Gly Gly Ser Val
305 310 315 320

Ser Ala Gly Phe Ser Asn Ser Asn Ser Ser Thr Val Ala Ile Asp His
325 330 335

Ser Leu Ser Leu Ala Gly Glu Arg Thr Trp Ala Glu Thr Met Gly Leu
340 345 350

Asn Thr Ala Asp Thr Ala Arg Leu Asn Ala Asn Ile Arg Tyr Val Asn
355 360 365

Thr Gly Thr Ala Pro Ile Tyr Asn Val Leu Pro Thr Thr Ser Leu Val
370 375 380

Leu Gly Lys Asn Gln Thr Leu Ala Thr Ile Lys Ala Lys Glu Asn Gln
385 390 395 400

Leu Ser Gln Ile Leu Ala Pro Asn Asn Tyr Tyr Pro Ser Lys Asn Leu
 405 410 415

Ala Pro Ile Ala Leu Asn Ala Gln Asp Asp Phe Ser Ser Thr Pro Ile
 420 425 430

Thr Met Asn Tyr Asn Gln Phe Leu Glu Leu Glu Lys Thr Lys Gln Leu
 435 440 445

Arg Leu Asp Thr Asp Gln Val Tyr Gly Asn Ile Ala Thr Tyr Asn Phe
 450 455 460

Glu Asn Gly Arg Val Arg Val Asp Thr Gly Ser Asn Trp Ser Glu Val
 465 470 475 480

Leu Pro Gln Ile Gln Glu Thr
 485

<210> 6

<211> 1461

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：用来编码 SEQ ID NO: 5 的 DNA 序列

<400> 6

gaagttaaac aggagaaccg gttattaaat gaatcagaat caagttccca ggggttacta 60
 ggatactatt ttagtgattt gaattttcaa gcacccatgg tggttacttc ttctactaca 120
 ggggatttat ctattcctag ttctgagtta gaaaatattc catcgaaaa ccaatatttt 180
 caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct 240
 acttccgctg ataatcatgt aacaatgtgg gtagatgacc agaagtgat taataaagct 300
 tctaattcta acaaaatcag attagaaaaa ggaagattat atcaaataaa aattcaatat 360
 caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgatttcaa 420
 aataaaaaag aagtgatttc tagtgataac ttacaactgc cagaattaaa acaaaaatct 480
 tcgaactcaa gaaaaaagcg aagtacaagt gctggaccta cggttccaga ccgtgacaat 540
 gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga 600
 acttttcttt caccatggat ttctaattt catgaaaaga aaggattaac caaatataaa 660
 tcattctcctg aaaaatggag cacggcttct gatccgtaca gtgatttcga aaaggttaca 720
 ggacggattg ataagaatgt atcaccagag gcaagacacc cccttggtgc agcttatccg 780

```

attgtacatg tagatatgga gaatattatt ctctcaaaaa atgaggatca atccacacag 840
aatactgata gtgaaacgag aacaataagt aaaaatactt ctacaagtag gacacatact 900
agtgaagtac atggaaatgc agaagtgcac ggcgtgttct ttgatattgg tgggagtgtg 960
tctgcaggat ttagtaattc gaattcaagt acggctgcaa ttgatcattc actatctcta 1020
gcaggggaaa gaacttgggc tgaaacaatg ggtttaaata ccgctgatac agcaagatta 1080
aatgccataa ttagatatgt aaatactggg acggctccaa tctacaacgt gttaccaacg 1140
acttcgttag tgttaggaaa aaatcaaaca ctgcgcacaa ttaaagctaa ggaaaaccaa 1200
ttaagtcaaa tacttgcacc taataattat tacccttcta aaaacttggc gccaatcgca 1260
ttaaatgcac aagacgattt cagttctact ccaattacaa tgaattacaa tcaatttctt 1320
gagttagaaa aaacgaaaca attaagatta gatacggatc aagtatatgg gaatatagca 1380
acatacaatt ttgaaaatgg aagagtgagg gtggatacag gctcgaactg gagtgaagtg 1440
ttaccgcaaa ttcaagaaac a 1461

```

<210> 7

<211> 318

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：融合蛋白

<400> 7

```

Ser Ala Gly Pro Thr Val Pro Asp Arg Asp Asn Asp Gly Ile Pro Asp
  1             5             10            15

```

```

Ser Leu Glu Val Glu Gly Tyr Thr Val Asp Val Lys Asn Lys Arg Thr
      20             25            30

```

```

Phe Leu Ser Pro Trp Ile Ser Asn Ile His Glu Lys Lys Gly Leu Thr
      35             40            45

```

```

Lys Tyr Lys Ser Ser Pro Glu Lys Trp Ser Thr Ala Ser Asp Pro Tyr
      50             55            60

```

```

Ser Asp Phe Glu Lys Val Thr Gly Arg Ile Asp Lys Asn Val Ser Pro
      65             70            75            80

```

```

Glu Ala Arg His Pro Leu Val Ala Ala Tyr Pro Ile Val His Val Asp
      85             90            95

```

Met Glu Asn Ile Ile Leu Ser Lys Asn Glu Asp Gln Ser Thr Gln Asn
 100 105 110

Thr Asp Ser Glu Thr Arg Thr Ile Ser Lys Asn Thr Ser Thr Ser Arg
 115 120 125

Thr His Thr Ser Glu Val His Gly Asn Ala Glu Val His Ala Ser Phe
 130 135 140

Phe Asp Ile Gly Gly Ser Val Ser Ala Gly Phe Ser Asn Ser Asn Ser
 145 150 155 160

Ser Thr Val Ala Ile Asp His Ser Leu Ser Leu Ala Gly Glu Arg Thr
 165 170 175

Trp Ala Glu Thr Met Gly Leu Asn Thr Ala Asp Thr Ala Arg Leu Asn
 180 185 190

Ala Asn Ile Arg Tyr Val Asn Thr Gly Thr Ala Pro Ile Tyr Asn Val
 195 200 205

Leu Pro Thr Thr Ser Leu Val Leu Gly Lys Asn Gln Thr Leu Ala Thr
 210 215 220

Ile Lys Ala Lys Glu Asn Gln Leu Ser Gln Ile Leu Ala Pro Asn Asn
 225 230 235 240

Tyr Tyr Pro Ser Lys Asn Leu Ala Pro Ile Ala Leu Asn Ala Gln Asp
 245 250 255

Asp Phe Ser Ser Thr Pro Ile Thr Met Asn Tyr Asn Gln Phe Leu Glu
 260 265 270

Leu Glu Lys Thr Lys Gln Leu Arg Leu Asp Thr Asp Gln Val Tyr Gly
 275 280 285

Asn Ile Ala Thr Tyr Asn Phe Glu Asn Gly Arg Val Arg Val Asp Thr
 290 295 300

Gly Ser Asn Trp Ser Glu Val Leu Pro Gln Ile Gln Glu Thr
 305 310 315

<210> 8

<211> 954

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：用来编码 SEQ ID NO: 7 的 DNA 序列

<400> 8

```

agtgctggac ctacggttcc agaccgtgac aatgatggaa tccctgattc attagaggta 60
gaaggatata cggttgatgt caaaaataaa agaacttttc tttcaccatg gatttctaata 120
attcatgaaa agaaaggatt aaccaaatat aaatcatctc ctgaaaaatg gagcacggct 180
tctgatccgt acagtgtatt cgaaaagggt acaggacgga ttgataagaa tgtatcacca 240
gaggcaagac acccccttgt ggcagcttat ccgattgtac atgtagatat ggagaatatt 300
attctctcaa aaaatgagga tcaatccaca cagaatactg atagtgaac gagaacaata 360
agtaaaaata cttctacaag taggacacat actagtgaag tacatggaaa tgcagaagt 420
catgcgtcgt tctttgatat tgggtgggagt gtatctgcag gatttagtaa ttcgaattca 480
agtacggtcg caattgatca ttcactatct ctacgagggg aaagaacttg ggctgaaaca 540
atgggtttta ataccgctga tacagcaaga ttaaatgcca atattagata tgtaataact 600
gggacggctc caatctacaa cgtgttacca acgacttcgt tagtgtagg aaaaaatcaa 660
acactcgcga caattaaagc taaggaaaac caattaagtc aaatacttgc acctaatat 720
tattatcett ctaaaaactt ggcgccaatc gcattaaatg cacaagacga tttcagttct 780
actccaatta caatgaatta caatcaattt cttgagttag aaaaaacgaa acaattaaga 840
ttagatacgg atcaagtata tgggaatata gcaacataca attttgaaaa tggaagagtg 900
agggtggata caggctcgaa ctggagtga gtgttaccgc aaattcaaga aaca 954

```

<210> 9

<211> 426

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：融合蛋白

<400> 9

```

Ser Ala Gly Pro Thr Val Pro Asp Arg Asp Asn Asp Gly Ile Pro Asp
 1           5           10          15

Ser Leu Glu Val Glu Gly Tyr Thr Val Asp Val Lys Asn Lys Arg Thr
          20           25           30

```

Phe Leu Ser Pro Trp Ile Ser Asn Ile His Glu Lys Lys Gly Leu Thr
35 40 45

Lys Tyr Lys Ser Ser Pro Glu Lys Trp Ser Thr Ala Ser Asp Pro Tyr
50 55 60

Ser Asp Phe Glu Lys Val Thr Gly Arg Ile Asp Lys Asn Val Ser Pro
65 70 75 80

Glu Ala Arg His Pro Leu Val Ala Ala Tyr Pro Ile Val His Val Asp
85 90 95

Met Glu Asn Ile Ile Leu Ser Lys Asn Glu Asp Gln Ser Thr Gln Asn
100 105 110

Thr Asp Ser Glu Thr Arg Thr Ile Ser Lys Asn Thr Ser Thr Ser Arg
115 120 125

Thr His Thr Ser Glu Val His Gly Asn Ala Glu Val His Ala Ser Phe
130 135 140

Phe Asp Ile Gly Gly Ser Val Ser Ala Gly Phe Ser Asn Ser Asn Ser
145 150 155 160

Ser Thr Val Ala Ile Asp His Ser Leu Ser Leu Ala Gly Glu Arg Thr
165 170 175

Trp Ala Glu Thr Met Gly Leu Asn Thr Ala Asp Thr Ala Arg Leu Asn
180 185 190

Ala Asn Ile Arg Tyr Val Asn Thr Gly Thr Ala Pro Ile Tyr Asn Val
195 200 205

Leu Pro Thr Thr Ser Leu Val Leu Gly Lys Asn Gln Thr Leu Ala Thr
210 215 220

Ile Lys Ala Lys Glu Asn Gln Leu Ser Gln Ile Leu Ala Pro Asn Asn
225 230 235 240

Tyr Tyr Pro Ser Lys Asn Leu Ala Pro Ile Ala Leu Asn Ala Gln Asp
245 250 255

Asp Phe Ser Ser Thr Pro Ile Thr Met Asn Tyr Asn Gln Phe Leu Glu
260 265 270

Leu Glu Lys Thr Lys Gln Leu Arg Leu Asp Thr Asp Gln Val Tyr Gly
275 280 285

Asn Ile Ala Thr Tyr Asn Phe Glu Asn Gly Arg Val Arg Val Asp Thr
290 295 300

Gly Ser Asn Trp Ser Glu Val Leu Pro Gln Ile Gln Glu Thr Thr Ala
305 310 315 320

Arg Ile Ile Phe Asn Gly Lys Asp Leu Asn Leu Val Glu Arg Arg Ile
325 330 335

Ala Ala Val Asn Pro Ser Asp Pro Leu Glu Thr Thr Lys Pro Asp Met
340 345 350

Thr Leu Lys Glu Ala Leu Lys Ile Ala Phe Gly Phe Asn Glu Pro Asn
355 360 365

Gly Asn Leu Gln Tyr Gln Gly Lys Asp Ile Thr Glu Phe Asp Phe Asn
370 375 380

Phe Asp Gln Gln Thr Ser Gln Asn Ile Lys Asn Gln Leu Ala Glu Leu
385 390 395 400

Asn Ala Thr Asn Ile Tyr Thr Val Leu Asp Lys Ile Lys Leu Asn Ala
405 410 415

Lys Met Asn Ile Leu Ile Arg Asp Lys Arg
420 425

<210> 10

<211> 1278

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：用来编码 SEQ ID NO: 9 的 DNA 序列

<400> 10

```

agtgcgtggac ctacggttcc agaccgtgac aatgatggaa tccctgattc attagaggta 60
gaaggatata cggttgatgt caaaaataaa agaacttttc tttcaccatg gatttctaata 120
attcatgaaa agaaaggatt aaccaaatat aaatcatctc ctgaaaaatg gagcacggct 180
tctgatccgt acagtgattt cgaagagggt acaggacgga ttgataagaa tgtatcacca 240
gaggcaagac acccccttgt ggcagcttat ccgattgtac atgtagatat ggagaatatt 300
attctctcaa aaaatgagga tcaatccaca cagaatactg atagtgaac gagaacaata 360
agtaaaaata cttctacaag taggacacat actagtgaag tacatggaaa tgcagaagt 420
catgcgtcgt tctttgatat tgggtgggagt gtatctgcag gatttagtaa ttcgaattca 480
agtaacggctg caattgatca ttcactatct ctacgagggg aaagaacttg ggctgaaaca 540
atgggtttta ataccgctga tacagcaaga ttaaatgcca atattagata tgtaataact 600
gggacggctc caatctacaa cgtgttacca acgacttcgt tagtgtagg aaaaaatcaa 660
acactcgca caattaaagc taaggaaaac caattaagtc aaatacttgc acctaataat 720
tattatcctt ctaaaaactt ggccccaatc gcattaaatg cacaagacga ttcagttct 780
actccaatta caatgaatta caatcaattt cttgagttag aaaaaacgaa acaattaaga 840
ttagatacgg atcaagtata tgggaatata gcaacataca attttgaaaa tggaagagt 900
aggggtggata caggtcgaa ctggagtga gtgttaccgc aaattcaaga aacaactgca 960
cgtatcattt ttaatggaaa agatttaaata ctggtagaaa ggcggatagc ggoggttaat 1020
cctagtgate cattagaaac gactaaaccg gatatgacat taaaagaagc ccttaaaata 1080
gcatttgat ttaacgaacc gaatggaaac ttacaatata aagggaaga cataaccgaa 1140
tttgatttta atttcgatca acaaacatct caaaatatca agaatcagtt agcggaatta 1200
aacgcaacta acatatatac tgtattagat aaaatcaaat taaatgcaaa aatgaatatt 1260
ttaataagag ataaacgt

```

1278

<210> 11

<211> 595

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：融合蛋白

<400> 11

Glu Val Lys Gln Glu Asn Arg Leu Leu Asn Glu Ser Glu Ser Ser Ser

1

5

10

15

Gln Gly Leu Leu Gly Tyr Tyr Phe Ser Asp Leu Asn Phe Gln Ala Pro

20

25

30

Met Val Val Thr Ser Ser Thr Thr Gly Asp Leu Ser Ile Pro Ser Ser
 35 40 45

Glu Leu Glu Asn Ile Pro Ser Glu Asn Gln Tyr Phe Gln Ser Ala Ile
 50 55 60

Trp Ser Gly Phe Ile Lys Val Lys Lys Ser Asp Glu Tyr Thr Phe Ala
 65 70 75 80

Thr Ser Ala Asp Asn His Val Thr Met Trp Val Asp Asp Gln Glu Val
 85 90 95

Ile Asn Lys Ala Ser Asn Ser Asn Lys Ile Arg Leu Glu Lys Gly Arg
 100 105 110

Leu Tyr Gln Ile Lys Ile Gln Tyr Gln Arg Glu Asn Pro Thr Glu Lys
 115 120 125

Gly Leu Asp Phe Lys Leu Tyr Trp Thr Asp Ser Gln Asn Lys Lys Glu
 130 135 140

Val Ile Ser Ser Asp Asn Leu Gln Leu Pro Glu Leu Lys Gln Lys Ser
 145 150 155 160

Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro
 165 170 175

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

Thr Val Asp Val Lys Asn Lys Arg Thr Phe Leu Ser Pro Trp Ile Ser
 195 200 205

Asn Ile His Glu Lys Lys Gly Leu Thr Lys Tyr Lys Ser Ser Pro Glu
 210 215 220

Lys Trp Ser Thr Ala Ser Asp Pro Tyr Ser Asp Phe Glu Lys Val Thr
 225 230 235 240

Gly Arg Ile Asp Lys Asn Val Ser Pro Glu Ala Arg His Pro Leu Val
 245 250 255

Ala Ala Tyr Pro Ile Val His Val Asp Met Glu Asn Ile Ile Leu Ser
260 265 270

Lys Asn Glu Asp Gln Ser Thr Gln Asn Thr Asp Ser Glu Thr Arg Thr
275 280 285

Ile Ser Lys Asn Thr Ser Thr Ser Arg Thr His Thr Ser Glu Val His
290 295 300

Gly Asn Ala Glu Val His Ala Ser Phe Phe Asp Ile Gly Gly Ser Val
305 310 315 320

Ser Ala Gly Phe Ser Asn Ser Asn Ser Ser Thr Val Ala Ile Asp His
325 330 335

Ser Leu Ser Leu Ala Gly Glu Arg Thr Trp Ala Glu Thr Met Gly Leu
340 345 350

Asn Thr Ala Asp Thr Ala Arg Leu Asn Ala Asn Ile Arg Tyr Val Asn
355 360 365

Thr Gly Thr Ala Pro Ile Tyr Asn Val Leu Pro Thr Thr Ser Leu Val
370 375 380

Leu Gly Lys Asn Gln Thr Leu Ala Thr Ile Lys Ala Lys Glu Asn Gln
385 390 395 400

Leu Ser Gln Ile Leu Ala Pro Asn Asn Tyr Tyr Pro Ser Lys Asn Leu
405 410 415

Ala Pro Ile Ala Leu Asn Ala Gln Asp Asp Phe Ser Ser Thr Pro Ile
420 425 430

Thr Met Asn Tyr Asn Gln Phe Leu Glu Leu Glu Lys Thr Lys Gln Leu
435 440 445

Arg Leu Asp Thr Asp Gln Val Tyr Gly Asn Ile Ala Thr Tyr Asn Phe
450 455 460

Glu Asn Gly Arg Val Arg Val Asp Thr Gly Ser Asn Trp Ser Glu Val
465 470 475 480

Leu Pro Gln Ile Gln Glu Thr Thr Ala Arg Ile Ile Phe Asn Gly Lys
485 490 495

Asp Leu Asn Leu Val Glu Arg Arg Ile Ala Ala Val Asn Pro Ser Asp
500 505 510

Pro Leu Glu Thr Thr Lys Pro Asp Met Thr Leu Lys Glu Ala Leu Lys
515 520 525

Ile Ala Phe Gly Phe Asn Glu Pro Asn Gly Asn Leu Gln Tyr Gln Gly
530 535 540

Lys Asp Ile Thr Glu Phe Asp Phe Asn Phe Asp Gln Gln Thr Ser Gln
545 550 555 560

Asn Ile Lys Asn Gln Leu Ala Glu Leu Asn Ala Thr Asn Ile Tyr Thr
565 570 575

Val Leu Asp Lys Ile Lys Leu Asn Ala Lys Met Asn Ile Leu Ile Arg
580 585 590

Asp Lys Arg
595

<210> 12

<211> 1785

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：用来编码 SEQ ID NO: 11 的 DNA 序列

<400> 12

gaagttaaac aggagaaccg gttattaaat gaatcagaat caagttccca ggggttacta 60
ggatactatt ttagtgattt gaattttcaa gcacccatgg tggttacctc ttctactaca 120
ggggatttat ctattcctag ttctgagtta gaaaatattc catcgaaaa ccaatatttt 180
caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttget 240
acttccgctg ataatcatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct 300
tctaattcta acaaaatcag attagaaaaa ggaagattat atcaaataaa aattcaatat 360
caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgattctcaa 420
aataaaaaag aagtgatttc tagtgataac ttacaattgc cagaattaaa acaaaaatct 480

```

tcgaactcaa gaaaaaagcg aagtacaagt gctggaccta cggttccaga ccgtgacaat 540
gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga 600
actttttctt caccatggat ttctaattatt catgaaaaga aaggattaac caaatataaa 660
tcatctcctg aaaaatggag cacggcttct gatccgtaca gtgatttcga aaaggttaca 720
ggacggattg ataagaatgt atcaccagag gcaagacacc cccttggtggc agcttatccg 780
attgtacatg tagatatgga gaatattatt ctctcaaaaa atgaggatca atccacacag 840
aatactgata gtgaaacgag aacaataagt aaaaatactt ctacaagtag gacacatact 900
agtgaagtac atggaaatgc agaagtgc atgcgtcttct ttgatattgg tgggagtgt 960
tctgcaggat ttagtaattc gaattcaagt acggtcgcaa ttgatcattc actatctcta 1020
gcaggggaaa gaacttgggc tgaacaatg ggtttaaata ccgtgatac agcaagatta 1080
aatgccaata ttagatatgt aaatactggg acggctccaa tctacaacgt gttaccaacg 1140
acttcgtagg tgtaggaaa aaatcaaaca ctgcgcgcaa ttaaagctaa ggaaaaccaa 1200
ttaagtcaaa tacttgcacc taataattat tctccttcta aaaacttggc gccaatcgca 1260
ttaaatgcac aagacgattt cagttctact ccaattacaa tgaattacaa tcaatttctt 1320
gagttagaaa aaacgaaaca attaagatta gatacggatc aagtatatgg gaatatagca 1380
acatacaatt ttgaaaatgg aagagtgagg gtggatacag gctcgaactg gagtgaagt 1440
ttaccgcaaa ttcaagaaac aactgcacgt atcattttta atggaaaaga tttaaatctg 1500
gtagaaagge ggatagcggc ggttaatcct agtgatccat tagaaacgac taaaccggat 1560
atgacattaa aagaagccct taaaatagca tttggattta acgaaccgaa tggaaactta 1620
caatatcaag ggaaagacat aaccgaattt gatttttaatt tcgatcaaca aacatctcaa 1680
aatatcaaga atcagttagc ggaattaaac gcaactaaca tatatactgt attagataaa 1740
atcaaattaa atgcaaaaat gaatatttta ataagagata aacgt 1785

```

<210> 13

<211> 735

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：融合蛋白

<400> 13

```

Glu Val Lys Gln Glu Asn Arg Leu Leu Asn Glu Ser Glu Ser Ser Ser
  1             5             10             15

```

```

Gln Gly Leu Leu Gly Tyr Tyr Phe Ser Asp Leu Asn Phe Gln Ala Pro
      20             25             30

```

```

Met Val Val Thr Ser Ser Thr Thr Gly Asp Leu Ser Ile Pro Ser Ser
      35             40             45

```

Glu Leu Glu Asn Ile Pro Ser Glu Asn Gln Tyr Phe Gln Ser Ala Ile
 50 55 60

Trp Ser Gly Phe Ile Lys Val Lys Lys Ser Asp Glu Tyr Thr Phe Ala
 65 70 75 80

Thr Ser Ala Asp Asn His Val Thr Met Trp Val Asp Asp Gln Glu Val
 85 90 95

Ile Asn Lys Ala Ser Asn Ser Asn Lys Ile Arg Leu Glu Lys Gly Arg
 100 105 110

Leu Tyr Gln Ile Lys Ile Gln Tyr Gln Arg Glu Asn Pro Thr Glu Lys
 115 120 125

Gly Leu Asp Phe Lys Leu Tyr Trp Thr Asp Ser Gln Asn Lys Lys Glu
 130 135 140

Val Ile Ser Ser Asp Asn Leu Gln Leu Pro Glu Leu Lys Gln Lys Ser
 145 150 155 160

Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro
 165 170 175

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

Thr Val Asp Val Lys Asn Lys Arg Thr Phe Leu Ser Pro Trp Ile Ser
 195 200 205

Asn Ile His Glu Lys Lys Gly Leu Thr Lys Tyr Lys Ser Ser Pro Glu
 210 215 220

Lys Trp Ser Thr Ala Ser Asp Pro Tyr Ser Asp Phe Glu Lys Val Thr
 225 230 235 240

Gly Arg Ile Asp Lys Asn Val Ser Pro Glu Ala Arg His Pro Leu Val
 245 250 255

Ala Ala Tyr Pro Ile Val His Val Asp Met Glu Asn Ile Ile Leu Ser
 260 265 270

Lys Asn Glu Asp Gln Ser Thr Gln Asn Thr Asp Ser Gln Thr Arg Thr
275 280 285

Ile Ser Lys Asn Thr Ser Thr Ser Arg Thr His Thr Ser Glu Val His
290 295 300

Gly Asn Ala Glu Val His Ala Ser Phe Phe Asp Ile Gly Gly Ser Val
305 310 315 320

Ser Ala Gly Phe Ser Asn Ser Asn Ser Ser Thr Val Ala Ile Asp His
325 330 335

Ser Leu Ser Leu Ala Gly Glu Arg Thr Trp Ala Glu Thr Met Gly Leu
340 345 350

Asn Thr Ala Asp Thr Ala Arg Leu Asn Ala Asn Ile Arg Tyr Val Asn
355 360 365

Thr Gly Thr Ala Pro Ile Tyr Asn Val Leu Pro Thr Thr Ser Leu Val
370 375 380

Leu Gly Lys Asn Gln Thr Leu Ala Thr Ile Lys Ala Lys Glu Asn Gln
385 390 395 400

Leu Ser Gln Ile Leu Ala Pro Asn Asn Tyr Tyr Pro Ser Lys Asn Leu
405 410 415

Ala Pro Ile Ala Leu Asn Ala Gln Asp Asp Phe Ser Ser Thr Pro Ile
420 425 430

Thr Met Asn Tyr Asn Gln Phe Leu Glu Leu Glu Lys Thr Lys Gln Leu
435 440 445

Arg Leu Asp Thr Asp Gln Val Tyr Gly Asn Ile Ala Thr Tyr Asn Phe
450 455 460

Glu Asn Gly Arg Val Arg Val Asp Thr Gly Ser Asn Trp Ser Glu Val
465 470 475 480

Leu Pro Gln Ile Gln Glu Thr Thr Ala Arg Ile Ile Phe Asn Gly Lys
485 490 495

Asp Leu Asn Leu Val Glu Arg Arg Ile Ala Ala Val Asn Pro Ser Asp
500 505 510

Pro Leu Glu Thr Thr Lys Pro Asp Met Thr Leu Lys Glu Ala Leu Lys
515 520 525

Ile Ala Phe Gly Phe Asn Glu Pro Asn Gly Asn Leu Gln Tyr Gln Gly
530 535 540

Lys Asp Ile Thr Glu Phe Asp Phe Asn Phe Asp Gln Gln Thr Ser Gln
545 550 555 560

Asn Ile Lys Asn Gln Leu Ala Glu Leu Asn Ala Thr Asn Ile Tyr Thr
565 570 575

Val Leu Asp Lys Ile Lys Leu Asn Ala Lys Met Asn Ile Leu Ile Arg
580 585 590

Asp Lys Arg Phe His Tyr Asp Arg Asn Asn Ile Ala Val Gly Ala Asp
595 600 605

Glu Ser Val Val Lys Glu Ala His Arg Glu Val Ile Asn Ser Ser Thr
610 615 620

Glu Gly Leu Leu Leu Asn Ile Asp Lys Asp Ile Arg Lys Ile Leu Ser
625 630 635 640

Gly Tyr Ile Val Glu Ile Glu Asp Thr Glu Gly Leu Lys Glu Val Ile
645 650 655

Asn Asp Arg Tyr Asp Met Leu Asn Ile Ser Ser Leu Arg Gln Asp Gly
660 665 670

Lys Thr Phe Ile Asp Phe Lys Lys Tyr Asn Asp Lys Leu Pro Leu Tyr
675 680 685

Ile Ser Asn Pro Asn Tyr Lys Val Asn Val Tyr Ala Val Thr Lys Glu
690 695 700

Asn Thr Ile Ile Asn Pro Ser Glu Asn Gly Asp Thr Ser Thr Asn Gly
705 710 715 720

Ile Lys Lys Ile Leu Ile Phe Ser Lys Lys Gly Tyr Glu Ile Gly
 725 730 735

<210> 14
 <211> 2208
 <212> DNA
 <213> 人工序列

<220>
 <223> 人工序列的描述: 用来编码 SEQ ID NO: 13 的 DNA 序列

<400> 14
 gaagttaaac aggagaaccg gttattaaat gaatcagaat caagttccca ggggttacta 60
 ggatactatt ttagtgattt gaattttcaa gcacccatgg tggttacctc ttctactaca 120
 ggggatttat ctattcctag ttctgagtta gaaaatattc catcgaaaa ccaatatttt 180
 caatctgcta ttgtgtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct 240
 acttccgctg ataatcatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct 300
 tctaattcta acaaaatcag attagaaaaa ggaagattat atcaaataaa aattcaatat 360
 caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgatttctca 420
 aataaaaaag aagtgatttc tagtgataac ttacaattgc cagaattaaa acaaaaatct 480
 tcgaactcaa gaaaaaagcg aagtacaagt gctggacctc cggttccaga ccgtgacaat 540
 gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga 600
 acttttcttt caccatggat ttctaattat catgaaaaga aaggattaac caaatataaa 660
 tcatctcctg aaaaatggag cacggcttct gatccgtaca gtgatttcga aaaggttaca 720
 ggacggattg ataagaatgt atcaccagag gcaagacacc cccttgtggc agcttatccg 780
 attgtacatg tagatatgga gaatattatt ctctcaaaaa atgaggatca atccacacag 840
 aatactgata gtgaaacgag aacaataagt aaaaatactt ctacaagtag gacacatact 900
 agtgaagtac atggaaatgc agaagtgcac gcgtcgttct ttgatattgg tgggagtgtg 960
 tctgcaggat ttagtaattc gaattcaagt acggtcgcaa ttgatcattc actatctcta 1020
 gcaggggaaa gaacttgggc tgaacaatg ggtttaaata ccgctgatac agcaagatta 1080
 aatgccataa ttagatatgt aaatactggg acggctccaa tctacaacgt gttaccaacg 1140
 acttcgttag tgttaggaaa aatcaaaca ctgcgcgcaa ttaaagctaa ggaaaacca 1200
 ttaagtcaaa tacttgcccc taataattat tatccttcta aaaacttggc gccaatcgca 1260
 ttaaattgac aagacgattt cagttctact ccaattacaa tgaattacaa tcaatttctt 1320
 gagttagaaa aaacgaacaa attaagatta gatacggatc aagtatatgg gaatatagca 1380
 acatacaatt ttgaaaatgg aagagtgagg gtggatacag gctcgaaact gagtgaagt 1440
 ttaccgcaaa ttcaagaaac aactgcacgt atcatTTTTA atggaagaaga tttaaactctg 1500
 gtagaaaggc ggatagcggc ggttaatcct agtgatecat tagaaacgac taaaccggat 1560
 atgacattaa aagaagccct taaaatagca ttgtgattta acgaaccgaa tggaaactta 1620
 caatatcaag ggaaagacat aaccgaattt gatttttaatt tcgatcaaca aacatctcaa 1680
 aatatcaaga atcagtttagc ggaattaaac gcaactaaca tatatactgt attagataaa 1740

```

atcaaattaa atgcaaaaat gaatatttta ataagagata aacgttttca ttatgataga 1800
aataacatag cagttggggc ggatgagtca gtagttaagg aggtcatag agaagtaatt 1860
aatcgteaa cagagggatt attgttaaatt attgataagg atataagaaa aatattatca 1920
ggttatattg tagaaattga agatactgaa gggcttaaag aagttataaa tgacagatat 1980
gatatgttga atatttctag ttacggcaa gatggaaaaa catttataga ttttaaaaaa 2040
tataatgata aattaccgtt atatataagt aatcccaatt ataaggtaaa tgtatatgct 2100
gttactaaag aaaacactat tattaatcct agtgagaatg gggatactag taccaacggg 2160
atcaagaaaa ttttaatctt ttctaaaaaa ggctatgaga taggataa 2208

```

<210> 15

<211> 140

<212> PRT

<213> 炭疽芽孢杆菌(Bacillus anthracis)

<400> 15

```

Phe His Tyr Asp Arg Asn Asn Ile Ala Val Gly Ala Asp Glu Ser Val
  1             5             10            15

```

```

Val Lys Glu Ala His Arg Glu Val Ile Asn Ser Ser Thr Glu Gly Leu
      20             25            30

```

```

Leu Leu Asn Ile Asp Lys Asp Ile Arg Lys Ile Leu Ser Gly Tyr Ile
      35             40            45

```

```

Val Glu Ile Glu Asp Thr Glu Gly Leu Lys Glu Val Ile Asn Asp Arg
      50             55            60

```

```

Tyr Asp Met Leu Asn Ile Ser Ser Leu Arg Gln Asp Gly Lys Thr Phe
      65             70            75            80

```

```

Ile Asp Phe Lys Lys Tyr Asn Asp Lys Leu Pro Leu Tyr Ile Ser Asn
      85             90            95

```

```

Pro Asn Tyr Lys Val Asn Val Tyr Ala Val Thr Lys Glu Asn Thr Ile
      100            105            110

```

```

Ile Asn Pro Ser Glu Asn Gly Asp Thr Ser Thr Asn Gly Ile Lys Lys
      115            120            125

```

```

Ile Leu Ile Phe Ser Lys Lys Gly Tyr Glu Ile Gly
      130            135            140

```

<210> 16

<211> 423

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：结构域 4 的 DNA 编码序列

<400> 16

```
tttcattatg atagaaataa catagcagtt ggggcggatg agtcagtagt taaggaggct 60
catagagaag taattaattc gtcaacagag ggattattgt taaatattga taaggatata 120
agaaaaatat tatcaggtta tattgtagaa attgaagata ctgaagggtt taaagaagtt 180
ataaatgaca gatatgatat gttgaatatt tctagtttac ggcaagatgg aaaaacattt 240
atagatttta aaaaatataa tgataaatta cgttatata taagtaatcc caattataag 300
gtaaattgtat atgctgttac taaagaaaac actattatta atcctagtga gaatggggat 360
actagtacca acgggatcaa gaaaatttta atcttttcta aaaaaggcta tgagatagga 420
taa                                                                                     423
```

***Escherichia coli* [gbtct]: 14457 个 CDS (4541860 个密码子)**

字段: [三联体] [频率: 每千个] ([数目])

UUU 22.0 (100128)	UCU 9.3 (42367)	UAU 16.7 (75774)	UGU 5.2 (23461)
UUC 16.5 (74885)	UCC 8.9 (40365)	UAC 12.3 (55847)	UGC 6.3 (28747)
UUA 13.8 (62823)	UCA 7.9 (35837)	UAA 2.0 (9006)	UGA 1.0 (4428)
UUG 13.3 (60322)	UCG 8.7 (39546)	UAG 0.3 (1172)	UGG 14.5 (65630)
CUU 11.3 (51442)	CCU 7.2 (32678)	CAU 12.7 (57585)	CGU 20.7 (93997)
CUC 10.6 (48147)	CCC 5.4 (24383)	CAC 9.6 (43743)	CGC 21.1 (96053)
CUA 4.0 (18067)	CCA 8.5 (38663)	CAA 14.8 (67129)	CGA 3.7 (16607)
CUG 50.9 (231373)	CCG 22.3 (101467)	CAG 28.8 (130898)	CGG 5.7 (25751)
AUU 29.9 (135873)	ACU 9.5 (43256)	AAU 18.7 (84846)	AGU 9.1 (41544)
AUC 24.6 (111878)	ACC 22.7 (103121)	AAC 21.6 (98018)	AGC 15.6 (70867)
AUA 5.3 (24233)	ACA 7.9 (35995)	AAA 34.4 (156169)	AGA 2.7 (12345)
AUG 27.2 (123604)	ACG 14.0 (63696)	AAG 11.4 (51685)	AGG 1.6 (7423)
GUU 19.1 (86572)	GCU 16.2 (73677)	GAU 32.3 (146794)	GGU 25.1 (114185)
GUC 14.8 (67356)	GCC 25.0 (113412)	GAC 19.3 (87759)	GGC 28.6 (130043)
GUA 11.2 (51020)	GCA 20.6 (93390)	GAA 39.5 (179460)	GGA 8.6 (39036)
GUG 25.5 (115687)	GCG 32.2 (146264)	GAG 18.5 (83804)	GGG 11.1 (50527)

编码 GC 51.37% 第1个字母 GC 58.50% 第2个字母 GC 40.70% 第三个字母 GC 54.90%

***Bacillus anthracis* [gbtct]: 180 个 CDS (52031 个密码子)**

字段: [三联体] [频率: 每千个] ([数目])

UUU 33.5 (1745)	UCU 17.3 (902)	UAU 34.4 (1792)	UGU 5.1 (319)
UUC 10.2 (530)	UCC 5.3 (275)	UAC 9.4 (490)	UGC 2.1 (107)
UUA 44.2 (2301)	UCA 14.0 (730)	UAA 2.3 (118)	UGA 0.5 (24)
UUG 11.3 (589)	UCG 3.6 (188)	UAG 0.7 (37)	UGG 9.8 (511)
CUU 14.7 (763)	CCU 10.1 (525)	CAU 16.8 (873)	CGU 10.9 (567)
CUC 3.7 (195)	CCC 2.7 (141)	CAC 4.6 (239)	CGC 2.6 (137)
CUA 13.2 (686)	CCA 14.9 (773)	CAA 33.7 (1752)	CGA 6.8 (353)
CUG 4.7 (242)	CCG 4.6 (237)	CAG 10.4 (542)	CGG 1.8 (95)
AUU 44.6 (2322)	ACU 14.6 (761)	AAU 44.6 (2321)	AGU 16.5 (861)
AUC 11.8 (616)	ACC 5.2 (269)	AAC 13.7 (711)	AGC 5.1 (266)
AUA 24.9 (1295)	ACA 25.9 (1350)	AAA 69.5 (3614)	AGA 13.8 (720)
AUG 23.8 (1240)	ACG 8.1 (419)	AAG 23.5 (1223)	AGG 4.3 (226)
GUU 19.9 (1036)	GCU 17.9 (930)	GAU 39.7 (2068)	GGU 17.3 (900)
GUC 5.2 (268)	GCC 4.7 (244)	GAC 8.8 (456)	GGC 5.4 (279)
GUA 26.8 (1395)	GCA 22.6 (1178)	GAA 55.7 (2897)	GGA 20.2 (1049)
GUG 9.7 (507)	GCG 7.1 (368)	GAG 19.3 (1003)	GGG 8.9 (461)

编码 GC 33.59% 第1个字母 GC 44.51% 第2个字母 GC 31.07% 第三个字母 GC 25.20%

图 1

```

1  AAGCTTCATA TGGAAGTAAA GCAAGAGAAC CGTCTGCTGA ACGAATCTGA ATCCAGCTCT
61  CAGGGCCTGC TTGGTTACTA TTTCTCTGAC CTGAACTTCC AAGCACCGAT GGTTGTAACC
121 AGCTCTACCA CTGGCGATCT GTCCATCCCG TCTAGTGAAC TTGAGAACAT TCCAAGCGAG
181 AACCAGTATT TCCAGTCTGC AATCTGGTCC GGTTTTATCA AAGTCAAGAA ATCTGATGAA
241 TACACGTTTG CCACCTCTGC TGATAACCAC GTAACCATGT GGGTTGACGA TCAGGAAGTG
301 ATCAACAAAG CATCCAATC CAACAAAATT CGTCTGGAAA AAGGCCGTCT GTATCAGATC
361 AAGATTCACT ACCAACGCGA GAACCCGACT GAAAAAGGCC TGGACTTTAA ACTGTATTGG
421 ACTGATTCTC AGAACAAGAA AGAAGTGATC AGCTCTGACA ATCTGCAACT GCCGGAATTG
481 AAACAGAAAA GCTCCAATC TCGTAAGAAA CGTTCACCA GCGCTGGCCC GACCGTACCA
541 GATCGCGACA ACGATGGTAT TCCGGACTCT CTGGAAGTTG AAGGCTACAC GGTGATGTA
601 AAGAACAAAC GTACCTTCCT TAGTCCGTGG ATCTCCAATA TTCACGAGAA GAAAGGTCTG
661 ACCAAATACA AATCCAGTCC GGAATAATGG TCCACTGCAT CTGATCCGTA CTCTGACTTT
721 GAGAAAGTGA CCGTCTGTAT CGACAAGAAC GTCTCTCCGG AAGCACGCCA TCCACTGGTT
781 GCTGCGTATC CGATCGTACA TGTTGACATG GAAAACATCA TTTTGTCCAA GAACGAAGAC
841 CAGTCCACTC AGAACACTGA CTCTGAAACT CGTACCATCT CCAAGAACAC CTCCACGTCT
901 CGTACTCACA CCAGTGAAGT ACATGGTAAC GCTGAAGTAC ACGCCTCTTT CTTTGACATC
961 GCGGCTCTG TTAGCGCTGG CTTCTCCAAC TCTAATTCTT CTACTGTTGC CATTGATCAC
1021 TCTCTGAGTC TGGCTGGCGA ACGTACCTGG GCAGAGACCA TGGGTCTTAA CACTGCTGAT
1081 ACCGCGCGTC TGAATGCTAA CATTGCTAC GTCAACACTG GTACGGCACC GATCTACAAC
1141 GTACTGCCAA CCACCAGCCT GGTCTGGGT AAGAACCAGA CTCTTGCGAC CATCAAAGCC
1201 AAAGAGAACC AACTGTCTCA GATTCTGGCA CCGAATAACT ACTATCCTTC CAAGAAGCTG
1261 GCTCCGATCG CACTGAACGC ACAGGATGAC TTCTCTTCCA CTCCGATCAC CATGAACCTAC
1321 AACCAGTTCC TGGAACCTGA GAAGACCAAA CAGCTGCGTC TTGACACTGA CCAAGTGTAC
1381 GGTAACATCG CGACCTACAA CTTTGAGAAC GGTGCGCTCC GCGTTGACAC AGGCTCTAAT
1441 TGGTCTGAAG TACTGCCTCA GATTGAGGAA ACCACCGCTC GTATCATCTT CAACGGTAAA
1501 GACCTGAACC TGGTTGAACG TCGTATTGCT GCTGTGAACC CGTCTGATCC ATTAGAGACC
1561 ACCAAACCGG ATATGACTCT GAAAGAAGCC CTGAAGATCG CCTTTGGCTT CAACGAGCCG
1621 AACGGTAATC TTCAGTACCA AGGTAAAGAC ATCACTGAAT TTGACTTCAA CTTTGATCAG
1681 CAGACCTCTC AGAATATCAA GAACCAACTG GCTGAGCTGA ACGCGACCAA TATCTATACG
1741 GTACTCGACA AGATCAAAC GAACGCGAAA ATGAACATTC TGATTGCGGA CAAACGTTTC
1801 CACTACGATC GTAATAACAT CGCTGTTGGC GCTGATGAAT CTGTTGTGAA AGAAGCGCAT
1861 CGCGAAGTCA TCAACTCCAG CACCGAAGGC CTGCTTCTGA ACATCGACAA AGACATTCTG
1921 AAGATCCTGT CTGGTTACAT TGTTGAGATC GAAGACACCG AAGGCCTGAA AGAAGTGATC
1981 AATGATCGTT ACGACATGCT GAACATCAGC TCTCTGCGTC AAGATGGTAA GACGTTCAAT
2041 GACTTCAAGA AATACAACGA CAACTTCCG CTGTATATCT CTAATCCGAA CTACAAAGTG
2101 AACGTTTACG CTGTTACCAA AGAGAACACC ATCATCAATC CATCTGAGAA CGGCGATACC
2161 TCTACCAACG GTATCAAGAA GATTCTGATC TTCTCCAAGA AAGGTTACGA GATCGGTTAA
2221 TAGGATCC

```

(SEQ ID No 1)

图 2

```

1 EVKQENRLLN ESESSSQGLL GYYFSDLNFQ APMVVTSSST GDLSIPSEL ENIPSENQYF
61 QSAIWSGFIK VKKSDEYTFA TSADNHVTMW VDDQEVINKA SNSNKIRLEK GRLYQIKIQY
121 QRENPTKGL DFKLYWTDQ NKKEVISSDN LQPELKQKS SNSRKRSTS AGPTVPDRDN
181 DGIPDSLEVE GYTVDVKNKR TFLSPWISNI HEKKGLTKYK SSPEKWSTAS DPYSDFEKVT
241 GRIDKNVSPE ARHPLVAA

```

(Seq ID No 3)

```

1 gaagttaaac aggagaaccg gttattaaat gaatcagaat caagttccca ggggttacta
61 ggatactatt ttagtgattt gaattttcaa gcacccatgg tggttacctc ttctactaca
121 ggggattttat ctattcctag ttctgagtta gaaaatattc catcggaaaa ccaatatttt
181 caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct
241 acttccgctg ataatcatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct
301 tctaattcta acaaaatcag attagaaaaa ggaagattat atcaataaaa aattcaatat
361 caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgattctcaa
421 aataaaaaag aagtgatttc tagtgataac ttacaattgc cagaattaaa acaaaaatct
481 tcgaactcaa gaaaaaagcg aagtacaagt gctggacctc cggttccaga ccgtgacaat
541 gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga
601 acttttcttt caccatggat ttctaataat catgaaaaga aaggattaac caaatataaa
661 tcatctcctg aaaaatggag cacggcttct gatccgtaca gtgatttcga aaaggttaca
721 ggacggattg ataagaatgt atcaccagag gcaagacacc cccttggtggc agct

```

(Seq ID No 4)

```

1 EVKQENRLLN ESESSSQGLL GYYFSDLNFQ APMVVTSSST GDLSIPSEL ENIPSENQYF
61 QSAIWSGFIK VKKSDEYTFA TSADNHVTMW VDDQEVINKA SNSNKIRLEK GRLYQIKIQY
121 QRENPTKGL DFKLYWTDQ NKKEVISSDN LQPELKQKS SNSRKRSTS AGPTVPDRDN
181 DGIPDSLEVE GYTVDVKNKR TFLSPWISNI HEKKGLTKYK SSPEKWSTAS DPYSDFEKVT
241 GRIDKNVSPE ARHPLVAAYP IVHVDMENII LSKNEDQSTQ NTDSETRTIS KNTSTSRHT
301 SEVHGNAEVH ASFFDIGGSV SAGFSNSNSS TVAIDHSLSL AGERTWAETM GLNTADTARL
361 NANIRYVNTG TAPIYNVLP TSLVLGKNQT LATIKAKENQ LSQILAPNNY YPSKNLAPIA
421 LNAQDDFSST PITMNYNQFL ELEKTKQLRL DTDQVYGNIA TYNFENGRVR VDTGSNWSEV
481 LPQIQET

```

(SEQ ID No 5)

图 3

```

1 gaagttaaacc aggagaaccg gttatttaaatt gaatcagaat caagttccca ggggttacta
61 ggatactatt ttagtgattt gaattttcaa gcacccatgg tgggttacttc ttctactaca
121 ggggattttat ctatttcctag ttctgagtta gaaaatattc catcggaaaa ccaatatttt
181 caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct
241 acttccgctg ataatcatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct
301 tctaattcta acaaaatcag attagaaaaa ggaagattat atcaaatata aattcaatat
361 caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgatttctcaa
421 aataaaaaag aagtgatttc tagtgataac ttacaactgc cagaattaaa acaaaaatct
481 tcgaactcaa gaaaaaagcg aagtacaagt gctggaccta cgggttccaga ccgtgacaat
541 gatggaatcc ctgatttcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga
601 acttttcttt caccatggat ttctaattat catgaaaaga aaggattaac caaatataaa
661 tcactctctg aaaaatggag cacggcttct gatccgtaca gtgatttcga aaaggttaca
721 ggacggattg ataagaatgt atcaccagag gcaagacacc cccttggtgc agcttatccg
781 attgtacatg tagatatgga gaatattatt ctctcaaaaa atgaggatca atccacacag
841 aatactgata gtgaaacgag aacaataagt aaaaataact ctacaagtag gacacatact
901 agtgaagtac atggaaatgc agaagtgcac gcgtcgttct ttgatattgg tgggagtgtg
961 tctgcaggat ttagtaattc gaattcaagt acggtcgcga ttgatcattc actatctcta
1021 gcaggggaaa gaacttgggc tgaaacaatg ggtttaaata ccgctgatac agcaagatta
1081 aatgccataa ttagatatgt aaatactggg acggctccaa tctacaacgt gttaccaacg
1141 acttcgttag tgtaggaaa aaatcaaaaca ctgcgcgaca ttaaagctaa ggaaaacca
1201 ttaagtcaaa tacttgacac taataattat tatccttcta aaaacttggc gccaatcgca
1261 ttaaatgcac aagacgattt cagttctact ccaattacaa tgaattacaa tcaatttctt
1321 gagttagaaa aaacgaaaca attaagatta gatacggatc aagtatatgg gaatatagca
1381 acatacaatt ttgaaaatgg aagagtggag gtggatacag gctcgaactg gagtgaagtg
1441 ttaccgcaa ttcaagaaac a

```

(SEQ ID No 6)

```

1 SAGPTVPDRD NDGIPDSLEV EGYTVDVKNK RTFLSPWISN IHEKKGLTKY KSSPEKWSTA
61 SDPYSDFEKV TGRIDKNVSP EARHPLVAAY PIVHVDMENI ILSKNEDQST QNTDSETRTI
121 SKNTSTSRTH TSEVHGNAEV HASFFDIGGS VSAGESNSNS STVAIDHSLs LAGERTWAET
181 MGLNTADTAR INANIRYVNT GTAPIYNVLP TTSLVLGKNQ TLATIKAKEN QLSQILAPNN
241 YYPskNLAPI ALNAQDDFSS TPITMNYNQF LELEKTKQLR LDTDQVYGNi ATYNFENG RV
301 RVDtGSNWSE VLPQIQET

```

(SEQ ID No 7)

图 3 (续)

```

1 agtgctggac ctacggttcc agaccgtgac aatgatggaa tccctgattc attagaggta
61 gaaggatata cggttgatgt caaaaataaa agaacttttc tttcaccatg gattttcta
121 attcatgaaa agaaaggatt aaccaaataa aaatcatctc ctgaaaaatg gagcacggct
181 tctgatccgt acagtgattt cgaaaagggt acaggacgga ttgataagaa tgtatcacca
241 gaggcaagac accccottgt ggcagcttat ccgattgtac atgtagatat ggagaatatt
301 attctctcaa aaaatgagga tcaatccaca cagaatactg atagtgaac gagaacaata
361 agtaaaaata cttctacaag taggacacat actagtgaag tacatggaaa tgcagaagtg
421 catgcgtcgt tctttgatat tgggtggagt gtatctgcag gatttagtaa ttcgaattca
481 agtacggtcg caattgatca ttcactatct ctagcagggg aaagaacttg ggctgaaaca
541 atgggtttta ataccgctga tacagcaaga ttaaagcca atattagata tgtaaatact
601 gggacggctc caatctacaa cgtgttacca aogacttcgt tagtgttagg aaaaaatcaa
661 acactcgcg caattaaagc taaggaaaac caattaagtc aaatacttgc acctaatat
721 tattatcctt ctaaaaactt ggcgccaatc gcattaaatg cacaagacga tttcagttct
781 actccaatta caatgaatta caatcaattt cttgagttag aaaaaacgaa acaattaaga
841 ttagatacgg atcaagtata tgggaatata gcaacatata attttgaaaa tgggaagagt
901 aggggtgata caggctcgaa ctggagtga gtgttaccgc aaattcaaga aaca

```

(SEQ ID No 8)

```

1 SAGPTVPDRD NDGIPDSLEV EGYTVDVKNK RTFLSPWISN IHEKKGLTKY KSSPEKWSTA
61 SDPYSDFEKV TGRIDKNVSP EARHPLVAAY PIVHVDMENI ILSKNEDQST QNTDSETRTI
121 SKNTSTSRTH TSEVHGNAEV HASFFDIGGS VSAGFSNSNS STVAIDHSL S LAGERTWAET
181 MGLNTADTAR LNANIRYVNT GTAPIYNVLP TTSLVLGKNQ TLATIKAKEN QLSQILAPNN
241 YPSKLNAPI ALNAQDDFSS TPITMNYNQF LELEKTKQLR LDTDQVYGN I ATYNFENGRV
301 RVD TGSNWSE VLPQIQETTA RIIFNGKDLN LVERRIAAVN PSDPLETTKP DMTLKEALKI
361 AFGFNEPNGN LQYQKGDITE FDFNFDQOTS QNIKNQLAEL NATNIYTVLD KIKLNAKMNI
421 LIRDKR

```

(SEQ ID No 9)

图 3 (续)

```

1 agtgctggac ctacggttcc agaccgtgac aatgatggaa tccctgattc attagaggta
61 gaaggatata cggttgatgt caaaaataaa agaacttttc tttcaccatg gatttctaata
121 attcatgaaa agaaaggatt aaccaaatat aaatcatctc ctgaaaaatg gagcacggct
181 tctgatccgt acagtgattt cgaaaagggt acaggacgga ttgataagaa tgtatcacca
241 gaggcaagac acccccttgt gccagcttat ccgattgtac atgtagatat ggagaatatt
301 attctctcaa aaaatgagga tcaatccaca cagaatactg atagtgaac gagaacaata
361 agtaaaaata cttctacaag taggacacat actagtgaag tacatggaaa tgcagaagtg
421 catgctcgt tctttgatat tggtagggagt gtatctgcag gatttagtaa ttcgaattca
481 agtacggctc caattgatca ttcactatct ctacgagggg aaagaacttg ggctgaaaca
541 atgggtttta ataccgctga tacagcaaga ttaaattgcca atattagata tgtaaatact
601 gggacggctc caatctacaa cgtgttacca acgacttcgt tagtgttagg aaaaaatcaa
661 aactcgcga caattaaagc taaggaaaac caattaagtc aaataacttg acctaataat
721 tattatcctt ctaaaaactt ggcgccaatc gcattaaatg cacaagacga tttcagttct
781 actccaatta caatgaatta caatcaattt cttgagttag aaaaaacgaa acaattaaga
841 ttagatacgg atcaagtata tgggaatata gcaacataca attttgaaaa tggaagagtg
901 agggtaggata caggctcgaa ctggagtga gtgttaccgc aaattcaaga aacaactgca
961 cgtatcattt ttaatggaaa agatttaaat ctggtagaaa ggcggatagc ggcggttaat
1021 cctagtgatc cattagaaac gactaaaccg gatatgacat taaaagaagc ccttaaaata
1081 gcatttggat ttaacgaacc gaatggaaac ttacaatatc aagggaaaga cataaccgaa
1141 tttgatttta atttcgatca acaaacatct caaaatatca agaatcagtt agcgggaatta
1201 aacgcaacta acatatatac tgtattagat aaaatcaaat taaatgcaaa aatgaatatt
1261 ttaataagag ataaacgt

```

(SEQ ID No 10)

```

1 EVKQENRLLN ESESSSQGLL GYFSDLNFO APMVVTSSST GDLSIPSEL ENIPSENQYF
61 QSAIWSGFIK VKKSDEYTFE TSADNHVTMW VDDQEVINKA SNSNKIRLEK GRLYQIKIQY
121 QRENPTKGL DFKLYWTDQ NKKEVISSDN LQLPELKQKS SNSRKKRSTS AGPTVPDRDN
181 DGIPDSLEVE GYTVDVKNKR TFLSPWISNI HEKKGLTKYK SSPEKWSTAS DPYSDFEKVT
241 GRIDKNVSPE ARHPLVAAYP IVHVDMENII LSKNEDQSTQ NTDSETRTIS KNTSTSRHTT
301 SEVHGNAEVH ASFFDIGGSV SAGFSNSNSS TVAIDHSLSL AGERTWAETM GLNTADTARL
361 NANIRYVNTG TAPIYNVLP TSLVLGKNQT LATIKAKENQ LSQILAPNNY YPSKNLAPIA
421 LNAQDDFSST PITMNYNQFL ELEKTKQLRL DTDQVYGNIA TYNFENGRVR VDTGSNWSEV
481 LPQIQETAR IIFNGKDLNL VERRIAAVNP SDPLETTKPD MTLKEALKIA FGFNEPNGNL
541 QYQKGDITEF DFNFDQOTSQ NIKNQLAELN ATNIYTVLDK IKLNAMNII IRDKR

```

(SEQ ID No 11)

图 3 (续)

```

1  gaagttaaac aggagaaccg gttatttaa at gaatcagaat caagttccca ggggttacta
61 ggatactatt ttagtgattt gaatttttcaa gcacccatgg tggttacctc ttctactaca
121 ggggatttat ctattcctag ttctgagtta gaaaatattc catcggaata ccaatatttt
181 caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct
241 acttccgctg ataatcatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct
301 tctaattcta acaaaatcag attagaaaaa ggaagattat atcaataaaa aattcaatat
361 caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgatttctcaa
421 aataaaaaag aagtgatttc tagtgataac ttacaattgc cagaattaaa acaaaaatct
481 tcgaactcaa gaaaaaagcg aagtacaagt gctggaccta cggttccaga ccgtgacaat
541 gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaaga
601 acttttcttt caccatggat ttctaattat catgaaaaga aaggattaac caaatataaa
661 tcatctcctg aaaaatggag cagggttctt gatccgtaca gtgatttcga aaagggttaca
721 ggacggattg ataagaatgt atcaccagag gcaagacacc cccttggtgc agcttatccg
781 attgtacatg tagatatgga gaatattatt ctctcaaaaa atgaggatca atccacacag
841 aatactgata gtgaaacgag aacaataagt aaaaataact ctacaagtag gacacatact
901 agtgaagtac atggaatgc agaagtgc at ggcgtcttct ttgatattgg tgggagtgt
961 tctgcaggat ttagtaattc gaattcaagt acggctcga ttgatcattc actatctcta
1021 gcaggggaaa gaacttgggc tgaacaatg ggtttaata ccgctgatac agcaagatta
1081 aatgccaata ttagatatgt aaatactggg acggctccaa tctacaacgt gttaccaacg
1141 acttcgtagg tgtaggaaa aaatcaaaca ctgcgcgaca ttaaagctaa ggaaaaccac
1201 ttaagtcaaa tacttgcacc taataattat tatccttcta aaaacttggc gccaatcgca
1261 ttaaatgcac aagacgattt cagttctact ccaattacaa tgaattacaa tcaatttctt
1321 gagttagaaa aaacgaaaca attaagatta gatacggatc aagtatatgg gaatatagca
1381 acatacaatt ttgaaaatgg aagagtgagg gtggatacag gctcgaactg gagtgaagt
1441 ttaccgcaaa ttcaagaaac aactgcacgt atcattttta atggaaaaga tttaaatctg
1501 gtagaaaggc ggatagcggc ggtaaatcct agtgatccat tagaaacgac taaaccggat
1561 atgacattaa aagaagccct taaaatagca tttggattta acgaaccgaa tggaaactta
1621 caatatcaag ggaaagacat aaccgaattt gattttaatt tcgatcaaca aacatctcaa
1681 aatatcaaga atcagttagc ggaattaaac gcaactaaca tatatactgt attagataaa
1741 atcaaattaa atgcaaaaat gaatatttta ataagagata aacgt

```

(SEQ ID No 12)

```

1  EVKQENRLLN ESESSSQGLL GYYFSDLNFQ APMVVTSSTT GDLSIPSEL ENIPSENQYF
61 QSAIWSGFIK VKKSDEYTFA TSADNHVTMW VDDQEVINKA SNSNKIRLEK GRLYQIKIQY
121 QRENPTKGL DFKLYWTD SQ NKKEVISSDN LQPELKQKS SNSRKKRSTS AGPTVPDRDN
181 DGIPDSLEVE GYTVDVKNKR TFLSPWISNI HEKKGLTKYK SSPEKWSTAS DPYSDFEKT
241 GRIDKNVSPE ARHPLVAAYP IVHVDMENII LSKNEDQSTQ NTDSQTRTIS KNTSTSRHT
301 SEVHGNAEVH ASFFDIGGSV SAGFSNSNSS TVAIDHSLSL AGERTWAETM GLNTADTARL
361 NANIRYVNTG TAPIYNVLP TSLVLGKNQT LATIKAKENQ LSQILAPNNY YPSKNLAPIA
421 LNAQDDFSST PITMNYNQFL ELEKTKQLRL DTDQVYGNIA TYNFENGVR VDTGSNWSEV
481 LPQIQETTAR IIFNGKDLNL VERRIAAVNP SDPLETTKPD MTLKEALKIA FGFNEPNGNL
541 QYQGDITEF DFNFDQOTSQ NIKNQLAELN ATNIYTVLDK IKLNAKMNIL IRDKRFHYDR
601 NNIAGVADES VVKEAHREVI NSSTEGLLL IDKDIRKILS GYIVEIEDTE GLKEVINDRY
661 DMLNISSLRQ DGKTFIDFKK YNDKLPYIS NPNYKVNVA VTKENTIINP SENGDTSTNG
721 IKKILIFSKK GYEIG

```

(SEQ ID No 13)

图 3 (续)

```

1  gaagttaaac aggagaaccg gttatttaa at gaatcagaat caagttccca ggggttacta
61  ggatactatt ttagtgattt gaattttcaa gcacccatgg tggttaccto ttctactaca
121  ggggatttat ctattcctag ttctgagtta gaaaatatto catcgaaaa ccaatatatt
181  caatctgcta tttggtcagg atttatcaaa gttaagaaga gtgatgaata tacatttgct
241  acttccgctg ataatcatgt aacaatgtgg gtagatgacc aagaagtgat taataaagct
301  tctaattcta acaaaatcag attagaaaaa ggaagattat atcaaataaa aattcaatat
361  caacgagaaa atcctactga aaaaggattg gatttcaagt tgtactggac cgattctcaa
421  aataaaaaag aagtgatttc tagtgataac ttacaattgc cagaattaaa acaaaaatct
481  tcgaactcaa gaaaaaagcg aagtacaagt gctggaccta cggttccaga ccgtgacaat
541  gatggaatcc ctgattcatt agaggtagaa ggatatacgg ttgatgtcaa aaataaaga
601  acttttcttt caccatggat ttctaataat catgaaaaga aaggattaac caaatataaa
661  tcactctcctg aaaaatggag cacggcttct gatccgtaca gtgatttoga aaaggttaca
721  ggaaggattg ataagaatgt atcaccagag gcaagacacc cccttgtggc agcttatccg
781  attgtacatg tagatatgga gaattattat ctctcaaaaa atgaggatca atccacacag
841  aatactgata gtgaaacgag aacaataagt aaaaatactt ctacaagtag gacacatact
901  agtgaagtac atggaaatgc agaagtgcgt gcgtcgttct ttgatattgg tgggagtgtg
961  tctgcaggat ttagtaattc gaattcaagt acggtcgcaa ttgatcattc actatctcta
1021  gcaggggaaa gaacttgggc tgaaacaatg ggtttaataa ccgctgatac agcaagatta
1081  aatgccaaata ttagatatgt aaatactggg acggctccaa tctacaacgt gttaccaacg
1141  acttcgttag tgtaggaaa aaatcaaaca ctgcgcacaa ttaaagctaa ggaaaacca
1201  ttaagtcaaa tacttgcacc taataattat tatccttcta aaaacttggc gccaatcgca
1261  ttaaatgcac aagacgattt cagttctact ccaattacaa tgaattacaa tcaatttctt
1321  gagttagaaa aaacgaaaca attagatta gatacggatc aagtatatgg gaatatagca
1381  acatacaatt ttgaaaatgg aagagtgagg gtggatacag gctcgaactg gagtgaagtg
1441  ttaccgcaaa ttcaagaaac aactgcacgt atcattttta atggaaaaga tttaaatctg
1501  gtagaaaggc ggatagcggc ggttaatcct agtgatccat tagaaacgac taaaccggat
1561  atgacattaa aagaagccct taaaatagca tttggattta acgaaccgaa tggaaactta
1621  caatatcaag ggaaagacat aaccgaattt gatttttaatt tcgatcaaca aacatctcaa
1681  aatatcaaga atcagtttagc ggaattaaac gcaactaaca tatatactgt attagataaa
1741  atcaaattaa atgcaaaaat gaatttttta ataagagata aacgttttca ttatgataga
1801  aataacatag cagttggggc ggatgagtcg tagttaagg aggctcatag agaagtaatt
1861  aattcgtcaa cagagggatt attgttaa attgataagg atataagaaa aatattatca
1921  ggttatattg tagaaattga agatactgaa gggcttaaag aagttataaa tgacagatat
1981  gatatgttga atatttctag ttacggcaa gatggaaaaa catttataga ttttaaaaaa
2041  tataatgata aattaccgtt atatataagt aatcccaatt ataaggtaaa tgtatatgct
2101  gttactaaag aaaacactat tattaatcct agtgagaatg gggatactag taccaacggg
2161  atcaagaaaa ttttaatctt ttctaaaaaa ggctatgaga taggataa

```

(SEQ ID No 14)

图 3 (续)

1 FHYDRNNIAV GADESVVKEA HREVINSSTE GLLLNIDKDI RKILSGYIVE IEDTEGLKEV
61 INDRYDMLNI SSLRQDGKTF IDFKKYNDKL PLYISNPNYK VNVYAVTKEN TIINPSENGD
121 TSTNGIKKIL IFSKKGYEIG

(SEQ ID No 15)

1 ttctcattatg atagaaataa catagcagtt gggggcggatg agtcagtagt taaggaggct
61 catagagaag taattaattc gtcaacagag ggattattgt taaatattga taaggatata
121 agaaaaatat tatcagggtta tattgtagaa attgaagata ctgaagggct taaagaagtt
181 ataaatgaca gatatgatat gttgaatatt tctagttttac ggcaagatgg aaaaacattt
241 atagattttta aaaaatataa tgataaatta ccgttatata taagtaatcc caattataag
301 gtaaagtgtat atgctgtttac taaagaaaac actattatta atcctagtga gaatggggat
361 actagtagca acgggatcaa gaaaatttta atcttttcta aaaaaggcta tgagatagga
421 taa

(SEQ ID No 16)

图 4

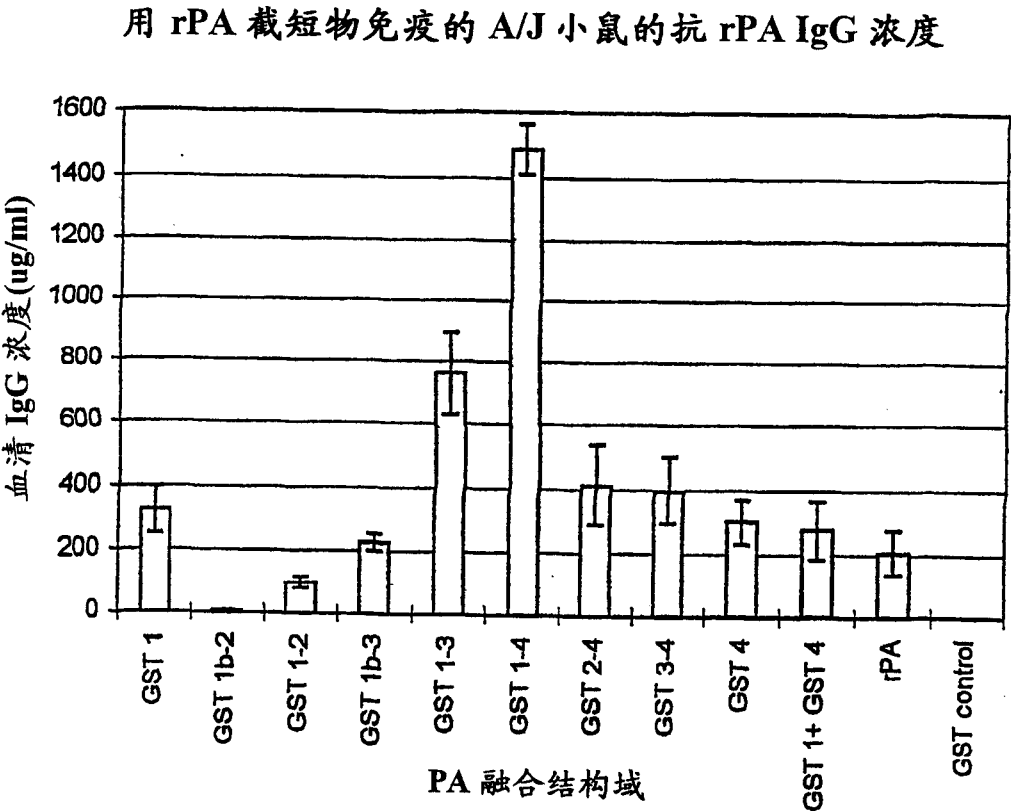


图 5