

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/31

C07K 14/22 A61K 39/095

C07K 16/12

[12] 发明专利申请公开说明书

[21] 申请号 00805181.X

[43] 公开日 2002 年 10 月 16 日

[11] 公开号 CN 1375006A

[22] 申请日 2000.1.19 [21] 申请号 00805181.X

[30] 优先权

[32]1999.1.22 [33]GB [31]9901368.2

[32]1999.1.28 [33]GB [31]9901944.0

[32]1999.1.29 [33]GB [31]9902086.9

[32]1999.2.15 [33]GB [31]9903417.5

[32]1999.2.16 [33]GB [31]9903535.4

[86] 国际申请 PCT/EP00/00428 2000.1.19

[87] 国际公布 WO00/43519 英 2000.7.27

[85] 进入国家阶段日期 2001.9.18

[71] 申请人 史密丝克莱恩比彻姆生物有限公司

地址 比利时里克森萨特

[72] 发明人 J·L·吕勒

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 刘 玥

权利要求书 3 页 说明书 76 页 附图页数 0 页

[54] 发明名称 新型化合物

[57] 摘要

本发明提供 BASB047、BASB054、BASB068 和 BASB069 多肽，编码 BASB047、BASB054、BASB068 和 BASB069 多肽的多核苷酸 以及通过重组技术产生这样的多肽的方法。还提供诊断、预防和治疗用途。

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

权 利 要 求 书

1. 一种分离的多肽, 它包含与选自 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8 的氨基酸序列有至少 85% 同一性的氨基酸序列。
2. 权利要求 1 要求保护的分离多肽, 其中所述氨基酸序列与选自 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8 的氨基酸序列有至少 95% 同一性。
3. 权利要求 1 要求保护的多肽, 所述多肽包含选自 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8 的氨基酸序列。
4. 一种分离的多肽, 它为 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6 或 SEQ ID NO: 8 的多肽。
5. 权利要求 1 到 4 任一项要求保护的多肽的免疫原性片段, 其中所述免疫原性片段的免疫原性活性基本上与 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6 或 SEQ ID NO: 8 的多肽的免疫原性活性相同。
6. 一种分离的多核苷酸或与该分离的多核苷酸互补的核苷酸序列, 所述分离的多核苷酸包含这样的核苷酸序列: 它编码分别在 SEQ ID NO: 2、4、6 或 8 的全长上与 SEQ ID NO: 2、4、6 或 8 的氨基酸序列有至少 85% 同一性的多肽。
7. 一种分离的多核苷酸或与该分离的多核苷酸互补的核苷酸序列, 所述分离的多核苷酸包含这样的核苷酸序列: 它在整个编码区上与编码 SEQ ID NO: 2、4、6 或 8 多肽的核苷酸序列有至少 85% 同一性。
8. 一种分离的多核苷酸或与该分离的多核苷酸互补的核苷酸序列, 所述分离的多核苷酸包含分别在 SEQ ID NO: 1、3、5 或 7 的全长上与 SEQ ID NO: 1、3、5 或 7 的核苷酸序列有至少 85% 同一性的核苷酸序列。
9. 权利要求 6 到 8 任一项要求保护的分离的多核苷酸, 其中与 SEQ ID NO: 1、3、5 或 7 的同一性至少为 95%。

10. 一种分离的多核苷酸, 它包含编码 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6 或 SEQ ID NO: 8 的多肽的核苷酸序列。
11. 一种分离的多核苷酸, 它包含 SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5 或 SEQ ID NO: 7 的多核苷酸。
- 5 12. 一种分离的多核苷酸, 它包含编码 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6 或 SEQ ID NO: 8 的多肽的核苷酸序列, 所述核苷酸序列可以通过在严格杂交条件下使用具有 SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、SEQ ID NO: 7 的序列或其片段的标记探针筛选合适文库而获得。
- 10 13. 一种表达载体或重组活微生物, 它们包含依照权利要求 6 到 12 任一项的分离多核苷酸。
14. 一种包含权利要求 13 的表达载体的宿主细胞或所述宿主细胞的亚细胞部分或膜, 它们表达一种分离多肽, 所述分离多肽包含与选自 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8 的氨基酸序列有至少 85% 同一性的氨基酸序列。
- 15 15. 一种产生多肽的方法, 所述多肽包含与选自 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8 的氨基酸序列有至少 85% 同一性的氨基酸序列, 所述方法包括: 在足以产生所述多肽的条件下培养权利要求 14 的宿主细胞, 然后从培养基中回收所述多肽。
- 20 16. 一种表达权利要求 6 到 12 任一项的多核苷酸的方法, 所述方法包括: 用包含至少一种所述多核苷酸的表达载体转化宿主细胞, 然后在足以表达任何一种所述多核苷酸的条件下培养所述宿主细胞。
17. 一种疫苗组合物, 它包含有效量的权利要求 1 到 5 任一项的多肽以及一种药学上可接受的载体。
- 25 18. 一种疫苗组合物, 它包含有效量的权利要求 6 到 12 任一项的多肽以及一种药学上可接受的载体。
19. 依照权利要求 17 或 18 中任一项的疫苗组合物, 其中所述组合物包含至少一种其它脑膜炎奈瑟氏球菌抗原。

20. 一种抗体，它为权利要求 1 到 5 任一项要求保护的多肽或免疫片段的免疫特异性抗体。

21. 一种诊断脑膜炎奈瑟氏球菌感染的方法，所述方法包括用怀疑受到这样的感染的动物的生物样品，鉴定权利要求 1 到 5 任一项要求保护的多肽或该多肽的免疫特异性抗体的含量。

22. 一种组合物在制备用于在动物体内产生免疫应答的药物中的用途，所述组合物包含免疫有效量的权利要求 1 到 5 任一项要求保护的多肽。

23. 一种组合物在制备用于在动物体内产生免疫应答的药物中的用途，所述组合物包含免疫有效量的权利要求 6 到 12 任一项要求保护的多核苷酸。

24. 一种用于治疗患有脑膜炎奈瑟氏球菌病的病人的治疗组合物，该组合物包含至少一种针对权利要求 1 到 5 的多肽的抗体以及一种合适的药用载体。

15

说明书

新型化合物

5 发明领域

本发明涉及多核苷酸(本文称为“BASB047 多核苷酸”、“BASB054 多核苷酸”、“BASB068 多核苷酸”和“BASB069 多核苷酸”)、由它们编码的多肽(本文分别称为“BASB047”、“BASB054”、“BASB068”和“BASB069”或分别称为“BASB047 多肽”、“BASB054 多肽”、“BASB068 多肽”和“BASB069 多肽”)、用于产生它们的重组材料和方法。另一方面,本发明涉及这样的多肽和多核苷酸的使用方法,包括针对细菌感染的疫苗。又一方面,本发明涉及检测某些病原体感染的诊断测定。

15 发明背景

脑膜炎奈瑟氏球菌(*Neisseria meningitidis*)(脑膜炎球菌)是从人类上呼吸道分离出来的格兰氏阴性菌。它偶尔导致侵入性细菌疾病,如菌血症和脑膜炎。脑膜炎球菌疾病的发生呈现地理季节性差异和年度差异(Schwartz, B., Moore, P. S., Broome, C. V.; Clin. Microbiol. Rev. 2 (增刊), S18-S24, 1989)。温带国家大多数疾病是由B型血清群菌株引起,发生率1-10/100,000/年总人口,有时更高(Kaczmarek, E. B. (1997), Commun. Dis. Rep. Rev. 7: R55-9, 1995; Scholten, R. J. P. M., Bijlmer, H. A., Poolman, J. T.等人, Clin. Infect. Dis. 16: 237-246, 1993; Cruz, C., Pavez, G., Aguilar, E.等人, Epidemiol. Infect. 105: 119-126, 1990)。

中非主要流行A型血清群脑膜炎球菌感染,有时达到1000/100,000/年的水平(Schwartz, B., Moore, P. S., Broome, C. V. Clin. Microbiol. Rev. 2 (增刊), S18-S24, 1989)。整体来看,几乎所有脑膜炎

球菌疾病的病例都是由 A 型、B 型、C 型、W-135 型和 Y 型血清群脑膜炎球菌引起, 并且可获得四价 A、C、W-135、Y 多糖疫苗(Armand, J., Arminjon, F., Maynard, M. C., Lafaix, C., J. Biol. Stand. 10: 335-339, 1982).

5 目前通过将所述多糖疫苗化学缀合到载体蛋白而进行改进(Lieberman, J. M., Chiu, S. S., Wong, V. K.等人, JAMA 275: 1499-1503, 1996).

还没有 B 型血清群疫苗可用, 因为发现 B 型荚膜多糖是没有免疫原性的, 很可能是因为它与宿主成分具有结构相似性(Wyle, F. A., Artenstein, M. S., Brandt, M. L.等人, J. Infect. Dis. 126: 514-522, 1972; Finne, J. M., Leinonen, M., Makela, P. M. Lancet ii.: 355-357, 1983).

许多年来已经开始并进行发展基于脑膜炎球菌外膜的疫苗(de Moraes, J. C., Perkins, B., Carnargo, M. C.等人 Lancet 340: 1074-1078, 1992; Bjune, G., Hoiby, E. A. Gronnesby, J. K.等人 338: 1093-1096, 1991). 这样的疫苗已经证实在较大的儿童(>4 岁)和青少年中具有 57% - 85%的有效率。

在这些疫苗中存在许多细菌外膜成分, 如 PorA、PorB、Rmp、Opc、Opa、FrpB, 这些成分对所观察到的保护的作用仍需要进一步的确定。使用动物抗体或人类抗体已经确定其它细菌外膜成分可能与诱导保护性免疫有关, 如 TbpB 和 NspA (Martin, D., Cadieux, N., Hamel, J., Brodeux, B. R., J. Exp. Med. 185: 1173-1183, 1997; Lissolo, L., Maitre-Wilmotte, C., Dumas, p.等人, Inf. Immun. 63: 884-890, 1995). 保护性免疫的机制将涉及抗体介导的杀菌活性和调理素性吞噬作用(opsonophagocytosis).

25 已经使用一个菌血症模型来组合所有抗体介导的机制(Saukkonen, K., Leinonen, M., Abdillahi, H. Poolman, J. T. Vaccine 7: 325-328, 1989). 一般认为晚期补体成分介导的杀菌机制对于对抗脑膜炎球菌疾病的免疫是关键性的(Ross, S. C., Rosenthal P. J., Berberic, H. M.,

多核苷酸，包括多核糖核苷酸。

多肽

5 本发明的一个方面提供脑膜炎奈瑟氏球菌的多肽，在本文中称为“BASB047”、“BASB054”、“BASB068”和“BASB069”以及“BASB047 多肽”、“BASB054 多肽”、“BASB068 多肽”和“BASB069 多肽”以及它们在生物学上、诊断上、预防上、临床上或治疗上有用的变异体，以及包含它们的组合物。

本发明还提供：

10 (a) 一种分离的多肽，所述多肽包含与 SEQ ID NO: 2 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的氨基酸序列；

15 (b) 一种由分离的多核苷酸编码的多肽，所述多核苷酸包含在 SEQ ID NO: 1 的全长上与 SEQ ID NO: 1 有至少 85%同一性、更优选有至少 90%同一性、更优选有至少 95%同一性、甚至更优选有至少 97-99%同一性或完全相同的核苷酸序列；

20 (c) 一种由包含多核苷酸序列的分离多核苷酸编码的多肽，所述多核苷酸序列编码与 SEQ ID NO: 2 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的多肽。

SEQ ID NO: 2 提供的 BASB043 多肽是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB047 多肽。

25 本发明还提供 BASB047 多肽的免疫原性片段，即 BASB047 多肽的连续部分，所述部分具有与包含 SEQ ID NO: 2 的氨基酸序列的多肽相同或基本相同的免疫原性活性。这就是说，所述片段(如果必要，当缀和到一种载体上时)能够引起识别 BASB047 多肽的免疫反应。这样的免疫原性片段可以包括例如缺乏 N-末端前导序列和/或跨

膜结构域和/或 C-末端锚着结构域的 BASB047 多肽。在一个优选的方面，依照本发明的 BASB047 的免疫原性片段基本包含多肽的所有细胞外结构域，所述多肽在 SEQ ID NO: 2 的全长上与 SEQ ID NO: 2 的多肽具有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性。

本发明还提供：

(a) 一种分离的多肽，所述多肽包含与 SEQ ID NO: 4 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的氨基酸序列；

(b) 一种由分离的多核苷酸编码的多肽，所述分离的多核苷酸包含在 SEQ ID NO: 3 的全长上与 SEQ ID NO: 3 有至少 85%同一性、更优选有至少 90%同一性、更优选有至少 95%同一性、甚至更优选有至少 97-99%同一性或完全相同的多核苷酸序列；

(c) 一种由包含多核苷酸序列的分离多核苷酸编码的多肽，所述多核苷酸序列编码与 SEQ ID NO: 4 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的多肽。

SEQ ID NO: 4 提供的 BASB054 多肽是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB054 多肽。

本发明还提供 BASB054 多肽的免疫原性片段，即 BASB054 多肽的连续部分，所述部分具有与包含 SEQ ID NO: 4 的氨基酸序列的多肽相同或基本相同的免疫原性活性。这就是说，所述片段(如果必要，当缀和到一种载体上时)能够引起识别所述 BASB054 多肽的免疫反应。这样的免疫原性片段可以包括例如缺乏 N-末端前导序列和/或跨膜结构域和/或 C-末端锚着结构域的 BASB054 多肽。在一个优选的方面，依照本发明的 BASB054 的免疫原性片段基本包含多肽的所有细胞外结构域，所述多肽在 SEQ ID NO: 4 的全长上与 SEQ ID NO: 4

的多肽具有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性。

本发明还提供:

5 (a) 一种分离的多肽, 所述多肽包含与 SEQ ID NO: 6 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的氨基酸序列;

10 (b) 一种由分离的多核苷酸编码的多肽, 所述多核苷酸包含在 SEQ ID NO: 5 的全长上与 SEQ ID NO: 5 有至少 85%同一性、更优选有至少 90%同一性、更优选有至少 95%同一性、甚至更优选有至少 97-99%同一性或完全相同的核苷酸序列;

15 (c) 一种由包含多核苷酸序列的分离多核苷酸编码的多肽, 所述多核苷酸序列编码与 SEQ ID NO: 6 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的多肽。

SEQ ID NO: 6 提供的 BASB068 多肽是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB068 多肽。

20 本发明还提供 BASB068 多肽的免疫原性片段, 即 BASB068 多肽的连续部分, 所述部分具有与包含 SEQ ID NO: 6 的氨基酸序列的多肽相同或基本相同的免疫原性活性。这就是说, 所述片段(如果必要, 当缀和到一种载体上时)能够引起识别所述 BASB068 多肽的免疫反应。这样的免疫原性片段可以包括例如缺乏 N-末端前导序列和/或跨膜结构域和/或 C-末端锚着结构域的 BASB068 多肽。在一个优选的方面, 依照本发明的 BASB068 的免疫原性片段基本包含多肽的所有细胞外结构域, 所述多肽在 SEQ ID NO: 6 的全长上与 SEQ ID NO: 6 的多肽具有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性。

本发明还提供:

(a) 一种分离的多肽, 所述多肽包含与 SEQ ID NO: 8 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的氨基酸序列;

5 (b) 一种由分离的多核苷酸编码的多肽, 所述分离的多核苷酸包含在 SEQ ID NO: 7 的全长上与 SEQ ID NO: 7 有至少 85%同一性、更优选有至少 90%同一性、更优选有至少 95%同一性、甚至更优选有至少 97-99%同一性或完全相同的多核苷酸序列;

10 (c) 一种由包含多核苷酸序列的分离多核苷酸编码的多肽, 所述多核苷酸序列编码与 SEQ ID NO: 8 的氨基酸序列有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性或完全相同的多肽。

SEQ ID NO: 8 提供的 BASB069 多肽是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB069 多肽。

15 本发明还提供 BASB069 多肽的免疫原性片段, 即 BASB069 多肽的连续部分, 所述部分具有与包含 SEQ ID NO: 8 的氨基酸序列的多肽相同或基本相同的免疫原性活性。这就是说, 所述片段(如果必要, 当缀和到一种载体上时)能够引起识别所述 BASB069 多肽的免疫反应。这样的免疫原性片段可以包括例如缺乏 N-末端前导序列和/或跨膜结构域和/或 C-末端锚着结构域的 BASB069 多肽。在一个优选的方面, 依照本发明的 BASB069 的免疫原性片段基本包含多肽的所有细胞外结构域, 所述多肽在 SEQ ID NO:8 的全长上与 SEQ ID NO: 8 的多肽具有至少 85%同一性、更优选有至少 90%同一性、甚至更优选有至少 95%同一性、最优选有至少 97-99%同一性。

25 片段是具有与本发明的任何多肽的部分而非全部氨基酸序列完全相同的氨基酸序列的多肽。与 BASB047、BASB054、BASB068 和 BASB069 多肽一样, 片段可以是独立的, 或者包含于更大的多肽内, 在其中它们形成部分或区, 最好是在单个更大多肽中的单个连续区。

5 优选的片段包括例如具有 SEQ ID NO: 2、4、6、8 或它们的变
异体的一部分的截短多肽，例如包括氨基末端氨基酸序列和/或羧基
末端氨基酸序列的连续系列残基。也优选由宿主细胞中产生或在宿
主细胞中产生的本发明多肽的降解形式。还优选结构或功能属性特
征片段，例如包含以下区的片段： α -螺旋和 α -螺旋形成区、 β -折叠和
 β -折叠形成区、转角和转角形成区、螺旋和螺旋形成区、亲水区、疏
水区、 α 两亲区、 β 两亲区、柔性区、表面形成区、底物结合区和高
抗原性指数区。

10 其它优选片段包括包含具有 SEQ ID NO: 2、4、6、8 的氨基酸
序列的至少 15、20、30、40、50 或 100 个连续氨基酸的氨基酸序列
的分离多肽，或者包括具有 SEQ ID NO: 2、4、6、8 的氨基酸序列截
短或缺失的至少 15、20、30、40、50 或 100 个连续氨基酸的氨基酸
序列的分离多肽。

15 可以使用本发明的各多肽片段通过肽合成制备相应全长多肽；
因此，这些片段可以用作产生本发明的全长多肽的中间体。

特别优选其中几个即 5-10 个、1-5 个、1-3 个、1-2 个或 1 个氨
基酸以任何组合取代、缺失或添加的变异体。

20 本发明多肽或免疫原性片段可以是“成熟”蛋白形式，或者可
以是更大的蛋白如前体或融合蛋白的一部分。包括包含以下序列的
额外氨基酸序列通常是有利的：分泌序列或前导序列、原序列(pro-
sequence)、有助于纯化的序列如多组氨酸残基、或在重组产生过程
中有利稳定性的其它序列。此外，也可以考虑添加外源多肽或脂质
尾或多核苷酸序列以增加最终分子的免疫原性潜力。

25 一方面，本发明涉及通过遗传工程制备可溶性融合蛋白，所述
融合蛋白包含本发明多肽或其片段，以及不同免疫球蛋白亚族(IgG、
IgM、IgA、IgE)重链或轻链的恒定区的不同部分。作为免疫球蛋白，
优选的是人类 IgG(特别是 IgG1)的重链的恒定部分，其中在绞链区融
合。在一个特定实施方案中，可以通过加入可用凝血因子 Xa 切除的

切割序列就可除去 Fc 部分。

此外，本发明涉及通过遗传工程制备这些融合蛋白的方法，以及这些融合蛋白在药物筛选、诊断和治疗中的应用。本发明的再一方面还涉及编码这样的融合蛋白的多核苷酸。在国际专利申请号
5 WO94/29458 和 WO94/22914 中可以找到融合蛋白技术实例。

所述蛋白也可以化学缀合或表达为重组融合蛋白，以使得与未融合的蛋白相比，在表达系统中高水平表达。所述融合配偶体可以协助提供 T 辅助细胞表位(免疫融合配偶体)，最好是被人类识别的 T 辅助细胞表位，或者所述配偶体协助以比天然重组蛋白更高产量表达所述蛋白(表达增强物)。最好所述融合配偶体既是免疫融合配偶体，又是表达增强配偶体。
10

融合配偶体包括 B 型流感嗜血杆菌(*Haemophilus influenzae B*)的 D 蛋白以及流感病毒的非结构蛋白 NS1(血凝素)。另一种免疫融合配偶体是称为 LytA 的蛋白。最好使用 LYTA 分子的 C 末端部分。LytA 产生自肺炎链球菌(*Streptococcus pneumoniae*)，它合成 N-乙酰-L-丙氨酸酰胺酶(酰胺酶 LytA) (lytA 基因编码 {Gene, 43(1986)第 265-272 页})即特异性降解肽聚糖主链中某些键的自溶素。LytA 蛋白的 C 末端结构域对胆碱或一些胆碱类似物如 DEAE 具有亲和性。已经利用该特性开发用于表达融合蛋白的大肠杆菌 C-LytA 表达质粒。已经描述了纯化在其氨基末端包含 C-LytA 片段的杂种蛋白 {Biotechnology:
15 10, (1992)第 795-798 页}。可利用存在于 LytA 分子 C 末端的起始于残基 178 的重复部分，例如残基 188-305。

本发明还包括上面提到的多肽的变异体，即通过保守氨基酸取代与参比物不同的多肽，其中一个残基由另一个具有相似特性的残基取代。通常这样的取代发生在 Ala、Val、Leu 和 Ile 之间；Ser 和 Thr 之间；酸性残基 Asp 和 Glu 之间；Asn 和 Gln 之间；碱性残基 Lys 和 Arg 之间；或芳香族残基 Phe 和 Tyr 之间。
25

可以以任何合适的方法制备本发明多肽。这样的多肽包括分离

的天然多肽、重组产生的多肽、合成产生的多肽或通过这些方法的组合产生的多肽。用于制备这样的多肽的方法是本领域众所周知的。

5 最优选本发明的多肽从脑膜炎奈瑟氏球菌获得，然而，最好从同一分类属的其它物种获得。本发明的多肽也可以获自例如同同一分类科或分类目的生物。

多核苷酸

本发明目的之一是提供编码 BASB047 多肽的多核苷酸，尤其是编码本文中命名为 BASB047 的多肽的多核苷酸。

10 在一个本发明尤其优选的实施方案中，所述多核苷酸包含一个编码 BASB047 多肽的区段或其变异体，该区段包含包括一个完整基因的 SEQ ID NO: 1 所示的序列。

SEQ ID NO: 1 提供的 BASB047 多核苷酸是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB047 多核苷酸。

15 本发明的再一方面提供编码和/或表达 BASB047 多肽和多核苷酸、尤其是脑膜炎奈瑟氏球菌的 BASB047 多肽和多核苷酸的分离核酸分子，所述分离的核酸分子包括例如未加工的 RNA、核酶 RNA、mRNA、cDNA、基因组 DNA、B-DNA 和 Z-DNA。本发明的其它实施方案包括在生物学上、诊断上、预防上、临床上或治疗上有用的多核苷酸和多肽，和它们的变异体，以及包含上述多核苷酸、多肽及其变异体的组合物。

本发明的另一方面涉及编码具有 SEQ ID NO: 2 的推导氨基酸序列的 BASB047 多肽的分离多核苷酸以及与其密切相关的多核苷酸及其变异体，所述分离的多核苷酸包括至少一个全长基因。

25 在另一个本发明尤其优选的实施方案中为包含 SEQ ID NO: 2 的氨基酸序列或由 SEQ ID NO: 2 的氨基酸序列组成的脑膜炎奈瑟氏球菌的 BASB047 多肽或其变异体。

使用本文提供的信息，例如 SEQ ID NO: 1 所示的多核苷酸序列，

可以使用标准克隆和筛选方法获得本发明的编码 BASB047 多肽的多核苷酸，所述克隆和筛选方法例如以下这样的方法：使用脑膜炎奈瑟氏球菌细胞作为起始材料，从细菌中克隆和测序染色体 DNA 片段，然后获得全长的克隆。例如，为获得本发明的多核苷酸序列，如 SEQ ID NO: 1 所示的多核苷酸序列，通常用得自某一部分序列的最好 17 单体或更长的放射性标记寡核苷酸，探测在大肠杆菌或一些其它合适宿主中的脑膜炎奈瑟氏球菌的染色体 DNA 克隆文库。然后使用严格杂交条件鉴别带有与所述探针相同 DNA 的克隆。使用根据原始多肽序列或多核苷酸序列设计的测序引物进行杂交而鉴定各个克隆，对克隆进行测序，因此有可能在两个方向延伸所述多核苷酸序列以确定全长的基因序列。例如，比较方便的是使用由质粒克隆制备的变性双链 DNA 进行这样的测序。合适的方法描述于 Maniatis, T., Fritsch, E. F. 和 Sambrook 等人，*MOLECULAR CLONING: A LABORATORY MANUAL*，第二版；Cold Spring Harbor Laboratory Press, Cold Spring harbor, New York (1989)。 (特别参阅杂交筛选 1.90 和测序变性的双链 DNA 模板 13.70)。也可进行直接基因组 DNA 测序以获得全长的基因序列。作为本发明的实例，在从脑膜炎奈瑟氏球菌获得的 DNA 文库中发现了 SEQ ID NO: 1 所示的各多核苷酸。

此外，SEQ ID NO: 1 所示 DNA 序列包含一个可读框，该可读框编码具有 SEQ ID NO: 2 所示氨基酸残基大致数目的蛋白质，该蛋白质的推导分子量可以使用本领域技术人员众所周知的氨基酸残基分子量值计算。

在核苷酸 1 的起始密码子和起始于 SEQ ID NO: 1 的核苷酸 1201 的终止密码子之间的 SEQ ID NO: 1 多核苷酸编码 SEQ ID NO: 2 的多肽。

另一方面，本发明提供分离的多核苷酸，所述多核苷酸包括以下多核苷酸或由以下多核苷酸组成：

(a) 一种在 SEQ ID NO: 1 的全长上与 SEQ ID NO: 1 有至少 85%

同一性、更优选至少 90%同一性、更优选至少 95%同一性、甚至更优选至少 97-99%同一性或完全相同的多核苷酸序列；或

5 (b) 编码一种多肽的多核苷酸序列，所述多肽在 SEQ ID NO: 2 的全长上与 SEQ ID NO: 2 的氨基酸序列有至少 85%同一性、更优选至少 90%同一性、更优选至少 95%同一性、甚至更优选至少 97-99%同一性或 100%相同。

10 编码本发明的多肽的多核苷酸，包括脑膜炎奈瑟氏球菌以外物种的同系物和直向同源物(ortholog)，可以通过包括以下步骤的方法获得：在严格杂交条件下(例如，使用 45-65℃范围温度和 0.1-1% SDS 浓度)，使用包含或由 SEQ ID NO: 1 的序列或其片段组成的标记探针或可检测探针筛选合适文库；然后分离包含所述多核苷酸序列的全长基因和/或基因组克隆。

15 本发明提供在其全长上与 SEQ ID NO: 1 中的编码序列(可读框)相同的多核苷酸序列。本发明还提供编码成熟多肽或其片段的编码序列本身以及在可读框中还有另一编码序列的编码成熟多肽或片段的编码区，所述另一编码序列如编码分泌序列或前导序列、前蛋白(pre-protein)序列、原蛋白(pro-protein)序列或前原蛋白(prepro-protein)序列的序列。本发明的多核苷酸还可以包含至少一个非编码序列，包括例如但不限于至少一个非编码 5'序列和 3'序列，如转录但不翻译的序列、终止信号(例如依赖于 rho 的终止信号和不依赖于 rho 的终止信号)、核糖体结合位点、Kozak 序列、稳定 mRNA 的序列、内含子和聚腺苷酸化信号。所述多核苷酸序列也可以包括编码额外氨基酸的额外编码序列。例如，可以编码有助于纯化融合多肽的标记序列。在本发明的某些实施方案中，所述标记序列为 pQE 载体(Qiagen, 20 Inc.)提供并在 Gentz 等人, *Proc. Natl. Acad. Sci. USA* 86: 821-824 (1989) 中描述的六组氨基酸，或者 HA 肽标记(Wilson 等人, *Cell* 37: 767 (1984))，这两种标记可以用于纯化与其融合的多肽序列。本发明的多核苷酸还包括但不限于包含结构基因和控制基因表达的与其天然

结合的序列的多核苷酸。

5 编码 SEQ ID NO: 2 的 BASB047 多肽的核苷酸序列可以与 SEQ ID NO: 1 的核苷酸 1 到 1200 包含的多肽编码序列相同。或者它可以是由于遗传密码的丰余性(简并性)也编码 SEQ ID NO: 2 的多肽的序列。本文所用的术语“编码多肽的多核苷酸”包括包含编码本发明的多肽的序列的多核苷酸, 所述多肽优选为细菌多肽, 更优选具有 SEQ ID NO: 2 所示氨基酸序列的脑膜炎奈瑟氏球菌 BASB047 多肽。该术语也包括这样的多核苷酸: 所述多核苷酸包含编码所述多肽的单个连续区或多个不连续区(例如, 整合噬菌体、整合插入序列、整合载体序列、整合转座子序列或由于 RNA 编辑或基因组 DNA 重建而间断的多核苷酸)以及额外的区段, 所述额外的区段也可包含编码序列和/或非编码序列。

10 本发明还涉及本文描述的多核苷酸的变体, 所述变体编码具有 SEQ ID NO: 2 的推导氨基酸序列的多肽的变体。本发明的多核苷酸的片段可以用于例如合成本发明的全长多核苷酸。

15 其它特别优选的实施方案为编码 BASB047 变异体的多核苷酸, 所述 BASB047 变体具有 SEQ ID NO: 2 的 BASB047 多肽的氨基酸序列, 其中若干个、几个、5 到 10 个、1 到 5 个、1 到 3 个、2 个、1 个或没有氨基酸残基以任何组合发生取代、修饰、缺失和/或添加。其中特别优选不改变 BASB047 多肽的特性和活性的沉默取代、添加和缺失。

20 本发明其它优选的实施方案为在它们的全长上与编码具有 SEQ ID NO: 2 所示氨基酸序列的 BASB047 多肽的多核苷酸至少 85% 相同的多核苷酸, 以及与这样的多核苷酸互补的多核苷酸。在该方面, 在它们的全长上与上面所述多核苷酸至少 90% 相同的多核苷酸是特别优选的, 在这些特别优选的多核苷酸中, 至少 95% 相同的多核苷酸是尤其优选的。此外, 在至少 95% 相同的多核苷酸中, 至少 97% 相同的多核苷酸是高度优选的, 其中至少 98% 和至少 99% 相同的多

核苷酸是尤其高度优选的，而至少 99% 相同的多核苷酸是更优选的。

优选的实施方案为编码基本保留与 SEQ ID NO: 1 的 DNA 编码的成熟多肽同样的生物功能或活性的多肽的多核苷酸。

5 依照本发明某些优选实施方案，提供与 BASB047 多核苷酸序列杂交、尤其是在严格条件下杂交的多核苷酸，例如 SEQ ID NO: 1 的各多核苷酸。

10 本发明还涉及与本文提供的多核苷酸序列杂交的多核苷酸。在这一方面，本发明尤其涉及在严格条件下与本文描述的多核苷酸杂交的多核苷酸。本文所用的术语“严格条件”和“严格杂交条件”是指仅当序列间存在至少 95% 同一性、最好至少 97% 同一性时才发生杂交。严格杂交条件的具体实例是在包含下列组分的溶液中于 42 °C 温育过夜：50% 甲酰胺，5x SSC (150 mM NaCl, 15 mM 柠檬酸三钠)，50 mM 磷酸钠 (pH 7.6)，5x Denhardt 氏溶液，10% 葡聚糖硫酸酯以及 20 微克/ml 变性剪切鲑鱼精子 DNA；然后在约 65 °C 下在 0.1 x SSC 中洗涤杂交支持物。杂交和洗涤条件是众所周知的，并在下面文献中有范例：Sambrook 等人, Molecular Cloning: A Laboratory Manual, 第 15 二版, Cold Spring Harbor, N. Y., (1989)，特别是其中的第 11 章。溶液杂交也可以用于本发明提供的多核苷酸序列。

20 本发明还提供包含或由多核苷酸序列组成的多核苷酸，所述多核苷酸序列如下获得：用具有 SEQ ID NO: 1 所示的所述多核苷酸序列或其片段的序列的探针在严格杂交条件下筛选包括所述完整基因的合适文库中的 SEQ ID NO: 1 所示的多核苷酸序列；然后分离所述多核苷酸序列。用于获得这样的多核苷酸的片段包括例如在本文其它地方充分描述的探针和引物。

25 如本文中其它地方关于本发明的多核苷酸测定所述，例如本发明的多核苷酸可以用作分离编码 BASB047 的全长 cDNA 和基因组克隆的针对 RNA、cDNA 和基因组 DNA 的杂交探针，或者可以用作分离与 BASB047 基因具有高度同一性、特别高的序列同一性的其它基

因的 cDNA 和基因组克隆的针对 RNA、cDNA 和基因组 DNA 的杂交探针。这样的探针一般具有至少 15 个核苷酸残基或碱基对。优选这样的探针具有至少 30 个核苷酸残基或碱基对，并可以具有至少 50 个核苷酸残基或碱基对。尤其优选的探针具有至少 20 个核苷酸残基或碱基对而且少于 30 个核苷酸残基或碱基对。

使用 SEQ ID NO: 1 提供的 DNA 序列进行筛选可以分离 BASB047 基因的编码区，以合成寡核苷酸探针。然后使用具有与本发明的基因的序列互补的序列的标记寡核苷酸，筛选 cDNA 文库、基因组 DNA 文库或 mRNA 文库，以确定所述探针结合所述文库中的哪些成员。

本发明目的之一是提供编码 BASB054 多肽的多核苷酸，尤其是编码本文中命名为 BASB054 的多肽的多核苷酸。

在一个本发明尤其优选的实施方案中，所述多核苷酸包含一个编码 BASB054 多肽的区段或其变异体，该区段包含包括一个完整的基因的 SEQ ID NO: 3 所示序列。

SEQ ID NO: 3 提供的 BASB054 多核苷酸是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB054 多核苷酸。

本发明的再一方面提供编码和/或表达 BASB054 多肽和多核苷酸、尤其是脑膜炎奈瑟氏球菌 BASB054 多肽和多核苷酸的分离的核酸分子，所述分离的核酸分子包括例如未加工的 RNA、核酶 RNA、mRNA、cDNA、基因组 DNA、B-DNA 和 Z-DNA。本发明的其它实施方案包括在生物学上、诊断上、预防上、临床上或治疗上有用的多核苷酸和多肽和它们的变异体，以及包含上述多核苷酸、多肽及其变异体的组合物。

本发明的再一方面涉及编码具有 SEQ ID NO: 4 的推导氨基酸序列的 BASB054 多肽的分离多核苷酸和与之密切相关的多核苷酸以及它们的变异体，所述分离的多核苷酸包括至少一个全长基因。

本发明另一个尤其优选的实施方案为包含或由 SEQ ID NO: 4 的氨基酸序列组成的脑膜炎奈瑟氏球菌的 BASB054 多肽或其变异体。

使用本文提供的信息，例如 SEQ ID NO: 3 所示的多核苷酸序列，可以使用标准克隆和筛选方法获得本发明的编码 BASB054 多肽的多核苷酸，所述克隆和筛选方法例如以下这样的方法：使用脑膜炎奈瑟氏球菌细胞作为起始材料，从细菌中克隆和测序染色体 DNA 片段，
5 然后获得全长的克隆。例如，为了获得本发明的多核苷酸序列，如 SEQ ID NO: 3 所示的多核苷酸序列，通常用得自某一部分序列的最好 17 单体或更长的放射性标记寡核苷酸，探测在大肠杆菌或一些其它合适宿主中的脑膜炎奈瑟氏球菌的染色体 DNA 克隆文库。然后使用严格杂交条件鉴别带有与所述探针相同 DNA 的克隆。使用根据原始多肽序列或多核苷酸序列设计的测序引物进行杂交而鉴定各个克隆，
10 对克隆进行测序，因此有可能在两个方向延伸所述多核苷酸序列以确定全长基因序列。例如，比较方便的是使用由质粒克隆制备的变性双链 DNA 进行这样的测序。合适的方法描述于 Maniatis, T., Fritsch, E. F. 和 Sambrook 等人, *MOLECULAR CLONING: A LABORATORY MANUAL*, 第二版; Cold Spring Harbor Laboratory Press, Cold Spring harbor, New York (1989)。 (特别参阅杂交筛选 1.90 和测序变性的双链 DNA 模板 13.70)。也可进行直接基因组 DNA 测序以获得全长的基因序列。作为本发明的实例，由脑膜炎奈瑟氏球菌获得的 DNA 文库中发现了 SEQ ID NO: 3 所示的各多核苷酸。

20 此外， SEQ ID NO: 3 所示 DNA 序列包含一个可读框，该可读框编码具有 SEQ ID NO: 4 所示的大致氨基酸残基数目的蛋白质，该蛋白质的推导分子量可以使用本领域技术人员众所周知的氨基酸残基分子量值计算。

25 在核苷酸 1 的起始密码子和起始于 SEQ ID NO: 3 的核苷酸 2407 的终止密码子之间的 SEQ ID NO: 3 多核苷酸编码 SEQ ID NO: 4 的多肽。

再一方面，本发明提供分离的多核苷酸，所述多核苷酸包括以下多核苷酸序列或由以下多核苷酸序列组成：

(a) 在 SEQ ID NO: 3 的全长上与 SEQ ID NO: 3 有至少 85%同一性、更优选至少 90%同一性、更优选至少 95%同一性、甚至更优选至少 97-99%同一性或完全相同的多核苷酸序列；或

5 (b) 一种编码多肽的多核苷酸序列，所述多肽在 SEQ ID NO: 4 的全长上与 SEQ ID NO: 4 的氨基酸序列有至少 85%同一性、更优选至少 90%同一性、更优选至少 95%同一性、甚至更优选至少 97-99%同一性或 100%相同。

10 编码本发明的多肽的多核苷酸，包括脑膜炎奈瑟氏球菌以外物种的同系物和直向同源物，可以通过包括以下步骤的方法获得：在严格杂交条件下(例如，使用 45-65°C 范围温度和 0.1-1% SDS 浓度)，使用包含或由 SEQ ID NO: 3 的序列或其片段组成的标记探针或可检测探针筛选合适文库；然后分离包含所述多核苷酸序列的全长基因和/或基因组克隆。

15 本发明提供在其全长上与 SEQ ID NO: 3 中的编码序列(可读框)相同的多核苷酸序列。本发明还提供编码成熟多肽或其片段的编码序列本身以及在可读框中还有另一编码序列的编码成熟多肽或片段的编码序列，所述另一编码序列例如编码分泌序列或前导序列、前蛋白(pre-protein)序列、原蛋白(pro-protein)序列或前原蛋白(prepro-protein)序列的序列。本发明的多核苷酸还可以包含至少一个非编码
20 序列，包括例如但不限于至少一个非编码 5'序列和 3'序列，如转录但不翻译的序列、终止信号(例如依赖于 rho 的终止信号和不依赖于 rho 的终止信号)、核糖体结合位点、Kozak 序列、稳定 mRNA 的序列、内含子和聚腺苷酸化信号。所述多核苷酸序列也可以包括编码额外氨基酸的额外编码序列。例如，可以编码有助于纯化融合多肽的标记
25 序列。在本发明的某些实施方案中，所述标记序列为 pQE 载体(Qiagen, Inc.)提供并在 Gentz 等人, *Proc. Natl. Acad. Sci. USA* 86: 821-824 (1989)中描述的六组氨酸肽, 或者 HA 肽标记(Wilson 等人, *Cell* 37: 767 (1984)), 这两种标记可以用于纯化与其融合的多肽序列。本发明

的多核苷酸还包括但不限于包含结构基因和控制基因表达的与其天然结合的序列的多核苷酸。

5 编码 SEQ ID NO: 4 的 BASB054 多肽的核苷酸序列可以与 SEQ ID NO: 3 的核苷酸 1 到 2406 包含的多肽编码序列相同。或者它可以是由于遗传密码的冗余性(简并性)也编码 SEQ ID NO: 4 的多肽的序列。本文所用的术语“编码多肽的多核苷酸”包括包含编码本发明的多肽的序列的多核苷酸, 所述多肽优选为细菌多肽, 更优选具有 SEQ ID NO: 4 所示氨基酸序列的脑膜炎奈瑟氏球菌 BASB054 多肽。该术语也包括这样的多核苷酸: 所述多核苷酸包含编码所述多肽的
10 单个连续区或多个不连续区(例如被整合噬菌体、整合插入序列、整合载体序列、整合转座子序列或由于 RNA 编辑或基因组 DNA 重组而间断的多核苷酸)以及额外的区段, 所述额外的区段也可包含编码序列和/或非编码序列。

15 本发明还涉及本文描述的多核苷酸的变体, 所述变体编码具有 SEQ ID NO: 4 的推导氨基酸序列的多肽的变体。本发明的多核苷酸的片段可以用于例如合成本发明的全长多核苷酸。

其它特别优选的实施为编码 BASB054 变异体的多核苷酸, 所述 BASB054 变体具有 SEQ ID NO: 4 的 BASB054 多肽的氨基酸序列, 其中若干个、几个、5 到 10 个、1 到 5 个、1 到 3 个、2 个、
20 1 个或没有氨基酸残基以任何组合发生取代、修饰、缺失和/或添加。其中特别优选不改变 BASB054 多肽的特性和活性的沉默取代、添加和缺失。

本发明其它优选的实施为在它们的全长上与编码具有 SEQ ID NO: 4 所示氨基酸序列的 BASB054 多肽的多核苷酸至少 85% 相同的
25 的多核苷酸, 以及与这样的多核苷酸互补的多核苷酸。在这一方面, 在它们的全长上与上面所述多核苷酸至少 90% 相同的多核苷酸是特别优选的, 在这些特别优选的多核苷酸中, 至少 95% 相同的多核苷酸是尤其优选的。此外, 在至少 95% 相同的多核苷酸中, 至少 97%

相同的多核苷酸是高度优选的，其中至少 98%和至少 99%相同的多核苷酸是尤其高度优选的，而至少 99%相同的多核苷酸是更优选的。

优选的实施方案为编码基本保留与 SEQ ID NO: 3 的 DNA 编码的成熟多肽同样的生物功能或活性的多肽的多核苷酸。

5 依照本发明某些优选实施方案，提供与 BASB054 多核苷酸序列杂交、尤其是在严格条件下杂交的多核苷酸，例如 SEQ ID NO: 3 的各多核苷酸。

10 本发明还涉及与本文提供的多核苷酸序列杂交的多核苷酸。在这一方面，本发明尤其涉及在严格条件下与本文描述的多核苷酸杂交的多核苷酸。本文所用的术语“严格条件”和“严格杂交条件”是指仅当序列间存在至少 95%同一性、最好至少 97%同一性时才发生杂交。严格杂交条件的具体实例是在包含下列组分的溶液中于 42 °C 温育过夜：50%甲酰胺，5x SSC (150 mM NaCl, 15 mM 柠檬酸三钠)，50 mM 磷酸钠(pH 7.6)，5x Denhardt 氏溶液，10%葡聚糖硫酸酯
15 以及 20 微克/ml 变性剪切鲑鱼精子 DNA；然后在约 65 °C 下在 0.1 x SSC 中洗涤杂交支持物。杂交和洗涤条件是众所周知的，并在下面文献中有范例：Sambrook 等人, Molecular Cloning: A Laboratory Manual, 第二版, Cold Spring Harbor, N. Y., (1989)，特别是其中的第 11 章。溶液杂交也可用于本发明提供的多核苷酸序列。

20 本发明还提供包含或由多核苷酸序列组成的多核苷酸，所述多核苷酸序列如下获得：用具有 SEQ ID NO: 3 所示的所述多核苷酸序列或其片段的序列的探针在严格杂交条件下筛选包括 SEQ ID NO: 3 所示多核苷酸序列的完整基因的合适文库；然后分离所述多核苷酸序列。用于获得这样的多核苷酸的片段包括例如在本文其它地方充
25 分描述的探针和引物。

如本文中其它地方关于本发明的多核苷酸测定所述，例如本发明的多核苷酸可以用作针对 RNA、cDNA 和基因组 DNA 的杂交探针，以分离编码 BASB054 的全长 cDNA 和基因组克隆的或者分离与

BASB054 基因具有高度同一性、特别高度序列同一性的其它基因的 cDNA 和基因组克隆。这样的探针一般具有至少 15 个核苷酸残基或碱基对。优选这样的探针具有至少 30 个核苷酸残基或碱基对，并可以具有至少 50 个核苷酸残基或碱基对。尤其优选的探针具有至少 20 个核苷酸残基或碱基对而且少于 30 个核苷酸残基或碱基对。

使用 SEQ ID NO: 3 提供的 DNA 序列进行筛选可以分离 BASB047 基因的编码区，以合成寡核苷酸探针。然后使用具有与本发明的基因的序列互补的序列的标记寡核苷酸，筛选 cDNA 文库、基因组 DNA 文库或 mRNA 文库，以确定所述探针结合所述文库中的哪些成员。

本发明目的之一是提供编码 BASB068 多肽的多核苷酸，尤其是编码本文中命名为 BASB068 的多肽的多核苷酸。

在一个本发明尤其优选的实施方案中，所述多核苷酸包含一个编码 BASB068 多肽的区段或其变异体，该区段包含包括一个完整的基因的 SEQ ID NO: 5 所示序列。

SEQ ID NO: 5 提供的 BASB068 多核苷酸是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB068 多核苷酸。

本发明的再一方面提供编码和/或表达 BASB068 多肽和多核苷酸、尤其是脑膜炎奈瑟氏球菌 BASB068 多肽和多核苷酸的分离的核酸分子，所述分离的核酸分子包括例如未加工的 RNA、核酶 RNA、mRNA、cDNA、基因组 DNA、B-DNA 和 Z-DNA。本发明的其它实施方案包括在生物学上、诊断上、预防上、临床上或治疗上有用的多核苷酸和多肽和它们的变异体，以及包含上述多核苷酸、多肽及其变异体的组合物。

本发明的再一方面涉及编码具有 SEQ ID NO: 6 的推导氨基酸序列的 BASB068 多肽的分离多核苷酸和与之密切相关的多核苷酸以及它们的变异体，所述分离的多核苷酸包括至少一个全长基因。

本发明另一个尤其优选的实施方案为包含或由 SEQ ID NO: 6 的氨基酸序列组成的脑膜炎奈瑟氏球菌的 BASB068 多肽或其变异体。

使用本文提供的信息，例如 SEQ ID NO: 5 所示的多核苷酸序列，可以使用标准克隆和筛选方法获得本发明的编码 BASB068 多肽的多核苷酸，所述克隆和筛选方法例如以下这样的方法：使用脑膜炎奈瑟氏球菌细胞作为起始材料，从细菌中克隆和测序染色体 DNA 片段，
5 然后获得全长的克隆。例如，为了获得本发明的多核苷酸序列，如 SEQ ID NO: 5 给出的多核苷酸序列，通常用得自某一部分序列的最好 17 单体或更长的放射性标记寡核苷酸，探测在大肠杆菌或一些其它合适宿主中的脑膜炎奈瑟氏球菌的染色体 DNA 克隆文库。然后使用严格杂交条件鉴别带有与所述探针相同 DNA 的克隆。使用根据原始多肽序列或多核苷酸序列设计的测序引物进行杂交而鉴定各个克隆，
10 对克隆进行测序，因此有可能在两个方向延伸所述多核苷酸序列以确定全长基因序列。例如，比较方便的是使用由质粒克隆制备的变性双链 DNA 进行这样的测序。合适的方法描述于 Maniatis, T., Fritsch, E. F. 和 Sambrook 等人, *MOLECULAR CLONING: A LABORATORY MANUAL*, 第二版; Cold Spring Harbor Laboratory Press, Cold Spring harbor, New York (1989)。 (特别参阅杂交筛选 1.90 和测序变性的双链 DNA 模板 13.70)。也可进行直接基因组 DNA 测序以获得全长的基因序列。作为本发明的实例，由脑膜炎奈瑟氏球菌获得的 DNA 文库中发现了 SEQ ID NO: 5 所示的各多核苷酸。

20 此外， SEQ ID NO: 5 所示 DNA 序列包含一个可读框，该可读框编码具有 SEQ ID NO: 6 所示大致氨基酸残基数目的蛋白质，该蛋白质的推导分子量可以使用本领域技术人员众所周知的氨基酸残基分子量值计算。

在核苷酸 1 的起始密码子和起始于 SEQ ID NO: 5 的核苷酸 2014 的终止密码子之间的 SEQ ID NO: 5 多核苷酸编码 SEQ ID NO: 6 的多肽。
25

再一方面，本发明提供分离的多核苷酸，所述多核苷酸包括以下多核苷酸序列或由以下多核苷酸序列组成：

(a) 在 SEQ ID NO: 5 的全长上与 SEQ ID NO: 5 有至少 85% 同一性、更优选至少 90% 同一性、更优选至少 95% 同一性、甚至更优选至少 97-99% 同一性或完全相同的多核苷酸序列；或

5 (b) 一种编码多肽的多核苷酸序列，所述多肽在 SEQ ID NO: 6 的全长上与 SEQ ID NO: 6 的氨基酸序列有至少 85% 同一性、更优选至少 90% 同一性、更优选至少 95% 同一性、甚至更优选至少 97-99% 同一性或 100% 相同。

10 编码本发明的多肽的多核苷酸，包括脑膜炎奈瑟氏球菌以外物种的同系物和直向同源物，可以通过包括以下步骤的方法获得：在严格杂交条件下(例如，使用 45-65°C 范围温度和 0.1-1% SDS 浓度)，使用包含或由 SEQ ID NO: 5 的序列或其片段组成的标记探针或可检测探针筛选合适文库；然后分离包含所述多核苷酸序列的全长基因和/或基因组克隆。

15 本发明提供在其全长上与 SEQ ID NO: 5 中的编码序列(可读框)相同的多核苷酸序列。本发明还提供编码成熟多肽或其片段的编码序列本身以及在可读框中还有另一编码序列的编码成熟多肽或片段的编码序列，所述另一编码序列例如编码分泌序列或前导序列、前蛋白(pre-protein)序列、原蛋白(pro-protein)序列或前原蛋白(prepro-protein)序列的序列。本发明的多核苷酸还可以包含至少一个非编码序列，包括例如但不限于至少一个非编码 5' 序列和 3' 序列，如转录但不翻译的序列、终止信号(例如依赖于 rho 的终止信号和不依赖于 rho 的终止信号)、核糖体结合位点、Kozak 序列、稳定 mRNA 的序列、内含子和聚腺苷酸化信号。所述多核苷酸序列也可以包括编码额外氨基酸的额外编码序列。例如，可以编码有助于纯化融合多肽的标记序列。在本发明的某些实施方案中，所述标记序列为 pQE 载体(Qiagen, Inc.)提供并在 Gentz 等人, *Proc. Natl. Acad. Sci. USA* 86: 821-824 (1989) 中描述的六组氨酸肽，或者 HA 肽标记(Wilson 等人, *Cell* 37: 767 (1984))，这两种标记可以用于纯化与其融合的多肽序列。本发明

20

25

的多核苷酸还包括但不限于包含结构基因和控制基因表达的与其天然结合的序列的多核苷酸。

5 编码 SEQ ID NO: 6 的 BASB068 多肽的核苷酸序列可以与 SEQ ID NO: 5 的核苷酸 1 到 2013 包含的多肽编码序列相同。或者它可以是由于遗传密码的丰余性(简并性)也编码 SEQ ID NO: 6 的多肽的序列。本文所用的术语“编码多肽的多核苷酸”包括包含编码本发明的多肽的序列的多核苷酸, 所述多肽优选为细菌多肽, 更优选具有 SEQ ID NO: 6 所示氨基酸序列的脑膜炎奈瑟氏球菌 BASB068 多肽。该术语也包括这样的多核苷酸: 所述多核苷酸包含编码所述多肽的
10 单个连续区或多个不连续区(例如被整合噬菌体、整合插入序列、整合载体序列、整合转座子序列或由于 RNA 编辑或基因组 DNA 重组而间断的多核苷酸)以及额外的区段, 所述额外的区段也可包含编码序列和/或非编码序列。

15 本发明还涉及本文描述的多核苷酸的变体, 所述变体编码具有 SEQ ID NO: 6 的推导氨基酸序列的多肽的变体。本发明的多核苷酸的片段可以用于例如合成本发明的全长多核苷酸。

其它特别优选的实施为编码 BASB068 变异体的多核苷酸, 所述 BASB068 变体具有 SEQ ID NO: 6 的 BASB068 多肽的氨基酸序列, 其中若干个、几个、5 到 10 个、1 到 5 个、1 到 3 个、2 个、
20 1 个或没有氨基酸残基以任何组合发生取代、修饰、缺失和/或添加。其中特别优选不改变 BASB068 多肽的特性和活性的沉默取代、添加和缺失。

本发明其它优选的实施为在它们的全长上与编码具有 SEQ ID NO: 6 所示氨基酸序列的 BASB068 多肽的多核苷酸至少 85% 相同的
25 的多核苷酸, 以及与这样的多核苷酸互补的多核苷酸。在这一方面, 在它们的全长上与上面所述多核苷酸至少 90% 相同的多核苷酸是特别优选的, 在这些特别优选的多核苷酸中, 至少 95% 相同的多核苷酸是尤其优选的。此外, 在至少 95% 相同的多核苷酸中, 至少 97%

相同的多核苷酸是高度优选的，其中至少 98%和至少 99%相同的多核苷酸是尤其高度优选的，而至少 99%相同的多核苷酸是更优选的。

优选的实施实施方案为编码基本保留与 SEQ ID NO: 5 的 DNA 编码的成熟多肽同样的生物功能或活性的多肽的多核苷酸。

5 依照本发明某些优选实施方案，提供与 BASB068 多核苷酸序列杂交、尤其是在严格条件下杂交的多核苷酸，例如 SEQ ID NO: 5 的各多核苷酸。

10 本发明还涉及与本文提供的多核苷酸序列杂交的多核苷酸。在这一方面，本发明尤其涉及在严格条件下与本文描述的多核苷酸杂交的多核苷酸。本文所用的术语“严格条件”和“严格杂交条件”是指仅当序列间存在至少 95%同一性、最好至少 97%同一性时才发生杂交。严格杂交条件的具体实例是在包含下列组分的溶液中于 42℃ 温育过夜：50%甲酰胺，5x SSC (150 mM NaCl, 15 mM 柠檬酸三钠)，50 mM 磷酸钠(pH 7.6)，5x Denhardt 氏溶液，10%葡聚糖硫酸酯以及 20 微克/ml 变性剪切鲑鱼精子 DNA；然后在约 65℃ 下在 0.1 x SSC 中洗涤杂交支持物。杂交和洗涤条件是众所周知的，并在下面文献中有范例：Sambrook 等人，Molecular Cloning: A Laboratory Manual, 第二版，Cold Spring Harbor, N. Y., (1989)，特别是其中的第 11 章。溶液杂交也可用于本发明提供的多核苷酸序列。

20 本发明还提供包含或由多核苷酸序列组成的多核苷酸，所述多核苷酸序列如下获得：用具有 SEQ ID NO: 5 所示的所述多核苷酸序列或其片段的序列的探针在严格杂交条件下筛选包括 SEQ ID NO: 5 所示多核苷酸序列的完整基因的合适文库；然后分离所述多核苷酸序列。用于获得这样的多核苷酸的片段包括例如在本文其它地方充分描述的探针和引物。

25 如本文中其它地方关于本发明的多核苷酸测定所述，例如本发明的多核苷酸可以用作针对 RNA、cDNA 和基因组 DNA 的杂交探针，以分离编码 BASB068 的全长 cDNA 和基因组克隆的或者分离与

BASB068 基因具有高度同一性、特别高度序列同一性的其它基因的 cDNA 和基因组克隆。这样的探针一般具有至少 15 个核苷酸残基或碱基对。优选这样的探针具有至少 30 个核苷酸残基或碱基对，并可以具有至少 50 个核苷酸残基或碱基对。尤其优选的探针具有至少 20 个核苷酸残基或碱基对而且少于 30 个核苷酸残基或碱基对。

使用 SEQ ID NO: 5 提供的 DNA 序列进行筛选可以分离 BASB068 基因的编码区，以合成寡核苷酸探针。然后使用具有与本发明的基因的序列互补的序列的标记寡核苷酸，筛选 cDNA 文库、基因组 DNA 文库或 mRNA 文库，以确定所述探针结合所述文库中的哪些成员。

本发明目的之一是提供编码 BASB069 多肽的多核苷酸，尤其是编码本文中命名为 BASB069 的多肽的多核苷酸。

在一个尤其优选的本发明实施方案中，所述多核苷酸包含一个编码 BASB069 多肽的区段或其变体，该区段包含包括一个完整的基因的 SEQ ID NO: 7 所示序列。

SEQ ID NO: 7 提供的 BASB069 多核苷酸是脑膜炎奈瑟氏球菌菌株 ATCC13090 的 BASB069 多核苷酸。

本发明的再一方面提供编码和/或表达 BASB069 多肽和多核苷酸、尤其是脑膜炎奈瑟氏球菌 BASB069 多肽和多核苷酸的分离的核酸分子，所述分离的核酸分子包括例如未加工的 RNA、核酶 RNA、mRNA、cDNA、基因组 DNA、B-DNA 和 Z-DNA。本发明的其它实施方案包括在生物学上、诊断上、预防上、临床上或治疗上有用的多核苷酸和多肽和它们的变体，以及包含上述多核苷酸、多肽及其变异体的组合物。

本发明的再一方面涉及编码具有 SEQ ID NO: 8 的推导氨基酸序列的 BASB069 多肽的分离多核苷酸和与之密切相关的多核苷酸以及它们的变体，所述分离的多核苷酸包括至少一个全长基因。

本发明另一个尤其优选的实施方案为包含或由 SEQ ID NO: 8 的氨基酸序列组成的脑膜炎奈瑟氏球菌的 BASB069 多肽或其变体。

使用本文提供的信息，例如 SEQ ID NO: 7 所示的多核苷酸序列，可以使用标准克隆和筛选方法获得本发明的编码 BASB069 多肽的多核苷酸，所述克隆和筛选方法例如以下这样的方法：使用脑膜炎奈瑟氏球菌细胞作为起始材料，从细菌中克隆和测序染色体 DNA 片段，
5 然后获得全长的克隆。例如，为了获得本发明的多核苷酸序列，如 SEQ ID NO: 7 给出的多核苷酸序列，通常用得自某一部分序列的最好 17 单体或更长的放射性标记寡核苷酸，探测在大肠杆菌或一些其它合适宿主中的脑膜炎奈瑟氏球菌的染色体 DNA 克隆文库。然后使用严格杂交条件鉴别带有与所述探针相同 DNA 的克隆。使用根据原始多肽序列或多核苷酸序列设计的测序引物进行杂交而鉴定各个克隆，
10 对克隆进行测序，因此有可能在两个方向延伸所述多核苷酸序列以确定全长基因序列。例如，比较方便的是使用由质粒克隆制备的变性双链 DNA 进行这样的测序。合适的方法描述于 Maniatis, T., Fritsch, E. F. 和 Sambrook 等人, *MOLECULAR CLONING: A LABORATORY MANUAL*, 第二版; Cold Spring Harbor Laboratory Press, Cold Spring harbor, New York (1989)。 (特别参阅杂交筛选 1.90 和测序变性的双链 DNA 模板 13.70)。也可进行直接基因组 DNA 测序以获得全长的基因序列。作为本发明的实例，由脑膜炎奈瑟氏球菌获得的 DNA 文库中发现了 SEQ ID NO: 7 所示的各多核苷酸。

20 此外， SEQ ID NO: 7 所示 DNA 序列包含一个可读框，该可读框编码具有 SEQ ID NO: 8 所示大致氨基酸残基数目的蛋白质，该蛋白质的推导分子量可以使用本领域技术人员众所周知的氨基酸残基分子量值计算。

在核苷酸 1 的起始密码子和起始于 SEQ ID NO: 7 的核苷酸 2014
25 的终止密码子之间的 SEQ ID NO: 7 多核苷酸编码 SEQ ID NO: 8 的多肽。

再一方面，本发明提供分离的多核苷酸，所述多核苷酸包括以下多核苷酸序列或由以下多核苷酸序列组成：

(a) 在 SEQ ID NO: 7 的全长上与 SEQ ID NO: 7 有至少 85%同一性、更优选至少 90%同一性、更优选至少 95%同一性、甚至更优选至少 97-99%同一性或完全相同的多核苷酸序列; 或

5 (b) 一种编码多肽的多核苷酸序列, 所述多肽在 SEQ ID NO: 8 的全长上与 SEQ ID NO: 8 的氨基酸序列有至少 85%同一性、更优选至少 90%同一性、更优选至少 95%同一性、甚至更优选至少 97-99%同一性或 100%相同。

10 编码本发明的多肽的多核苷酸, 包括脑膜炎奈瑟氏球菌以外物种的同系物和直向同源物, 可以通过包括以下步骤的方法获得: 在严格杂交条件下(例如, 使用 45-65°C 范围温度和 0.1-1% SDS 浓度), 使用包含或由 SEQ ID NO: 7 的序列或其片段组成的标记探针或可检测探针筛选合适文库; 然后分离包含所述多核苷酸序列的全长基因和/或基因组克隆。

15 本发明提供在其全长上与 SEQ ID NO: 7 中的编码序列(可读框)相同的多核苷酸序列。本发明还提供编码成熟多肽或其片段的编码序列本身以及在可读框中还有另一编码序列的编码成熟多肽或片段的编码序列, 所述另一编码序列例如编码分泌序列或前导序列、前蛋白(pre-protein)序列、原蛋白(pro-protein)序列或前原蛋白(prepro-protein)序列的序列。本发明的多核苷酸还可以包含至少一个非编码序列, 包括例如但不限于至少一个非编码 5'序列和 3'序列, 如转录但不翻译的序列、终止信号(例如依赖于 rho 的终止信号和不依赖于 rho 的终止信号)、核糖体结合位点、Kozak 序列、稳定 mRNA 的序列、内含子和聚腺苷酸化信号。所述多核苷酸序列也可以包括编码额外氨基酸的额外编码序列。例如, 可以编码有助于纯化融合多肽的标记序列。在本发明的某些实施方案中, 所述标记序列为 pQE 载体 (Qiagen, Inc.)提供并在 Gentz 等人, *Proc. Natl. Acad. Sci. USA* 86: 821-824 (1989)中描述的六组氨酸肽, 或者 HA 肽标记(Wilson 等人, *Cell* 37: 767 (1984)), 这两种标记可以用于纯化与其融合的多肽序列。本发明

25

的多核苷酸还包括但不限于包含结构基因和控制基因表达的与其天然结合的序列的多核苷酸。

5 编码 SEQ ID NO: 8 的 BASB069 多肽的核苷酸序列可以与 SEQ ID NO: 7 的核苷酸 1 到 2073 包含的多肽编码序列相同。或者它可以是由于遗传密码的冗余性(简并性)也编码 SEQ ID NO: 8 的多肽的序列。本文所用的术语“编码多肽的多核苷酸”包括包含编码本发明的多肽的序列的多核苷酸, 所述多肽优选为细菌多肽, 更优选具有 SEQ ID NO: 8 所示氨基酸序列的脑膜炎奈瑟氏球菌 BASB069 多肽。该术语也包括这样的多核苷酸: 所述多核苷酸包含编码所述多肽的
10 单个连续区或多个不连续区(例如被整合噬菌体、整合插入序列、整合载体序列、整合转座子序列或由于 RNA 编辑或基因组 DNA 重组而间断的多核苷酸)以及额外的区段, 所述额外的区段也可包含编码序列和/或非编码序列。

15 本发明还涉及本文描述的多核苷酸的变体, 所述变体编码具有 SEQ ID NO: 8 的推导氨基酸序列的多肽的变体。本发明的多核苷酸的片段可以用于例如合成本发明的全长多核苷酸。

其它特别优选的实施为编码 BASB069 变异体的多核苷酸, 所述 BASB069 变体具有 SEQ ID NO: 8 的 BASB069 多肽的氨基酸序列, 其中若干个、几个、5 到 10 个、1 到 5 个、1 到 3 个、2 个、
20 1 个或没有氨基酸残基以任何组合发生取代、修饰、缺失和/或添加。其中特别优选不改变 BASB069 多肽的特性和活性的沉默取代、添加和缺失。

本发明其它优选的实施为在它们的全长上与编码具有 SEQ ID NO: 8 所示氨基酸序列的 BASB069 多肽的多核苷酸至少 85% 相同的
25 的多核苷酸, 以及与这样的多核苷酸互补的多核苷酸。在这一方面, 在它们的全长上与上面所述多核苷酸至少 90% 相同的多核苷酸是特别优选的, 在这些特别优选的多核苷酸中, 至少 95% 相同的多核苷酸是尤其优选的。此外, 在至少 95% 相同的多核苷酸中, 至少 97%

相同的多核苷酸是高度优选的，其中至少 98%和至少 99%相同的多核苷酸是尤其高度优选的，而至少 99%相同的多核苷酸是更优选的。

优选的实施方案为编码基本保留与 SEQ ID NO: 7 的 DNA 编码的成熟多肽同样的生物功能或活性的多肽的多核苷酸。

5 依照本发明某些优选实施方案，提供与 BASB069 多核苷酸序列杂交、尤其是在严格条件下杂交的多核苷酸，例如 SEQ ID NO: 7 的各多核苷酸。

10 本发明还涉及与本文提供的多核苷酸序列杂交的多核苷酸。在这一方面，本发明尤其涉及在严格条件下与本文描述的多核苷酸杂交的多核苷酸。本文所用的术语“严格条件”和“严格杂交条件”是指仅当序列间存在至少 95%同一性、最好至少 97%同一性时才发生杂交。严格杂交条件的具体实例是在包含下列组分的溶液中于 42 °C 温育过夜：50%甲酰胺，5x SSC (150 mM NaCl, 15 mM 柠檬酸三钠)，50 mM 磷酸钠(pH 7.6)，5x Denhardt 氏溶液，10%葡聚糖硫酸酯
15 以及 20 微克/ml 变性剪切鲑鱼精子 DNA；然后在约 65 °C 下在 0.1 x SSC 中洗涤杂交支持物。杂交和洗涤条件是众所周知的，并在下面文献中有范例：Sambrook 等人, Molecular Cloning: A Laboratory Manual, 第二版, Cold Spring Harbor, N. Y., (1989)，特别是其中的第 11 章。溶液杂交也可用于本发明提供的多核苷酸序列。

20 本发明还提供包含或由多核苷酸序列组成的多核苷酸，所述多核苷酸序列如下获得：用具有 SEQ ID NO: 7 所示的所述多核苷酸序列或其片段的序列的探针在严格杂交条件下筛选包括 SEQ ID NO: 7 所示多核苷酸序列的完整基因的合适文库；然后分离所述多核苷酸序列。用于获得这样的多核苷酸的片段包括例如在本文其它地方充
25 分描述的探针和引物。

如本文中其它地方关于本发明的多核苷酸测定所述，例如本发明的多核苷酸可以用作针对 RNA、cDNA 和基因组 DNA 的杂交探针，以分离编码 BASB069 的全长 cDNA 和基因组克隆的或者分离与

BASB069 基因具有高度同一性、特别高度序列同一性的其它基因的 cDNA 和基因组克隆。这样的探针一般具有至少 15 个核苷酸残基或碱基对。优选这样的探针具有至少 30 个核苷酸残基或碱基对，并可以具有至少 50 个核苷酸残基或碱基对。尤其优选的探针具有至少 20 个核苷酸残基或碱基对而且少于 30 个核苷酸残基或碱基对。

使用 SEQ ID NO: 7 提供的 DNA 序列进行筛选可以分离 BASB069 基因的编码区，以合成寡核苷酸探针。然后使用具有与本发明的基因的序列互补的序列的标记寡核苷酸，筛选 cDNA 文库、基因组 DNA 文库或 mRNA 文库，以确定所述探针结合所述文库中的哪些成员。

对于本领域技术人员来说，有几种获得全长 cDNA 或延伸短 cDNA 的方法可利用并且是众所周知的，例如基于 cDNA 末端快速扩增 (RACE) 的方法(见例如 Frohman 等人, PNAS USA 85, 8998-9002, 1988)。最新改进的该技术例如 Marathon™ 技术(Clontech Laboratory Inc.)明显简化了对更长 cDNA 的搜索。在 Marathon™ 技术中，从由选定组织提取的 mRNA 制备 cDNA，并将“接头”序列连接到每个末端。然后使用混合的基因特异性寡核苷酸引物和接头特异性寡核苷酸引物，进行核酸扩增(PCR)以扩增所述 cDNA “缺失的” 5' 端。然后使用“嵌套”引物重复所述 PCR 反应，“嵌套”引物即是设计在所扩增的产物内退火的引物(通常是在所述接头序列的远 3' 端退火的接头特异性引物和已知基因序列的远 5' 退火的基因特异性引物)。然后通过 DNA 测序分析该反应的产物，随后通过直接连接所述产物到存在的 DNA 产生完整的序列，或使用新的序列信息设计 5' 引物来进行另一次全长 PCR，构造全长 DNA。

如本文关于多核苷酸测定的进一步讨论，本发明的多核苷酸和多肽可以用作例如发现对疾病、尤其是人类疾病治疗和诊断的研究试剂和材料。为得自 SEQ ID NO: 1-8 序列的寡核苷酸的本发明多核苷酸可以用于如本文所述的方法，但最好用于 PCR，以确定本发明鉴定的多核苷酸是否在感染组织细菌中完整地或部分地转录。已知

这样的序列也可以用于诊断已经获得病原体的感染阶段以及感染类型。

5 本发明还提供编码多肽的多核苷酸，所述多肽是成熟蛋白 + 额外氨基末端氨基酸或羧基末端氨基酸，或是在所述成熟多肽内部氨基酸（例如当成熟形式蛋白具有一条以上多肽链时）。这样的序列可能在从前体蛋白加工为成熟形式中起作用、可使得能进行蛋白质转运、可以延长或缩短蛋白的半衰期、或可有助于蛋白测定或生产操作。一般体内情况：细胞酶从所述成熟蛋白加工除去所述额外氨基酸。

10 对于本发明每个多核苷酸，都提供与之互补的多核苷酸。优选这些互补多核苷酸与每一个它们与之互补的多核苷酸完全互补。

前体蛋白可是所述多肽的无活性形式，前体蛋白具有与一个或多个原序列(prosequence)融合的多肽的成熟形式。当除去原序列时，这样的无活性前体通常被激活。在激活前可以除去部分或所有的原序列。这样的前体一般称为原蛋白。

15 除用标准的 A、G、C、T/U 代表核苷酸外，还可用字母“N”描述本发明的某些多核苷酸。“N”表示四种 DNA 或 RNA 核苷酸中的任何一种均可以存在于 DNA 或 RNA 序列中这样的指定位置上，但优选 N 不是这样的核酸：当同时考虑相邻核苷酸位置、以正确读框阅读时，具有在这样的可读框内产生成熟前终止密码子的作用。

20 总之，本发明的多核苷酸可以编码成熟蛋白、成熟蛋白 + 前导序列(可以称之为前蛋白)、编码具有并非前蛋白的前导序列的一个或多个原序列的成熟蛋白的前体、或编码前原蛋白，即具有前导序列和一个或多个原序列的原蛋白的前体，其中所述前导序列和原序列一般在产生所述多肽的活性成熟形式的加工步骤中除去。

25 按照本发明的一个方面，提供本发明的多核苷酸用于治疗或预防目的、尤其是遗传免疫接种的用途。

在遗传免疫接种中使用本发明的多核苷酸时，最好使用合适的

5 传送方法,如直接肌肉注射质粒 DNA (Wolff 等, *Hum Mol Genet* (1992) 1: 363. Manthorpe 等, *Hum. Gene Ther.* (1983) 4: 419)、传送与特异性蛋白载体复合的 DNA (Wu 等, *J Biol Chem.* (1989) 264: 16985)、与磷酸钙共沉淀 DNA (Benvenisty 和 Reshef, *PNAS USA*, (1986) 83: 9551)、
10 将 DNA 包入各种形式的脂质体内(Kaneda 等, *Science* (1989) 243: 375)、粒子轰击(Tang 等, *Nature* (1992) 356: 152, Eisenbraun 等, *DNA Cell Biol* (1993) 12:791)以及用克隆的逆转录病毒载体进行体内感染 (Seeger 等, *PNAS USA* (1984) 81: 5849)。

10 载体、宿主细胞、表达系统

本发明还涉及包含一种或多种本发明多核苷酸的载体、用本发明的载体遗传工程存在的宿主细胞以及通过重组技术产生本发明的多肽。也可以使用无细胞翻译系统由本发明 DNA 构建物得到的 RNA 产生这样的蛋白。

15 可以通过本领域技术人员所熟知的方法,从包含表达系统的遗传工程宿主细胞制备本发明的重组多肽。因此,再一方面,本发明涉及包含一种或多种本发明多核苷酸的表达系统、用这样的表达系统遗传工程操作的宿主细胞以及通过重组技术产生本发明的多肽。

20 为了重组产生本发明的多肽,可以遗传工程操作宿主细胞以引入本发明的表达系统或其部分或多核苷酸。可以采用许多标准实验室手册描述的方法将多核苷酸引入所述宿主细胞,所述实验室手册如 Davis 等, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986)和 Sambrook 等, *MOLECULAR CLONING: A LABORATORY MANUAL*, 第二版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989), 所述方法如磷酸钙转染、DEAE-葡聚糖介导的转染、转位、
25 显微注射、阳离子型脂质介导的转染、电穿孔、转导、scrape loading、弹道引入(ballistic introduction)及感染。

合适宿主的典型实例包括细菌细胞,如链球菌(streptococci)、葡

5 萄球菌(staphylococci)、肠球菌(enterococci)、大肠杆菌(*E. coli*)、链霉菌属(streptomyces)、蓝细菌(cyanobacteria)、枯草杆菌(*Bacillus subtilis*)、粘膜炎莫拉氏菌(*Moraxella catarrhalis*)、流感嗜血杆菌(*Haemophilus influenzae*)和脑膜炎奈瑟氏球菌(*Neisseria meningitidis*)
10 细胞; 真菌细胞, 如酵母、克鲁维酵母菌属(*Kluveromyces*)、酵母菌属(*Saccharomyces*)、担子菌(basidiomycete)、白假丝酵母(*Candida albicans*)和曲霉属(*Aspergillus*)的细胞; 昆虫细胞, 如果蝇 S2 (*Drosophila* S2)和贪夜蛾 Sf9 (*Spodoptera*) Sf9 的细胞; 动物细胞, 如 CHO、COS、HeLa、C127、3T3、BHK、293、CV-1 和 Bowes 黑素瘤细胞; 以及植物细胞, 如裸子植物或被子植物的细胞。

可以使用各种各样表达系统产生本发明的多肽。这样的载体其中包括染色体衍生的载体、附加体衍生的载体和病毒衍生的载体, 例如从细菌质粒衍生的载体、从噬菌体衍生的载体、从转座子衍生的载体、从酵母附加体衍生的载体、从插入元件衍生的载体、从酵母染色体元件衍生的载体、从病毒衍生的载体以及从它们的组合衍生的载体, 如那些从质粒和噬菌体遗传元件衍生的载体(如粘粒和噬菌粒), 其中所述病毒如杆状病毒、乳多空病毒(如 SV40)、痘苗病毒、腺病毒、禽痘病毒、假狂犬病病毒、小核糖核酸病毒、逆转录病毒和甲病毒属(alphavirus)。所述表达系统构建物可以包含调节以及引发
15 20 表达的控制区。一般来说, 在这一方面, 可以使用任何适于在宿主中维持、繁殖或表达多核苷酸和/或表达多肽的系统或载体用于表达。可以通过多种众所周知的常规技术中的任何一种将合适的 DNA 序列插入所述表达系统中, 所述常规技术如 Sambrook 等, *MOLECULAR CLONING, A LABORATORY MANUAL*, (见上文)所介绍的技术。

25 在真核生物重组表达系统中, 为将翻译后的蛋白分泌入内质网腔、分泌入壁膜间隙或分泌入胞外环境, 可以将合适的分泌信号引入所表达的多肽。这些信号对于所述多肽可以是内源的, 或者它们可以是异源信号。

可以用众所周知的方法由重组细胞培养物回收和纯化本发明多肽，所述方法包括硫酸铵沉淀或乙醇沉淀、酸提取、阴离子交换层析或阳离子交换层析、磷酸纤维素层析、疏水作用层析、亲和层析、羟磷灰石层析和凝集素层析。最优选使用离子金属亲和层析(IMAC)进行纯化。当所述多肽在胞内合成、分离和或纯化过程中变性时，
5 可以使用众所周知的重折叠蛋白技术重建活性构象。

所述表达系统也可以是活的重组微生物，例如病毒或细菌。可以将目的基因插入活的重组病毒或细菌的基因组。用该活的载体接种和体内感染将使所述抗原在体内表达并诱导免疫反应。用于该目的
10 的病毒和细菌可以是例如：痘病毒(如痘苗病毒、禽痘病毒、金丝雀痘病毒)、甲病毒属(新培斯病毒、塞姆利基森林病毒、委内瑞拉马脑炎病毒)、腺病毒、腺伴随病毒、小核糖核酸病毒(脊髓灰质炎病毒、鼻病毒)、疱疹病毒(水痘-带状疱疹病毒等)、李斯特菌属、沙门氏菌属、志贺菌属、奈瑟氏球菌属、BCG。这些病毒和细菌可以是有毒力的，或者以各种方式减毒获得活疫苗。这样的活疫苗也构成本发明的一部分。
15

诊断测定、预后测定、血清分型测定以及突变测定

本发明还涉及使用本发明的 BASB047、BASB054、BASB068
20 或 BASB069 多核苷酸及多肽作为诊断试剂的用途。检测真核生物、尤其是哺乳动物、特别是人类中的 BASB047、BASB054、BASB068 或 BASB069 多核苷酸和/或多肽将提供用于疾病诊断、疾病分期或传染性生物对药物的反应的诊断方法。可以用各种众所周知的技术以及本文提供的方法在核酸或氨基酸水平，对真核生物、尤其是哺乳
25 动物而且特别是人类、尤其是感染或怀疑感染包含 BASB047、BASB054、BASB068 或 BASB069 基因或蛋白的生物的人进行检测。

可以从推测感染和/或感染的个体的机体材料获得用于预后、诊断或其它分析的多肽和多核苷酸。其中任何一种来源的多核苷酸、

尤其是 DNA 或 RNA, 可直接用于检测, 或者使用 PCR 或任何其它
扩增技术酶促扩增后进行分析。也可以按同样方法使用 RNA (尤其是
mRNA)、cDNA 和基因组 DNA。可以利用扩增, 通过分析选定的生物
多核苷酸的基因型, 对个体中存在的传染性生物或居留生物
5 (resident organism)的种类和株系进行特征鉴定。与从相关生物选出的
参比序列的基因型相比, 以扩增产物大小变化可以检测缺失或插入,
其中所述相关生物最好是同一属的不同物种或同一物种的不同株
系。使扩增的 DNA 与标记的 BASB047、BASB054、BASB068 或
BASB069 多核苷酸序列杂交, 可以鉴定点突变。可以通过 DNA 酶
10 或 RNA 酶消化(分别对于 DNA 或 RNA 而言)、或通过检测解链温度
或复性动力学的差异, 区别完全配对序列或显著配对序列与不完全
错配或更显著错配双链体。也可以通过与参比序列相比, 以多核苷
酸片段在凝胶上的电泳迁移率变化检测多核苷酸序列的差异。这可
以使用或不使用变性试剂进行。也可以通过直接的 DNA 测序或 RNA
15 测序检测多核苷酸差异。参阅例如 Myers 等, *Science*, 230: 1242
(1985)。也可以通过核酸酶保护分析(如 RNA 酶、V1 和 S1 保护分析)
或化学切割方法揭示在特定位点的序列改变。参阅例如 Cotton 等, *Proc.*
Natl. Acad. Sci., USA, 85: 4397-4401 (1985)。

在另一实施方案中, 可以构建包含 BASB047、BASB054、
20 BASB068 或 BASB069 核苷酸序列或其片段的寡核苷酸探针的阵列,
以进行例如遗传突变的有效筛选、血清分型、分类学分类或鉴定。
阵列技术方法是众所周知的并具有普遍应用性, 而且可以用于阐明
分子遗传学中的多种问题, 包括基因表达、遗传连锁、以及遗传变
异性(参阅例如 Chee 等, *Science*, 274: 610 (1996))。

25 因此再一方面, 本发明涉及诊断试剂盒, 其中包括:

(a) 本发明的多核苷酸, 最好是 SEQ ID NO: 1、3、5、7 的核苷
酸序列或其片段;

(b) 与(a)的核苷酸序列互补的核苷酸序列;

(c) 本发明的多肽, 最好是 SEQ ID NO: 2、4、6、8 的多肽或其片段; 或

(d) 针对本发明的多肽的抗体, 最好是针对 SEQ ID NO: 2、4、6、8 的多肽的抗体。

5 人们知道, 在任何这样的试剂盒中, (a)、(b)、(c)或(d)可以构成实质性组分。这样的试剂盒将尤其在诊断疾病或诊断对疾病的易感性方面有用。

10 本发明还涉及将本发明的多核苷酸用作诊断试剂。检测与疾病或致病性有关的突变形式的本发明多核苷酸(最好是 SEQ ID NO: 1、3、5 或 7), 将提供可以帮助或确定疾病诊断、预测疾病病程、确定疾病阶段、或对疾病的易感性的诊断工具, 所述疾病源于所述多核苷酸的表达不足、过量表达或表达变化。可以用多种技术如本文别处描述的技术, 在多核苷酸水平上检测在这样的多核苷酸上带有突变的生物、尤其是传染性生物。

15 也可以采用多种技术, 在多核苷酸或多肽水平上检测在本发明的多核苷酸和/或多肽上带有突变或多态性(等位基因变异)的生物细胞, 以进行例如血清分型。例如可以使用 RT-PCR 检测 RNA 突变。尤其优选使用 RT-PCR 结合自动化检测系统, 例如 GeneScan。RNA、cDNA 或基因组 DNA 也可以用于同样目的 PCR。例如, 可以使用与
20 编码 BASB047、BASB054、BASB068 或 BASB069 多肽的多核苷酸互补的 PCR 引物鉴定和分析突变。

25 本发明还提供从 5'和/或 3'端切除的 1、2、3 或 4 个核苷酸的引物。其中这些引物可以用于扩增由个体得到的样品(如机体材料)分离出的 BASB047、BASB054、BASB068 或 BASB069 DNA 和/或 RNA。所述引物可以用于扩增从感染个体分离的多核苷酸, 以便随后可以采用各种技术研究所述多核苷酸以阐明所述多核苷酸序列。这样, 可以检测到所述多核苷酸序列中的突变, 并用于诊断和/或预测感染或其阶段或病程, 或者用于对所述传染性病原体进行血清分型和/或

分类。

5 本发明还提供用于诊断疾病、优选是细菌感染、更优选是脑膜炎奈瑟氏球菌导致的感染的方法，包括确定得自个体的样品(如机体材料)中具有 SEQ ID NO: 1、3、5 或 7 的序列的多核苷酸的表达水平增加。可以使用任何一种本领域众所周知的用于定量多核苷酸的方法，测定 BASB047、BASB054、BASB068 或 BASB069 多核苷酸表达的
10 增加或降低，所述方法例如扩增、PCR、RT-PCR、RNA 酶保护、RNA 印迹法、光谱测定法和其它杂交方法。

此外，根据本发明的用于检测与正常对照组织样品相比，
10 BASB047、BASB054、BASB068 或 BASB069 多肽过量表达的诊断检测，可以用于例如检测是否存在感染。可以用来测定从宿主得到的样品(如机体材料)中 BASB047、BASB054、BASB068 或 BASB069 多肽水平的测定技术，是本领域技术人员熟知的。这样的检测方法包括放射免疫测定、竞争性结合测定、蛋白质印迹分析、抗体夹心
15 测定、抗体检测和 ELISA 测定。

本发明的多核苷酸可以用作多核苷酸阵列、最好是高密度阵列或网格的成分。这些高密度阵列对于诊断和预测目的特别有用。例如，有一组点，每个点包括不同的基因，并且还包含一种或多种本发明多核苷酸，该组点可以利用从机体样品获得或衍生的探针用于
20 探测例如使用杂交或核酸扩增，确定个体中特定多核苷酸序列或相关序列的存在与否。这样一种存在可能说明存在病原体、尤其是脑膜炎奈瑟氏球菌，并可以用于诊断和/或预测疾病或病程。优选包括大量 SEQ ID NO: 1、3、5 或 7 的多核苷酸序列的变异体的网格。还优选包括大量编码 SEQ ID NO: 2、4、6 或 8 的多肽序列的多核苷酸的
25 变异体的网格。

抗体

本发明的多肽和多核苷酸或其变异体、或表达它们的细胞可以

用作免疫原，产生分别对这样的多肽或多核苷酸的免疫专一性抗体。

在本发明某些优选的实施方案中，提供抗 BASB047、BASB054、BASB068 或 BASB069 多肽或多核苷酸的抗体。

可以使用常规方法，给予动物(最好不是人类)本发明的多肽和/或多核苷酸、或二者之一或二者带有表位的片段、二者之一或二者的类似物、或表达之一或二者的细胞，获得针对本发明的多肽或多核苷酸产生的抗体。为了制备单克隆抗体，可以使用提供由连续细胞系培养物产生的抗体的本领域任何技术。实例包括各种技术，如下文文献中的技术：Kohler, G.和 Milstein, C., *Nature* 256: 495- 497 (1975); Kozbor 等, *Immunology Today* 4: 72 (1983); Cole 等, *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, 第 77-96 页, Alan R. Liss, Inc. (1985).

可以改进用于制备单链抗体的技术(美国专利第 4,946,778 号)，以产生针对本发明的多肽或多核苷酸的单链抗体。此外，可以使用转基因小鼠、或其它生物或动物(如其它哺乳动物)表达对本发明的多肽或多核苷酸免疫专一性的人源化抗体。

另一方面，可以利用噬菌体展示技术，或者从根据带有抗 BASB047、BASB054、BASB068 或 BASB069 而筛选出的人淋巴细胞的 PCR 扩增的 v 基因的所有组成成分中、或者从原初文库(naive library)中，选择对本发明的多肽有结合活性的抗体基因(McCafferty 等, (1990). *Nature* 348, 552-554; Marks 等, (1992) *Biotechnology* 10, 779-783)。也可以通过例如链改组提高这些抗体的亲和力(Clackson 等, (1991) *Nature* 352: 628)。

可以使用上述抗体来分离或鉴定表达本发明的多肽或多核苷酸的克隆，以采用例如亲和层析纯化所述多肽或多核苷酸。

因此，其中可以使用抗 BASB047、BASB054、BASB068 或 BASB069 多肽或 BASB047、BASB054、BASB068 或 BASB069 多核苷酸的抗体来治疗感染，尤其是细菌感染。

多肽变体包括在抗原上、表位上或免疫上等价的变体，这些构成本发明一个具体的方面。

5 最好修饰所述抗体及其变体，使其在个体中的免疫原性较低。例如，假如所述个体是人，那么最好“人源化”所述抗体，其中将杂交瘤产生的抗体的一个或多个 complementarity 决定区移植到人单克隆抗体中，例如 Jones 等(1986), *Nature* 321, 522-525 或 Tempest 等, (1991) *Biotechnology* 9, 266-273 中所述。

拮抗剂和激动剂 - 测定和分子

10 也可以使用本发明的多肽和多核苷酸，评估小分子底物与在例如细胞、无细胞制备物、化学文库以及天然产物混合物中的配体的结合。这些底物和配体可以是天然底物和配体，或者可以是结构模拟物或功能模拟物。参阅例如 Coligan 等, *Current Protocols in Immunology* 1(2): 第 5 章(1991)。

15 使用与候选化合物直接或间接结合的标记，筛选方法就可以测量所述候选化合物与所述多肽或多核苷酸的结合、或所述候选化合物与带有所述多肽或多核苷酸的细胞或膜的结合、或所述候选化合物与带有所述多肽的融合蛋白的细胞或膜的结合。或者，所述筛选方法涉及与标记竞争物的竞争。此外，使用对于包含所述多肽或多核苷酸的细胞合适的检测系统，这些筛选方法可以测试所述候选化合物是否导致由于所述多肽或多核苷酸的激活或抑制而产生的信号。一般在一种已知激动剂的存在下测定激活的抑制物，并观察所述候选化合物的存在对所述激动剂的激活作用的影响。组成型活性多肽和/或组成型表达的多肽和多核苷酸可以用于筛选逆向激动剂或抑制剂的筛选方法：在缺乏激动剂或抑制物的情况下，根据具体情况测试所述候选化合物是否导致对所述多肽或多核苷酸的激活的抑制。此外，所述测试方法可以仅包括以下步骤：将候选化合物与含有本发明的多肽或多核苷酸的溶液混合，形成混合物，测定所述混

20

25

合物中 BASB047、BASB054、BASB068 或 BASB069 多肽和/或多核苷酸活性，并将所述混合物的 BASB047、BASB054、BASB068 或 BASB069 多肽和/或多核苷酸活性与标准品活性比较。也可以使用融合蛋白，例如如前文所述由 Fc 部分和 BASB047、BASB054、BASB068 或 BASB069 多肽制成的融合蛋白，进行高通量筛选测定，以鉴定本发明多肽的拮抗剂、以及系统发育上和/或功能上相关多肽的拮抗剂 (参阅 D. Bennett 等, *J Mol Recognition*, 8: 52-58 (1995); 和 K. Johanson 等, *J Biol Chem*, 270 (16): 9459-9471 (1995))。

也可以使用所述多核苷酸、多肽以及结合本发明多肽和/或与本发明多肽相互作用的抗体来构建筛选方法，以测试添加的化合物对细胞中 mRNA 和/或多肽的产生的影响。例如，使用单克隆抗体和多克隆抗体，通过本领域已知的标准方法，可以构建 ELISA 测定以测定多肽的分泌水平或细胞结合水平。这可以用于从适当操作的细胞或组织中发现可能抑制或增强多肽产生的物质(也分别称为拮抗剂或激动剂)。

本发明还提供筛选化合物的方法，以鉴定增强(激动剂)或阻断(拮抗剂)BASB047、BASB054、BASB068 或 BASB069 多肽或多核苷酸的作用的化合物，尤其是鉴定制菌和/或杀菌的化合物。所述筛选方法可涉及高通量技术。例如，为了筛选激动剂或拮抗剂，在存在或不存在可能是 BASB047、BASB054、BASB068 或 BASB069 激动剂或拮抗剂的候选分子的情况下，将包含 BASB047、BASB054、BASB068 或 BASB069 多肽的合成反应混合物、细胞区室(如膜、胞外被膜或细胞壁、或它们中任何一种的制备物)与所述多肽的标记底物或配体一起温育。候选分子激动或拮抗 BASB047、BASB054、BASB068 或 BASB069 多肽的能力表现为结合标记配体减少或所述底物产生的产物减少。无作用的结合分子，即不诱导 BASB047、BASB054、BASB068 或 BASB069 多肽作用的分子很有可能是良好的拮抗剂。很好结合并且根据具体情况增加底物产生产物的速率、增加信号转导或增加化

学通道活性的分子是激动剂。使用报道系统，可以根据具体情况检测增强底物产生产物、信号转导或化学通道活性的速率或水平。在这一方面可能有用的报告系统包括但不限于比色法、转化成产物的标记底物、应答 BASB047、BASB054、BASB068 或 BASB069 多核苷酸或多肽的活性变化的报告基因以及本领域已知的结合测定。

测定 BASB047、BASB054、BASB068 或 BASB069 激动剂的另一个实例是竞争性测定，即将 BASB047、BASB054、BASB068 或 BASB069 和潜在激动剂与 BASB047、BASB054、BASB068 或 BASB069 结合分子、重组 BASB047、BASB054、BASB068 或 BASB069 结合分子、天然底物或配体、或底物或配体模拟物在合适条件下混合，以进行竞争性抑制分析。可以采用如放射性或显色化合物来标记 BASB047、BASB054、BASB068 或 BASB069，以便可以准确测定与结合分子结合或转化成产物的 BASB047、BASB054、BASB068 或 BASB069 分子数量，以评估所述潜在拮抗剂的有效性。

潜在拮抗剂其中包括结合本发明的多核苷酸和/或多肽并因此抑制或消除其活性或表达的小有机分子、肽、多肽以及抗体。潜在拮抗剂还可以是小有机分子、肽、多肽，如结合在结合分子(如结合分子)上相同位点的密切相关的蛋白或抗体，而不诱导 BASB047、BASB054、BASB068 或 BASB069 所诱导的活性，因此使 BASB047、BASB054、BASB068 或 BASB069 多肽和/或多核苷酸不能结合而阻止 BASB047、BASB054、BASB068 或 BASB069 多肽和/或多核苷酸的作用或表达。

潜在拮抗剂包括这样的小分子：所述小分子结合并占据所述多肽的结合位点，因此防止其结合到细胞结合分子，从而阻止正常生物活性。小分子的实例包括但不限于小有机分子、肽或肽样分子。其它潜在拮抗剂包括反义分子(参阅 Okano, *J. Neurochem.* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), 关于这些

分子的介绍)。优选的潜在拮抗剂包括与 BASB047、BASB054、BASB068 或 BASB069 相关的化合物以及 BASB047、BASB054、BASB068 或 BASB069 的变异体。

再一方面，本发明涉及遗传工程制备的可溶性融合蛋白，该融合蛋白包含本发明的多肽或其片段，以及不同亚类免疫球蛋白(IgG、IgM、IgA、IgE)的重链或轻链的恒定区的不同部分。作为免疫球蛋白优选人类 IgG (尤其是 IgG1)的重链的恒定部分，其中在绞链区融合。在一个具体实施方案中，可以通过插入能用凝血因子 Xa 切割的切割序列就可除去 Fc 部分。此外，本发明涉及采用遗传工程制备这些融合蛋白的方法，以及所述融合蛋白用于药物筛选、诊断和治疗的应用。本发明的另一方面还涉及编码这样的融合蛋白的多核苷酸。融合蛋白技术的实例可见于国际专利申请第 WO94/29458 号和第 WO94/22914 号。

本文提供的每一种多核苷酸序列都可以用于发现和研制抗菌化合物。表达后，所编码的蛋白可以用作筛选抗细菌药物的靶。此外，相应的 mRNA 上编码所述编码蛋白氨基末端区的多核苷酸序列、或 Shine-Delgarno 或其它有利于翻译的序列可以用于构建控制目的编码序列表达的反义序列。

本发明还提供使用本发明的多肽、多核苷酸、激动剂或拮抗剂来干扰一种病原体或多种病原体与真核宿主、最好是哺乳动物宿主之间导致感染后续结果的初始物理作用。具体来说，本发明的分子可以用于：防止细菌、尤其是革兰氏阳性和/或革兰氏阴性细菌附着到留置装置上真核生物、最好是哺乳动物的细胞外基质蛋白，或附着到伤口中的细胞外基质蛋白；阻断介导组织损伤的真核生物、最好是哺乳动物的细胞外基质蛋白与细菌 BASB047、BASB054、BASB068 或 BASB069 蛋白间的细菌附着，和/或；阻止植入留置装置或其它外科技技术以外引发的感染发病机理的正常进行。

按照本发明的再一方面，提供 BASB047、BASB054、BASB068

或 BASB069 激动剂和拮抗剂，最好是制菌或杀菌的激动剂和拮抗剂。

本发明的拮抗剂和激动剂可以用于例如预防、抑制和/或治疗疾病。

5 再一方面，本发明涉及本发明多肽的 mimotope。mimotope 是与天然肽足够相似(序列上或结构上)的肽序列，能够被识别天然肽的抗体识别；或者当偶联到合适载体时，能够产生识别天然肽的抗体。

10 可以通过添加、缺失或取代选定的氨基酸而设计肽 mimotope 用于特定用途。因此，可以为了易于缀合于蛋白载体而修饰所述肽。例如，一些化学缀合方法最好包括一个末端半胱氨酸。此外，与蛋白载体缀合的肽最好在所述肽缀合末端远处包括一个疏水末端，以便所述肽游离的未结合末端保持与载体蛋白的表面结合。因此使所述肽呈现出这样的构象：所述构象与在完整的天然分子中发现的所述肽的构象最密切相似。例如，可以改变所述肽使其具有一个 N 末端半胱氨酸和一个 C 末端疏水性酰胺化尾。或者，可以进行一个或多个氨基酸的 D-立体异构体的添加或取代，以产生有益的衍生物，
15 例如增强所述肽的稳定性。

或者，使用如噬菌体展示技术(EP 0 552 267 B1)的技术，用本身能够结合本发明多肽的抗体，可以鉴定肽 mimotope。该技术产生了大量模拟天然肽的结构并因此能够结合抗天然肽抗体的肽序列，但
20 所述肽序列本身并不一定与所述天然多肽具有显著的序列同一性。

疫苗

25 本发明的另一方面涉及在个体、特别是哺乳动物、最好是人类诱导免疫应答的方法，包括用足以产生抗体和/或 T 细胞免疫应答的 BASB047、BASB054、BASB068 或 BASB069 多核苷酸和/或多肽或它们的片段或变体接种所述个体，以保护所述个体免受感染，尤其是细菌感染，更特别是脑膜炎奈瑟氏球菌感染。还提供的是这样的方法：用这种方法产生的所述免疫应答减慢细菌的复制。而本发

5 明的再一方面涉及在个体中诱导免疫应答的方法，包括传递核酸载体、序列或核酶给这样的个体，来指导 BASB047、BASB054、BASB068 或 BASB069 多核苷酸和/或多肽、或其片段或变异体的表达，以在体内表达 BASB047、BASB054、BASB068 或 BASB069 多核苷酸和/或多肽、或其片段或变异体，诱导免疫应答，如产生抗体和/或 T 细胞免疫应答(包括例如产生细胞因子的 T 细胞或细胞毒性 T 细胞)，来保护所述个体(最好是人类)抵抗疾病，而不论该疾病是否已经在所述个体内产生。给予所述基因的一个实例是以颗粒的包被物或者非颗粒包被物加速其进入所需细胞。这样的核酸载体可以包括 DNA、RNA、
10 核酶、修饰核酸、DNA/RNA 杂种、DNA-蛋白质复合物或 RNA-蛋白质复合物。

本发明的再一方面涉及免疫组合物，当所述免疫组合物引入个体、最好是人体时，由于所述免疫组合物能够在所述个体中诱导免疫应答，因此在所述个体中诱导针对 BASB047、BASB054、BASB068
15 或 BASB069 多核苷酸和/或由其编码的多肽的免疫应答，其中所述组合物包含重组 BASB047、BASB054、BASB068 或 BASB069 多核苷酸和/或由其编码的多肽，和/或所述组合物包括编码和表达所述 BASB047、BASB054、BASB068 或 BASB069 多核苷酸的抗原的 DNA 和/或 RNA、由其编码的多肽、或其它本发明的多肽。所述免疫应答
20 可以用于治疗或为预防目的，并采用抗体免疫和/或细胞免疫的形式，如 CTL 或 CD4+ T 细胞产生的细胞免疫。

BASB047、BASB054、BASB068 或 BASB069 多肽或其片段可以与辅蛋白(co-protein)或化学部分融合，所述化学部分本身可能产生抗体或不产生抗体，但能够稳定所述第一种蛋白并产生具有抗原性和/或免疫原性的特性、最好是保护特性的融合或修饰蛋白。因此，
25 融合的重组蛋白最好还包含一种抗原性辅蛋白，或任何其它增溶所述蛋白并有利于其生产和纯化的较大辅蛋白，其中所述抗原性辅蛋白如流感嗜血杆菌的脂蛋白 D、谷胱甘肽-S-转移酶(GST)或 β -半乳

糖苷酶。此外，所述辅蛋白可以作为佐剂在提供对接受所述蛋白生物的免疫系统的普通性刺激的意义起作用。所述辅蛋白可以连接到第一种蛋白的氨基端或者羧基端。

5 在依照本发明的疫苗组合物中，BASB047、BASB054、BASB068 或 BASB069 多肽和/或多核苷酸或其片段、或 mimotope、或变体可以为一种载体，例如上面描述的活重组载体，例如活细菌载体。

10 BASB047、BASB054、BASB068 或 BASB069 多肽的非活载体也是合适的，例如细菌膜外小泡或“囊泡(bleb)”。OM 囊泡获自格兰氏阴性细菌双层膜的外膜，并且根据文献在许多种格兰氏阴性细菌中存在(Zhou, L 等人, 1998. *FEMS Microbiol. Lett.* 163: 223-228)，包括沙眼衣原体(*C. trachomatis*)和鸚鵡热衣原体(*C. psittaci*)。据报道产生囊泡的细菌病原体的不完全一览表包括：百日咳博德特氏菌(*Bordetella pertussis*)、布氏疏螺旋体(*Borrelia burgdorferi*)、马尔他布鲁氏菌(*Brucella melitensis*)、羊氏布鲁氏菌(*Brucella ovis*)、大肠杆菌(*Escherichia coli*)、流感嗜血杆菌、嗜肺军团杆菌(*Legionella pneumophila*)、淋病奈瑟氏球菌(*Neisseria gonorrhoeae*)、脑膜炎奈瑟氏球菌、绿脓杆菌(*Pseudomonas aeruginosa*)和小肠结肠炎耶尔森氏菌(*Yersinia enterocolitica*)。

20 囊泡具有能以其天然构象提供外膜蛋白的优势，因此对于疫苗尤其有用。也可以通过工程操作所述细菌而改善囊泡用于疫苗的应用，以便改善一个或多个外膜分子的表达。因此可以引入或正调节(如通过改变启动子)例如所需免疫原性蛋白如 BASB047、BASB054、BASB068 或 BASB069 多肽在外膜上的表达。或者或此外，可以下调节或者非相关(如非保护性抗原或免疫显性但变化的蛋白)、或者有害(如毒性分子如 LPS 或自身免疫反应潜在诱导剂)的外膜分子的表达。25 这些方法在下面详细讨论。

BASB047、BASB054、BASB068 或 BASB069 基因的非编码侧翼区包含在所述基因的表达中重要的调节元件。这种调节作用可发

生于转录和翻译水平。可以通过 DNA 测序获得在所述基因的可读框上游或下游的这些区段的序列。该序列信息使得可以确定潜在调节基序，例如不同的启动子元件、终止子序列、诱导序列元件、阻抑物、相变元件、SD 序列、具有涉及调节的潜在二级结构的区以及其它类型的调节基序或序列。

该序列信息使得能够调节 BASB047、BASB054、BASB068 或 BASB069 基因的天然表达。通过改变启动子、SD 序列、潜在阻抑物或操纵基因元件或任何其它涉及的元件，可以成功正调节所述基因的表达。同样，通过相似类型的改进可以完成表达的下调节。或者，通过改变相变序列，可以将所述基因的表达置于相变控制之下，或者可以使其与这种调节作用无关。在另一种方法中，可以将所述基因的表达置于一种或多种诱导元件的控制下，以允许进行调节表达。这样的调节实例包括但不限于通过温度变化、加入诱导物底物如选定的糖或它们的衍生物、微量元素、维生素、辅因子、金属离子等等进行诱导。

可以采用几种不同方法引入如上所述的修饰。通过随机诱变并随后筛选所需表型，可以在体内进行涉及基因表达的序列的修饰。另一种方法包括分离目的区段并通过随机诱变或定点置换、插入或缺失突变而进行修饰。随后可以将所述修饰区段通过同源重组再引入细菌基因组中，并且评估对基因表达的影响。在另一种方法中，可以使用目的区段的序列信息取代或缺失所有或部分天然调节序列。在这种情况下，分离目标调节区并进行修饰，以包含另一基因的调节元件、不同基因调节元件的组合、合成的调节区或任何其它调节区，或者分离目标调节区并进行修饰以缺失所述野生型调节序列的选定部分。随后可以将这些修饰序列再引入细菌，并通过同源重组进入基因组中。可以用于正调节基因表达的优选启动子的不完全一览表包括脑膜炎奈瑟氏球菌或淋病奈瑟氏球菌的启动子 porA、porB、lbpB、tbpB、p110、lst、hpuAB；ompCD、copB、lbpB、ompE、

UspA1; UspA2; 粘膜炎莫拉氏菌的 TbpB; 流感嗜血杆菌的 p1、p2、p4、p5、p6、lpD、tbpB、D15、Hia、Hmw1、Hmw2。

5 在一个实施例中，可以通过将所述基因的启动子换为一个更强的启动子(通过分离所述基因的上游序列，体外改变该序列，然后通过同源重组再引入基因组)而调节所述基因的表达。可以用所述细菌和由所述细菌脱落(或产生)的外膜小泡获得正调节的表达。

10 在其它实施例中，可以使用所述方法产生具有疫苗应用改善特性的重组细菌株。这些株可以是但不限于：减毒株、选定抗原表达增加的株、干扰免疫反应的基因失去作用(或表达降低)的株、免疫显性蛋白的表达受到调节的株、外膜小泡的脱落受到调节的株。

15 因此，本发明还提供 BASB047、BASB054、BASB068 或 BASB069 基因的修饰上游区，所述修饰上游区包含改变位于外膜的 BASB047、BASB054、BASB068 或 BASB069 蛋白表达水平的异源调节元件。依照本发明这一方面的上游区包括 BASB047、BASB054、BASB068 或 BASB069 基因上游的序列。所述上游区起始于 BASB047、BASB054、BASB068 或 BASB069 基因的直接上游，通常延伸到从所述基因的 ATG 起始密码子上游 1000 bp 以内的位置。在位于多顺反子序列(操纵子)中的基因的情况下，所述上游区可以紧接于目的基因之前起始，或在所述操纵子中第一个基因之前。依照本发明的修
20 饰上游区最好在 ATG 上游 500 bp 到 700 bp 之间的某个位置包含一个异源启动子。

25 因此，本发明以修饰的细菌囊泡提供 BASB047、BASB054、BASB068 或 BASB069 多肽。本发明还提供能够产生非活的基于膜的囊泡载体的修饰的宿主细胞。本发明还提供包含 BASB047、BASB054、BASB068 或 BASB069 基因的核酸载体，所述 BASB047、BASB054、BASB068 或 BASB069 基因具有包含一个异源调节元件的上游修饰区。

本发明还提供制备依照本发明的宿主细胞和细菌囊泡的方法。

本发明提供组合物，尤其是疫苗组合物，以及包含本发明的多肽和/或多核苷酸以及免疫刺激性 DNA 序列的方法，所述免疫刺激性 DNA 序列如 Sato, Y.等 Science 273: 352 (1996)中所述序列。

5 本发明还提供使用所述多核苷酸或其特定片段的方法，其中在脑膜炎奈瑟氏球菌感染的动物模型中证实，在所述遗传免疫实验中使用的多核苷酸构建物中，所述多核苷酸或其特定片段编码细菌细胞表面蛋白的非可变区。这样的实验对于鉴定能够激发预防性或治疗性免疫应答的蛋白质表位特别有用。相信该方法将使得此后能够制备特别有用的单克隆抗体，以开发针对在哺乳动物、尤其是人类
10 细菌感染(尤其是脑膜炎奈瑟氏球菌感染)的预防剂或治疗方法，其中所述单克隆抗体得自成功抵御或清除感染的动物的要求器官。

本发明还包括疫苗制剂，所述疫苗制剂包括本发明的免疫原性重组多肽和/或多核苷酸以及合适的载体，如药学上可接受的载体。由于所述多肽和多核苷酸可能在胃中分解，因此它们中的每一种最好胃肠外给予，包括例如皮下给予、肌肉给予、静脉内给予或皮内
15 给予。适于胃肠外给予的制剂包括水性和非水性无菌注射溶液，所述溶液可以包含抗氧化剂、缓冲剂、制菌化合物和使得所述制剂与个体的体液(最好是血液)等渗的溶质；以及水性和非水性无菌悬浮液，所述悬浮液可以包括悬浮剂或增稠剂。所述制剂可以盛装在单
20 位剂量或多剂量容器(如密封安瓿和管形瓶)中，并可以在冻干条件下保存，只需要在使用前加入无菌液体载体。

本发明的疫苗制剂也可以包括增强所述制剂的免疫原性的佐剂系统。所述佐剂系统最好增强 TH1 型反应。

免疫反应可以主要分为两大类：体液免疫应答或细胞介导的免疫应答(传统上其特征分别在于保护性抗体机制和效应细胞机制)。这两类免疫反应称为 TH1 型反应(细胞免疫反应)以及 TH2 型免疫反应(体液免疫反应)。
25

非常典型的 TH1 型免疫反应特征在于产生抗原特异性单倍型限

制性细胞毒性 T 淋巴细胞以及天然杀伤细胞反应。在小鼠中，TH1 型反应通常特征在于产生 IgG2a 亚型抗体，而在人类中这些抗体相当于 IgG1 型抗体。TH2 型免疫反应特征在于产生各种各样的免疫球蛋白同种型，在小鼠包括 IgG1、IgA 和 IgM。

5 可以认为产生这两种类型免疫反应的驱动力是细胞因子。高水平的 TH1 型细胞因子往往促进诱导对给定抗原的细胞介导的免疫反应，而高水平的 TH2 型细胞因子倾向于促进诱导对抗原的体液免疫反应。

10 TH1 型免疫反应和 TH2 型免疫反应的区别并不是绝对的。实际上，有个别人支持将免疫反应描述为主要是 TH1 型或主要是 TH2 型的免疫反应。然而，按照 Mosmann 和 Coffman 描述鼠 CD4+ve T 细胞克隆时采用的细胞因子家族分类来考虑细胞因子家族通常是合适的(Mosmann, T.R. 和 Coffman, R.L. (1989) *TH1 和 TH2 细胞: 不同模式的淋巴因子分泌产生不同的功能特性. Annual Review of Immunology, 7, 第 145-173 页*)。传统上，TH1 型反应与 T 淋巴细胞产生的 IFN- γ 和 IL-2 细胞因子有关。其它常常与诱导 TH1 型免疫反应有关的细胞因子如 IL-12 并不是 T 细胞产生的。相反，TH2 型反应与 IL-4、IL-5、IL-6 和 IL-13 分泌有关。

20 已知某些疫苗佐剂特别适于刺激 TH1 型或 TH2 型细胞因子反应。传统上，在免疫接种或感染后免疫反应的 TH1: TH2 平衡的最佳指标包括在体外用抗原再刺激后直接测量 T 细胞产生的 TH1 或 TH2 细胞因子，和/或测量抗原特异性抗体反应的 IgG1: IgG2a 比例。

因此，TH1 型佐剂是在体外用抗原再刺激时优先刺激分离的 T 细胞群体产生高水平 TH1 型细胞因子而且促进产生 CD8+ 细胞毒性 T 淋巴细胞以及与 TH1 型同种型有关的抗原特异性免疫球蛋白反应的佐剂。

国际专利申请号 WO 94/00153 和 WO 95/17209 中描述了能够优先刺激 TH1 细胞反应的佐剂。

3 脱氧酰化单磷酸脂质 A (3D-MPL)就是一种这样的佐剂。这可以从 GB 2220211 (Ribi)得知。在化学上它是具有 4、5 或 6 个酰化链的 3 脱氧酰化单磷酸脂质 A 的混合物，由 Ribi Immunochem, Montana 生产。3 脱氧酰化单磷酸脂质 A 的优选形式公开于欧洲专利 0 689 454 B1 (SmithKline Beecham Biologicals SA)。

3D-MPL 的颗粒最好小得足以使其可以无菌滤过 0.22 微米的膜 (欧洲专利号 0 689 454)。3D-MPL 剂量范围为每剂 10 μ g-100 μ g、最好 25-50 μ g，其中所述抗原的剂量范围通常为每剂 2-50 μ g。

另一种优选的佐剂包含 QS21，即一种从 Quillaja Saponaria Molina 的树皮得到的 Hplc 纯化无毒部分。可选地将 QS21 与 3 脱氧酰化单磷酸脂质 A (3D-MPL)混合，可选地还与另一种载体混合。

美国专利号 5,057,540 公开了制备 QS21 的方法。

以前已经描述了包含 QS21 的非反应性佐剂制剂 (WO 96/33739)。已经证实包含 QS21 和胆固醇的这类制剂当与抗原一起配制时是成功的 TH1 刺激佐剂。

TH1 细胞反应优先刺激物的其它佐剂包括免疫调制性寡核苷酸，例如在 WO 96/02555 中公开的非甲基化 CpG 序列。

还考虑组合诸如上述的不同 TH1 刺激佐剂，以提供属于 TH1 细胞反应优选刺激物的佐剂。例如 QS21 可以与 3D-MPL 一起配制。

QS21: 3D-MPL 的比例通常约为 1: 10 到 10: 1；最好是 1: 5 到 5: 1，而实际上常常是 1: 1。最佳协同作用的最优范围是 2.5: 1 到 1: 1 的 3D-MPL: QS21。

依照本发明的疫苗组合物中最好还存在一种载体。所述载体可以是水包油乳剂，或者是一种铝盐，如磷酸铝或氢氧化铝。

优选的水包油乳剂包含一种可代谢的油如鲨烯、 α -生育酚和 Tween 80。在一个特别优选的方面，用这样的乳剂将本发明疫苗组合物的抗原与 QS21 和 3D-MPL 混合。此外所述水包油乳剂可以含有 Span 85 和/或卵磷脂和/或三辛酸甘油酯。

给予人时，通常疫苗中的 QS21 和 3D-MPL 含量范围为每剂 1 μg -200 μg ，如 10-100 μg ，最好是 10 μg -50 μg 。而水包油乳剂通常含有 2-10%的鲨烯、2-10%的 α -生育酚和 0.3-3%的 Tween 80。鲨烯： α -生育酚的比例最好是等于或小于 1，因为这样的比例提供更稳定的乳浊液。Span 85 含量也可以为 1%。在某些情况下，本发明的疫苗另外含有一种稳定剂是有利的。

无毒的水包油乳剂最好在水溶液载体中含有一种无毒的油，如角沙烷或角鲨烯，以及一种乳化剂如 Tween 80。所述水溶液载体可以是例如磷酸缓冲盐溶液。

WO 95/17210 中描述了一种特别有效的佐剂制剂，该制剂为包含 QS21、3D-MPL 和生育酚的水包油乳剂。

本发明还提供多价疫苗组合物，所述多价疫苗组合物包含本发明的疫苗制剂以及其它抗原，尤其是用于治疗癌症、自身免疫性疾病和相关病症的抗原。这样的一种多价疫苗组合物可以包括如前面所述的 TH-1 诱导佐剂。

虽然已经参考某些 BASB047、BASB054、BASB068 或 BASB069 多肽和多核苷酸介绍了本发明，但应当知道，本发明包括天然多肽和多核苷酸的片段，以及带有添加、缺失或取代的相似多肽和多核苷酸，其中所述添加、缺失或取代基本不影响所述重组多肽或多核苷酸的免疫原性特性。

所述抗原也可以完整细菌(死细菌或活细菌)形式或亚细胞组分给予，但是这些可能性不包括脑膜炎奈瑟氏球菌本身。

组合物、试剂盒以及使用方法

本发明的再一方面，提供用于给予单细胞生物或多细胞生物的包含 BASB047、BASB054、BASB068 或 BASB069 多核苷酸和/或 BASB047、BASB054、BASB068 或 BASB069 多肽的组合物。

本发明还涉及包含本文所讨论的多核苷酸和/或多肽或它们的激

5 动剂或拮抗剂的组合物。本发明的多肽和多核苷酸可以与非无菌或
无菌的一种载体或多种载体组合使用以用于细胞、组织或生物，所
述载体如适于给予个体的药用载体。这样的组合物包括例如媒介添
加物(media additive)或治疗有效量的本发明的多肽和/或多核苷酸以及
药学上可接受的载体或赋形剂。这样的载体可以包括但不限于盐水、
缓冲盐水、葡萄糖、水、甘油、乙醇以及它们的组合物。所述制剂
应当适合所述给药方式。本发明还涉及诊断用和药用包装和试剂盒，
所述包装和试剂盒包括一个或多个容器，所述容器中装有前文所述
本发明的组合物的一种或多种成份。

10 本发明的多肽、多核苷酸和其它化合物可以单独使用或与其它
化合物如治疗化合物联合使用。

所述药用组合物可以以任何有效、方便的方式给予，所述方式
其中包括例如局部应用、口服、肛门、阴道、静脉内、腹膜内、肌
内、皮下、鼻内或皮内途径给予。

15 在治疗或用作预防药时，可以将所述活性药物以注射组合物给
予个体，例如为无菌水性分散体，所述分散体最好是等渗的。

另一方面，本发明提供药用组合物，所述药用组合物包括治疗
有效量的多肽和/或多核苷酸(如可溶形式的本发明多肽和/或多核苷
酸)、激动剂肽或拮抗剂肽或小分子化合物和药学可接受的载体或赋
形剂。这样的载体包括但不限于盐水、缓冲盐水、葡萄糖、水、甘
油、乙醇以及它们的组合物。本发明还涉及药用包装和试剂盒，所
述药用包装和试剂盒包括一个或多个容器，所述容器中装有前文所
述本发明的组合物的一种或多种成份。

20 本发明的多肽、多核苷酸和其它化合物可以单独使用或与其它
化合物如治疗化合物结合使用。

25 可以改变所述组合物以适应给药途径，所述给药途径如系统性
途径或口服途径。系统性给药的优选形式包括注射，一般是通过静
脉内注射。可以使用其它注射途径，如皮下注射、肌内注射或腹膜

内注射。用于系统性给药的其它可选方法包括使用渗透剂经粘膜和经皮给药，所述渗透剂如胆汁酸盐或梭链孢酸或其它去污剂。此外，如果本发明的多肽或其它化合物可以配制为肠溶制剂或胶囊化制剂中，那么口服给药也是可以的。这些化合物也可以是以药膏、糊剂、
5 凝胶、溶液、粉剂等形式表面和/或局部给予。

为了给予哺乳动物、尤其是人类，预期所述活性药物的每日剂量为 0.01 mg/kg 到 10 mg/kg，一般是 1 mg/kg 左右。在任何情况下，医师将确定最适于个体的实际剂量，并且该剂量将随着特定个体的年龄、体重和反应而不同。上面的剂量代表一般情况。当然，个别
10 情况更高或更低剂量范围是有益的，而这些情况也属于本发明范畴。

所需的剂量范围取决于选择的肽、给药途径、制剂的性质、患者病症的性质以及执业医师的判断。然而，合适的剂量范围为 0.1-100 $\mu\text{g/kg}$ 患者体重。

疫苗组合物方便地采用注射形式。可以使用常规佐剂增强免疫应答。用于疫苗接种的合适单位剂量是 0.5-5 微克/kg 的抗原，并且最好以 1-3 周的间隔给予这样的剂量 1-3 次。使用所述剂量范围，对于本发明的化合物将不会观察到阻止将它们给予合适个体的不利毒理学作用。
15

然而，考虑到可用的化合物的多样性以及各种给药途径的不同效率，预期所需剂量将有很大变化。例如，预期口服将比通过静脉内注射给予需要更高剂量。使用本领域熟知的优化标准经验方法，可以调整这些剂量水平上的变化。
20

序列数据库、有形媒介序列以及算法

多核苷酸和多肽序列构成有价值的信息资源，使用这样的资源可确定它们的 2 维和 3 维结构，并鉴定具有相似同源性的其它序列。通过将序列储存在计算机可读媒介中，然后使用已知大分子结构程序的存储数据，或采用众所周知的检索工具如 GCG 程序包检索序列
25

数据库，最有利于这些方法。

本发明还提供的是字符序列或字符串(string)、尤其是基因序列或所编码的蛋白序列的分析方法。优选的序列分析方法包括例如序列同源性分析的方法(如同一性和相似性分析)、DNA、RNA 和蛋白质结构分析、序列装配、种系发生分析、序列基序分析、确定可读框、核酸碱基调入(nucleic acid base calling)、密码子选择分析、核酸碱基修整(nucleic acid base trimming)以及序列色谱峰分析。

为进行同源性鉴定提供基于计算机的方法。该方法包括下列步骤：以计算机可读媒介提供包含本发明多核苷酸序列的第一个多核苷酸序列；将所述第一个多核苷酸序列与至少一个第二多核苷酸或多肽序列比较，鉴定同源性。

还为进行同源性分析提供基于计算机的方法，所述方法包括下列步骤：在计算机可读媒介中提供包含本发明多肽的序列的第一个多肽序列；将所述第一个多肽序列与至少一种第二多核苷酸或多肽序列比较，鉴定同源性。

在本说明书中引用的所有出版物和参考资料(包括但不限于专利和专利申请)通过引用完整地结合到本文中，就象具体单独地指出各个出版物或参考资料完全通过引用结合到本文中。同时以出版物和参考资料的上述方式将本申请要求获得优先权的任何专利申请通过引用完整地结合到本文中。

定义

“同一性”，如本领域所知，是指根据情况，通过比较序列而确定的两种或两种以上多肽序列或两种或两种以上多核苷酸序列之间的关系。在本领域内，“同一性”也指根据情况，通过匹配所述多肽或多核苷酸序列的字符串而确定的多肽或多核苷酸的序列之间的序列相似程度。采用已知方法可以容易地计算出“同一性”，所述方法包括但不限于以下文献中描述的方法：*Computational Molecular*

Biology, Lesk, A. M.编辑, Oxford University Press, New York, 1988;
Biocomputing: Informatics and Genome Projects, Smith, D. W.编辑,
Academic Press, New York, 1993; *Computer Analysis of Sequence Data*,
第1部分, Griffin, A. M.和Griffin, H. G.编辑, Humana Press, New Jersey,
5 1994; *Sequence Analysis in Molecular Biology*, von Heine, G., Academic
Press, 1987; 和 *Sequence Analysis Primer*, Gribskov, M.和 Devereux, J.,
编辑, M Stockton Press, New York, 1991; 以及 Carillo, H.和 Lipman, D.,
SIAM *J. Applied Math.*, 48: 1073 (1988)。设计测定同一性的方法用来
获得所测试序列之间的最大匹配。此外, 测定同一性的方法可用公
10 众可得到的计算机程序编纂。用于测定两序列间同一性的计算机程
序方法包括但不限于: GCG 程序包中的 GAP 程序(Devereux, J.等,
Nucleic Acids Research 12(1): 387 (1984)), BLASTP、BLASTN (Altschul,
S. F.等, *J. Molec. Biol.* 215: 403-410 (1990))以及 FASTA (Pearson 和
Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444-2448 (1988)。公众可以从
15 NCBI 和其它来源公开地得到 BLAST 程序系列(*BLAST Manual*,
Altschul, S.等, NCBI NLM NIH Bethesda, MD 20894; Altschul, S.等, *J.*
Mol. Biol. 215: 403-410 (1990)。也可以使用众所周知的 Smith
Waterman 算法测定同一性。

用于多肽序列比较的参数包括以下参数:

20 算法: Needleman 和 Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

比较矩阵: Henikoff 和 Henikoff 的 BLOSSUM62, *Proc. Natl. Acad.*
Sci. USA. 89: 10915-10919 (1992)

空位罚分: 8

空位长度罚分: 2

25 可使用这些参数的程序可以“空位(gap)”程序从 Genetics
Computer Group, Madison WI 公开获得。上述参数是进行多肽比较(同
时对末端空位没有罚分)的默认参数。

用于多核苷酸比较的参数包括下列参数:

算法: Needleman 和 Wunsch, J. Mol Biol. 48: 443-453 (1970)

比较矩阵: 匹配 = +10, 错配 = 0

空位罚分: 50

空位长度罚分: 3

5 可从 Genetics Computer Group, Madison WI 得到“空位”程序。
这些参数是进行核酸比较的默认参数。

根据情况, 下面的(1)和(2)提供多核苷酸和多肽的“同一性”优选意义。

10 (1) 多核苷酸实施方案还包括分离的多核苷酸, 所述分离的多核苷酸包括与 SEQ ID NO: 1 的参比序列有至少 50、60、70、80、85、90、95、97 或 100% 同一性的多核苷酸序列, 其中所述多核苷酸序列可以与 SEQ ID NO: 1 的参比序列相同, 或者与所述参比序列相比, 包括多至某一整数的核苷酸改变, 其中所述改变选自: 至少一个核苷酸缺失、取代(包括碱基转换和碱基颠换)或插入, 并且其中所述改变可以发生在所述参比核苷酸序列的 5' 或 3' 末端位置或两个末端位置间的任何地方, 或者各自单独散置于参比序列的核苷酸, 或者散置于参比序列的一个或多个连续组, 而且所述核苷酸改变的数量如下确定: SEQ ID NO: 1 核苷酸总数乘以定义同一性百分率的整数(除以 100), 随后从所述 SEQ ID NO: 1 核苷酸总数减去该乘积, 或:

20
$$n_n \leq x_n - (x_n \cdot y)$$

其中 n_n 是核苷酸改变的数目, x_n 是 SEQ ID NO: 1 核苷酸总数, y 是 0.50(50%)、0.60(60%)、0.70(70%)、0.80(80%)、0.85(85%)、0.90(90%)、0.95(95%)、0.97(97%)、1.00(100%)等, 而 \bullet 为乘法运算符号, 其中将 x_n 和 y 的任何非整数乘积从 x_n 减去之前, 将该乘积下舍到最近的整数。编码 SEQ ID NO: 2 多肽的多核苷酸序列改变可能在该编码序列中产生无义、错义或移码突变, 并因此在这样的改变后改变由所述多核苷酸编码的多肽。

举例来说, 本发明的多核苷酸序列可以与 SEQ ID NO: 1 的参比

序列相同，即可以 100%相同，或者与所述参比序列比较，该序列包括多至某一整数的核酸改变，以致同一性百分率小于 100%同一性。这样的改变选自：至少一个核苷酸缺失、取代(包括碱基转换和碱基颠换)或插入，其中所述改变可以发生在所述参比多核苷酸序列的 5' 或 3' 末端位置或那些末端位置间的任何地方，或者各自单独散置于参比序列的核苷酸，或者散置于参比序列的一个或多个连续组。如下确定给定百分率的核酸改变的数目：SEQ ID NO: 1 核酸总数乘以定义同一性百分率的整数(除以 100)，随后从所述 SEQ ID NO: 1 核酸总数减去该乘积，或：

$$n_n \leq x_n - (x_n \cdot y)$$

其中 n_n 是核酸改变的数目， x_n 是 SEQ ID NO: 1 中核酸的总数， y 是例如 0.70(70%)、0.80(80%)、0.85(85%) 等等，而 \cdot 是乘法运算符号，其中将 x_n 和 y 的任何非整数乘积从 x_n 减去之前，将该乘积下舍入为最接近的整数。

(2) 多肽实施方案还包括分离的多肽，所述分离的多肽包括与 SEQ ID NO: 2 的多肽参比序列有至少 50、60、70、80、85、90、95、97 或 100%同一性的多肽，其中所述多肽序列可以与 SEQ ID NO: 2 的参比序列相同，或者与所述参比序列相比，包括多至某一整数的氨基酸改变，其中所述改变选自：至少一个氨基酸缺失、取代(包括保守取代和非保守取代)或插入，并且所述改变可以发生在所述参比多肽序列的氨基末端或羧基末端位置或两个末端位置间的任何地方，或者各自单独散置于参比序列的氨基酸中，或者散置于参比序列的一个或多个连续的组，其中如下确定所述氨基酸改变的数量：SEQ ID NO: 2 氨基酸总数乘以定义同一性百分率的整数(除以 100)，然后从所述 SEQ ID NO: 2 氨基酸总数中减去该乘积，或：

$$n_a \leq x_a - (x_a \cdot y)$$

其中 n_a 是氨基酸改变的数目， x_a 是 SEQ ID NO: 2 中氨基酸的总数， y 是 0.50(50%)、0.60(60%)、0.70(70%)、0.80(80%)、0.85(85%)、

0.90(90%)、0.95(95%)、0.97(97%)或 1.00(100%)，而 \cdot 是乘法运算符号，其中将 x_a 和 y 的任何非整数乘积从 x_a 减去之前，将该乘积下舍入为最接近的整数。

5 举例来说，本发明的多肽序列可以与 SEQ ID NO: 2 的参比序列相同，即可以 100%相同，或者与所述参比序列比较，该序列包括多至某一整数的氨基酸改变，以致同一性百分率小于 100%同一性。这样的改变选自：至少一个氨基酸缺失、取代(包括保守取代和非保守取代)或插入，并且所述改变可以发生在所述参比多肽序列的氨基末端或羧基末端位置或这两个末端位置间的任何地方，或者各自单独
10 散置于参比序列的氨基酸中，或者散置于参比序列的一个或多个连续的组，其中如下确定给定百分率同一性的氨基酸改变的数量：SEQ ID NO: 2 氨基酸总数乘以定义同一性百分率的整数(除以 100)，然后从所述 SEQ ID NO: 2 氨基酸总数中减去该乘积，或：

$$n_a \leq x_a - (x_a \cdot y)$$

15 其中 n_a 是氨基酸改变的数目， x_a 是 SEQ ID NO: 2 的氨基酸总数， y 是例如 0.70(70%)、0.80(80%)、0.85(85%)等等，而 \cdot 是乘法运算符号，其中将 x_a 和 y 的任何非整数乘积从 x_a 减去之前，将该乘积下舍入为最接近的整数。

“个体”，当在本文内用以指生物时，是指多细胞真核生物，
20 包括但不限于后生动物、哺乳动物、羊科动物(ovoid)、牛科动物(bovid)、猿、灵长类动物和人类。

“分离”是指自然状态的“人为”改变，即如果发生分离，那么其自然环境已改变、或已经从其自然环境中分离出来、或者两者兼而有之。例如作为本文使用的此术语，在活的生物中天然存在的
25 多核苷酸或多肽并不是“分离的”，但与其天然状态共存的物质分开的相同多核苷酸或多肽是“分离的”。此外，通过转化、遗传操作或通过任何其它重组方法引入生物中的多核苷酸或多肽是“分离的”，即使它仍然存在于所述生物中，而所述生物可以是活的或非

活的。

“多核苷酸”一般指任何多核糖核苷酸或多脱氧核糖核苷酸，它们可以是包括单链和双链区的未修饰 RNA 或 DNA 或修饰的 RNA 或 DNA。

5 “变异体”是指与参比多核苷酸或多肽不同但保持原有基本特性的多核苷酸或多肽。典型多核苷酸变异体的核苷酸序列与另一参比多核苷酸不同。变异体核苷酸序列的变化可以改变或不改变参比多核苷酸编码的多肽的氨基酸序列。如下文所论述，核苷酸改变可能导致参比序列编码的多肽的氨基酸取代、添加、缺失、融合
10 和截短。典型多肽变异体的氨基酸序列与另一参比多肽不同。一般来说，差异是有限的，以便所述参比多肽和所述变异体的序列在总体上非常相似并在许多区是相同的。变异体和参比多肽可能由于任何组合的一个或多个取代、添加、缺失而氨基酸序列不同。取代或插入的氨基酸残基可以是或可以不是由遗传密码编码的氨基酸残基。
15 多核苷酸或多肽的变异体可以是天然存在变异体如等位基因变异体，或者所述变异体可以是未知的天然变异体。可以通过诱变技术或通过直接合成制备多核苷酸和多肽的非天然变异体。

“疾病”指任何由细菌感染引起或与细菌感染有关的疾病，包括例如上呼吸道感染以及侵入性细菌疾病，例如菌血症和脑膜炎。

20

实施例

使用标准技术实施下面的实施例，所述技术除另外详细描述外，对于本领域的技术人员是熟知的常规技术。这些实施例是说明性的，而不是不限制本发明。

25

实施例 1

脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB047 基因

脑膜炎萘瑟氏球菌菌株 ATCC13090 的 BASB047 基因示于 SEQ

5 ID NO: 1. 示于 SEQ ID NO: 2 的 BASB047 多核苷酸序列翻译显示与大肠杆菌 ferric aerobactin 受体前体(阴沟肠道菌素受体)非常相似。所述 BASB047 多肽包含具有前导信号序列特征的一个序列。可以在 SEQ ID NO: 2 显示的序列中在残基 25 以后切除该信号序列。BASB047 具有涉及铁摄入的外膜蛋白的特征。

实施例 2

脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB054 基因

10 脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB047 基因示于 SEQ ID NO: 3。示于 SEQ ID NO: 4 的 BASB054 多核苷酸序列的翻译显示与大肠杆菌有机溶剂耐受蛋白非常相似。所述 BASB054 多肽包含具有前导信号序列特征的一个序列。可以在 SEQ ID NO: 4 多肽的残基 22 以后切除该信号序列。

15 实施例 3

脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB068 基因

20 脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB068 基因示于 SEQ ID NO: 5。示于 SEQ ID NO: 6 的 BASB068 多核苷酸序列的翻译显示与大肠杆菌 AIDA-I 蛋白非常相似。所述 BASB068 多肽包含具有前导信号序列特征的一个序列。可以在 SEQ ID NO: 6 多肽的残基 34 以后切除该信号序列。BASB068 具有外膜蛋白的特征。

实施例 4

脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB069 基因

25 脑膜炎萘瑟氏球菌菌株 ATCC13090 中的 BASB069 基因示于 SEQ ID NO: 7。示于 SEQ ID NO: 8 的 BASB069 多核苷酸序列的翻译显示与大肠杆菌 AIDA-I 非常相似。所述 BASB069 多肽包含具有前导信号序列特征的一个序列。可以在 SEQ ID NO: 8 序列的残基 31 以后切

除该信号序列。BASB069 具有外膜蛋白的特征。

多核苷酸和多肽序列

5 SEQ ID NO: 1

菌株 ATCC13090 的脑膜炎奈瑟氏球菌 BASB047 多核苷酸序列

```
ATGCGTCATTCCCCTATTTTCAATGGTTATCTTTGCCTTTACTAAGTGTGGCAGTAACT
CAGCAGTTGTACGCTCAACCCAATGAGTCATTACCAACGGTTGAATTAGAGCCTGTGGTT
ATTACCATTGATAAGAGCGGTATGGCACTTGCCAATCGTATCACGCAAATGCCCATACC
ACCAAAGTTATTTATGAAGAGCAAATTCAGAGCAAGCAACAGGCTCTCGACAGCTTGCC
GATGTGATGGCACAGCTCATTCCAAGTTTGGGGTAAGTAGTGGCACTACCAGTAACTTT
GGGCAAACCATGCACGGTCTGCAAGTGAATTTTGTAAATGGCGTGCCTTTGACAGGT
TCGCGAGACATCTCTAGACAGCTTAATAGTATCAATCCCAATCAAGTGGCTAGAATTGAA
GTTTTATCAGGAGCAACCAGTATTTATGGGTCTGGAGCAACAGGCGGTTTGATTAATATC
GTTACTAAGTCTGATTTGGAAGAGGAGCAATTTGAAACCCGCATCGGTGTACATGGTAGT
AAATTATCCAGTGAAGGTATCGGTTATCAGGTAGGTGAGAGTGTAGCAGGTGTCAGCGAA
AATGGTAATGTCTTGCACGACTTGATGTGACTATCGCACCACAGGAGGGGCATTTGAT
GCTAACGGTAAACGCATCGCTCTGAGCCTGCCCAAAGTATAAGCAAGACAGCAAAGC
CTAAGTGTCAATACAAATGTTGATTGGCAACTTGACGACAAGCAAATATCAATCTGGCA
TTGACGCATTATAACGACAAACAGATACCGATTATGCACCTGATTATGGTAATCGCCTT
GCGGTGTTGTTTGGAGAAAAGCCTTCATTAATGCCATCAAAGGCTTATCATTATCAGAA
CAGCCAAAACCCAAAAGCACCTTTAATATCAACTATCATCATGATGATTTGTGGGGT
AACACCATCAATACCAATGCTTATATCGCAGAGAGAAAGGCAGATTTTATCCCTTGT
GCCCGTTTTTCGATCGCCAAAGCCCTGCCTATTTTACAAAGCATGAATTTGCCATCAGCC
ACTTTGGATGCTTATACCAAGGCTCCACAAGCTCGCGCCTATGGGGTGTACAAATCCGAA
TCTAAGGCAGAGGTACTAGGGCGTGTCCCTAATTTGAATAAGCCAAAAGAGCCCTATTT
TAA
```

SEQ ID NO: 2

根据 SEQ ID NO: 1 多核苷酸序列推导出的脑膜炎奈瑟氏球菌

10 BASB047 多肽序列

```
MRSHYFQWLSLPLLSVAVTQQLYAQPNESLPTVELEPVVITIDKSGMALANRITQMPHT
TKVIYEEQIQEQATGSRQLADVMAQLIPSLGVSSGTTSNFGQTMHGRQVQFLLNGVPLTG
SRDLSRQLNSINPNQVARIIVLSGATS IYSGATGGLINIVTKSDLEEQFETRIGVHGS
KLSSEGIGYQVQSVAGVSENGNVLARLDVDYRTTGAFDANGKRIAPEPAQTDKQDSKS
LSVNTNVDWQLDDKQINLALTHYNDKQDQTDYAPDYGNRLAVLFGKPSLNAIKGLSLSE
QPKTTKSTFNINYHDDDLWGNTINTNAYRREKGRFYPFVAPFSIAKALPILQSMNLPSA
TLDAYTKAPQARAYGVLQSESKAEVLRVPLNKNPKRALF
```

SEQ ID NO: 3

菌株 ATCC13090 的脑膜炎奈瑟氏球菌 BASB054 多核苷酸序列

```
TTGGCTCGTTTATTTCACTCAAACCACTGGTGCTGGCATTGGGCTTCTGCTTCGGCAGC
CATTGCGCCGCCCGGATGCCGTTGCGGGCGGAGGAAACGGACAATCCGACCGCCGGAGGA
AGTGTTCGGAGCGTGTCCGAACCCATGCAGCCTGCCGGCCTGAGCCTCGGTTTCGACCTGC
CTGTTTTGCAGTAACGAAAAGCGGCAAACCCGAAAAAACCGAATCTGCCGTCAAAGGAAGC
```

GGCGAAGGGCCTGTGCCGAAAACACACGCGAATTGTGCGCCGACAAGGTGGAAGGGCAG
TCGCAAGGTCAAGGTACGCGCGGAGGGCGGCGTCTGTTGTCGAACGCAACCGGACGACCCTT
AATGCCGACTGGGCGGATTACGACCAGTCCGGGCGACACCGTTACCGTAGGGCGACCGGTT
GCCCTCCAACAGGACGGTACGCTGATTCCGGGCGAAACCTGACCTACAATCTCGAGCAG
CAGACCGGCGAAGCGCACAACTCCGCATGGAACCGAACAAAGGCGGACGGCGGCTGCAA
AGCGTCAGCCGACCGCCGAAATGTTGGGCGAAGGGCATTACAAACTGACGGAAACCCAA
TTCAACACCTGTTCCGCGCGGATGCCGGCTGGTATGTCAAGGCAGCCTCTGTGCGAAGCC
GATCGGGA AAAAGGCATAGGCGTTGCCAAACACGCCGCTTCGTGTTCCGGCGGCGTTCCT
ATTTCTACACCCCTTGGGCGGACTTCCCGCTTGACGGCAACCGCAAAGCGGCGCTGCTT
GTTCCCTCACTGTCCGCGGTTCCGACGGCGTTCCTTTCCGTTCCCTATTATTTC AAC
CTTGCCCCAATCTCGATGCCACGTTCCGCGCCAGCGTGATCGGCGAACGCGGCGCGGTC
TTGACGGGCGAGTACGCTACCTGCGGCGGATTATGCCGGCCAGTCCGACCTGACCTGG
CTGCCGACGACAAGAAAAGCGGCGAGGAATAACCGCTATCAGGCGAAATGGCAGCATCGG
CAGGACATTTCCGACACGCTTCAGGCGGGTGTGATTTCAACCAAGTCTCCGACAGCGGC
TACTACCGGACTTTTACGGCAACAAAGAAATCGCCGGCAACGTCAACCTCAACCGCCGT
GTATGGCTGGATTATGGCGGCGGGCGGGCGGGCAGCCTGAATGCCGGCCTTTCGGTT
CTGAAATACCAGACGCTGGCAAACCAAAGCGGCTACAAAGACAAACCGTATGCCCTGATG
CCGCGCTTTCGCGCGATTGCGCAAAAACCGGCGGGCGCAAATCGGCGTGTCCGCA
CAATTTACCCGCTTACGCCACGACAGCCGCAAGACGGCAGCCGCTCGTCTATCC
GACATCAAATGGGATTTACGCAACAGCTGGGGTTACGTCCGTCCCAAACCTCGGACTGCAC
GCCACCTATTACAGCCTCAACCGCTTCGGCAGCCAAGAAGCCCGACGCGTCAGCCGCACT
CTACCCATCGTCAACATCGACAGCGGCATGACCTTCGAACGCAATACGCGGATGTTCCGG
GGAGAAGTCTGCAAACCTCGAGCCGCGCCTGTTCTACAACCTATATTCTGCAAATCC
CAAACGACCTGCCCAATTTGATTTCGTCCGAAAGCAGCTTCGGCTACGGGCGAGCTTTTT
CGTGAAAACCTCTATTACGGCAACGACAGGATTAACACCGCAAACAGCCTTTCCGCGGCC
GTGCAAAGCCGTATTTTGGACGGCGCGACGGGGCGAGAGCGTTTCCGCGCGGCGATCGGG
CAGAAATCTACTTCAA AAAACGACGAGTCATGCTTGACGGCAGTGTCCGCAAAAACCG
CGCAGCCGTTCCGACTGGGTGGCATTCCGCTCCAGCGGCATCGGCGAGCCGCTTCATCCTC
GACAGCAGCATCCACTACAACCAAACGACAAACCGCGCCGAGAACTACGCCGTCCGGTGA
AGCTACCGTCCCGCACAGGGCAAAGTGCTGAACGCCCGCTACAAATACGGGCGCAACGAA
AAAATCTACCTGAAGTCGACGGTTCCTATTTTACGACAAACTCAGCCAGCTCGACCTG
TCCGCACAATGGCCCTGACCGCAACCTGTCCGCCGTCTCGTTACAACCTACGGTTTT
GAAGCCAAAACCGATAGAGGTGCTGGCGGGTGC GGAATACAAAAGCAGTTGCGGCTGC
TGGGCGCGGGCGTGTACGCCAACGCTACGTTACCGGCGAAAACACCTACAAAACGCT
GTC TTTTTCTCACTTCACTGAAAGACCTCAGCAGTGTCCGCGAAACCCCGCAGACAGG
ATGGATGTCGCCGTTCCCGCTATATCCCCGCCACTCTCTTCCGCGGACGCAACAAA
CGGCCCTGA

SEQ ID NO: 4

根据 SEQ ID NO: 3 多核苷酸序列推导出的脑膜炎奈瑟氏球菌

BASB054 多肽序列

MARLFS LKPLVLALGF CFGTHCAAADAVAAEETDNPTAGGSVRSVSEPMQPAGLSLGSTC
LFC SNESGKPEKTESAVKSGGEGVPENHTRIVADKVEGQSQVKVRAEGGVVVERNRTTL
NADWADYDQSGDTVTVGDRFALQQDGT LIRGETLTYNLEQQTGEAHNVRMETEQGGRRLQ
SVSR TAEMLGEGHYKLTETQFN TCSAGDAGWYVKAASVEADREKIGVAKHAAAFVFGGVP
IFYPWADFP L DGNRKSGLLVPSLSAGSDGVSLSVPY YFN LAPNL DATFAPS V IGERGAV
FDGQVRYLR PDYAGQSDLTWLP HDKKSGRNRYQAKWQHRHDSDTLQAGVDFNQVSDSG
YYRDFYGNKEIAGNVNLNRRVWLDYGGRAAGGSLNAGLSVLKYQTLANQSGYKDKPYALM
PRLSADWRKNTGRAQIGVSAQFTRF SHDSRQDGSRLVVYPDIKWDFSN SWGYVRPKLGLH
ATYYS LNRFGSQEARRVSR TLP I VNIDSGMTFERNTRMFGGEVLQ TLEPRLFYNYI PAKS

QNDLPNFDSSSESSFGYGQLFRENLYYGNDRINTANSLSAVQSRILDGATGAERFRAGIG
QKFFYFNDAVMDGSGVKKPRSRSDWVAFASSGIGSRFILDSSIHYNQMDKRAENYAVGA
SYRPAQGXVNLNARYKYGRNEKIYLKSDGSYFYDKLSQLDLSAQWPLTRNLSAVVRYNYGF
EAKKPIEVLAGA EYKSSCGCWGAGVYAQRVVTGENTYKNAVFFSLQLKDLSSVGRNPADR
MDVAVPGYI PAHSL SAGRNRKP

SEQ ID NO: 5

菌株 ATCC13090 的脑膜炎奈瑟氏球菌 BASB068 多核苷酸序列

ATGAACTCGAAGCAAGCAAGCAAGCAAGCAAGCAGAAGTTAAAAAATCATTATTATA
AGTCTATTTTTTCTATTCTTTATACCTCTCCGCTTTTGGCTGTTGATTACGTTTACGAC
AAAACCAAGCTCACTGATGATGAAATTACCCGCTTAAAAAACTCCGCGATAGAAATAGT
GAATATTGGAAAGAAGAACTTATCACATAAAAAGTAACAACCGAGTTTATCCAAACATT
CCCGCATTATCCCTAAACATCCTTTTCGATCCATTTCGAAAAATCAATAATTCAAAAAGG
ATTTCTTTTATGACAAAGAATACACTGAAGATTACCTTGTTGGTTTTGCTCAAGGTTTA
GGCGTTGCAAAAAGAAATGGGGAAACAGAAAAACCAATACGGCAATATTTAAGGAATGT
TTAAACACTGGGAAATATAGTGATGATACTTGCAAATCTCAACAATCTATTCCCTACAGTA
AGAAGTGATATTTTCGCCCTAAATACGAAAATAAAAAATAGCCATATCAACAGCGAAAT
TTCGCTGTTCGGTAATTATACAAAATTGATGTACTCAGCCCAACATCATTCTATTTGGTCA
GACATCTCTATCCAAATCAGAATTATCTCTTGACGTTGATAACTCACACGTTTATCGGA
CAAACGATTGATTTGGGAGCATTAGAATTAACAAAATCTCTATGGGAACCCCGTTGGAA
TCTAATATTGATTATTTAATAACAAAAATGCTGAAATTCGTTTTAACACTAAAAGTGAA
AGTTTACTCGTAAAAGAAGATTATGCTGGCGGAGCTCGTTTTTCGTTTTGCTTACGGTCTA
AAAGATAAAGTCCCTGAAACACCAATTTAARCTTTTGAAAAAATATAACTGGCACATCA
GATATTATTTTTGAGGGAAAAGCGTTGGATAATTTAAAAACCTAGACGGGCATCAAATT
ATCAAAGTAAATGGCACAGCAGATAAACACGCATTCCGCTCTTTCTGGCAAACACCAAAG
GGAATTTATACGCTTTCTTTACAACAACGCCAGAGGGCTTTTTTACCAAAGTGCAAGAA
CGCGATGATATTGCGATTTATGCACAACAAGCCCAAGCCGCAATACCTTATTCGCCCTTG
CGTTTAAACGACAAAACAGCGATATTTTTGACCGCACTTTACCGCGCAAAGGCTTGTGG
TTACGTTGATTGACGGACATTCAGCCAATGGGTACAAGGCAAAAACGGCACCAGTAGAA
GGAAATCGTAAAGGCATACAACCTGGTGGCGATGTTTTCTCATTGCAAATCACAATAT
CAACTTTCCGTTGGCTTAATGGGCGGACAAGCAGAACAACGCAGTACTTTCCGCAACCCA
GATACAGACAATCTTACAACGGGAAATGTGAAAGGCTTTGGTGCAGGCGTTTACGCCACT
TGGCATCAGCTTCAGGACAAACAGACAGGTGCGTATGCGGATAGCTGGGTACAATATCAA
CGTTTTCCGCCACCGTATCAACACTGAAGATGGTACAGAACGTTTTACTTCAAAGGTATT
ACTGCCTCAATTGAAGCAGGTTACAACGCTTTATTGGCGGAACAGGTAATAAAAAGGGC
AACAGCCTTCGTGTTTTACCTACAACCACAGGCGCAATTGACTTATTTGGGGGTAAACGGA
AAATTCAGCGATAGCGAAAATGCCACGTGAATTTACTTGGCTCTCGCCAATTACAAAGC
CGAGTGGGCGTTCAAGCTAAAGCTCAATTTGCTTTCACTAATGGCGTCACTTTCCAACCA
TTTGTTCCGTCAATTCATCTACCAACAAAACCTTTCCGGGTAGAAATGGACGGAGAA
CGTCCGAGTGATAAACAACAAAACCGCGATTGAAAGCCAAATTAGGCGTTGCGGTAAAATT
AAATCTCACTTAACTTTACAAGCAACATTCAACCGCCAAACAGGCAAACATCATCAAGCT
AAACAAGGCGCATTGAATTTACAGTGGACGTTTTAA

5 SEQ ID NO: 6

根据 SEQ ID NO: 5 多核苷酸序列推导出的脑膜炎奈瑟氏球菌
BASB068 多肽序列

MKLEASKQASKQKFKKSFIIISLFFSILYTSPLLAVDYVYDKTKLTDDEITRLKCLRDRNS

EYWKEETYHIKSNRVPNIPALFPKHPFDPFENINNSKRISFYDKEYTEDYLVGPAQGL
GVAKRNGETEKPIRQYFKECLNTGKYSDDTCKSQSIPTVRSDFALNTKIKNSHINSEI
FAVGNYYTKLMSAQHHSIWSEHLYSNSELSLDVDNSHVIGQIDLGALLETNSLWEPRWN
SNIDYELITKNAEIRFNTKSESLLVKEDYAGGARFRFAYGLKDKVPETPILTFEKNTGTS
DIIIFEGKALDNLKHLDDGHQIIKVNVTADKHAFLRSLGKHQKGIYTLSSLQQRPEGFFTKVQE
RDDIAIYAQQAQAANTLFFALRLNDKNSDIFDRTLPRKGLWLRVIDGHSSQWVQGKTAPVE
GNRKGIQLGDDVFSLQNHNYQLSVGLMGGQAEQRSTFRNPDNDLTTGNVKGFGAGVYAT
WHQLQDKQTGAYADSWVYQFRFRHRINTEDGTERFSTKGITASIEAGYNALLAEHVTKKG
NSLRVYLQPPAQLTYLGVNGKFSSENAHVNLLGSRQLQSRVGVQAKAQFAFTNGVTFQP
FVSVNSIYQOKPFGVEMDGERRVINNKTAIESQLGVAVKIKSHLTLQATFNRQTGKHHQA
KQGALNLQWTF

SEQ ID NO: 7

菌株 ATCC13090 的脑膜炎奈瑟氏球菌 BASB069 多核苷酸序列

ATGACACTGAAAGCAAGCAAGCAAGCAAGCGGTTCGGGTTAATCTATTAACATTATCTGTT
TTATCGCTGTTTTCGACGCCATATGTTTGTGGTTCGGATGCGTACGATCCCGTCAAAGAA
GCCGAGATTA AAAACAAATTTATTTTAGAAGCGGCGGAAGACAGAAATCCCACGTTTGG
CGCGGCCCGTGCAGCATATCTTTTGATTGCTTCGGTATGTTTCAGAGCTCAGCTTGGTTCA
AATACTCGTTCCTACCAAATCGGCGACGATGCCGATTTTTTCATTTTCAGACAAGCCGAAA
CCCCGCACTTCCCATTTATTTTCCAGCGGTA AAAACCGATCAAATTCATCCGAATATGGG
TATGACGAAATCAATATCCAAGGTA AAAACTACAATAGCGGCATACTCGCCGTCGATAAT
ATGCCCGTTGTTAAGAAATATATTACAGATACTTACGGGGATAATTTAAAGGATGCGGTT
AAGAAGCAATTACAGGATTTATACAAAACAAGACCCGAAGCTTGGGAAGAAAATAAAAA
CGGACTGAGGAGGCGTATATAGAACAGCTTGGACCAAATTTAGTATACTCAAACAGAAA
AACCCCGATTTAATTAATAAATTTGGTAGAAGATTCCGTACTCACTCCTCATAGTAATACA
TCACAGACTAGTCTCAACAACATCTTCAATAAAAAATTACACGTCAAATCGAAAACAAA
TCCCACGTGCGCCGACAGGTGTTGGAACGACCAAGATGACGCTGAAAGATTCCTTTGG
GAACCGCGCCGCAATCCGACATCCATACGCTGGAACTTCCGATAATGCCCGCATCCGC
CTGAACACGAAAGATGAAAACTGACCGTCCATAAAGCGTATCAGGGCGGCGCGGATTT
CTGTTCCGCTACGACGTGCGGAGTCCGACGAACCCGCCCTGACCTTTGAACAAAACGTC
AGCGGAAAATCCGGCGTGGTTTTGGAACGCCGGCCGGAATCTGAAAACGCTCGACGGG
CGCAAATGATTGCGGCGGAAAAGGCAGACCCTAATTCGTTTGCGTTTAAACAAAATTAC
CGGCAGGACTGTACGAATTATGCTCAAGCAATGCGAAGGCGGATTTGCTTGGGTGTG
CAGCGTTTGGCTATCCCTGAGGCGGAAGCGGTTTTATATGCCCAACAGGCTTATGCGGCA
AATACTTTGTTTGGGCTGCGTGCCGCCGACAGGGGCGACGACGTGTATGCCGCCGATCCG
TCCCGTCAA AAATTTGTTGCTGCGCTTCA TCGGCGGCCGGTTCGCATCAAATATACGGGGC
GGCGCGGCTGCGGACGGGCGGCGCAAAGGCGTGCAATCGGCGGTGAGGTGTTTGTACGG
CAAATGAAGGCAGTCCGCTGGCAATCGGCGTGATGGGCGGCAGGGCTGGCCAGCACGCA
TCAGTCAACGGCAAAGGCGGTGCGGCAGGCAGTTATTTGCATGGTTATGGCGGGGGTGT
TATGCTGCGTGGCATCAGTTGCGCGATAAACAAACGGGTGCGTATTTGGACGGCTGGTTG
CAATACCAACGTTTCAAACACCGCATCAATGATGAAAACCGTGCGGAACGCTACAAAACC
AAAGTTGGACGGCTTCTGTGCAAGGCGGCTACAACGCGCTTGTGGCGGAAGGCGTTGTC
GGAAAAGGCAATAATGTGCGGTTTTACCTGCAACCGCAGGCGCAGTTTACCTACTTGGGC
GTAAACGGCGCTTTACCGACAGCGAGGGACGCGGTCGGACTGCTCGGCAGCGGTCAG
TGGCAAAGCCGCGCATTCCGGCAAAAACCCGTTTTGCTTTGCGTAACGGTGTCAAT
CTTCAGCCTTTTGCCGCTTTTAAATGTTTTGCACAGGTCAAATCTTTCCGCGTGGAAATG
GACCGCGAAAAACAGACGCTGGCAGGACGAGCGGCTCGAAGGGCGGTTCCGCATTGAA
GCCGGTTGGAAGGCCATATGTCCGCACGCATCGGATACGGCAAAGGACGGACGGCGAC
AAAGAAGCCGCAATGTCGCTCAAATGGCTGTTTTGA

SEQ ID NO: 8

根据 SEQ ID NO: 7 多核苷酸序列推导出的脑膜炎奈瑟氏球菌
BASB069 多肽序列

MTLKASKQASGRVNLTLTSLVLSLFCPTPYVCGSDAYDPVKEAEIKNKFILEAAEDRNSHW
RGPCSI SFDCFGMFRAQLGSNTRSTKIGDDADFSFSDKPKPGTSHYFSSGKTDQNSSEYG
YDEINI QGKNYNSGILAVDNMPVVKYITD TYGDNLKDAVKKQLQDLYKTRPEAW EENK
RTEEAYIEQLGPKFSILKQKNPDLINKLVEDSVLTPHSNTSQTSLNFI FNKCLHVKIENK
SHVAGQVLELTKMTLKDSLWEPRRHSDIHTLETS DNARIRLNTKDEKLTVHKAYQGGADF
LFGYDVRESDEPALTFEQNVSGKSGVVLERRPENLKTLDGRKLI AAEKADPNSFAFKQNY
RQGLYELLKQCEGGFCLGVQRLAIP EAEAVLYAQQAYAANTL FGLRAADRGDDVYAADP
SRQKLWLRFIGGRSHQNI RGGAAADGRRKGVQIGGEVFVRQNEGSRLAIGVMGGRAGQHA
SVNGKGAAGSYLHG YGGGVYAAWHQLRDKQTGAYLDGWLQYQRFKHRINDENRAERYKT
KGWTASVEGGYNALV AEGVVGKGNVRFYLQPQAQFTYLG VNGGFTDSEGTA VGLLGSQ
WQSRAGIRAKTRFALRNGVNLQPF AAFNVLHRSKSGVEMDGEKQTLAGRTALEGRFGIE
AGWKGHMSARIGYGRRTDGDKEAALS LKWLF

保藏材料

5 包含脑膜炎奈瑟氏球菌血清群 B 菌株的保藏物于 1997 年 6 月 22 日保藏在美国典型培养物保藏中心(本文称为“ATCC”), 保藏号为 13090。所述保藏物描述为脑膜炎奈瑟氏球菌(Albrecht 和 Ghon)并且用脑膜炎奈瑟氏球菌分离株构建的 1.5-2.9 kb 冻干插入片段文库。在 Int. Bull. Bacteriol. Nomencl. Taxon. 8: 1-15 (1958)中介绍了该保藏物。

脑膜炎奈瑟氏球菌菌株保藏物在本文内称为“保藏菌株”或“保藏菌株的 DNA”。

10 所述保藏菌株包含全长 BASB047、BASB054、BASB068 和 BASB069 基因。包含在所述保藏菌株内的多核苷酸的序列以及其编码的任何多肽的氨基酸序列在与本文序列的任何描述发生冲突的情况下, 都作为对照(controlling)。

15 所述保藏菌株已经根据国际承认用于专利程序的微生物保存布达佩斯条约的条款保藏。在专利颁发时, 所述保藏菌株将不可撤回并且无限制或无条件地发放给公众。提供所述保藏菌株仅仅是为了方便本领域的技术人员, 而不是承认需要保藏物能够实现(enablement), 例如根据 35 U.S.C. § 112 的要求。

说明书的核苷酸和氨基酸序列

序列表

5	<110> SmithKline Beecham Biologicals S. A.	
	<120> 新型化合物	
10	<130> BM45352	
	<160> 8	
	<170> FastSEQ for Windows Version 3.0	
15	<210> 1	
	<211> 1203	
	<212> DNA	
	<213> 脑膜炎奈瑟氏菌	
20	<400> 1	
	atgcgtcatt cccactat tcaatggta tctttgcct tactaagtgt ggcagtaact 60	
	cagcagttgt acgctcaacc caatgagta ttaccaacgg ttgaattaga gcctgtggtt 120	
	attaccattg ataagagcgg tatggcaact gccaatcgta tcacgcaaat gcccataacc 180	
25	accaaagtta tttatgaaga gcaaattcaa gagcaagcaa caggctctcg acagcttgcc 240	
	gatgtgatgg cacagctcat tccaagttg ggggtaagta gtggcactac cagtaacttt 300	
	gggcaaacca tgcacggtcg tcaagtgeaa tttttgtaa atggcgtgcc ttgacaggt 360	
	tcgcgagaca tctctagaca gettaatagt atcaatccca atcaagtgc tagaattgaa 420	
	gttttatcag gagcaaccag tatttatggg tctggagcaa caggcggttt gattaatata 480	
30	gttactaagt ctgatttggg agaggagcaa ttgaaaccc gcacggtgt acatggtagt 540	
	aaattatcca gtgaaggtat cggttatcag gtaggtcaga gtgtagcagg tgcagcgaa 600	
	aatggtaatg tccttgacg acttgatgtc gactatcga ccacaggagg ggcatttgat 660	
	gctaaccgta aacgcacgc tcctgagcct gcccaactg ataagcaaga cagcaaaagc 720	
	ctaagtgtca atacaatgt tgattggcaa cttgacgaca agcaaaatat caatctggca 780	
35	ttgacgcatt ataacgacaa acaagatacc gattatgcac ctgattatgg taatgcctt 840	
	gggtgtgtt ttggagaaaa gccttcatta aatgcatca aaggcttate attatcagaa 900	
	cagccaaaaa ccacaaaag cacctttaat atcaactate atcatgatga tttgtggggt 960	
	aacaccatca ataccaatgc ttattatcgc agagagaaag gcagatttta tcctttggt 1020	
	gcccgtttt cgatcgccaa agccctgect atttacaaa gcataatgtt gccatcagcc 1080	
40	actttgatg cttataccaa ggetccacaa gctcgcgct atgggtgtt acaatccgaa 1140	
	tctaaggcag aggtactagg gcgtgtccct aatttgaata agcccaaaag agccctattt 1200	
	taa 1203	
	<210> 2	
45	<211> 400	
	<212> PRT	
	<213> 脑膜炎奈瑟氏菌	
	<400> 2	
50	Met Arg His Ser His Tyr Phe Gln Trp Leu Ser Leu Pro Leu Leu Ser	
	1 5 10 15	
	Val Ala Val Thr Gln Gln Leu Tyr Ala Gln Pro Asn Glu Ser Leu Pro	
	20 25 30	
	Thr Val Glu Leu Glu Pro Val Val Ile Thr Ile Asp Lys Ser Gly Met	

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

<210> 3

<211> 2409

<212> DNA

50

<213> 脑膜炎奈瑟氏菌

<400> 3

ttggctcggtt tattttcaact caaaccactg gtgctggcat tgggtttctg cttcggcacg	60
cattgcgccg ccgccgatgc cgttgccggcg gaggaaacgg acaatccgac cgcgggagga	120
agtgttcgga gcgtgtccga acceatgcag cctgccggcc tgagcctcgg ttcgacctgc	180

ctgttttgca gtaacgaaag cggcaaaccc gaaaaaaccc aatctgccgt caaaggaagc 240
ggcgaagggc ctgtgcccgaa aaaccacacg cgaattgtcg ccgacaaggt ggaagggcag 300
tcgcaggta aggtacgcgc ggagggcggc gtcgtgtcg aacgcaaccg gacgaccctt 360
aatgocgact ggccggatta cgaccagtcg ggcgacaccg ttaccgtagg cgaccggttc 420
5 gcectccaac aggacggtac gctgattcgg ggcgaaacce tgacctaca tctcgagcag 480
cagaccggcg aagcgcacaa cgtccgcgat gaaaccgaac aaggcggacg cgggctgcaa 540
agcgtcagcc gcaccgcgga aatgttgggc gaagggcatt acaactgac ggaaccccaa 600
tcaaacacct gttccgcggc cgatgccggc tggtagtca aggcagcctc tctcgaagcc 660
gatcgggaaa aaggcatagg cgttgccaaa caecgcgect tcgtgttcgg cggcgttctt 720
10 attttctaca ccccttgggc ggacttccc cttgacggca accgcaaaag cggcctgctt 780
gttccctcac tgcgcggcgg ttccgacggc gtttccctt cegtcccta ttattcaac 840
cttgccccca atctcgatgc cagcttcggc cccagcgtga tcggcgaacg cggcgcggtc 900
tttgacgggc aggtacgcta cctgcggcgc gattatgccg gccagtcoga cctgacctgg 960
ctgcgcgacg acaagaaaag cggcaggaat aaccgctatc aggcgaaatg gcagcatcgg 1020
15 cagcacattt ccgacacgct tcagggcggt gtcgatttca accaagtctc cgacagcggc 1080
tactaccgcg acttttacgg caacaaagaa atcgcggcga acgtcaacct caaccgcctt 1140
gtatggctgg attatggcgg cagggcggcg ggcggcagcc tgaatgccgg cctttcggtt 1200
ctgaaatacc agacgctggc aaaccaaagc ggctacaaag acaaacgta tgcctgatg 1260
ccgcgccttt ccgcgattg gcgcaaaaac accgcagggc cgcaaatcgg cgtgtccgca 1320
20 caatttacc gttcagcca cgacagccgc caagacggca gccgcctcgt cgtctatccc 1380
gacatcaaat gggatttcag caacagctgg ggttacgtcc gtcccaact cggactgcac 1440
gccacctatt acagcctcaa ccgcttcggc agccaagaag cccgacgct cagccgcact 1500
ctaccatcgc tcaacatcga cagcggcatg acctcgaac gcaatacgc gatgttcggc 1560
ggagaagtcc tgcaaacctc cgagccgcgc ctgttctaca actatattcc tgccaaatcc 1620
25 caaaacgacc tgcccaattt tgattcgtcg gaaagcagct tcggctacgg gcagctttt 1680
cgtgaaaacc tctattacgg caacgacagg attaacaccg caaacgcct ttcgcgcgc 1740
gtgcaaagcc gtattttgga cggcgcgacg ggggcagagc gttccgcgc cggcatcggg 1800
cagaaattct actcaaaaa cgacgcagtc atgcttgac gcagtgctgg caaaaaaccg 1860
cgcagccgtt ccgactgggt ggcattcgc tccagcggca tcggcagccg cttcctcctc 1920
30 gacagcagca tccactaca ccaaacgac aaacgcgcg agaactacgc cgtcgggtgca 1980
agctaccgtc ccgcacaggg caaagtctg aacgccgct acaatacgg gcgcaacgaa 2040
aaaaatctacc tgaagtccga cggttctat ttttaogaca aactcagcca gctcgacctg 2100
tcgcacaaat ggcccctgac gcgcaacctg tcggccgctg tcggttaca ctacggtttt 2160
gaagccaaaa aaccgataga ggtgctggcg ggtgcggaat acaaaagcag ttgcggctgc 2220
35 tggggcgcgg gcgtgtacgc ccaacgctac gttaccggcg aaaacaccta caaaaacgt 2280
gtctttttct cacttcagtt gaaagacctc agcagtgctg gcagaaacct cgcagacagg 2340
atggatgtcg ccgttcccgg ctatatcccc gccactctc tttccgcggc acgcaacaaa 2400
cggcctga 2409

40 <210> 4
<211> 802
<212> PRT
<213> 脑膜炎奈基氏菌

45 <400> 4
Met Ala Arg Leu Phe Ser Leu Lys Pro Leu Val Leu Ala Leu Gly Phe
1 5 10 15
Cys Phe Gly Thr His Cys Ala Ala Ala Asp Ala Val Ala Ala Glu Glu
20 25 30
50 Thr Asp Asn Pro Thr Ala Gly Gly Ser Val Arg Ser Val Ser Glu Pro
35 40 45
Met Gln Pro Ala Gly Leu Ser Leu Gly Ser Thr Cys Leu Phe Cys Ser
50 55 60
Asn Glu Ser Gly Lys Pro Glu Lys Thr Glu Ser Ala Val Lys Gly Ser
55 65 70 75 80

caaacgattg atttgggagc attagaatta acaaattctc tatgggaacc cgttggaac 720
 tctaattattg attatttaat acaaaaaaat gctgaaatte gttttaacac taaaagttaa 780
 agtttactcg taaaagaaga ttatgctggc ggagctcgtt ttcgttttgc ttacggtcta 840
 aaagataaag tccttgaaac accaatttta acttttgaaa aaaatataac tggcacatca 900
 5 gatattattht ttgagggaaa agcgttggat aattttaaac acctagacgg gcatcaaatt 960
 atcaaagtaa atggcacagc agataaacac gcattccgctc tttctggcaa acacaaaag 1020
 ggaatttata cgctttcttt acaacaacgc ccagagggtc tttttaccaa agtgaagaa 1080
 cgcgatgata ttgcgattta tgcacaacaa gcccaagccg ccaataacctt attcgccttg 1140
 cgtttaaacg acaaaaacag cgatattttt gaccgcactt taccgcgcaa aggcttgtgg 1200
 10 ttacgtgtga ttgaaggaca ttccagccaa tgggtacaag gcaaaaacgc accagtagaa 1260
 ggaaatcgta aaggcataca acttgggtggc gatgttttct cattgcaaaa tcacaactat 1320
 caactttccg ttggcttaat gggcggacaa gcagaacaac gcagtacttt ccgcaaccca 1380
 gatacagaca atcttacaac gggaaatgtg aaaggetttg gtgcaggcgt ttacgccact 1440
 tggcatcagc ttcaggacaa acagacaggt gcgtatggcg atagctgggt acaatatcaa 1500
 15 cgtttccgcc accgtatcaa cactgaagat ggtacagaac gttttacttc aaaaggtatt 1560
 actgectcaa ttgaagcagg ttacaacgct ttattggcgg aacacgtaac taaaaggggc 1620
 aacagccttc gtgtttacct acaaccacag gcgcaattga cttatttggg gtaaacgga 1680
 aaattcagcg atagcgaaaa tgcccacgtg aatttacttg gctctcgcca attacaaagc 1740
 cgagtgggcg ttcaagctaa agctcaattt gctttacta atggcgctac ttccaacca 1800
 20 tttgttccg tcaattcaat ctaccaacaa aaacctttcg gggtagaaat ggacggagaa 1860
 cgtcgagtga taaacaacaa aaccgcgatt gaaagccaat taggcgttgc ggtaaaaatt 1920
 aaatctcact taactttaca agcaacatc aaccgcaaaa caggcaacaa tcatcaagct 1980
 aaacaaggcg cattgaattt acagtggacg ttttaa 2016

25 <210> 6
 <211> 671
 <212> PRT
 <213> 脑膜炎奈瑟氏菌

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

		195				200				205				
		Leu Ser	Leu Asp	Val Asp	Asn Ser	Ser His	Val Ile	Gly Gln	Thr Ile	Asp				
		210			215			220						
5		Leu Gly	Ala Leu	Glu Leu	Thr Asn	Ser Ser	Leu Trp	Glu Pro	Arg Trp	Asn				
		225			230			235		240				
		Ser Asn	Ile Asp	Tyr Leu	Ile Thr	Lys Asn	Ala Glu	Ile Arg	Phe Asn					
				245			250		255					
		Thr Lys	Ser Glu	Ser Leu	Leu Val	Lys Glu	Asp Tyr	Ala Gly	Gly Ala					
			260			265			270					
10		Arg Phe	Arg Phe	Ala Tyr	Gly Leu	Lys Asp	Lys Val	Pro Glu	Thr Pro					
			275			280		285						
		Ile Leu	Thr Phe	Glu Lys	Asn Ile	Thr Gly	Thr Ser	Asp Ile	Ile Phe					
		290			295			300						
		Glu Gly	Lys Ala	Leu Asp	Asn Leu	Lys His	Leu Asp	Gly His	Gln Ile					
15		305			310			315		320				
		Ile Lys	Val Asn	Gly Thr	Ala Asp	Lys His	Ala Phe	Arg Leu	Ser Gly					
				325			330		335					
		Lys His	Gln Lys	Gly Ile	Tyr Thr	Leu Ser	Leu Gln	Gln Arg	Pro Glu					
				340			345		350					
20		Gly Phe	Phe Thr	Lys Val	Gln Glu	Arg Asp	Asp Ile	Ala Ile	Tyr Ala					
			355			360		365						
		Gln Gln	Ala Gln	Ala Ala	Asn Thr	Leu Phe	Ala Leu	Arg Leu	Asn Asp					
		370			375			380						
		Lys Asn	Ser Asp	Ile Phe	Asp Arg	Thr Leu	Pro Arg	Lys Gly	Leu Trp					
25		385			390			395		400				
		Leu Arg	Val Ile	Asp Gly	His Ser	Ser Ser	Gln Trp	Val Gln	Gly Lys	Thr				
				405			410		415					
		Ala Pro	Val Glu	Gly Asn	Arg Lys	Gly Ile	Gln Leu	Gly Gly	Asp Val					
				420			425		430					
30		Phe Ser	Leu Gln	Asn His	Asn Tyr	Gln Leu	Ser Val	Gly Leu	Met Gly					
			435			440		445						
		Gly Gln	Ala Glu	Gln Arg	Ser Thr	Phe Arg	Asn Pro	Asp Thr	Asp Asn					
		450			455			460						
		Leu Thr	Thr Gly	Asn Val	Lys Gly	Phe Gly	Ala Gly	Val Tyr	Ala Thr					
35		465			470			475		480				
		Trp His	Gln Leu	Gln Asp	Lys Gln	Thr Gly	Ala Tyr	Ala Asp	Ser Trp					
				485			490		495					
		Val Gln	Tyr Gln	Arg Phe	Arg His	Arg Ile	Asn Thr	Glu Asp	Gly Thr					
				500			505		510					
40		Glu Arg	Phe Thr	Ser Lys	Gly Ile	Thr Ala	Ser Ile	Glu Ala	Gly Tyr					
			515			520		525						
		Asn Ala	Leu Leu	Ala Glu	His Val	Thr Lys	Lys Lys	Gly Asn	Ser Leu	Arg				
		530			535			540						
		Val Tyr	Leu Gln	Pro Gln	Ala Gln	Leu Thr	Tyr Leu	Gly Val	Asn Gly					
45		545			550			555		560				
		Lys Phe	Ser Asp	Ser Glu	Asn Ala	His Val	Asn Leu	Leu Gly	Ser Arg					
				565			570		575					
		Gln Leu	Gln Ser	Arg Val	Gly Val	Gln Ala	Lys Ala	Gln Phe	Ala Phe					
				580			585		590					
50		Thr Asn	Gly Val	Thr Phe	Gln Pro	Phe Val	Ser Val	Asn Ser	Ile Tyr					
			595			600		605						
		Gln Gln	Lys Pro	Phe Gly	Val Glu	Met Asp	Gly Glu	Arg Arg	Val Ile					
		610			615			620						
		Asn Asn	Lys Thr	Ala Ile	Glu Ser	Gln Leu	Gly Val	Ala Val	Lys Ile					
55		625			630			635		640				

Lys Ser His Leu Thr Leu Gln Ala Thr Phe Asn Arg Gln Thr Gly Lys
 645 650 655
 His His Gln Ala Lys Gln Gly Ala Leu Asn Leu Gln Trp Thr Phe
 660 665 670

5

<210> 7
 <211> 2076
 <212> DNA
 <213> 脑膜炎奈瑟氏菌

10

<400> 7

	atgacactga	aagcaagcaa	gcaagcaagc	ggtcgggtta	atctattaac	attatctggt	60
	ttatcgctgt	tttgcacgcc	atatgtttgt	ggttcggatg	cgtagcatcc	cgtaaaagaa	120
	gccgagatta	aaaacaaatt	tattttagaa	gcggcggaag	acagaaatc	ccacgtttgg	180
15	cgcgccccgt	gcagcatatc	ttttgattgc	ttcggtatgt	tcagagctca	gcttggttca	240
	aatactcgtt	ctaccaaaat	cggcgacgat	gccgattttt	cattttcaga	caagccgaaa	300
	cccggcactt	cccattattt	ttccagcgg	aaaaccgatc	aaaattcacc	cgaatatggg	360
	tatgacgaaa	tcaatatcca	aggtaaaaac	tacaatagcg	gcatactcgc	cgtcgataat	420
	atgcccgttg	ttaagaaata	tattacagat	acttacgggg	ataatttaa	ggatgcggtt	480
20	aagaagcaat	tacaggattt	atacaaaaca	agaccgaag	cttggaaga	aaataaaaa	540
	cggactgagg	aggcgtatat	agaacagctt	ggaccaaaat	ttagtatact	caaacagaaa	600
	aaccccgatt	taattaataa	attggtagaa	gattccgtac	tcactcctca	tagtaataca	660
	tcacagacta	gtctcaacaa	catcttcaat	aaaaaattac	acgtcaaaat	cgaaaacaaa	720
	tcccacgtcg	ccggacaggt	ggtggaactg	accaagatga	cgctgaaaga	ttccctttgg	780
25	gaaccgcgcc	gccattccga	catccatacg	ctggaaactt	ccgataatgc	ccgcatecgc	840
	ctgaacacga	aagatgaaaa	actgaccgtc	cataaagcgt	atcagggcgg	cgcggtttc	900
	ctgttcggct	acgacgtgcg	ggagtcggac	gaaccgcgcc	tgacctttga	acaaaacgtc	960
	agcggaaaat	ccggcgtggt	tttgaacgc	cgcccgaaa	atctgaaaac	gctcgacggg	1020
	cgcaactga	ttgcggcgga	aaaggcagac	cctaattcgt	ttgcgtttaa	acaaaattac	1080
30	cgccagggac	tgtacgaatt	attgctcaag	caatgcgaag	gcggattttg	cttgggtgtg	1140
	cagcgtttgg	ctatccctga	ggcggaaagc	gttttatatg	cccaacaggc	ttatgcggca	1200
	aatactttgt	ttggcgtcgc	tgccgccgac	aggggcgacg	acgtgtatgc	cgccgatccg	1260
	tcccgtaaaa	aattgtggct	gcgcttcacc	ggcggccggt	cgcatcaaaa	tatacggggc	1320
	ggcgcggctg	cggacggggc	gcgcaaagcg	gtgcaaatcg	gcggtgaggt	gtttgtacgg	1380
35	caaaatgaag	gcagtcggct	ggcaatcggc	gtgatggggc	gcagggctgg	ccagcacgca	1440
	tcagtaaacg	gcaaaggcgg	tgcggcagcg	agttatttgc	atggttatgg	cggggtgttt	1500
	tatgctgcgt	ggcatcagtt	gcgcgataaa	caaacgggtg	cgtatttggg	cgctggttg	1560
	caataccaac	gtttcaaaaca	ccgcatcaat	gatgaaaacc	gtgcggaacg	ctacaaaacc	1620
	aaaggttggg	cggtctctgt	cgaaggcggc	tacaacgcgc	ttgtggcgga	aggcgttctc	1680
40	ggaaaaggca	ataatgtgcg	gtttacctg	caaccgcagc	cgagtttac	ctacttgggc	1740
	gtaaacggcg	gctttaccga	cagcgagggg	acggcgctcg	gactgctcgg	cagcggtcag	1800
	tggcaaaagc	gcgccggcat	tcgggcaaaa	accgctttg	ctttcgtaa	cggtgtcaat	1860
	cttcagcctt	ttgccgcttt	taatgtttg	cacaggtcaa	aatcttctgg	cgtggaaatg	1920
	gacggcgaag	aacagacgct	ggcaggcagc	acggcgctcg	aagggcggtt	cggcattgaa	1980
45	gccggttggg	aaggccatat	gtccgcacgc	atcggatagc	gcaaaaggac	ggacggcgac	2040
	aaagaagccg	cattgtcgtc	caaatggctg	ttttga			2076

50

<210> 8
 <211> 691
 <212> PRT
 <213> 脑膜炎奈瑟氏菌

<400> 8

55

Met Thr Leu Lys Ala Ser Lys Gln Ala Ser Gly Arg Val Asn Leu Leu
 1 5 10 15

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

专利名称(译)	新型化合物		
公开(公告)号	CN1375006A	公开(公告)日	2002-10-16
申请号	CN00805181.X	申请日	2000-01-19
[标]申请(专利权)人(译)	史密斯克莱恩比彻姆生物有限公司		
申请(专利权)人(译)	史密斯克莱恩比彻姆生物有限公司		
当前申请(专利权)人(译)	史克必成生物制品S.A.		
[标]发明人	JL吕勒		
发明人	J·L·吕勒		
IPC分类号	G01N33/53 A61K39/00 A61K39/095 A61K39/395 C07K14/22 C07K16/12 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/31 C12P21/02 C12R1/19 G01N33/569		
CPC分类号	C07K14/22		
代理人(译)	刘玥		
优先权	1999002086 1999-01-29 GB 1999001944 1999-01-28 GB 1999001368 1999-01-22 GB 1999003535 1999-02-16 GB 1999003417 1999-02-15 GB		
外部链接	Espacenet SIPO		

摘要(译)

本发明提供BASB047、BASB054、BASB068和BASB069多肽,编码BASB047、BASB054、BASB068和BASB069多肽的多核苷酸以及通过重组技术产生这样的多肽的方法。还提供诊断、预防和治疗用途。