



(12) 发明专利申请

(10) 申请公布号 CN 102770444 A

(43) 申请公布日 2012. 11. 07

(21) 申请号 201180011318. 4

(22) 申请日 2011. 02. 28

(30) 优先权数据

1003333. 0 2010. 02. 26 GB

(85) PCT申请进入国家阶段日

2012. 08. 27

(86) PCT申请的申请数据

PCT/IB2011/000562 2011. 02. 28

(87) PCT申请的公布数据

W02011/104632 EN 2011. 09. 01

(71) 申请人 诺华有限公司

地址 瑞士巴塞尔

(72) 发明人 D·迈文 C·D·日努多

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 韦东

(51) Int. Cl.

C07K 14/315 (2006. 01)

A61K 39/09 (2006. 01)

权利要求书 2 页 说明书 45 页

序列表 13 页 附图 8 页

(54) 发明名称

免疫原性蛋白和组合物

(57) 摘要

本发明提供用于治疗 and 阻止无乳链球菌 (*Streptococcus agalactiae*) (B 群链球菌 ;GBS) 的蛋白和组合物。

1. 一种多肽,所述多肽包括:

i) SEQ ID NO:1 的至少 7 个连续氨基酸片段,其中所述片段包括 SEQ IDNO:4 的氨基酸序列的表位;

ii) 具有与 SEQ ID NO:1 至少 90% 相同性的氨基酸序列中至少 7 个连续氨基酸片段,其中所述片段包括与 SEQ ID NO:4 氨基酸序列的表位具有至少 90% 相同性的表位;

iii) SEQ ID NO:5 的至少 7 个连续氨基酸片段,其中所述片段包括 SEQ IDNO:8 的氨基酸序列的表位;

iv) 具有与 SEQ ID NO:5 至少 90% 相同性的氨基酸序列中至少 7 个连续氨基酸片段,其中所述片段包括与 SEQ ID NO:8 氨基酸序列的表位具有至少 90% 相同性的表位。

2. 如权利要求 1 所述的多肽,其特征在于,所述多肽由以下片段组成:

i) SEQ ID NO:1 的片段,其中所述片段含有 SEQ ID NO:4 的氨基酸序列;

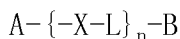
ii) 与 SEQ ID NO:1 具有 90% 相同性的氨基酸序列的片段,其中所述片段包括与 SEQ ID NO:4 具有至少 90% 相同性的氨基酸序列;

iii) SEQ ID NO:5 的片段,其中所述片段含有 SEQ ID NO:8 的氨基酸序列;

iv) 与 SEQ ID NO:5 具有 90% 相同性的氨基酸序列的片段,其中所述片段包括与 SEQ ID NO:8 具有至少 90% 相同性的氨基酸序列。

3. 如权利要求 1 或 2 所述的多肽,其特征在于,所述多肽包括或由以下片段组成:i) SEQ ID NO:1 的片段,所述片段包括 SEQ ID NO:4 的氨基酸序列;或 ii) SEQ ID NO:5 的片段,所述片段包括 SEQ ID NO:8 的氨基酸序列。

4. 一种多肽,所述多肽包含氨基酸序列:



其中:各 X 是权利要求 1-3 中任一项定义的多肽;L 是任选的接头氨基酸序列;A 是任选的 N 末端氨基酸序列;B 是任选的 C 末端氨基酸序列;n 是 1 或更大的整数。

5. 如权利要求 1-4 中任一项所述的多肽,其特征在于,所述多肽引起包含抗体的抗体应答,所述抗体结合具有 SEQ ID NO:1 氨基酸序列的野生型 GBS 蛋白和具有 SEQ ID NO:5 氨基酸序列的野生型 GBS 蛋白的氨基酸序列。

6. 如权利要求 1-5 中任一项所述的多肽,其特征在于,所述多肽引起包含抗体的抗体应答,该抗体与具有氨基酸序列 SEQ ID NO:9 的野生型 GBS 蛋白(株系 CJB111)、具有氨基酸序列 SEQ ID NO:13 的野生型 GBS 蛋白(株系 515)、具有氨基酸序列 SEQ ID NO:17 的野生型 GBS 蛋白(株系 NEM316)、具有氨基酸序列 SEQ ID NO:21 的野生型 GBS 蛋白(株系 DK21)、具有氨基酸序列 SEQ IDNO:25 的野生型 GBS 蛋白(株系 CJB110)结合。

7. 一种包括糖部分和载体蛋白部分的偶联物,所述载体蛋白部分包括如权利要求 1-6 中任一项所述的多肽。

8. 一种编码权利要求 1-6 中任一项所述的多肽的核酸。

9. 一种包括权利要求 1-6 中任一项所述的多肽、权利要求 7 所述的偶联物或权利要求 8 所述的核酸的免疫原性组合物。

10. 权利要求 1-6 中任一项所述的多肽、权利要求 7 所述的偶联物、权利要求 8 所述的核酸或权利要求 9 所述的免疫原性组合物在治疗中的应用。

11. 权利要求 1-6 中任一项所述的多肽、权利要求 7 所述的偶联物、权利要求 8 所述的

核酸或权利要求 9 所述的免疫原性组合物在治疗或预防 GBS 引起的疾病和 / 或感染,包括在治疗或预防脑膜炎中的应用。

12. 一种在哺乳动物中治疗或预防 GBS 引起的疾病和 / 或感染, 优选脑膜炎的方法, 所述方法包括给予有效量的权利要求 1-6 中任一项所述的多肽、权利要求 7 所述的偶联物、权利要求 8 所述的核酸或权利要求 9 所述的免疫原性组合物。

13. 一种表达权利要求 1-6 中任一项所述的多肽的细菌。

免疫原性蛋白和组合物

技术领域

[0001] 本发明提供用于治疗 and 阻止无乳链球菌 (*Streptococcus agalactiae*) (B 群链球菌 ;GBS) 的蛋白和组合物。

技术背景

[0002] 革兰氏阳性菌无乳链球菌(或“B 群链球菌”,简称“GBS”)会在免疫受损个体和新生儿中引起严重的疾病,菌血症和脑膜炎。有两种新生儿感染。第一种(早发,通常在出生 5 天内)表现为菌血症和肺炎。婴儿穿过产道发生垂直收缩。GBS 在约 25% 年轻女性的阴道中繁殖,通过有该菌繁殖的母亲阴道分娩婴儿中约 1% 会受到感染。致死率 50-70%。第二种是脑膜炎,出生后 10-60 天发生。若怀孕女性用 III 型荚膜接种,则婴儿会被动免疫,迟发性脑膜炎发病率会降低但不是完全消除。

[0003] “GBS”中的“B”表示兰氏分类法(Lancefield classification),其基于稀酸中可溶性糖(称为 C 糖)的抗原性。兰氏分类鉴定到 13 种 C 糖,称为 A-0,其可在血清学上区分。最常感染人类的生物体在 A、B、D 和 G 群中。B 群中株系基于其多糖荚膜的结构被分为 10 个血清型(Ia、Ib、II、III、IV、V、VI、VII、VIII 和 XI)

[0004] 已经进行研究开发基于蛋白和基于多糖的针对 GBS 的疫苗,但目前没有市售可得的 GBS 疫苗。因此需要针对无乳链球菌(*S. agalactiae*) 感染的有效疫苗。

[0005] 本发明目的是提供能用于开发所述疫苗的蛋白和免疫原性组合物。

发明内容

[0006] GBS 的菌毛结构被认为是感兴趣的疫苗候选物。GBS 具有三种菌毛变体,各由不同毒力岛(pathogenicity island)PI-1、PI-2a 和 PI-2b 编码 [1, 2]。各毒力岛由 5 种基因组成,它们编码:菌毛主干蛋白(BP);2 种辅助蛋白(AP1 和 AP2);和 2 种涉及菌毛组装的分选酶蛋白。

[0007] 所有 GBS 株系载有这 3 种毒力岛中至少一种且这些毒力岛编码的菌毛结构蛋白(BP、AP1 和 AP2) 的序列通常非常保守。毒力岛 2a(API-2a) 编码的辅助蛋白 1 (AP1) 序列(本文也称为 GBS67) 在 GBS 株系中不同。至少存在 2 个 GBS67 蛋白家族。

[0008] 原始 'GBS67' (SAG1408) 序列在参考文献 147 中注释为细胞壁表面锚定家族蛋白(参见 GI:22534437)。2603 菌株中发现的全长 GBS67 的氨基酸序列在本文中列为 SEQ ID NO:1。GBS 株系 CJB111、515 和 NEM316 表达 GBS67 序列,其与 2603 株系的 GBS67 序列属于同一家族。CJB111、515 和 NEM316 株系中发现的全长 GBS67 的氨基酸序列在本文中列为 SEQ ID NO:9、13 和 17。

[0009] H36B 株系中存在 GBS67(SAI1512) 变体。该 'GBS67' (SAG1408) 序列变体在参考文献 3 中注释为细胞壁表面锚定家族蛋白(参见 GI:77405751)。H36B 菌株中发现的全长 GBS67 的氨基酸序列在本文中列为 SEQ ID NO:5。GBS 株系 DK21 和 CJB110 表达 GBS67 序列,其与 H36B 株系的 GBS67 序列属于同一家族。DK21 和 CJB110 株系中发现的全长 GBS67

的氨基酸序列在本文中列为 SEQ ID NO:21 和 25。

[0010] 如本文所述,针对 2603 株系和相关株系中所发现全长 GBS67 的氨基酸序列产生的血清对表达 H36B 株系和相关株系中所发现全长 GBS67 的氨基酸序列的 GBS 株系有活性,反之亦然。因此全长 GBS67 提供针对表达两种家族任一种中 GBS67 变体的 GBS 株系的交叉保护。

[0011] 本发明目前从 GBS 的 2603 株系和 GBS 的 H36B 株系中成功鉴定到含有负责交叉保护表位的全长 GBS67 序列片段。

[0012] 2603 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文作为 SEQ ID NO:3 给出。SEQ ID NO:3 的氨基酸序列是位于 2603 株系中 GBS67 序列(SEQ ID NO:1)的氨基酸 218-615 的 398 个氨基酸片段。

[0013] 2603 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文作为 SEQ ID NO:4 给出。SEQ ID NO:4 的氨基酸序列是位于 2603 株系中 GBS67 序列(SEQ ID NO:1)的氨基酸 616-866 的 251 个氨基酸片段。

[0014] H36B 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:7 给出。SEQ ID NO:7 的氨基酸序列是位于 H36B 株系中 GBS67 序列(SEQ ID NO:5)的氨基酸 218-610 的 393 个氨基酸片段。

[0015] H36B 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:8 给出。SEQ ID NO:8 的氨基酸序列是位于 H36B 株系中 GBS67 序列(SEQ ID NO:5)的氨基酸 611-861 的 251 个氨基酸片段。

[0016] 还在表达与 GBS 株系 2603 的 GBS67 为同一家族的 GBS67 的 GBS 株系即 GBS 株系 CJB111、515 和 NEM316 中和表达与 GBS 株系 H36B 为同一家族的 GBS67 的 GBS 株系即 DK21 和 CJB110 中鉴定到相应片段。

[0017] CJB111 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:11 给出。SEQ ID NO:11 的氨基酸序列是位于 CJB111 株系中 GBS67 序列(SEQ ID NO:9)的氨基酸 218-615 的 398 个氨基酸片段。

[0018] CJB111 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文作为 SEQ ID NO:12 给出。SEQ ID NO:12 的氨基酸序列是位于 CJB111 株系中 GBS67 序列(SEQ ID NO:9)的氨基酸 616-866 的 251 个氨基酸片段。

[0019] 515 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:15 给出。SEQ ID NO:15 的氨基酸序列是位于 515 株系中 GBS67 序列(SEQ ID NO:13)的氨基酸 218-615 的 398 个氨基酸片段。

[0020] 515 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:16 给出。SEQ ID NO:16 的氨基酸序列是位于 515 株系中 GBS67 序列(SEQ ID NO:13)的氨基酸 616-866 的 251 个氨基酸片段。

[0021] NEM316 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:19 给出。SEQ ID NO:19 的氨基酸序列是位于 NEM316 株系中 GBS67 序列(SEQ ID NO:17)的氨基酸 218-615 的 398 个氨基酸片段。

[0022] NEM316 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:20 给出。SEQ ID NO:20 的氨基酸序列是位于 NEM316 株系中 GBS67 序列(SEQ ID

NO:17)的氨基酸 616-866 的 251 个氨基酸片段。

[0023] DK21 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:23 给出。SEQ ID NO:23 的氨基酸序列是位于 DK21 株系中 GBS67 序列 (SEQ ID NO:21) 的氨基酸 218-610 的 393 个氨基酸片段。

[0024] DK21 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:24 给出。SEQ ID NO:24 的氨基酸序列是位于 DK21 株系中 GBS67 序列 (SEQ ID NO:21) 的氨基酸 611-861 的 251 个氨基酸片段。

[0025] CJB110 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:27 给出。SEQ ID NO:27 的氨基酸序列是位于 CJB110 株系中 GBS67 序列 (SEQ ID NO:25) 的氨基酸 218-610 的 393 个氨基酸片段。

[0026] CJB110 株系中发现的含有负责交叉保护表位的 GBS67 序列片段在本文中作为 SEQ ID NO:28 给出。SEQ ID NO:28 的氨基酸序列是位于 CJB110 株系中 GBS67 序列 (SEQ ID NO:25) 的氨基酸 611-861 的 251 个氨基酸片段。

[0027] GBS67 多肽

[0028] GBS67 的片段和这些片段的表位可用于替代免疫原性组合物中的全长 GBS67 以治疗或阻止 GBS。

[0029] GBS672603

[0030] 因此,根据本发明的一个方面,提供包括或由下述组成的多肽:

[0031] i) SEQ ID NO:1 的至少 t 个连续氨基酸片段,其中所述片段包括 SEQ IDNO:3 和 / 或 SEQ ID NO:4 的氨基酸序列的表位;

[0032] ii) 具有与 SEQ ID NO:1 至少 a% 相同性的氨基酸序列中至少 t 个连续氨基酸片段,其中所述片段包括与 SEQ ID NO:3 和 / 或 SEQ ID NO:4 氨基酸序列的表位具有至少 b% 相同性的表位;

[0033] 本发明该方面的多肽可包括或由下述组成:

[0034] i) SEQ ID NO:1 的至少 t 个连续氨基酸片段,其中所述片段包括 SEQ IDNO:3 和 / 或 SEQ ID NO:4 的氨基酸序列;

[0035] ii) 与 SEQ ID NO:1 有 a% 相同性的氨基酸序列中至少 t 个连续氨基酸片段,其中所述片段包括与 SEQ ID NO:3 和 / 或 SEQ ID NO:4 有至少 b% 相同性的氨基酸序列。

[0036] 本发明该方面的多肽可包括或由以下组成:SEQ ID NO:1 的至少 t 个连续氨基酸片段,所述片段包括 SEQ ID NO:3 和 / 或 SEQ ID NO:4 的氨基酸序列。

[0037] GBS67 H36B

[0038] 根据本发明的另一方面,提供包括或由下述组成的多肽:

[0039] i) SEQ ID NO:5 的至少 u 个连续氨基酸片段,其中所述片段包括 SEQ IDNO:7 和 / 或 SEQ ID NO:8 的氨基酸序列的表位;

[0040] ii) 具有与 SEQ ID NO:5 至少 c% 相同性的氨基酸序列中至少 u 个连续氨基酸片段,其中所述片段包括与 SEQ ID NO:7 和 / 或 SEQ ID NO:8 氨基酸序列的表位具有至少 d% 相同性的表位;

[0041] 本发明该方面的多肽可包括或由下述组成:

[0042] i) SEQ ID NO:5 的至少 u 个连续氨基酸片段,其中所述片段包括 SEQ IDNO:7 和 /

或 SEQ ID NO:8 的氨基酸序列；

[0043] ii) 与 SEQ ID NO:5 有 c% 相同性的氨基酸序列中至少 u 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:7 和 / 或 SEQ ID NO:8 具有至少 d% 相同性的氨基酸序列。

[0044] 本发明该方面的多肽可包括或由以下组成：SEQ ID NO:5 的至少 u 个连续氨基酸片段，所述片段包括 SEQ ID NO:7 和 / 或 SEQ ID NO:8 的氨基酸序列。

[0045] GBS67 CJB111

[0046] 根据本发明的另一方面，提供包括或由下述组成的多肽：

[0047] i) SEQ ID NO:9 的至少 v 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:11 和 / 或 SEQ ID NO:12 的氨基酸序列的表位；

[0048] ii) 具有与 SEQ ID NO:9 至少 e% 相同性的氨基酸序列中至少 v 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:11 和 / 或 SEQ ID NO:12 氨基酸序列的表位具有至少 f% 相同性的表位；

[0049] 本发明该方面的多肽可包括或由下述组成：

[0050] i) SEQ ID NO:9 的至少 v 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:11 和 / 或 SEQ ID NO:12 的氨基酸序列；

[0051] ii) 与 SEQ ID NO:9 有 e% 相同性的氨基酸序列中至少 v 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:11 和 / 或 SEQ ID NO:12 具有至少 f% 相同性的氨基酸序列。

[0052] 本发明该方面的多肽可包括或由以下组成：SEQ ID NO:9 的至少 v 个连续氨基酸片段，所述片段包括 SEQ ID NO:11 和 / 或 SEQ ID NO:12 的氨基酸序列。

[0053] GBS67 515

[0054] 根据本发明的另一方面，提供包括或由下述组成的多肽：

[0055] i) SEQ ID NO:13 的至少 w 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:15 和 / 或 SEQ ID NO:16 的氨基酸序列的表位；

[0056] ii) 具有与 SEQ ID NO:13 至少 g% 相同性的氨基酸序列中至少 w 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:15 和 / 或 SEQ ID NO:16 氨基酸序列的表位具有至少 h% 相同性的表位；

[0057] 本发明该方面的多肽可包括或由下述组成：

[0058] i) SEQ ID NO:13 的至少 w 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:15 和 / 或 SEQ ID NO:16 的氨基酸序列；

[0059] ii) 与 SEQ ID NO:13 有 g% 相同性的氨基酸序列中至少 w 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:15 和 / 或 SEQ ID NO:16 具有至少 h% 相同性的氨基酸序列。

[0060] 本发明该方面的多肽可包括或由以下组成：SEQ ID NO:13 的至少 w 个连续氨基酸片段，所述片段包括 SEQ ID NO:15 和 / 或 SEQ ID NO:16 的氨基酸序列。

[0061] GBS67 NEM316

[0062] 根据本发明的另一方面，提供包括或由下述组成的多肽：

[0063] i) SEQ ID NO:17 的至少 x 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:19 和 / 或 SEQ ID NO:20 的氨基酸序列的表位；

[0064] ii) 具有与 SEQ ID NO:17 至少 i% 相同性的氨基酸序列中至少 x 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:19 和 / 或 SEQ ID NO:20 氨基酸序列的表位具有至少

j% 相同性的表位；

[0065] 本发明该方面的多肽可包括或由下述组成：

[0066] i) SEQ ID NO:17 的至少 x 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:19 和 / 或 SEQ ID NO:20 的氨基酸序列；

[0067] ii) 与 SEQ ID NO:17 有 i% 相同性的氨基酸序列中至少 x 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:19 和 / 或 SEQ ID NO:20 具有至少 j% 相同性的氨基酸序列。

[0068] 本发明该方面的多肽可包括或由以下组成：SEQ ID NO:17 的至少 x 个连续氨基酸片段，所述片段包括 SEQ ID NO:19 和 / 或 SEQ ID NO:20 的氨基酸序列。

[0069] GBS67 DK21

[0070] 根据本发明的另一方面，提供包括或由下述组成的多肽：

[0071] i) SEQ ID NO:21 的至少 y 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:23 和 / 或 SEQ ID NO:24 的氨基酸序列的表位；

[0072] ii) 具有与 SEQ ID NO:21 至少 k% 相同性的氨基酸序列中至少 y 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:23 和 / 或 SEQ ID NO:24 氨基酸序列的表位具有至少 l% 相同性的表位；

[0073] 本发明该方面的多肽可包括或由下述组成：

[0074] i) SEQ ID NO:21 的至少 y 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:23 和 / 或 SEQ ID NO:24 的氨基酸序列；

[0075] ii) 与 SEQ ID NO:21 有 k% 相同性的氨基酸序列中至少 y 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:23 和 / 或 SEQ ID NO:24 具有至少 l% 相同性的氨基酸序列。

[0076] 本发明该方面的多肽可包括或由以下组成：SEQ ID NO:21 的至少 y 个连续氨基酸片段，所述片段包括 SEQ ID NO:23 和 / 或 SEQ ID NO:24 的氨基酸序列。

[0077] GBS67 CJB110

[0078] 根据本发明的另一方面，提供包括或由下述组成的多肽：

[0079] i) SEQ ID NO:25 的至少 z 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:27 和 / 或 SEQ ID NO:28 的氨基酸序列的表位；

[0080] ii) 具有与 SEQ ID NO:25 至少 m% 相同性的氨基酸序列中至少 z 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:27 和 / 或 SEQ ID NO:28 氨基酸序列的表位具有至少 n% 相同性的表位；

[0081] 本发明该方面的多肽可包括或由下述组成：

[0082] i) SEQ ID NO:25 的至少 z 个连续氨基酸片段，其中所述片段包括 SEQ IDNO:27 和 / 或 SEQ ID NO:28 的氨基酸序列；

[0083] ii) 与 SEQ ID NO:25 有 m% 相同性的氨基酸序列中至少 z 个连续氨基酸片段，其中所述片段包括与 SEQ ID NO:27 和 / 或 SEQ ID NO:28 具有至少 n% 相同性的氨基酸序列。

[0084] 本发明该方面的多肽可包括或由以下组成：SEQ ID NO:25 的至少 z 个连续氨基酸片段，所述片段包括 SEQ ID NO:27 和 / 或 SEQ ID NO:28 的氨基酸序列。

[0085] “表位”表示被免疫系统识别并引发免疫应答的多肽部分。本发明的多肽能诱导针对表达 GBS67 肽变体的 GBS 株系的交叉保护。因此，本发明的多肽在给予对象时，能引起包含抗体的抗体应答，该抗体与具有氨基酸序列 SEQID NO:1 的野生型 GBS 蛋白（株系

2603) 和具有氨基酸序列 SEQ ID NO:5 的野生型 GBS 蛋白 (株系 H36B) 结合。因此本发明的多肽能与 SEQ ID NO:1 和 SEQ ID NO:5 竞争结合针对 SEQ ID NO:1 或 SEQ ID NO:5 所产生的抗体。

[0086] 本发明的多肽在给予对象时,通常还能引起包含抗体的抗体应答,该抗体与具有氨基酸序列 SEQ ID NO:9 的野生型 GBS 蛋白 (株系 CJB111)、具有氨基酸序列 SEQ ID NO:13 的野生型 GBS 蛋白 (株系 515)、具有氨基酸序列 SEQ ID NO:17 的野生型 GBS 蛋白 (株系 NEM316)、具有氨基酸序列 SEQ ID NO:21 的野生型 GBS 蛋白 (株系 DK21)、具有氨基酸序列 SEQ ID NO:25 的野生型 GBS 蛋白 (株系 CJB110) 结合。因此本发明的多肽还能与这些具有 SEQ ID NO:9、13、17、21 或 25 的野生型 GBS 蛋白竞争结合针对这些蛋白所产生的抗体。

[0087] 针对本发明多肽的抗体可用标准免疫方法容易地生成且这些抗体结合 SEQ ID NO:1、5、9、13、17、21 或 25 的野生型 GBS 蛋白的能力可用标准实验如 ELISA 实验评估。

[0088] 相似地,多肽竞争针对野生型 GBS 蛋白所产生抗体的能力可用本领域已知的竞争试验技术容易地测定,所述技术包括平衡方法如 ELISA、动力学方法如 **BIACORE®** 和流式细胞分析方法。与 SEQ ID NO:1、5、9、13、17、21 或 25 的野生型 GBS 蛋白竞争结合针对这些野生型 GBS 蛋白之一的抗体的多肽相较该多肽不存在时会使观察到的野生型 GBS 蛋白和抗体的总结合下降。通常,与具有 SEQ ID NO:1、5、9、13、17、21 或 25 的 GBS 蛋白所观察到的抗体结合相比,存在本发明多肽时,结合降低 10% 或更多、20% 或更多、30% 或更多、40% 或更多、60% 或更多,例如结合降低 70% 或更多。

[0089] 本发明多肽诱导针对表达 GBS67 蛋白变体的 GBS 株系的交叉保护还可在动物模型中确认,例如实施例中所描述的母体免疫模型,其中雌性小鼠用多肽免疫,其幼仔用表达 GBS67 蛋白变体的 GBS 株系攻击。

[0090] a 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。b 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。c 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。d 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。e 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。f 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。g 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。h 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。i 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。j 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。k 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。l 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。m 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。n 值是至少 75,例如 80、85、90、92、94、95、96、97、98、99 或更大。通常,a、b、c、d、e、f、g、h、i、j、k、l、m 和 n 为至少 90,例如至少 95。

[0091] t 值至少为 7,例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:1 所示的 2603 菌株的全长 GBS67 序列是 901 个氨基酸长度。因此 t 值还小于 901,例如小于 850、800、750、700、650、600、550、500、450。t 值可为 50-600、100-400、150-300、225-275,例如 120-150。

[0092] u 值至少为 7,例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、

26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:5 所示的 H36B 菌株的全长 GBS67 序列是 896 个氨基酸长度。因此 u 值还小于 896, 例如小于 860、850、800、750、700、650、600、550、500、450。u 值可为 50-600、100-400、150-300、225-275, 例如 120-150。

[0093] v 值至少为 7, 例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:9 所示的 CJB111 菌株的全长 GBS67 序列是 901 个氨基酸长度。因此 v 值还小于 901, 例如小于 860、850、800、750、700、650、600、550、500、450。v 值可为 50-600、100-400、150-300、225-275, 例如 120-150。

[0094] w 值至少为 7, 例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:13 所示的 515 菌株的全长 GBS67 序列是 901 个氨基酸长度。因此 w 值还小于 901, 例如小于 860、850、800、750、700、650、600、550、500、450。w 值可为 50-600、100-400、150-300、225-275, 例如 120-150。

[0095] x 值至少为 7, 例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:17 所示的 NEM316 菌株的全长 GBS67 序列是 901 个氨基酸长度。因此 x 值还小于 901, 例如小于 860、850、800、750、700、650、600、550、500、450。w 值可为 50-600、100-400、150-300、225-275, 例如 120-150。

[0096] y 值至少为 7, 例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:21 所示的 DK21 菌株的全长 GBS67 序列是 896 个氨基酸长度。因此 y 值还小于 896, 例如小于 860、850、800、750、700、650、600、550、500、450。y 值可为 50-600、100-400、150-300、225-275, 例如 120-150。

[0097] z 值至少为 7, 例如 8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400。如 SEQ ID NO:25 所示的 CJB110 菌株的全长 GBS67 序列是 896 个氨基酸长度。因此 z 值还小于 896, 例如小于 860、850、800、750、700、650、600、550、500、450。z 值可为 50-600、100-400、150-300、225-275, 例如 120-150。

[0098] 与 SEQ ID NO:1、5、9、13、17、21 和 25 相比, 本发明的多肽可包括一个或多个 (例如 1、2、5、4、5、6、7、8、9、10 等) 保守氨基酸置换, 即用另一个具有相关侧链的氨基酸置换一个氨基酸。这些保守氨基酸置换可位于 SEQ ID NO:1、5、9、13、17、21 和 25 的区域内, 分别对应于 SEQ ID NO:3 和 4、7 和 8、11 和 12、15 和 16、19 和 20、23 和 24、或 27 和 28。遗传编

码的氨基酸通常分为四类：(1) 酸性，即天冬氨酸、谷氨酸；(2) 碱性，即赖氨酸、精氨酸、组氨酸；(3) 非极性，即丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸；和(4) 极性不带电，即甘氨酸、天冬酰胺、谷氨酰胺、胱氨酸、丝氨酸、苏氨酸、酪氨酸。有时将苯丙氨酸、色氨酸和酪氨酸一起分类为芳族氨基酸。通常，这些家族中单个氨基酸的替换不会对生物活性产生重要影响。

[0099] 相对于 SEQ ID NO:1、5、9、13、17、21 和 25 的片段，本发明多肽可含有一个或多个（如 1、2、3、4、5、6、7、8、9、10 等）单一氨基酸缺失。相对于 SEQ ID NO:1、5、9、13、17、21 和 25 的片段，该多肽也可包含一个或多个（如 1、2、3、4、5、6、7、8、9 个等）插入（如各 1、2、3、4 或 5 个氨基酸）。这些缺失和插入可位于 SEQ ID NO:1、5、9、13、17、21 和 25 的区域内，分别对应于 SEQ ID NO:3 和 4、7 和 8、11 和 12、15 和 16、19 和 20、23 和 24、或 27 和 28。

[0100] 本发明的多肽可包括以下氨基酸序列：

[0101] (a) 与 SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:9、SEQ ID NO:13、SEQ ID NO:17、SEQ ID NO:21 或 SEQ ID NO:25 的片段相同（即 100% 相同性）；

[0102] (b) 与 SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:9、SEQ ID NO:13、SEQ ID NO:17、SEQ ID NO:21 或 SEQ ID NO:25 的片段共有序列相同性；

[0103] (c) 与 (a) 或 (b) 的序列相比，含有 1、2、3、4、5、6、7、8、9 或 10 个（或更多个）单氨基酸改变（缺失、插入、取代）的序列，这些改变可以位于不同位置或连续出现；和

[0104] (d) 用成对比对算法比对 SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:9、SEQ ID NO:13、SEQ ID NO:17、SEQ ID NO:21 或 SEQ ID NO:25 的片段时，从 N 末端到 C 末端的各 x 个氨基酸的移动窗口（从而对于延伸到 p 个氨基酸的比对来说， $p > x$ 时，存在 $p-x+1$ 个该窗口）具有至少 $x \cdot y$ 个相同的比对氨基酸，其中： x 选自 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200； y 选自 0.50, 0.60, 0.70, 0.75, 0.80, 0.85, 0.90, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99；且若 $x \cdot y$ 不是整数则将其四舍五入为最接近整数。优选的逐对比对算法是 Needleman-Wunsch 全局比对算法 [4]，使用默认参数（如缺口开放罚分 =10.0，缺口延伸罚分 =0.5，使用 EBLSUM62 积分矩阵）。用 EMBOSS 软件包中的 needle 工具能方便地实施这种算法 [5]。

[0105] 本发明的多肽可以杂交多肽的形式提供。所述杂交多肽可包括其他 GBS 或非 GBS 多肽序列。

[0106] 本发明还提供包含编码本发明多肽或杂交多肽的核苷酸序列的核酸。

[0107] 本发明还提供含有本发明的多肽、杂交多肽或核酸的免疫原性组合物。此类免疫原性组合物可用在治疗或阻止有关 GBS 的疾病或病症的方法中。

[0108] 本发明还提供表达本发明多肽或杂交多肽的细胞（通常为细菌）。

[0109] 杂交多肽

[0110] 本发明的多肽可与其他多肽作为单一多肽链组合表达（‘杂交’多肽或‘嵌合体’）。杂交多肽提供以下两个主要优点：首先，本身不稳定或者表达较差的多肽可以通过加入能够克服该问题的合适杂交伴侣得到改善；其次，商业生产得以简化，因为只需利用一次表达和纯化以生产可作抗原应用的两种多肽。

[0111] 杂交多肽可包括来自其他 GBS 抗原和 / 或其他非 GBS 抗原的序列。通常，所述杂交多肽包括来自其他 GBS 序列的序列，如其他菌毛亚基。这些其他 GBS 序列可为 GBS67 多

肽的 N 末端或 C 末端。不同的杂交多肽可以在单一制剂中混合在一起。

[0112] 杂交多肽可以式 $\text{NH}_2\text{-A-}\{-\text{X-L}\}_n\text{-B-COOH}$ 表示。

[0113] X 是本发明上述的 GBS67 多肽。如果野生型形式中 -X- 部分具有前导肽序列,那么在杂交蛋白中可以包含或者略去该序列。在一些实施方式中,前导肽可缺失,除了位于杂合蛋白 N- 末端的 -X- 部分,即保留 X_1 的前导肽,但略去 $X_2\cdots X_n$ 的前导肽。这相当于删除所有前导肽并使用 X_1 的前导肽作为 -A- 部分。

[0114] 在 $\{-\text{X-L}\}$ 的各个 n 值的情况下,接头氨基酸序列 -L- 可存在或不存在。例如,当 n=2 时,杂交体可以是 $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-L}_2\text{-COOH}$ 、 $\text{NH}_2\text{-X}_1\text{-X}_2\text{-COOH}$ 、 $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-COOH}$ 、 $\text{NH}_2\text{-X}_1\text{-X}_2\text{-L}_2\text{-COOH}$ 等。接头氨基酸序列 -L- 一般较短(如 20 个或更少的氨基酸,即 20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、1)。例子包括有利于克隆的短肽序列,聚-甘氨酸接头(即包含 Gly_n ,其中 n=2、3、4、5、6、7、8、9、10 或更高),组氨酸标签(即 His_n ,其中 n=3、4、5、6、7、8、9、10 或更高)。本领域技术人员显然了解其它合适的接头氨基酸序列。有用的接头是 GSGS(SEQ ID NO:29)、GSGGGG(SEQ ID NO:30) 或 GSGSGGGG(SEQ ID NO:31), Gly-Ser 二肽由 BamHI 限制位点形成,因而有助于克隆和操作, $(\text{Gly})_4$ 四肽是常用的多聚甘氨酸接头。其他合适接头,尤其用作最后一个 L_n 的接头是 Leu-Glu 二肽或 Gly-Ser。接头通常含至少一种甘氨酸残基以促进结构柔性,如 -L- 部分可含 1、2、3、4、5、6、7、8、9、10 或更多甘氨酸残基。所述甘氨酸可排列为在 Gly-Gly 二肽或更长的寡 Gly 序列(即 Gly_n ,其中 n=2、3、4、5、6、7、8、9、10 或更多)中包括至少 2 个连续甘氨酸。

[0115] -A- 是任选的 N 末端氨基酸序列。其一般较短(如 40 个或更少的氨基酸,即 40、39、38、37、36、35、34、33、32、31、30、29、28、27、26、25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、1 个)。例子包括指导蛋白运输的前导序列,或有利于克隆或纯化的短肽序列(如组氨酸标签,即 His_n ,其中 n=3、4、5、6、7、8、9、10 或更高)。其它合适的 N 末端氨基酸序列对本领域技术人员显而易见。如果 X_1 缺少其自身的 N- 末端甲硫氨酸,-A- 优选是提供 N- 末端甲硫氨酸的寡肽(例如,具有 1、2、3、4、5、6、7 或 8 个氨基酸),如 Met-Ala-Ser 或单个 Met 残基。在新生多肽中,-A- 部分可提供多肽的 N 末端甲硫氨酸(细菌中为甲酰甲硫氨酸,fMet)。可从新生 -A- 部分的 N 末端切割一种或多种氨基酸,然而本发明的成熟多肽中所述 -A- 部分不必包括 N 末端甲硫氨酸。

[0116] -B- 是任选的 C 末端氨基酸序列。其一般较短(如 40 个或更少的氨基酸,即 39、38、37、36、35、34、33、32、31、30、29、28、27、26、25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、1 个)。例子包括指导蛋白质运输的序列,有利于克隆或纯化的短肽序列(如包含组氨酸标签,即 His_n ,其中 n=3、4、5、6、7、8、9、10 或更高),或能提高蛋白质稳定性的序列。其他合适的 C 末端氨基酸序列对本领域技术人员显而易见,如谷胱甘肽 S 转移酶、硫氧还蛋白、金黄色葡萄球菌(*S. aureus*)蛋白 A 的 14kDa 片段、生物素化肽、麦芽糖结合蛋白、肠激酶标签等。

[0117] -A-、-B- 和 -L- 序列优选不包括与人多肽序列共有 10 或更多连续氨基酸的序列。

[0118] 在一些实施方式中,-L- 部分包含非 GBS67 抗原。在一些实施方式中,-A- 部分包含非 GBS67 抗原,且在一些实施方式中,-B- 部分包含非 GBS67 抗原。

[0119] 多肽

[0120] 可以多种方式制备本发明所用多肽,例如化学合成(全部或部分),用蛋白酶消化

较长多肽,由 RNA 翻译,由细胞培养物纯化(如通过重组表达),由生物体本身制备(如细菌培养后,或直接从患者制备)等。产生长度 <40 个氨基酸的肽的优选方法包括体外化学合成 [6, 7]。尤其优选固相肽合成,例如基于 tBoc 或 Fmoc [8] 化学的方法。也可部分或完全利用酶促合成 [9]。作为化学合成的替代方式,可利用生物合成,例如可通过翻译产生多肽。这一过程可以在体外或体内进行。生物学方法通常仅限于产生基于 L-氨基酸的多肽,但可通过操作翻译机制(如氨基酰基 tRNA 分子的翻译机制)引入 D-氨基酸(或其它非天然氨基酸,如碘化酪氨酸或甲基苯丙氨酸、叠氮基高丙氨酸等) [10]。然而,包含 D-氨基酸时,优选使用化学合成。多肽的 C 末端和 / 或 N 末端上可能有共价修饰。

[0121] 多肽可采取各种形式(如天然多肽、融合多肽、糖基化多肽、非糖基化多肽、脂化多肽、非脂化多肽、磷酸化多肽、非磷酸化多肽、肉豆蔻酰化多肽、非肉豆蔻酰化多肽、单体、多聚体、颗粒、变性多肽等)。

[0122] 多肽优选以纯化或基本纯化的形式提供,即基本不含其它多肽(如不含天然产生的多肽)、特别是不含其它肺炎球菌或宿主细胞多肽,多肽的纯度通常为至少约 50 重量%,通常至少约 90%,即组合物中少于约 50%,更优选少于约 10%(如 5% 或以下)由其它表达多肽构成。

[0123] 多肽可与固体支持物结合。多肽可包含可检测标记(如放射性或荧光标记,或生物素标记)。

[0124] 术语“多肽”指任何长度的氨基酸聚合物。所述聚合物可以是线性或支链聚合物,可以包含经修饰的氨基酸,可以为非氨基酸中断。该术语也包括天然修饰或通过人工介入修饰的氨基酸聚合物;例如,二硫键形成、糖基化、脂化、乙酰化、磷酸化或任何其它操作或修饰,如与标记组分偶联。该定义也包括,例如,含有一个或多个氨基酸类似物(包括例如,非天然氨基酸等)以及本领域已知其它修饰的多肽。多肽可以以单链或结合链的形式产生。多肽可以是天然或非天然糖基化的(即该多肽的糖基化模式不同于相应天然产生多肽的糖基化模式)。

[0125] 本发明提供产生本发明多肽的方法,该方法包括在诱导多肽表达的条件下培养本发明宿主细胞。虽然可以在链球菌 (*Streptococcus*) 中表达多肽,但本发明通常使用异源宿主进行表达。异源宿主可以是原核(例如细菌)或真核生物。所述宿主通常可以是大肠杆菌 (*E. coli*),其他合适的宿主包括枯草杆菌 (*Bacillus subtilis*)、霍乱弧菌 (*Vibrio cholerae*)、伤寒沙门菌 (*Salmonella typhi*)、鼠伤寒沙门菌 (*Salmonella typhimurium*)、乳酰胺奈瑟菌 (*Neisseria lactamica*)、灰色奈瑟菌 (*Neisseria cinerea*)、分枝杆菌 (*Mycobacteria*) (例如结核分支杆菌 (*M. tuberculosis*))、酵母等。

[0126] 本发明还提供产生本发明多肽的方法,其中所述多肽部分或完全用化学方式合成。

[0127] 本发明还提供含有两种或多种本发明多肽的组合物。

[0128] 核酸

[0129] 本发明还提供包含编码本发明多肽或杂交多肽的核苷酸序列的核酸。

[0130] 例如,本发明提供含编码多肽的核苷酸的核酸,所述多肽包含选自下组的氨基酸序列或由其组成:SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:23、

SEQ ID NO:24、SEQ ID NO:27 或 SEQID NO:28。

[0131] 本发明也提供所含核苷酸序列与这些核苷酸序列有序列相同性的核酸。这类核酸包括使用替代密码子编码相同氨基酸的核酸。具体地,核酸可包括为了在特定微生物如大肠杆菌(*E. coli*)中表达而优化的替代密码子。

[0132] 本发明也提供可与这些核酸杂交的核酸。可以在不同“严谨性”的条件下进行杂交反应。增加杂交反应严谨性的条件在本领域众所周知且已发表。相关条件的例子包括(按严谨性提高的顺序):孵育温度 25° C、37° C、50° C、55° C 和 68° C;缓冲液浓度 10x SSC、6x SSC、1x SSC、0.1x SSC(其中 SSC 是 0.15M NaCl 和 15mM 柠檬酸盐缓冲液)和使用其它缓冲液体系的等同条件;甲酰胺浓度 0%、25%、50% 和 75%;孵育时间 5 分钟至 24 小时;1 个、2 个或更多洗涤步骤;洗涤孵育时间 1、2 或 15 分钟;以及洗涤溶液 6x SSC、1x SSC、0.1x SSC 或去离子水。杂交技术和其优化方法为本领域熟知(例如,参见参考文献 11 和 222 等)。

[0133] 本发明包括含有这些序列的互补序列的核酸(例如,用于反义或检测,或用作引物)。

[0134] 本发明所述核酸可采取各种形式(如单链、双链、载体、引物、探针、标记等)。本发明核酸可以是环状或分枝状核酸,但通常是线性核酸。除非另有说明或要求,利用核酸的本发明任何实施方式可采用双链形式和构成该双链形式的两条互补单链形式中的每条链。引物、探针以及反义核酸通常是单链。

[0135] 本发明核酸优选以纯化或基本纯化的形式提供,即基本不含其它核酸(如不含天然产生的核酸),特别是不含其它 GBS 或宿主细胞核酸,通常其纯度至少为约 50 重量%,通常至少为约 90%。本发明核酸优选为 GBS 核酸。

[0136] 可以多种方式制备本发明的核酸,例如,通过完全或部分化学合成(例如 DNA 的亚磷酰胺合成)、通过用核酸酶(例如限制性酶)消化较长核酸、通过连接较短核酸或核苷酸(例如使用连接酶或聚合酶)、由基因组或 cDNA 文库制备等。

[0137] 本发明核酸可连接于固体支持物(如珠、平板、滤器、膜、玻片、微阵列支持物、树脂等)。可用例如放射性或荧光标记、或生物素标记来标记本发明核酸。这在将核酸用于检测技术时,例如核酸是引物或探针时特别有用。

[0138] 术语“核酸”通常包括任何长度的核苷酸聚合形式,包含脱氧核糖核苷酸、核糖核苷酸和/或其类似物。它包括 DNA、RNA、DNA/RNA 杂交体。它也包括 DNA 或 RNA 类似物,如含有经修饰主链(如肽核酸(PNA)或硫代磷酸酯)或经修饰碱基的类似物。因此,本发明包括 mRNA、tRNA、rRNA、核酶、DNA、cDNA、重组核酸、分枝核酸、质粒、载体、探针、引物等。本发明核酸采取 RNA 形式时,它可能具有或不具有 5' 帽。

[0139] 本发明核酸可以是载体的一部分,即设计用于转导/转染一种或多种细胞类型的核酸构建物的一部分。载体可以是,例如,设计用于分离、增殖和复制所插入核苷酸的“克隆载体”,设计用于在宿主细胞中表达核苷酸序列的“表达载体”,设计用于产生重组病毒或病毒样颗粒的“病毒载体”,或具有一种以上载体类型的属性的“穿梭载体”。优选的载体是质粒。“宿主细胞”包括单个细胞或细胞培养物,它可能是或已经是外源核酸的受体。宿主细胞包括单个宿主细胞的后代,由于天然、偶然或有意的突变和/或改变,这些后代与原始亲本细胞不必完全相同(形态或总 DNA 互补方面)。宿主细胞包括用本发明核酸体内或体外转染或感染的细胞。

[0140] 应理解核酸是 DNA 时, RNA 序列中的“U”被 DNA 中的“T”所替代。相似地, 应理解核酸是 RNA 时, DNA 中的“T”被 RNA 序列中的“U”所替代。

[0141] 涉及核酸的术语“互补”或“互补的”指沃森克里克碱基配对。因此, C 的互补物是 G, G 的互补物是 C, A 的互补物是 T(或 U), T(或 U) 的互补物是 A。也能使用诸如 I(嘌呤肌昔)等碱基, 例如与嘧啶(C 或 T) 互补。

[0142] 本发明的核酸可用于例如: 体外或体内产生多肽; 作为检测生物样品中核酸的杂交探针; 产生核酸的额外拷贝; 产生核酶或反义寡核苷酸; 作为单链 DNA 引物或探针; 或作为形成三链的寡核苷酸。

[0143] 本发明提供产生本发明核酸的方法, 其中所述核酸部分或完全用化学方式合成。

[0144] 本发明提供包含本发明核苷酸序列的载体(如克隆或表达载体)和用这种载体转化的宿主细胞。

[0145] 免疫原性组合物

[0146] 本发明的多肽和杂交多肽可用作免疫原性组合物的活化组分。该免疫原性组合物可用作疫苗。这些疫苗可以是预防性(即预防感染)或治疗性(即治疗感染)疫苗, 但一般是预防性疫苗。

[0147] 因此, 组合物可以是药学上可接受的。它们通常还包含抗原以外的组分, 例如它们一般包含一种或多种药物载体和/或赋形剂。对这类组分的充分讨论参见参考文献[217]。

[0148] 组合物通常以水性形式给予哺乳动物。然而在给药前, 该组合物可以是非水性形式。例如, 虽然一些疫苗制备成水性形式然后以水性形式填装、分销和给药, 但其他疫苗在制备过程中冻干, 并在使用时重建成水性形式。因此, 本发明组合物可被干燥, 例如冻干制剂。

[0149] 该组合物可含有防腐剂, 如硫柳汞或 2-苯氧乙醇。然而, 疫苗优选应基本不含(即小于 $5 \mu\text{g/ml}$) 含汞物质, 如不含硫柳汞。更优选无汞的疫苗。特别优选不含防腐剂的疫苗。

[0150] 为了控制张度, 优选包含生理盐如钠盐。优选氯化钠(NaCl), 它的浓度可以是 1-20mg/ml, 例如约 $10 \pm 2\text{mg/ml}$ NaCl。可以存在的其它盐包括氯化钾、磷酸二氢钾、无水磷酸氢二钠、氯化镁、氯化钙等。

[0151] 组合物的渗透压通常为 200mOsm/kg-400mOsm/kg, 优选为 240-360mOsm/kg, 更优选为 290-310mOsm/kg。

[0152] 组合物可含有一种或多种缓冲剂。常用缓冲剂包括: 磷酸盐缓冲剂; Tris 缓冲剂; 硼酸盐缓冲剂; 琥珀酸盐缓冲剂; 组氨酸缓冲剂(具体是有氢氧化铝佐剂); 或柠檬酸盐缓冲剂。包含的缓冲剂一般是 5-20mM。

[0153] 组合物的 pH 通常为 5.0-8.1, 更常为 6.0-8.0, 例如 6.5-7.5, 或者 7.0-7.8。

[0154] 该组合物优选无菌。该组合物优选无热原, 如每剂量含有 <1EU(内毒素单位, 标准量度), 优选每剂量 <0.1EU。该组合物优选不含谷蛋白。

[0155] 所述组合物可含有一次免疫的物质, 或者可含有多次免疫的物质(即‘多剂量’药盒)。多剂量配置优选含有防腐剂。作为多剂量组合物中包含防腐剂的替代方案(或补充方案), 所述组合物可包含在装有无菌接头以取出物质的容器中。

[0156] 人疫苗的给药剂量体积一般为约 0.5ml, 但可将一半剂量(即约 0.25ml) 给予儿童。

[0157] 本发明免疫原性组合物也可包含一种或多种免疫调节剂。优选地,一种或多种免疫调节剂包括一种或多种佐剂。佐剂可包括 TH1 佐剂和 / 或 TH2 佐剂,详述见下。

[0158] 可用于本发明组合物的佐剂包括但不限于:

[0159] A. 含有矿物质的组合物

[0160] 适用作本发明佐剂的含有矿物质的组合物包括矿物盐,例如铝盐和钙盐。本发明包括矿物盐,例如氢氧化物(如羟基氧化物)、磷酸盐(如羟基磷酸盐、正磷酸盐)、硫酸盐等[如参见参考文献 12 的第 8 和 9 章],或不同无机化合物的混合物,这些化合物可采取任何合适的形式(如凝胶、晶体、无定形等),优选具有吸附性。也可将含有矿物质的组合物制成金属盐的颗粒。

[0161] 称为“氢氧化铝”的佐剂一般是羟基氧化铝盐(通常至少部分为晶体)。可采用红外(IR)光谱将式 $AlO(OH)$ 代表的羟基氧化铝与其它铝化合物,如氢氧化铝 $Al(OH)_3$ 区别开,具体区别是 $1070cm^{-1}$ 处存在吸收条带和 $3090-3100cm^{-1}$ 处存在强肩[参考文献 12 的第 9 章]。半峰高处衍射带的宽度(WHH)反映了氢氧化铝佐剂的结晶程度,结晶不佳的颗粒因晶体尺寸较小而显示更强的谱线增宽。表面积随 WHH 的增加而增加,WHH 值较大的佐剂显示吸附抗原的能力较强。氢氧化铝佐剂通常呈纤维形态(例如,如透射电子显微图中所见)。氢氧化铝佐剂的 pI 通常约 11,即在生理 pH 下佐剂本身具有表面正电荷。据报道,pH 7.4 时氢氧化铝佐剂的吸附容量为 1.8-2.6 毫克蛋白质 / 毫克 Al^{+++} 。

[0162] 称为“磷酸铝”的佐剂一般是羟基磷酸铝,也常常含有少量硫酸盐(即羟基磷酸硫酸铝)。可通过沉淀获得这些佐剂,沉淀期间的反应条件和浓度影响磷酸根取代所述盐中羟基的程度。羟基磷酸盐的 PO_4/Al 摩尔比通常为 0.3-1.2。羟基磷酸盐因存在羟基而有别于严格的 $AlPO_4$ 。例如, $3164cm^{-1}$ 的 IR 光谱带(例如,当加热至 $200^{\circ}C$ 时)表明存在结构性羟基[参考文献 12 的第 9 章]。

[0163] 磷酸铝佐剂的 PO_4/Al^{3+} 摩尔比通常为 0.3-1.2,优选为 0.8-1.2,更优选为 0.95 ± 0.1 。磷酸铝通常是无定形的,尤其是羟基磷酸盐。典型的佐剂是 PO_4/Al 摩尔比为 0.84-0.92 的无定形羟基磷酸铝,包含 0.6mg Al^{3+}/ml 。磷酸铝通常是颗粒(如在透射电子显微图上观察到的板状形态)。抗原吸附后颗粒直径一般是 0.5-20 μm (如约 5-10 μm)。据报道,pH 7.4 时磷酸铝佐剂的吸附容量为 0.7-1.5 毫克蛋白质 / 毫克 Al^{+++} 。

[0164] 磷酸铝的零电点(PZC)与磷酸根取代羟基的程度逆相关,且这种取代程度可根据用于通过沉淀制备盐的反应条件和反应物浓度而变化。也通过改变溶液中游离磷酸根离子的浓度(更多磷酸根 = 更酸性 PZC)或加入缓冲剂如组氨酸缓冲剂(使 PZC 碱性更强)来改变 PZC。本发明所用的磷酸铝的 PZC 通常为 4.0-7.0,更优选为 5.0-6.5,例如约为 5.7。

[0165] 用于制备本发明组合物的铝盐悬浮液可以但不必定含有缓冲液(如磷酸盐或组氨酸或 Tris 缓冲液)。该悬浮液优选无菌且无热原。悬浮液可含有游离的水性磷酸根离子,如存在浓度为 1.0-20mM,优选 5-15mM,更优选约 10mM。该悬浮液也可含有氯化钠。

[0166] 在一个实施方式中,佐剂组分包括氢氧化铝和磷酸铝的混合物。在这种情况下,磷酸铝多于氢氧化铝,例如重量比为至少 2:1,例如, $\geq 5:1$ 、 $\geq 6:1$ 、 $\geq 7:1$ 、 $\geq 8:1$ 、 $\geq 9:1$ 等。

[0167] 给予患者的组合物中 Al^{+++} 的浓度优选小于 10mg/ml,例如 $\leq 5mg/ml$ 、 $\leq 4mg/ml$ 、 $\leq 3mg/ml$ 、 $\leq 2mg/ml$ 、 $\leq 1mg/ml$ 等。优选范围是 0.3-1mg/ml。优选最大值是 $<0.85mg/剂$ 。

[0168] B. 油乳剂

[0169] 适用作本发明佐剂的油乳剂组合物包含鲨烯-水乳剂,如 MF59[参考文献 12 的第 10 章;也参见参考文献 13](5% 鲨烯、0.5% 吐温 80 和 0.5% 司盘 85,用微流化床配制成亚微米颗粒)。也可以使用完全弗氏佐剂(CFA)和不完全弗氏佐剂(IFA)。

[0170] 已知各种合适的水包油乳剂,它们通常包括至少一种油和至少一种表面活性剂,所述油和表面活性剂是可生物降解(可代谢)和生物相容的。乳剂中的油滴直径通常小于 5 μm ,乳液宜包含具有亚微米直径的油滴,通过微流化床实现这种小尺寸以提供稳定乳剂。优选尺寸小于 220nm 的液滴,因为其可进行过滤灭菌。

[0171] 本发明可使用的油诸如来自动物(如鱼)或植物的油。植物油的来源包括坚果、种籽和谷物。最常见花生油、大豆油、椰子油和橄榄油是坚果油的示例。也可使用获自(例如)霍霍巴豆的霍霍巴油。种籽油包括红花油、棉花籽油、葵花籽油、芝麻籽油等。在谷物油中,最常见的是玉米油,但也可以使用其它谷类的油,如小麦、燕麦、黑麦、稻、画眉草、黑小麦等。可从坚果和种籽油开始,通过水解、分离和酯化合适物质制备甘油和 1,2-丙二醇的 6-10 碳脂肪酸酯,其不是种籽油中天然产生。来自哺乳动物乳汁的脂肪和油类是可代谢的,因此可以用于实施本发明。获得动物来源纯油所必需的分离、纯化、皂化和其它方法的过程为本领域熟知。大多数鱼类含有容易回收的可代谢油。例如,鳕鱼肝油、鲨鱼肝油和诸如鲸蜡的鲸油是可以用于本发明的几种鱼油的示例。通过生化途径用 5-碳异戊二烯单位合成许多支链油,其总称为萜类。鲨鱼肝油含有称为鲨烯的支链不饱和萜类化合物,即 2,6,10,15,19,23-六甲基-2,6,10,14,18,22-二十四碳六烯。其它优选油是生育酚(见下)。包含鲨烯的水包油乳液是特别优选的。可以使用油的混合物。

[0172] 表面活性剂可以按其“HLB”(亲水/亲脂平衡)分类。本发明优选的表面活性剂的 HLB 为至少 10,优选至少 15,更优选至少 16。可以与本发明一起使用的表面活性剂包括但不限于:聚氧乙烯去水山梨糖醇酯表面活性剂(通常称为吐温),特别是聚山梨酯 20 和聚山梨酯 80;以商品名 DOWFAX™ 出售的环氧乙烷(EO)、环氧丙烷(PO)和/或环氧丁烷(BO)的共聚物,如直链 EP/PO 嵌段共聚物;重复的乙氧基(氧-1,2-乙二基)数量不同的辛苯聚醇,特别感兴趣的是辛苯聚醇 9(曲通(Triton)X-100,或叔辛基苯氧基聚乙氧基乙醇);(辛基苯氧基)聚乙氧基乙醇(IGEPAL CA-630/NP-40);磷脂如磷脂酰胆碱(卵磷脂);衍生自十二烷醇、十六烷醇、十八烷醇和油醇的聚氧乙烯脂肪醚(称为苜泽表面活性剂),如三乙二醇单月桂基醚(苜泽 30);以及去水山梨糖醇酯(总称为司盘),如去水山梨糖醇三油酸酯(司盘 85)和去水山梨糖醇单月桂酸酯。乳液中包含的优选表面活性剂是吐温 80(聚氧乙烯脱水山梨糖醇单油酸酯)、司盘 85(脱水山梨糖醇三油酸酯)、卵磷脂和曲通 X100。如上所述,诸如吐温 80 等去污剂可提供下文实施例所见的热稳定性。

[0173] 可使用表面活性剂的混合物,如吐温 80/司盘 85 混合物。聚氧乙烯去水山梨糖醇酯如聚氧乙烯去水山梨糖醇单油酸酯(吐温 80)和辛苯聚醇如叔辛基苯氧基聚乙氧基乙醇(曲通 X-100)的组合也适合。另一种有用的组合包含月桂醇聚醚-9 加聚氧乙烯去水山梨糖醇酯和/或辛苯聚醇。

[0174] 优选的表面活性剂的含量(重量%)为:聚氧乙烯去水山梨糖醇酯(如吐温 80)0.01-1%,特别是约 0.1%;辛基-或壬基-苯氧基聚氧乙醇(如曲通 X100 或曲通系列的其它去污剂)0.001-0.1%,特别是 0.005-0.02%;聚氧乙烯醚(如月桂醇聚醚 9)0.1-20%,优选 0.1-10%,特别是 0.1-1% 或约 0.5%。

[0175] 本发明所用的具体水包油乳液佐剂包括但不限于：

[0176] • 鲨烯、吐温 80 (Tween 80) 和司盘 (Span) 85 的亚微米乳液。所述乳液的体积组成可以是约 5% 鲨烯、约 0.5% 聚山梨酯 80 和约 0.5% 司盘 85。以重量计, 这些比例为 4.3% 鲨烯、0.5% 聚山梨酯 80 和 0.48% 司盘 85。这种佐剂称为 ‘MF59’ [14-16], 参考文献 17 的第 10 章和参考文献 18 的第 12 章更详细地描述了该佐剂。MF59 乳液宜包含柠檬酸根离子, 如 10mM 柠檬酸钠缓冲液。

[0177] • 包含鲨烯、 α 生育酚和聚山梨酯 80 的乳液。这些乳液可含有 2-10% 鲨烯、2-10% 生育酚和 0.3-3% 吐温 80, 鲨烯 : 生育酚的重量比优选 ≤ 1 (例如 0.90), 因为这能使乳液更稳定。鲨烯和吐温 80 的体积比可以约为 5:2, 或者重量比约为 11:5。可通过将吐温 80 溶解于 PBS 产生 2% 溶液, 然后将 90ml 该溶液与 5gDL- α - 生育酚和 5ml 鲨烯的混合物混合, 随后使该混合物微流体化来制备一种这类乳液。得到的乳液可含有如平均直径为 100-250nm, 优选约 180nm 的亚微米油滴。

[0178] • 鲨烯、生育酚和曲通去污剂 (如曲通 X-100) 的乳液。该乳液也可包含 3d-MPL (见下)。所述乳液可包含磷酸盐缓冲液。

[0179] • 含有聚山梨酯 (如聚山梨酯 80)、曲通去污剂 (如曲通 X-100) 和生育酚 (如 α - 生育酚琥珀酸盐) 的乳液。该乳液可包含这三种组分, 其质量比约为 75:11:10 (如 750 μ g/ml 聚山梨酯 80、110 μ g/ml 曲通 X-100 和 100 μ g/ml 琥珀酸 α - 生育酚), 这些浓度应包括抗原中这些组分的贡献。所述乳液还可包含鲨烯。该乳液也可包含 3d-MPL (见下)。所述水相可包含磷酸盐缓冲液。

[0180] • 鲨烷、聚山梨酯 80 和泊洛沙姆 401 (“普流罗尼克™ L121” (“Pluronic™ L121”)) 的乳液。所述乳液可用 pH 7.4 的磷酸盐缓冲盐水配制。该乳液是一种有用的胞壁酰二肽递送载体, 且已与含苏氨酸基 -MDP 的 “SAF-1” 佐剂 [19] (0.05-1% Thr-MDP、5% 鲨烯、2.5% 普流罗尼克 L121 和 0.2% 聚山梨酯 80) 一起使用。也可不与 Thr-MDP 一起使用, 例如用 “AF” 佐剂 [20] (5% 鲨烯、1.25% 普流罗尼克 L121 和 0.2% 聚山梨酯 80)。优选微流体化。

[0181] • 含有鲨烯、水溶剂、聚氧乙烯烷基醚亲水性非离子型表面活性剂 (如聚氧乙烯 (12) 十六十八醚) 和疏水性非离子型表面活性剂 (如去水山梨糖醇酯或二缩甘露醇酯, 如去水山梨糖醇单油酸酯或 ‘司盘 80’) 的乳液。该乳液优选为热可逆的和 / 或其中至少 90% 油滴 (以体积计) 的尺寸小于 200nm [21]。该乳液也可含有以下一种或多种物质: 糖醇; 低温保护剂 (例如, 糖, 如十二烷基麦芽苷和 / 或蔗糖); 和 / 或烷基聚糖苷。这类乳液可冻干。

[0182] • 含有 0.5-50% 油、0.1-10% 磷脂和 0.05-5% 非离子型表面活性剂的乳液。如参考文献 22 所述, 优选的磷脂组分是磷脂酰胆碱、磷脂酰乙醇胺、磷脂酰丝氨酸、磷脂酰肌醇、磷脂酰甘油、磷脂酸、鞘磷脂和心磷脂。优选亚微米液滴尺寸。

[0183] • 不可代谢油 (如轻质矿物油) 和至少一种表面活性剂 (如卵磷脂、吐温 80 或司盘 80) 的亚微米水包油乳液。可包含添加剂, 例如 QuilA 皂苷、胆固醇、皂苷 - 亲脂偶联物 (如通过葡糖醛酸的羧基将脂族胺加到脱酰基皂苷上而产生的 GPI-0100, 如参考文献 23 所述)、二甲基双十八烷基溴化铵和 / 或 N, N- 双十八烷基 -N, N- 双 (2- 羟乙基) 丙二胺。

[0184] • 包含矿物油、非离子亲脂性乙氧基化脂肪醇和非离子亲水性表面活性剂 (例如, 乙氧基化脂肪醇和 / 或聚氧乙烯 - 聚氧丙烯嵌段共聚物) 的乳液 [24]。

[0185] • 包含矿物油、非离子亲水性乙氧基化脂肪醇和非离子亲脂性表面活性剂（例如，乙氧基化脂肪醇和 / 或聚氧乙烯 - 聚氧丙烯嵌段共聚物）的乳液 [24]。

[0186] • 皂苷（如 QuilA 或 QS21）和固醇（如胆固醇）结合成螺旋胶束的乳液 [25]。

[0187] 通常在传递给病人时混合组合物中的抗原和佐剂。可以在生产时或在递送时将该乳液与抗原临时混合。因此，在包装或出售的疫苗中该佐剂和抗原可分开保存，使用时配制成最终制剂。所述抗原通常是水性形式，从而最终通过混合两种液体制备疫苗。所述两种液体的混合体积比可变（例如 5:1-1:5），但通常约为 1:1。

[0188] C. 皂苷制剂 [参考文献 12 的第 22 章]

[0189] 皂苷制剂也可以用作本发明的佐剂。皂苷是在许多植物物种的树皮、叶、茎干、根甚至花中发现的甾醇糖苷和三萜糖苷的异质群体。已广泛研究了作为佐剂的得自皂皮树(*Quillaja saponaria* Molina) 树皮的皂苷。皂苷也可市售获自丽花菝葜(*Smilax ornata*) (墨西哥菝葜)、满天星(*Gypsophilla paniculata*) (婚纱花)和肥皂草(*Saponaria officinalis*)(皂根)。皂苷佐剂制剂包括纯化制剂如 QS21, 以及脂质制剂如 ISCOM。QS21 以商标 Stimulon™ 出售。

[0190] 已采用 HPLC 和 RP-HPLC 纯化皂苷组合物。已鉴定了用这些技术纯化的特定组分，包括 QS7、QS17、QS18、QS21、QH-A、QH-B 和 QH-C。所述皂苷优选 QS21。制备 QS21 的方法公开于参考文献 26。皂苷制剂也可包含甾醇，如胆固醇 [27]。

[0191] 皂苷和胆固醇的组合可用于形成称为免疫刺激复合物 (ISCOM) 的独特颗粒 [参考文献 12 第 23 章]。ISCOM 通常还包含磷脂，如磷脂酰乙醇胺或磷脂酰胆碱。任何已知的皂苷均可用于 ISCOM 中。ISCOM 优选包含 QuilA、QHA 和 QHC 中的一种或多种。参考文献 27-29 中进一步描述了 ISCOM。任选地，ISCOM 可不含其它去污剂 [30]。

[0192] 开发基于皂苷的佐剂的综述可参见参考文献 31 和 32。

[0193] D. 病毒体和病毒样颗粒

[0194] 病毒体和病毒样颗粒 (VLP) 也可以用作本发明的佐剂。这些结构通常包含一种或多种任选与磷脂组合或一起配制的病毒蛋白质。其通常无病原性，不能复制，且通常不含任何天然病毒基因组。所述病毒蛋白可重组生成或分离自全病毒。这些适用于病毒体或 VLP 的病毒蛋白包括衍生自流感病毒（例如 HA 或 NA）、乙肝病毒（例如核心蛋白或衣壳蛋白）、戊肝病毒、麻疹病毒、辛德比斯病毒、轮状病毒、口蹄疫病毒、逆转录病毒、诺沃克病毒、人乳头状瘤病毒、HIV、RNA- 噬菌体、Q β - 噬菌体（如外壳蛋白）、GA- 噬菌体、fr- 噬菌体、AP205 噬菌体和 Ty（如反转录转座子 Ty 蛋白 p1）的蛋白。VLP 在参考文献 33-38 中有进一步描述。病毒体在（例如）参考文献 39 中有进一步描述。

[0195] E. 细菌或微生物衍生物

[0196] 适用于本发明的佐剂包括细菌或微生物衍生物，如肠细菌脂多糖 (LPS) 的无毒衍生物、脂质 A 衍生物、免疫刺激性寡核苷酸和 ADP- 核糖基化毒素及其脱毒衍生物。

[0197] LPS 的无毒衍生物包括单磷酸脂质 A (MPL) 和 3-O- 脱酰基 MPL (3dMPL)。3dMPL 是 3 脱 -O- 酰基单磷酸脂质 A 与 4、5 或 6 条酰化链的混合物。3 脱 -O- 酰基单磷酸脂质 A 的优选“小颗粒”形式公开于参考文献 40。3dMPL 的这种“小颗粒”小到足以通过 0.22 μ m 膜过滤除菌 [40]。其它无毒 LPS 衍生物包括单磷酸脂质 A 模拟物，如氨基烷基氨基葡萄糖苷磷酸盐衍生物，例如 RC 529 [41, 42]。

[0198] 脂质 A 衍生物包括大肠杆菌的脂质 A 衍生物,如 OM-174。例如参考文献 43 和 44 中描述了 OM-174。

[0199] 适用作本发明佐剂的免疫刺激性寡核苷酸包括含 CpG 基序的核苷酸序列(含有通过磷酸键与鸟苷连接的非甲基化胞嘧啶的二核苷酸序列)。含回文结构或聚(dG)序列的双链 RNA 和寡核苷酸也显示具有免疫刺激性。

[0200] CpG 可以包含核苷酸修饰/类似物,如硫代磷酸酯修饰,且可以是双链或单链。参考文献 45、46 和 47 公开了可能的类似取代,例如用 2'-脱氧-7-脱氮鸟苷取代鸟苷。参考文献 48-53 中进一步讨论了 CpG 寡核苷酸的佐剂作用。

[0201] CpG 序列可针对 TLR9,例如基序 GTCGTT 或 TTCGTT[54]。CpG 序列可特异性诱导 Th1 免疫应答,如 CpG-A ODN,或更特异地诱导 B 细胞应答,如 CpG-B ODN。参考文献 55-57 中讨论了 CpG-A 和 CpG-B ODN。优选 CpG 为 CpG-A ODN。

[0202] CpG 寡核苷酸优选构建成 5' 末端可被受体识别。任选将两个 CpG 寡核苷酸序列的 3' 端相连接形成“免疫聚体”。参见例如参考文献 54 和 58-60。

[0203] 基于免疫刺激性寡核苷酸的特别有用的佐剂被称为 IC-31™[61]。因此,本发明使用的佐剂可以包含 (i) 和 (ii) 的混合物:(i) 含有至少一个(优选多个)CpI 基序(即胞嘧啶与肌苷相连形成二核苷酸)的寡核苷酸(例如 15-40 个核苷酸),和 (ii) 聚阳离子聚合物,如含有至少一个(优选多个)Lys-Arg-Lys 三肽序列的寡肽(如 5-20 个氨基酸)。所述寡核苷酸可以是包含 26-聚体序列 5'-(IC)₁₃-3' (SEQ ID NO:32) 的脱氧核苷酸。聚阳离子聚合物可以是含有 11-聚体氨基酸序列 KLKLLLLLKLK (SEQ ID NO:33) 的肽。

[0204] 细菌 ADP-核糖基化毒素及其去毒衍生物可以用作本发明的佐剂。优选所述蛋白衍生自大肠杆菌(大肠杆菌不耐热肠毒素“LT”)、霍乱菌(“CT”)或百日咳菌(“PT”)。参考文献 62 中描述了脱毒的 ADP-核糖基化毒素作为粘膜佐剂的应用,参考文献 63 中描述了其作为胃肠外佐剂的应用。所述毒素或类毒素优选全毒素形式,包含 A 和 B 亚基。A 亚基优选含有脱毒突变;B 亚基优选不突变。所述佐剂优选脱毒的 LT 突变体,如 LT-K63、LT-R72 和 LT-G192。ADP-核糖基化毒素及其脱毒衍生物,尤其是 LT-K63 和 LT-R72 作为佐剂的应用可参见参考文献 64-71。一种有用的 CT 突变体是 CT-E29H[72]。优选根据参考文献 73 中提出的 ADP-核糖基化毒素的 A 和 B 亚基排列对氨基酸取代基编号,该参考文献特定通过引用全文纳入本文。

[0205] F. 人免疫调节剂

[0206] 适用作本发明佐剂的人免疫调节剂包括细胞因子,如白介素(例如,IL-1、IL-2、IL-4、IL-5、IL-6、IL-7、IL-12[74]等)[75]、干扰素(例如干扰素-γ)、巨噬细胞集落刺激因子和肿瘤坏死因子。优选的免疫调节剂是 IL-12。

[0207] G. 生物粘着剂和粘膜粘着剂

[0208] 生物粘着剂和粘膜粘着剂也可以用作本发明的佐剂。合适的生物粘着剂包括酯化透明质酸微球[76]或粘膜粘着剂如聚(丙烯酸)交联衍生物、聚乙烯醇、聚乙烯吡咯烷酮、多糖和羧甲基纤维素。壳聚糖及其衍生物也可用作本发明的佐剂[77]。

[0209] H. 微粒

[0210] 微粒也可以用作本发明的佐剂。微粒(即直径为~100nm至~150μm,更优选直径~200nm~30μm,最优选直径~500nm~10μm的颗粒)由生物可降解的无毒材料(例如,聚

(α -羟酸)、聚羟基丁酸、聚原酸酯、聚酐、聚己内酯等)形成,优选聚丙交酯乙交酯共聚物,并任选经处理而具有带负电表面(例如用 SDS 处理)或带正电表面(例如用阳离子去污剂如 CTAB 处理)。

[0211] I. 脂质体(参考文献 12 的第 13 和 14 章)

[0212] 适用作佐剂的脂质体制剂的例子见参考文献 78-80 所述。

[0213] J. 聚氧乙烯醚和聚氧乙烯酯制剂

[0214] 适用于本发明的佐剂包括聚氧乙烯醚和聚氧乙烯酯 [81]。这种制剂还包括聚氧乙烯去水山梨糖醇酯表面活性剂和辛苯糖醇 [82] 以及聚氧乙烯烷基醚或酯表面活性剂和至少一种其它非离子表面活性剂如辛苯糖醇 [83] 的组合。优选的聚氧乙烯醚选自下组:聚氧乙烯-9-月桂醚(月桂醇聚醚 9)、聚氧乙烯-9-硬脂醚、聚氧乙烯-8-硬脂醚、聚氧乙烯-4-月桂醚、聚氧乙烯-35-月桂醚和聚氧乙烯-23-月桂醚。

[0215] K. 聚磷腈(PCPP)

[0216] PCPP 制剂参见例如参考文献 84 和 85。

[0217] L. 胞壁酰肽

[0218] 适用作本发明佐剂的胞壁酰肽的例子包括 N-乙酰基-胞壁酰-L-苏氨酸-D-异谷酰胺(thr-MDP)、N-乙酰基-正胞壁酰-L-丙氨酸-D-异谷酰胺(去甲-MDP)和 N-乙酰胞壁酰-L-丙氨酸-D-异谷氨酸酰胺-L-丙氨酸-2-(1'-2'-二棕榈酰-sn-甘油-3-羟基磷酰氧基)-乙胺 MTP-PE)。

[0219] M. 咪唑并喹诺酮(Imidazoquinolone)化合物

[0220] 适用作本发明佐剂的咪唑并喹诺酮化合物的例子包括咪唑莫特及其同系物(例如,“瑞喹莫德 3M”),其进一步描述于参考文献 86 和 87。

[0221] 本发明也可包含以上鉴定的一种或多种佐剂各方面的组合。例如,可将以下佐剂组合物用于本发明:(1)皂苷和水包油乳剂 [88];(2)皂苷(如 QS21)+无毒 LPS 衍生物(如 3dMPL) [89];(3)皂苷(如 QS21)+无毒 LPS 衍生物(如 3dMPL)+胆固醇;(4)皂苷(如 QS21)+3dMPL+IL12(任选+固醇) [90];(5)3dMPL 与(例如)QS21 和/或水包油乳剂的组合 [91];(6)SAF,含有 10% 鲨烯、0.4% 吐温 80TM、5% 普朗尼克-嵌段聚合物 L121 和 thr-MDP,或微流体化成为亚微米乳液或涡旋振荡产生粒度较大的乳液。(7)RibiTM佐剂系统(RAS)(RI 公司(Ribi Immunochem)),含有 2% 鲨烯、0.2% 吐温 80 和一种或多种细菌细胞壁组分,所述组分选自单磷酰脂质 A(MPL)、海藻糖二霉菌酸酯(TDM)和细胞壁骨架(CWS),优选 MPL+CWS(DetoxTM);和(8)一种或多种矿物盐(如铝盐)+LPS 的无毒衍生物(如 3dMPL)。

[0222] 用作免疫刺激剂的其它物质公开于参考文献 12 的第 7 章。

[0223] 特定在儿童中,使用氢氧化铝和/或磷酸铝佐剂有效,且抗原通常吸附于这些盐。也优选水包鲨烯乳液,特定在老年人中。有用的佐剂组合包括 Th1 和 Th2 佐剂的组合,如 CpG 和明矾或雷西莫特和明矾。可以使用磷酸铝和 3dMPL 的组合。

[0224] 本发明组合物可引起细胞介导的免疫应答以及体液免疫应答。

[0225] 通常认为两种 T 细胞类型 CD4 和 CD8 细胞是启动和/或增强细胞介导免疫和体液免疫所必需的。CD8T 细胞可表达 CD8 共受体,通常称为细胞毒性 T 淋巴细胞(CTL)。CD8T 细胞能够识别 MHC I 型分子上展示的抗原或与之相互作用。

[0226] CD4T 细胞可表达 CD4 共受体,通常称为 T 辅助细胞。CD4T 细胞能够识别结合 MHC

II 型分子的抗原性肽。与 MHC II 型分子相互作用后, CD4 细胞可分泌诸如细胞因子等因子。这些分泌的细胞因子可激活 B 细胞、细胞毒性 T 细胞、巨噬细胞和参与免疫应答的其它细胞。辅助 T 细胞或 CD4⁺ 细胞可进一步分为两个功能不同的亚组:即细胞因子和效应功能不同的 TH1 表型和 TH2 表型。

[0227] 活化的 TH1 细胞能增强细胞免疫(包括抗原特异性 CTL 生成增加),因而对响应胞内感染具有特定价值。活化的 TH1 细胞可分泌 IL-2、IFN- γ 和 TNF- β 中的一种或多种。TH1 免疫应答可通过激活巨噬细胞、NK(自然杀伤)细胞和 CD8 细胞毒性 T 细胞(CTL) 导致局部炎症反应。通过用 IL-12 刺激 B 和 T 细胞的生长, TH1 免疫应答也可用于放大免疫应答。TH1 刺激的 B 细胞可分泌 IgG2a。

[0228] 活化的 TH2 细胞提高抗体生成,因此对响应胞外感染具有价值。活化的 TH2 细胞可分泌 IL-4、IL-5、IL-6 和 IL10 中的一种或多种。TH2 免疫应答可引起产生 IgG1、IgE、IgA 和用于未来保护的记忆 B 细胞。

[0229] 增强的免疫应答可包括增强的 TH1 免疫应答和 TH2 免疫应答中的一种或多种。

[0230] TH1 免疫应答可包括以下一种或多种:CTL 增加,与 TH1 免疫应答相关的一种或多种细胞因子(如 IL-2、IFN γ 和 TNF- β) 增加,活化巨噬细胞增加,NK 活性增加,或者 IgG2a 生成增加。增强的 TH1 免疫应答优选包括 IgG2a 生成增加。

[0231] 可使用 TH1 佐剂引发 TH1 免疫应答。相对于不用佐剂的抗原免疫, TH1 佐剂通常引起 IgG2a 生成水平增加。适用于本发明的 TH1 佐剂可包括例如,皂苷制剂、病毒体和病毒样颗粒、肠细菌脂多糖(LPS)的无毒衍生物、免疫刺激性寡核苷酸。免疫刺激性寡核苷酸,如含有 CpG 基序的寡核苷酸是本发明所用的优选 TH1 佐剂。

[0232] TH2 免疫应答可包括以下一种或多种:与 TH2 免疫应答相关的一种或多种细胞因子(如 IL-4, IL-5, IL-6 和 IL-10) 增加,或者 IgG1、IgE、IgA 和记忆 B 细胞生成增加。增强的 TH2 免疫应答优选包括 IgG1 生成增加。

[0233] 可使用 TH2 佐剂引发 TH2 免疫应答。相对于不用佐剂的抗原免疫, TH2 佐剂通常引起 IgG1 生成水平增加。适用于本发明的 TH2 佐剂包括例如,含矿物质组合物、油乳剂和 ADP-核糖基化毒素和其脱毒衍生物。含矿物质组合物如铝盐是本发明使用的优选 TH2 佐剂。

[0234] 组合物可包括 TH1 佐剂和 TH2 佐剂的组合。这种组合物优选引起增强的 TH1 和增强的 TH2 应答,即 IgG1 和 IgG2a 的生成相对于不用佐剂的免疫均有增加。更优选地,相对于用单一佐剂的免疫(即,相对于只用 TH1 佐剂的免疫或只用 TH2 佐剂的免疫),包含 TH1 和 TH2 佐剂组合的组合物引起 TH1 增加和/或 TH2 免疫应答增强。

[0235] 所述免疫应答可以是 TH1 免疫应答和 TH2 免疫应答之一或两种。免疫应答优选提供增强的 TH1 应答和增强的 TH2 应答之一或两种。

[0236] 增强的免疫应答可以是全身免疫应答和粘膜免疫应答之一或两种。该免疫应答优选提供增强的全身免疫应答和增强的粘膜免疫应答之一或两种。优选粘膜免疫应答为 TH2 免疫应答。粘膜免疫应答优选包括 IgA 生成增加。

[0237] 链球菌(*Streptococcal*) 感染可影响机体的各个部分,因此可将本发明组合物制备成各种形式。例如,可将所述组合物制备为液体溶液或悬浮液形式的注射剂。也可制备适合在注射前溶解或悬浮于液体载剂的固体形式(如冻干组合物或喷雾冻干组合物)。可制

备所述组合物用于局部给药,例如作为油膏、乳膏或粉剂。该组合物可制备用于口服给药,如作为片剂或胶囊,喷雾剂,或糖浆剂(任选调味)。可将所述组合物制备为采用细粉或喷雾的肺部给药制剂,例如吸入剂。可将所述组合物制备为栓剂或阴道栓。所述组合物可以制成鼻部、耳部或眼部给药制剂,例如滴剂。组合物可以是药盒形式,设计成在临给予患者之前重建的合并组合物。这种药盒可包含一种或多种液体形式的抗原以及一种或多种冻干抗原。

[0238] 组合物在使用前临时制备(例如,以冻干形式提供组分)和以药盒形式提供时,该药盒可包括两个药瓶,或者可包括一个已填充的注射器和一个药瓶,所述注射器内容物用于在注射前再次激活药瓶内容物。

[0239] 用作疫苗的免疫原性组合物包含免疫有效量的抗原,以及需要的任何其它组分。“免疫有效量”指在一次剂量或一系列剂量的一部分中给予个体的对治疗或预防有效的量。该量根据待治疗个体的健康和身体状况、年龄、待治疗个体的分类群(例如,非人灵长类、灵长类等)、个体免疫系统合成抗体的能力、所需的保护程度、疫苗配方、治疗医生对医学情况的评估和其它相关因素而变化。预计所述量将落入可通过常规试验测定的相对较宽范围内。

[0240] 核酸免疫

[0241] 上述免疫原性组合物包括来自 GBS 的多肽抗原。然而,在所有情况下,可用编码这些多肽的核酸(通常是 DNA)替换多肽(和杂交多肽)抗原,以得到基于核酸免疫的组合物、方法和用途 [92-99]。

[0242] 编码免疫原的核酸在递送给病人后体内表达,然后所表达的免疫原刺激免疫系统。活性成分通常采用核酸载体形式,其包含:(i) 启动子;(ii) 可操作连接于启动子的免疫原编码序列;以及任选(iii) 选择标记物。优选的载体还可包含(iv) 复制起点;以及(v) 位于(ii) 下游并与其可操作连接的转录终止子。通常,(i) 和(v) 为真核的,(iii) 和(iv) 为原核的。

[0243] 优选的启动子是病毒启动子,如来自巨细胞病毒(CMV)的启动子。除了启动子,载体还可以包含与启动子功能性相互作用的转录调节序列(如增强子)。优选的载体包含立即早期 CMV 增强子/启动子,更优选的载体还包含 CMV 内含子 A。将该启动子可操作连接于编码免疫原的下游序列,从而使免疫原编码序列的表达在该启动子的控制之下。

[0244] 当使用标记物时,优选其在微生物宿主中(如原核生物、细菌、酵母中)具有功能。标记物优选原核选择标记物(如在原核启动子控制之下转录)。方便起见,典型的标记物为抗生素抗性基因。

[0245] 所述载体优选自主复制附加体或染色体外载体,如质粒。

[0246] 所述载体优选包括复制起点。该复制起点优选在原核生物中有活性而在真核生物中无活性。

[0247] 因此优选的载体包含用于选择载体的原核标记物、原核复制起点以及驱动免疫原编码序列转录的真核启动子。因此载体(a)在原核宿主中扩增并选择,但不进行多肽表达,而(b)在真核宿主中表达,但不进行扩增。这种配置对于核酸免疫载体来说是理想的。

[0248] 所述载体可以在编码序列下游包含真核转录终止序列。这能增强转录水平。当编码序列本身不包含聚腺苷酸化序列时,所述载体优选包含聚腺苷酸化序列。优选的聚腺苷

酸化序列来自于牛生长激素。

[0249] 所述载体可包括多克隆位点。

[0250] 除免疫原和标记物的编码序列外,载体可以包含第二个真核编码序列。载体还可在所述第二序列上游包含 IRES,以从与该免疫原相同的转录物中翻译第二条真核多肽。或者,免疫原编码序列可以在 IRES 的下游。

[0251] 所述载体可以包含非甲基化 CpG 基序,如非甲基化 DNA 序列,它们都在鸟苷之前有一个胞嘧啶,侧接有两个 5' 嘌呤和两个 3' 嘧啶。在其非甲基化形式中,已证明这些 DNA 基序是多种类型免疫细胞的强效刺激因子。

[0252] 载体可以靶向方式进行递送。例如,参考文献 100-105 中描述了受体介导的 DNA 递送技术。在基因治疗方案中,以约 100ng-200mg DNA 的范围局部给予包含核酸的治疗性组合物。在基因治疗方案中也能使用浓度范围约 500ng-50mg,约 1 μ g-2mg,约 5 μ g-500 μ g 以及约 20 μ g-100 μ g 的 DNA。诸如作用方法(如提高或抑制编码基因产物的水平)以及转化和表达效率等因素是影响最终功效所需剂量的考虑因素。当需要在更大面积组织上更强表达时,可能需要更大量的载体或在连续给药方案中重复给予相同量,或对不同相邻或紧密组织部分多次给药,以达到阳性治疗效果。在所有情况下,使用临床试验中的常规实验方法确定最优治疗效果的特定范围。

[0253] 可用基因递送运载体来递送载体。基因递送运载体可以是病毒或非病毒来源(通常参见参考文献 106-109)。

[0254] 本领域熟知用于递送所需核酸并在所需细胞中表达的基于病毒的载体。示例性基于病毒的载体包括但不限于:重组逆转录病毒(例如参考文献 110-120),基于 α 病毒的载体(如辛德毕斯病毒载体、西门利克森林病毒(ATCC VR-67;ATCC VR-1247)、罗斯河病毒(ATCC VR-373;ATCC VR-1246)和委内瑞拉马脑炎病毒(ATCC VR-923;ATCC VR-1250;ATCC VR 1249;ATCC VR-532);还可以使用这些病毒的杂交体或嵌合体)、痘病毒载体(如牛痘、鸟痘、金丝雀痘、改性的安卡拉牛痘等)、腺病毒载体和腺伴随病毒(AAV)载体(例如参见参考文献 121-126)。还能给予连接已灭活腺病毒的 DNA[127]。

[0255] 还能使用非病毒递送运载体和方法,包括但不限于:连接或不连接单独灭活腺病毒的聚阳离子浓缩 DNA[如 127],配体连接的 DNA[128],真核细胞递送运载体细胞[如参考文献 129-133]和核电荷中和或与细胞膜融合。还能使用裸露 DNA。参考文献 134 和 135 中描述了引入裸露 DNA 的示例性方法。参考文献 136-140 中描述了可作为基因递送运载体的脂质体(如免疫脂质体)。参考文献 141 和 142 中描述了其它方法。

[0256] 其它适用的非病毒递送包括机械递送系统,如参考文献 142 所述方法。而且,可通过光聚合水凝胶材料沉积或使用电离辐射来递送此编码序列和其表达产物[如参考文献 143 和 144]。其它可用于递送编码序列的常规基因递送方法包括例如,使用手持式基因传递颗粒枪[145]或使用离子辐射激活转移的基因[143 和 144]。

[0257] 使用 PLG(聚丙交酯乙交酯共聚物)微粒递送 DNA 是一个尤其优选的方法,如吸附于微粒上,微粒任选经处理成表面带负电(如用 SDS 处理)或带正电(如用阳离子去垢剂如 CTAB 处理)。

[0258] 治疗方法和所述疫苗给予

[0259] 本发明还提供了使哺乳动物产生免疫应答的方法,包括给予有效量的上述多肽、

杂交多肽、核酸或免疫原性组合物的步骤。所述免疫应答优选为保护性，并优选涉及抗体和 / 或细胞介导免疫。该方法可以引起增强的应答。

[0260] 本发明还提供上述多肽、杂交多肽、核酸或免疫原性组合物用作药物，如在哺乳动物中引起免疫应答。

[0261] 本发明还提供上述多肽、杂交多肽、核酸或免疫原性组合物在生产引起哺乳动物免疫应答的药物中的用途。

[0262] 借助这些应用和方法在哺乳动物中产生免疫应答后，能保护该哺乳动物抵御 GBS 引起的疾病和 / 或感染，例如抵御脑膜炎。

[0263] 本发明还提供了一种预填充本发明免疫原性组合物的递送装置。

[0264] 所述哺乳动物优选人。所述人可为青少年或成人。

[0265] 检测治疗性处理功效的一种方式包括在给予本发明组合物后监测 GBS 感染。一种检查预防性处理功效的方法包括用标准试验来测试免疫后血清；例如，可在调理吞噬实验 (OPKA) 中测试血清，其调理细菌的能力指示了保护效力。另一种检查预防性处理功效的方法涉及在 GBS 感染的动物模型，例如豚鼠或小鼠中的免疫后攻击。参考文献 146 中描述了一种此类模型。另一评价本发明组合物免疫原性的途径是重组表达多肽，通过免疫印迹和 / 或微阵列筛查患者血清或粘膜分泌物。多肽和患者样品间的阳性反应表明患者已经产生对受试多肽的免疫应答。该方法也可用于鉴定免疫显性抗原和 / 或抗原中的表位。

[0266] 本发明的组合物通常直接给予患者。可通过胃肠外注射（如皮下、腹膜内、静脉内、肌肉内或组织间隙），或通过粘膜，如经直肠、口腔（如片剂、喷雾）、阴道、局部、透皮或经皮、鼻内、眼、耳、肺或其它粘膜途径给药来完成直接递送。

[0267] 本发明可用于引起全身和 / 或粘膜免疫，优选引起增强的全身和 / 或粘膜免疫。

[0268] 增强的全身和 / 或粘膜免疫优先表现为提高的 TH1 和 / 或 TH2 免疫应答。增强的免疫应答优选包括 IgG1 和 / 或 IgG2a 和 / 或 IgA 生成增加。

[0269] 可以通过单剂量方案或多剂量方案进行给药。多剂量可以用于初次免疫方案和 / 或加强免疫方案。在多剂量方案中，可通过相同或不同途径给予各剂量，例如采用胃肠外初次免疫和粘膜加强免疫、粘膜初次免疫和胃肠外加强免疫等。一般以至少 1 周（例如约 2 周、约 3 周、约 4 周、约 6 周、约 8 周、约 10 周、约 12 周、约 16 周等）的间隔给予多个剂量。

[0270] 可用按照本发明制备的疫苗治疗儿童和成年人。因此，人患者可以小于 1 岁、小于 5 岁、1-5 岁、5-15 岁、15-55 岁或至少 55 岁。接受疫苗的优选患者是青少年（如 13-20 岁）、孕妇和老年人（如 ≥ 50 岁、 ≥ 60 岁和优选 ≥ 65 岁）。然而所述疫苗不仅适用于这些人群，还可用于更广泛的群体。

[0271] 可将本发明产生的疫苗与其它疫苗基本上同时（在健康护理专业人员或疫苗接种中心的同一用药咨询或就诊期间）给予患者，例如与风疹疫苗、水痘疫苗、白喉疫苗、破伤风疫苗、百日咳疫苗、DTP 疫苗、脊髓灰质炎病毒灭活疫苗、乙型肝炎病毒疫苗、脑膜炎球菌偶联疫苗（如四价 A-C-W135-Y 疫苗）、呼吸道合胞病毒疫苗、人乳头瘤病毒疫苗、流感病毒疫苗（包括流行性流感病毒疫苗）等基本上同时给药。

[0272] 还可将本发明疫苗与抗病毒化合物，具体是对流感病毒有活性的抗病毒化合物（如奥塞米韦和 / 或扎那米韦）基本上同时给予患者（如在健康护理专业人员的同一用药咨询或就诊期间）。这些抗病毒化合物包括神经氨酸酶抑制剂，如 (3R, 4R, 5S)-4-乙酰氨基

基-5-氨基-3(1-乙基丙氧基)-1-环己烯-1-羧酸或5-(乙酰基氨基)-4-[(氨基亚氨基甲基)-氨基]-2,6-脱水-3,4,5-三脱氧-D-甘油-D-半乳糖壬-2-烯酮酸,包括它们的酯(如乙酯)和盐(如磷酸盐)。优选的抗病毒化合物是(3R,4R,5S)-4-乙酰基氨基-5-氨基-3(1-乙基丙氧基)-1-环己烯-1-羧酸,乙酯和磷酸盐(1:1),也称为磷酸奥塞米韦(达菲(TAMIFLU)TM)。

[0273] 组合

[0274] 除了GBS67多肽片段以外,组合物可包括:(i)针对GBS蛋白,特别是针对GBS67以外的GBS蛋白引发抗体应答的一种或多种其他多肽;(ii)GBS的荚膜糖;和/或(iii)引发识别非GBS生物体上表位的抗体应答的一种或多种其他免疫原。

[0275] 与其他多肽抗原的组合

[0276] 上述GBS67多肽片段可与选自下组的一种或多种(即1、2、3、4、5、6、7、8、9或所有10)多肽抗原组合:(1)GBS80抗原;(2)GBS59抗原;(3)GBS1523抗原;(4)GBS 104抗原;(5)GBS1524抗原;(6)GBS3抗原;(7)SAN1485抗原;(8)GBS147抗原;(9)GBS328抗原;和/或(10)GBS84抗原。

[0277] 这些其他抗原可作为单独多肽加入。作为替代,其可作为杂交物加入,如GBS80-GBS1523杂交物。作为另一种替代,其可融合GBS67多肽片段以提供杂交多肽。

[0278] 任意这些组合也可包括一种或多种GBS荚膜糖,其通常偶联载体蛋白。下面提供所述糖和偶联的进一步信息。

[0279] GBS80

[0280] 原始'GBS67'(SAG0645)序列注释于参考文献147,为细胞壁表面锚定家族蛋白(参见GI:22533660)。出于参比目的,2603菌株中发现的全长GBS80的氨基酸序列在本文中作为SEQ ID NO:34。本发明所用的优选GBS80多肽包含某一氨基酸序列,该序列:(a)与SEQ ID NO:34具有60%或更高的相同性(如60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%或更高);和/或(b)包含SEQ ID NO:34的至少'n'个连续氨基酸的片段,其中'n'是7或更高(如8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250或更高)。这些GBS80蛋白包括SEQ ID NO:34的变体。

[0281] (b)的优选片段包括SEQ ID NO:34的表位。其他优选片段缺少SEQ IDNO:34C末端的一个或多个氨基酸(如1、2、3、4、5、6、7、8、9、10、15、20、25或更多个)和/或SEQ ID NO:34N末端的一个或多个氨基酸(如1、2、3、4、5、6、7、8、9、10、15、20、25或更多个),而保留SEQ ID NO:34的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0282] 野生型GBS 80在SEQ ID NO:34的氨基酸1-37包含N-末端前导或信号序列区域。GBS 80前导或信号序列区域的一个或多个氨基酸可被移除,如SEQ ID NO:35。野生型序列还在SEQ ID NO:34的氨基酸526-543包含C-末端跨膜区域。跨膜区域和/或胞质区域的一个或多个氨基酸可被移除,如SEQ IDNO:36。野生型GBS80在SEQ ID NO:34的氨基酸521-525包含指示细胞壁锚定的氨基酸基序。在一些重组宿主细胞系统中,移除该基序有助于重组GBS80多肽从所述宿主细胞中分泌。因此所述跨膜和/或胞质区域和细胞壁锚定基序可从GBS80中移除,如SEQ ID NO:37。或者,在一些重组宿主细胞系统中,用所述细胞壁锚定基序将重组表达多肽锚定在细胞壁上是有用的。表达多肽的胞外结构域在纯化中可被切割,或所述重组多肽可结合灭活的宿主细胞或最终组合物中的细胞膜,如SEQ ID NO:38。

野生型 GBS80 的特定免疫原性片段位于所述多肽的 N 末端方向,为 SEQ ID NO:39。

[0283] GBS59

[0284] GBS59 是毒力岛 2a(BP-2a) 编码的菌毛主干蛋白。出于参比目的,2603 菌株中发现的全长 GBS59 的氨基酸序列在本文中作为 SEQ ID NO:40。本发明所用的 GBS59 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:40 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和/或(b) 包含 SEQ ID NO:40 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS59 蛋白包括 SEQ ID NO:40 的变体。(b) 的优选片段包括 SEQ ID NO:40 的表位。其他优选片段缺少 SEQ ID NO:40C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:40 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0285] GBS59 变体存在于株系 H36B、515、CJB111、DK21 和 CJB110 中。出于参比目的,H36B、515、CJB111、CJB110 和 DK21 菌株中发现的全长 GBS59 的氨基酸序列在本文中作为 SEQ ID NO:41、42、43、44 和 45。本发明所用的优选 GBS59 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:41、42、43、44 或 45 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和/或(b) 包含 SEQ ID NO:41、42、43、44 或 45 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。(b) 的优选片段包含来自 SEQ ID NO:41、42、43、44 或 45 的表位。其他优选片段缺少 SEQ ID NO:41、42、43、44 或 45C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:41、42、43、44 或 45 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0286] GBS1523

[0287] 原始 'GBS 1523' (SAN1518; SpbI) 序列注释于参考文献 3,为细胞壁表面锚定家族蛋白(参见 GI:77408651)。出于参比目的,COH1 菌株中发现的全长 GBS1523 的氨基酸序列在本文中作为 SEQ ID NO:46。本发明所用的优选 GBS1523 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:46 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和/或(b) 包含 SEQ ID NO:46 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS1523 蛋白包括 SEQ ID NO:46 的变体。(b) 的优选片段包括 SEQ ID NO:46 的表位。其他优选片段缺少 SEQ ID NO:46C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:46 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0288] 野生型 GBS1523 在 SEQ ID NO:46 的氨基酸 1-29 包含 N-末端前导或信号序列区域,其可从片段中移除,如 SEQ ID NO:47。野生型序列在 SEQ ID NO:46 的氨基酸 468-472 包含指示细胞壁锚定的氨基酸基序(LPSTG)。在一些重组宿主细胞系统中,优选移除该基序以有助于重组多肽从所述细胞中分泌。或者优选使用所述细胞壁锚定基序将重组表达多肽锚

定在细胞壁上。表达多肽的胞外结构域可在纯化中进行切割,或所述重组多肽可结合灭活的宿主细胞或最终组合物中的细胞膜。还在 SEQ ID NO:46 的氨基酸 419-429 鉴定到含保守谷氨酸残基的 E 盒,在残基 423 处具有保守的谷氨酸。所述 E 盒基序可能对寡聚菌毛样结构的形成很重要,且因此 GBS1523 的有用片段可包括该保守谷氨酸残基。鉴定到 GBS1523 的突变,其中 SEQ ID NO:46 位置 41 处的谷氨酰胺(Q) 替换为赖氨酸(K),这是由于编码核酸序列中的密码子从 CAA 突变为 AAA。该替代可存在于 GBS1523 序列和 GBS1523 片段中(如 SEQ IDNO:48)。

[0289] 所述组合物包括 GBS80 和 GBS 1523 时,可使用杂交多肽。GBS80-GBS 1523 杂交物的示例见参考文献 148 且包括 SEQ ID NOS:49-52 的多肽。

[0290] GBS104

[0291] 原始 'GBS104' (SAG0649) 序列在参考文献 147 中注释为 '细胞壁表面锚定家族蛋白' (参见 GI:22533664)。出于参比目的,2603 菌株中发现的全长 GBS104 的氨基酸序列在本文中作为 SEQ ID NO:53。本发明所用的 GBS104 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:53 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和/或(b) 包含 SEQ ID NO:53 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS104 蛋白包括 SEQ ID NO:53 的变体。(b) 的优选片段包括 SEQ IDNO:40 的表位。其他优选片段缺少 SEQ ID NO:53C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:53 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0292] GBS1524

[0293] 出于参比目的,COH1 菌株中发现的全长 GBS1524(SAN1519) 的氨基酸序列在本文中作为 SEQ ID NO:54。本发明所用的优选 GBS1524 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:54 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和/或(b) 包含 SEQ ID NO:54 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS1524 蛋白包括 SEQ ID NO:54 的变体。(b) 的优选片段包括 SEQ ID NO:54 的表位。其他优选片段缺少 SEQ ID NO:54C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQID NO:54 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0294] GBS3

[0295] 原始 'GBS3' (SAG2603;BibA) 序列在参考文献 147 中标注为 '毒力蛋白' (参见 GI:22535109)。出于参比目的,2603 菌株中发现的全长 GBS3 的氨基酸序列在本文中作为 SEQ ID NO:55。本发明所用的优选 GBS3 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:55 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和/或(b) 包含 SEQ ID NO:55 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、

80、90、100、150、200、250 或更高)。这些 GBS3 蛋白包括 SEQ ID NO:55 的变体。(b) 的优选片段包括 SEQ ID NO:35 的表位。其他优选片段缺少 SEQ ID NO:55C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:55 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0296] 野生型 GBS3 在 SEQ ID NO:55 的氨基酸 1-36 包含 N-末端前导或信号序列区域,其可从片段中移除,如 SEQ ID NO:563。GBS3 还包括指示细胞壁锚定的氨基酸基序(LPXTG)、跨膜区域和胞质结构域(参见参考文献 149)。前导或信号序列区域、跨膜区域和胞质结构域、和细胞壁锚定基序都可从 GBS3 中移除,留下如下所示的包括卷曲螺旋和富含脯氨酸的区段(SEQ ID NO:57)。GBS3 的替代区段可包括:信号序列区域和卷曲螺旋区段(SEQ ID NO:58);卷曲螺旋区段(SEQ ID NO:59);或信号序列区域、卷曲螺旋区段和富含脯氨酸的区段(SEQ ID NO:60)。

[0297] GBS3 变体存在于 515 株系(SAL2118)、CJB111 株系(SAM1974)和 COH1 株系(SAN2207)中。515 株系、CJB111 株系和 COH1 株系中全长 GBS3 的参比氨基酸序列在本文中分别作为 SEQ ID NO:61、SEQ ID NO:62 和 SEQ ID NO:63。因此,本发明所用的 GBS3 多肽还可包含某一氨基酸序列,该序列:(a)与 SEQ ID NO:61、SEQ ID NO:62 或 SEQ ID NO:63 具有 60%或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%或更高);和/或(b)包含 SEQ ID NO:61、SEQ ID NO:62 或 SEQ ID NO:63 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS3 蛋白包括 SEQ ID NO:61、SEQ ID NO:62 或 SEQ ID NO:63 的变体。(b) 的优选片段包含来自 SEQ ID NO:61、SEQ ID NO:62 或 SEQ ID NO:63 的表位。其他优选片段缺少 SEQ ID NO:61、SEQ ID NO:62 或 SEQ ID NO:63C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和/或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:61、SEQ ID NO:62 或 SEQ ID NO:63 的至少一个表位。其它片段省去一个或多个蛋白质结构域。

[0298] 本发明包括使用与以上详述中讨论的 2603 株系中 GBS3 片段类似的 515、cjb111 和 coh1 株系中 GBS3 片段,如缺少 N-末端前导或信号序列区域;包括卷曲螺旋和富含脯氨酸的区段;包括信号序列区域和卷曲螺旋的区段;或包括信号序列区域、卷曲螺旋和富含脯氨酸的区段。

[0299] SAN1485

[0300] 原始 'SAN1485' 序列在参考文献 3 中注释为 '细胞壁表面锚定家族蛋白'(参见 GI:77408233)。出于参比目的,COH1 菌株中发现的全长 SAN1485 的氨基酸序列在本文中作为 SEQ ID NO:64。本发明所用的优选 SAN1485 多肽包含某一氨基酸序列,该序列:(a)与 SEQ ID NO:64 具有 60%或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%或更高);和/或(b)包含 SEQ ID NO:64 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 SAN1485 蛋白包括 SEQ ID NO:64 的变体。(b) 的优选片段包括 SEQ ID NO:64 的表位。其他优选片段缺少 SEQ ID NO:64C 末端的一个或

多个氨基酸（如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个）和 / 或 N 末端的一个或多个氨基酸（如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个），而保留 SEQ ID NO:64 的至少一个表位。其它片段省去一个或多个蛋白结构域。

[0301] GBS147

[0302] 原始 'GBS147' (SAG0416) 序列在参考文献 147 中标注为 '推定蛋白酶' (参见 GI:22533435)。出于参比目的,2603 菌株中发现的全长 GBS147 的氨基酸序列在本文中作为 SEQ ID NO:65。本发明所用的优选 GBS147 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:65 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和 / 或 (b) 包含 SEQ ID NO:65 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS147 蛋白包括 SEQ ID NO:65 的变体。(b) 的优选片段包括 SEQ ID NO:65 的表位。其他优选片段缺少 SEQ ID NO:65C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和 / 或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:65 的至少一个表位。

[0303] GBS328

[0304] 原始 'GBS328' (SAG1333) 序列在参考文献 147 中标注为 '5'-核苷酸酶家族蛋白' (参见 GI:22534359)。出于参比目的,2603 菌株中发现的全长 GBS328 的氨基酸序列在本文中作为 SEQ ID NO:66。本发明所用的优选 GBS328 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:66 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和 / 或 (b) 包含 SEQ ID NO:66 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS328 蛋白包括 SEQ ID NO:66 的变体。(b) 的优选片段包括 SEQ ID NO:66 的表位。其他优选片段缺少 SEQ ID NO:66C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和 / 或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:66 的至少一个表位。其它片段省去一个或多个蛋白结构域。

[0305] GBS84

[0306] 原始 'GBS84' (SAG0907) 序列在参考文献 147 中标注为 '推定脂蛋白' (参见 GI:22533929)。出于参比目的,2603 菌株中发现的全长 GBS84 的氨基酸序列在本文中作为 SEQ ID NO:67。本发明所用的优选 GBS84 多肽包含某一氨基酸序列,该序列:(a) 与 SEQ ID NO:67 具有 60% 或更高的相同性(如 60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或更高);和 / 或 (b) 包含 SEQ ID NO:67 的至少 'n' 个连续氨基酸的片段,其中 'n' 是 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200、250 或更高)。这些 GBS84 蛋白包括 SEQ ID NO:67 的变体。(b) 的优选片段包括 SEQ ID NO:67 的表位。其他优选片段缺少 SEQ ID NO:67C 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个)和 / 或 N 末端的一个或多个氨基酸(如 1、2、3、4、5、6、7、8、9、10、15、20、25 或更多个),而保留 SEQ ID NO:67 的至少一个表位。其它片段省去一个或多个蛋白结构域。

[0307] 与 GBS 糖的组合

[0308] GBS67 多肽片段可与通常偶联载体蛋白的一种或多种 GBS 荚膜糖组合。因此,本发明提供一种包含以下组合的免疫原性组合物:

[0309] (1) 上述讨论的 GBS67 多肽片段;和

[0310] (2) 一种或多种 GBS 荚膜糖。

[0311] 该组合的组分 (2) 所用的糖理想上作为含有糖部分和载体蛋白部分的偶联物存在。偶联物中的载体部分可为单一 GBS67 多肽片段、杂交 GBS67 多肽、非 GBS67 GBS 多肽或非 GBS 多肽。

[0312] 所述糖来自 GBS 的荚膜糖。糖可以是多糖,其大小是在从细菌纯化该糖期间形成,或者可以是这种多糖片段化产生的寡糖。

[0313] 组合物可包含来自以下一种或多种链球菌血清型的荚膜糖: Ia、Ib、Ia/c、II、III、IV、V、VI、VII 和 VIII。组合物可包括多种血清型如 2、3、4、5、6、7 或 8 种血清型。包括来自一种或多种血清型 Ia、Ib、II、III 和 V 的糖有用。这 5 种血清型的荚膜糖各包括:(a) 末端 N-乙酰基-神经氨酸 (NeuNAc) 残基(一般称为唾液酸),在所有情况下 2 → 3 连接至半乳糖残基;和 (b) 三糖核心内的 N-乙酰基-葡萄糖胺残基 (GlcNAc)。

[0314] 本发明所用糖可为天然形式,或也可被修饰。例如,糖可以比天然荚膜糖短,或可被化学修饰。例如,所述糖可去 O-乙酰化(部分或全部)、去 N-乙酰化(部分或全部)、N-丙酰化(部分或全部)等。可在偶联之前、期间或之后进行去乙酰化,但优选在偶联前进行。根据具体的糖,去乙酰化可影响或不影响免疫原性。参考文献 150 讨论了各种血清型中 GBS 糖上 O-乙酰化的关联,且在一些实施方式中,位置 7、8 和 / 或 9 的唾液酸残基的 O-乙酰化在偶联前、中和后都保留,例如通过保护 / 去保护、通过再乙酰化等。然而,本发明所用的 GBS 糖通常基本不发生位置 7、8 和 / 或 9 的唾液酸残基的 O-乙酰化。可通过常规试验评价去乙酰化等的效果。另一可能修饰是从糖中移除唾液酸残基如侧链末端的唾液酸 [151]。具体地,血清型 V 荚膜糖用于本发明时,其可如参考文献 [151] 所述通过去唾液酸苷化修饰。脱唾液酸化的 GBS 血清型 V 荚膜糖可通过以下方式制备:在温和酸性条件下(例如 0.1M 硫酸,80°C 持续 60 分钟)处理纯化的 GBS 血清型 V 荚膜糖或者用神经氨酸酶处理,如参考文献 151 所述。在另一实例中,可解聚全长多糖以提供用于本发明的较短片段,例如,通过在温和酸中水解、通过加热、通过尺寸选择色谱等。已报道链长影响 GBS 糖在兔子中的免疫原性 [152]。具体地,血清型 II 和 / 或 III 荚膜糖用于本发明时,其可如参考文献 153 所述解聚。该文献描述了通过温和去氨基切割为含末端还原的 2,5-脱水-D-甘露糖残基的抗原片段来部分解聚 II 型和 III 型荚膜糖。

[0315] 可通过已知技术纯化荚膜糖,如本文引用的参考文献所述,如参考文献 154。典型的方法包括碱提取、离心、过滤、RNA 酶 / DNA 酶处理、蛋白酶处理、浓缩、尺寸排阻色谱、超滤、阴离子交换色谱和进一步超滤。作为替代,可使用参考文献 155 所述的纯化过程。该过程涉及碱提取、乙醇 / CaCl₂ 处理、CTAB 沉淀和再溶。

[0316] 本发明不限于纯化自天然来源的糖,然而,可通过其它方法如全合成或部分合成获得所述糖。

[0317] 糖通常可与载体蛋白偶联。通常,与载体的共价偶联增强糖的免疫原性,因为偶联可以将糖由 T-非依赖性抗原转变为 T-依赖性抗原,由此能够引发免疫记忆。

[0318] 已广泛报道 GBS 糖的偶联,例如参见参考文献 156-163。GBS 糖偶联的典型现有技术方法涉及将纯化的糖还原性胺化到载体蛋白,例如破伤风类毒素 (TT) 或 CRM197 上 [157]。还原胺化涉及载体氨基酸侧链上的胺基与糖上的醛基。由于 GBS 荚膜糖在其天然形式中不包括醛基,通常是在偶联之前通过糖唾液酸残基的一部分发生氧化(如高碘酸盐氧化)产生醛基 [157, 164]。已显示此方法制备的偶联疫苗就各 GBS 血清型 Ia、Ib、II、III 和 V 而言在人体中安全且具有免疫原性 [165]。

[0319] 优选的载体蛋白是细菌毒素,如白喉或破伤风毒素、或其类毒素或突变体。这些常用于偶联疫苗。偶联物中的载体蛋白可以是或不是 (1) 中 GBS59 抗原之一。若不是 GBS59 抗原,其可为不同的 GBS 抗原。在一些实施方式中,尽管载体不是 GBS 抗原,其可以是例如细菌毒素或类毒素。

[0320] 常用载体蛋白是白喉或破伤风类毒素,或其突变体。还可使用毒素或类毒素的片段,例如破伤风类毒素的片段 C [166]。白喉毒素的 CRM197 突变体 [167-169] 对本发明特别有用。其他合适的载体蛋白包括,脑膜炎奈瑟球菌 (*N. meningitidis*) 外膜蛋白复合物 [170]、合成肽 [171、172]、热激蛋白 [173, 174]、百日咳蛋白 [175, 176]、细胞因子 [177]、淋巴因子 [187]、激素 [187]、生长因子、含有各种病原体衍生抗原的多种人 CD4⁺T 细胞表位的人造蛋白 [178] 如 N19 [179]、流感嗜血杆菌 (*H. influenzae*) 的 D 蛋白 [180-182]、铁摄取蛋白 [183]、艰难梭菌 (*C. difficile*) 的毒素 A 或 B [184]、重组金黄色葡萄球菌 (*P. aeruginosa*) 胞外蛋白 A (rEPA) [185] 等等。

[0321] 当组合物包含一种以上偶联物时,每种偶联物可使用相同的载体蛋白或不同的载体蛋白。

[0322] 在一些实施方式中,一种偶联物可携带来自多种血清型的糖 [186]。然而,每种偶联物通常包含来自一种血清型的糖。

[0323] 偶联物可包含过量载体 (w/w) 或过量糖 (w/w)。在一些实施方式中,偶联物可包含相同重量的载体和糖。例如,可使用糖:蛋白比 (w/w) 为 1:5-5:1 的偶联物,特定是 1:5-2:1 的比例。

[0324] 载体分子可直接与载体共价偶联,或通过接头偶联。可通过(例如)糖和载体之间的还原性胺化(如参考文献 187 和 188 所述),实现与蛋白质的直接连接。首先需要通过,例如氧化来激活糖。可用任何已知方法,如参考文献 189 和 190 所述的过程通过接头基团进行连接。优选的连接类型是己二酸接头,这种连接可通过以下方式形成:将游离 -NH₂ 基团与己二酸偶联(如通过胺化引入葡聚糖中)(例如,利用二酰亚胺活化),然后将蛋白质偶联于所得的糖-己二酸中间体 [191, 192]。另一种优选连接类型是羰基接头,这种连接可通过以下方式形成:使糖 CDI 的游离羟基发生反应 [193, 194],然后与蛋白质反应形成氨基甲酸酯连接。其它接头包括 β-丙酰胺基 [195]、硝基苯基-乙基胺 [196]、卤代酰基卤化物 [197]、糖苷键 [198]、6-氨基己酸 [199]、ADH [200]、C4-C12 部分 [201] 等。也采用碳二亚胺缩合反应 [202]。

[0325] 与非 GBS 抗原组合

[0326] GBS67 片段可与非 GBS 抗原联用。因此,本发明提供一种包含以下组合的免疫原性组合物:

[0327] (1) 上述讨论的 GBS67 多肽片段;和

[0328] (2) 一种或多种选自下组的抗原：白喉类毒素；破伤风类毒素；一种或多种百日咳抗原；乙型肝炎病毒表面抗原；灭活的脊髓灰质炎病毒抗原；C 群脑膜炎奈瑟球菌 (*Neisseria meningitidis*) 荚膜糖抗原的偶联物；Y 群脑膜炎奈瑟球菌荚膜糖抗原的偶联物；W135 群脑膜炎奈瑟球菌荚膜糖抗原的偶联物；A 群脑膜炎奈瑟球菌荚膜糖抗原的偶联物；一种或多种流感抗原；和一种或多种人乳头状瘤病毒抗原。

[0329] 可通过处理（例如用甲醛）来自白喉棒状杆菌 (*Corynebacterium diphtheriae*) 的白喉毒素获得白喉类毒素。例如，参考文献 203 的第 13 章更详细地公开了白喉类毒素。

[0330] 可通过处理（例如用甲醛）来自破伤风梭菌 (*Clostridium tetani*) 的破伤风毒素获得破伤风类毒素。参考文献 203 的第 27 章更详细地公开了破伤风类毒素。

[0331] 疫苗中的百日咳抗原是细胞(全细胞, Pw) 或无细胞(Pa)。本发明可用于分选百日咳抗原。已经详尽记载了细胞百日咳抗原的制备（参见例如参考文献 203 的第 21 章），例如，可通过加热灭活百日咳博德特菌 (*B. pertussis*) 的 I 期培养物获得该抗原。非细胞百日咳抗原包含特定的纯化百日咳博德特菌 (*B. pertussis*) 抗原，它们或是由天然细菌纯化，或是在重组宿主中表达后纯化。通常使用一种以上非细胞抗原，因此组合物可包含以下熟知和已良好鉴定的百日咳博德特菌抗原中的一种、两种或三种：(1) 脱毒的百日咳毒素（百日咳类毒素或“PT”）；(2) 丝状血细胞凝集素（“FHA”）；(3) 百日咳杆菌粘附素（也称为“69 千道尔顿外膜蛋白”）。在根据本发明使用之前，可用甲醛处理 FHA 和百日咳杆菌粘附素。可通过甲醛和 / 或戊二醛处理对 PT 脱毒，但是，作为此种化学脱毒方法的替代方案，它可以是酶活性经诱变降低的突变 PT [204]。可以使用的其他非细胞百日咳抗原包括菌毛（如凝集原 2 和 3）。

[0332] 乙型肝炎病毒表面抗原 (HBsAg) 是乙型肝炎病毒衣壳的主要组分。可通过在酵母如酿酒酵母 (*Saccharomyces cerevisiae*) 中重组表达，方便地产生这种抗原。

[0333] 灭活的脊髓灰质炎病毒 (IPV) 抗原由细胞培养物上生长的病毒制备并随后灭活（例如用甲醛）。因为脊髓灰质炎可由三类脊髓灰质炎病毒之一引起（如参考文献 203 的第 24 章所述），所以组合物可包含以下三类脊髓灰质炎病毒抗原：1 型脊髓灰质炎病毒（例如 Mahoney 毒株）、2 型脊髓灰质炎病毒（例如 MEF-1 毒株）和 3 型脊髓灰质炎病毒（例如 Saukett 毒株）。

[0334] 组合物的组分 (2) 中包含白喉类毒素、破伤风类毒素或非细胞百日咳抗原之一时，它通常包含所有这三种抗原，即组分 (2) 包含 D-T-Pa 组合。

[0335] 组合物的组分 (2) 中包含白喉类毒素、破伤风类毒素或细胞百日咳抗原之一时，它通常包含所有这三种抗原，即组分 (2) 包含 D-T-Pw 组合。

[0336] 人乳头瘤病毒抗原包括 L1 衣壳蛋白，其可装配形成称作病毒样颗粒 (VLP) 的结构。可通过在酵母细胞（例如酿酒酵母 (*S. cerevisiae*)) 或昆虫细胞（例如夜蛾 (*Spodoptera*) 细胞，如草地贪夜蛾 (*S. frugiperda*), 或果蝇 (*Drosophila*) 细胞) 中重组表达 L1 产生 VLP。在酵母细胞中，质粒载体可携带 L1 基因；在昆虫细胞中，杆状病毒载体可携带 L1 基因。更优选地，该组合物包含来自 HPV-16 和 HPV-18 毒株的 L1 VLP。已证明这种二价组合非常有效 [205]。除了 HPV-16 和 HPV-18 毒株外，也可能包含来自 HPV-6 和 HPV-11 毒株的 L1 VLP，产生四价组合。

[0337] 流感抗原可为当前流感病毒疫苗的形式。目前可以获得各种形式的流感病毒疫苗

(例如参见参考文献 [203] 的第 17 和 18 章)。疫苗通常基于活病毒、灭活病毒、重组血细胞凝集素或病毒体。灭活疫苗可基于全病毒颗粒、裂解病毒颗粒或基于纯化的表面抗原。本发明疫苗中的抗原可以采取活病毒的形式,或者更优选地,采取灭活病毒的形式。疫苗可为例如,三价疫苗(如包括来自 A/H1N1 菌株、A/H3N2 菌株和 B 菌株的血细胞凝集素)。在其他实施方式中,所述疫苗是单价疫苗(如包括来自 A/H1N1 菌株或 A/H5N1 菌株的血细胞凝集素)。所述疫苗可为含佐剂(如水包油乳液)或不含佐剂。

[0338] 人乳头瘤抗原为从重组 HPV 外壳蛋白组装的中空病毒样颗粒 (VLP),通常来自 HPV 16 型和 18 型,且任选地还来自 HPV 6 型和 11 型。

[0339] 抗体

[0340] GBS 抗原的抗体可用于被动免疫 [206]。因此本发明提供同时、分别或依次给予的抗体组合,其中所述组合包括以下至少两种:(a) 识别上述定义的第一氨基酸序列的抗体;(b) 识别上述定义的第二氨基酸序列的抗体;和/或(c) 识别上述定义的第三氨基酸序列的抗体;

[0341] 本发明还提供此类抗体组合在治疗中的应用。本发明还提供此类抗体组合在药物制造中的应用。本发明还提供一种治疗哺乳动物的方法,其包括给予哺乳动物有效量的该组合。如上就免疫原性组合物所述,这些方法和应用能够保护哺乳动物抵御 GBS 感染。

[0342] 术语“抗体”包括完整的免疫球蛋白分子及其能结合抗原的片段。这些包括杂交(嵌合)抗体分子 [207, 208];F(ab')₂ 和 F(ab) 片段和 Fv 分子;非共价异源二聚体 [209, 210];单链 Fv 分子 (sFv) [211];二聚和三聚抗体片段构建物;小抗体 [212, 213];人源化抗体分子 [214-216];和任何获自此类分子的功能性片段,以及通过非常规方法如噬菌体展示获得的抗体。优选所述抗体为单克隆抗体。获取单克隆抗体的方法为本领域熟知。优选人源化或完全人抗体。

[0343] 概述

[0344] 除非另有说明,本发明的实施将采用化学、生物化学、分子生物学、免疫学和药理学的常规方法,这些方法在本领域技术范围内。这些技术在文献中已有充分描述。参见例如参考文献 217-224 等。

[0345] 上文使用“GI”编号。GI 编号,或者“基因信息识别号”(GenInfo Identifier)是 NCBI 将序列加入其数据库时,连续对每一序列记录指定的一串数字。GI 编号与序列记录登录号没有相似之处。序列更新(如纠正或加入更多注释或信息)后,将接受新 GI 编号。因此与给定 GI 编号相关的序列是不变的。

[0346] 当本发明涉及“表位”时,该表位可以是 B 细胞表位和/或 T 细胞表位。可凭经验鉴定此类表位(如使用 PEPSCAN[225, 226] 或相似方法),或可对其进行预测(如使用詹姆森-沃尔夫抗原性指数 (Jameson-Wolf antigenic index) [227]、基于矩阵的方法 [228]、MAPITOPE [229]、TEPITOPE [230, 231]、神经网络 [232]、OptiMer 和 EpiMer [233, 234]、ADEPT [235]、T 位点 [236]、亲水性 [237]、抗原性指数 [238] 或参考文献 239-243 中公开的方法等)。表位是由抗体或 T 细胞受体的抗原结合位点识别并结合的抗原的某部分,它们也称作“抗原决定簇”。

[0347] 术语“包含”涵盖“包括”以及“由……组成”,例如,“包含”X 的组合物可以仅由 X 组成或可以包括其它物质,例如 X+Y。

[0348] 术语“基本”不排除“完全”，如“基本不含”Y 的组合物可能完全不含 Y。必要时，可以从本发明定义中删去术语“基本”。

[0349] 与数值 x 相关的术语“约”是可选的并表示例如 $x \pm 10\%$ 。

[0350] 除非另有说明，包括混合两种或多种组分的步骤的过程不要求任何特定的混合顺序。因此，组分可以任何顺序混合。有三种组分时，可将两种组分相互合并，然后可将合并物再与第三种组分混合等。

[0351] 抗体通常对于其靶标是特异性的。因此，其对靶标的亲和性高于无关对照蛋白，例如牛血清白蛋白。

[0352] 两个氨基酸序列间的序列相同性百分数表示进行比对时所比较的两条序列中相同氨基酸的百分数。利用本领域已知软件程序，例如参考文献 224 的 7.7.18 部分所描述的软件程序，可进行比对并确定同源性百分数或序列相同性百分数。优选的比对通过史密斯-沃特曼(Smith-Waterman)同源性搜索算法使用仿射缺口搜索确定，其中缺口开放罚 12 分，缺口延伸罚 2 分，BLOSUM 矩阵计 62 分。参考文献 245 中公开了史密斯-沃特曼同源性搜索算法。

[0353] 附图简要说明

[0354] 图 1:2603(SAG1408) 和 H37B(SAI1512) 的 GBS67 比对，显示片段 1(Fr1)、片段 2(Fr2) 和片段 3(Fr3) 的位置。

[0355] 图 2:2603(图 2A) 和 H36B(图 2B) 的片段 1(Fr1)、片段 2(Fr2) 和片段 3(Fr3) 的纯化。

[0356] 图 3: 针对 GBS672603(图 3A) 和 GBS67H36B(图 3B) 产生的多克隆抗体在 Western 印迹分析中识别两种变体的片段 3(Fr 3)，但不是片段 2(Fr 2) 或片段 1(Fr 1)。

[0357] 图 4:Western 印迹分析中，针对 2603 中片段 1 的抗体仅识别 2603 变体的重组片段 1 (1 2603) 和 2603 变体的全长 GBS67 (FL 2603)。H36B 的全长 GBS67 (FL H36B)、H36B 的片段 1、2 和 3 以及 2603 的片段 2 和 3 没有被识别。

[0358] 图 5A:Western 印迹分析中，针对 2603 中片段 2 的抗体识别 2603 变体的重组片段 2 (2 2603) 和 H36B 变体的重组片段 2 (2 2603)，以及 2603 变体的全长 GBS67 (FL 2603)。

[0359] 图 5B:Western 印迹分析中，针对 2603 中片段 3 的抗体识别 2603 变体的重组片段 3 (3 2603) 和 H36B 变体的重组片段 3 (3 2603)，以及 2603 变体的全长 GBS67 (FL 2603) 和 H36B 的全长 GBS67 (FL H36B)。

具体实施方式

[0360] GBS67 变体为交叉保护

[0361] 已鉴定 GBS67(AP1-2a) 的两种等位基因变体，一种在 GBS 株系 2603 中且一种在 GBS 株系 H36B 中。GBS 株系 2603 中鉴定的 GBS67 株系具有优势，是存在于 87% 的 GBS 株系中的变体。

[0362] 这两种 GBS67 变体中任一种都能赋予针对表达另一 GBS67 变体的 GBS 株系的交叉保护。例如，如下表 1 所述，用 2603 株系的 GBS67 (AP1-2a) 免疫的雌鼠幼仔得到保护免受表达 GBS67 2603 或 H36B 变体的 GBS 株系的攻击。

[0363] 表 1: 在 GBS 小鼠母体免疫 / 幼仔攻击模型中 GBS67 赋予交叉保护

[0364]

抗原	GBS 株系 (血清型)	等位基因变体	保护 %	统计显著性 p 值
AP-2a 2603 变体	CJB111 (V)	CJB111	69.6	<0.0001
	515 (Ia)	515	61.9	0.0018
	3050 (II)	2603	94.4	<0.0001
	5401 (II)	H36B	62.8	<0.0001
AP1-2a H36B 变体	515 (Ia)	515	57.4	<0.0001
	5401 (II)	H36B	58.7	<0.0001
	DK21 (II)	H36B	60.2	<0.0001

[0365]

[0366] 进行研究鉴定负责交叉保护的 GBS67 变体部分。

[0367] GBS67 三种片段的鉴定

[0368] AP1-2a(GBS67) 没有可得的晶体结构。2603 和 H36B 的 GBS67 变体二级结构的计算机分析鉴定到三种可能负责 GBS67 交叉保护活性的推定保守片段 (见表 2)。

[0369] 表 2 :GBS67 变体的保守片段

[0370]

片段	氨基酸残基	SEQIDNO	氨基酸残基数	理论 MW (kDa)
片段 12603	24-217	3	194	23.5
片段 22603	218-615	4	398	47.3
片段 32603	616-866	5	251	30.4
片段 1H36B	24-217	6	194	24
片段 2H36B	218-610	7	393	46.8
片段 3H36B	611-861	8	251	30.5

[0371] 图 1 显示表明 3 种片段位置的 GBS67 变体比对。

[0372] 这 6 种片段进行克隆并表达为 HK100 和 BL21 (DE3) 株系中的 His 标记蛋白。所有片段都过量表达且可溶。下表 3 显示获得的片段产量,图 2 中显示凝胶上分离的纯化片段。

[0373] 表 3a:GBS67 2603 片段的纯化

[0374]

片段	mg/ml	ml	总 mg
GBS67-26031His	2.600	6.0	15.602
GBS67-26032His	3.084	6.0	18.501
GBS67-26033His	3.099	5.0	15.495

[0375] 表 3a:GBS67 H36B 片段的纯化

[0376]

片段	mg/ml	ml	总 mg
GBS67-H36B 1 His	3.157	5.5	17.364
GBS67-H36B 2 His	5.153	3.0	15.459
GBS67-H36B 3 His	9.470	3.0	28.410

[0377] GBS67 片段的交叉保护活性评估

[0378] 针对 2603 和 H36B GBS67 变体产生的多克隆抗体能在 Western 印迹分析中识别来自两种变体的片段 3 (图 3),表明片段 3 含有负责诱导交叉保护的表位。

[0379] 后续 Western 印迹分析中,针对 2603GBS67 变体的片段 1 产生的抗体仅识别 2603 GBS67 的重组片段 1 和 2603 的全长 GBS67(图 4)。这些抗体不识别 H36B GBS67 变体的片段 1 或 H36B 的全长 GBS67。相反,针对 2603 GBS67 的片段 2 产生的抗体识别 H36B GBS67 的片段 2 和全长 H36B GBS67 (图 5A) 且针对 2603 GBS67 的片段 3 产生的抗体识别 H36B GBS67 的片段 3 和全长 H36B (图 5B)。

[0380] FAC 分析证明 GBS67 的片段 2 和 3 在 GBS 细菌的表面高度暴露 (见下表 4)。

[0381] 表 4: 片段 2 和 3 的表面暴露

[0382]

组	株系 / 血清型	FACS 暴露	株系 / 血清型	FACS 暴露
片段 12603	515(Ia)	-	5401(II)	-
片段 22603	515(Ia)	++	5401(II)	+++
片段 32603	515(Ia)	++	5401(II)	+++
GBS672603	515(Ia)	++	5401(II)	+++
GBS67H36B	515(Ia)	++	5401(II)	+++

[0383] 然后片段 1、2 和 3 诱导交叉保护的能力在母体免疫模型中体内测试。用 GBS67 2603 的片段 1、2 或 3,2603 或 H36B 的全长 GBS67,或 PBS 免疫雌性小鼠。然后用表达 GBS67H36B 变体的 5401GBS 株系攻击幼仔。结果示于下表 5。

[0384] 表 5:母体免疫模型结果

[0385]

组	抗原	mcg/剂	死亡 / 处理	存活率%
1	片段 1(2603)	20	54/60	10
2	片段 2(2603)	20	37/60	38

3	片段 3(2603)	20	22/45	51
4	GBS67(2603)	20	25/54	54
5	GBS67(H36B)	20	21/46	54
6	PBS	0	53/56	5

[0386] 这些结果显示 GBS67 2603 变体的片段 3 能和全长 GBS67 2603 或全长 GBS67H36B 一样赋予针对表达 GBS67H36B 变体的 5401 GBS 株系的相同交叉保护。

[0387] GBS67 2603 的片段 1、2 和 3 诱导针对表达 GBS67 2603 变体的 515 GBS 株系攻击的保护通过重复上述实验加以验证,除了所述幼仔是用 515 GBS 株系攻击,而不是 5401 GBS 株系。结果见示于下表 6:

[0388] 表 6:母体免疫模型结果

[0389]

组	抗原	死亡 / 处理	存活率 %
1	片段 1(2603)	34/39	12
2	片段 2(2603)	31/64	52
3	片段 3(2603)	28/40	30
4	GBS67(2603)	16/58	72
6	PBS	51/57	10

[0390] 在另一实验中,用 GBS67H36B 的片段 1、2 或 3, H36B 的全长 GBS67,或 PBS 免疫雌性小鼠。用表达 GBS67H36B 变体的 5401 GBS 株系攻击幼仔。结果示于下表 7。

[0391] 表 7:母体免疫模型结果

[0392]

组	抗原	死亡 / 处理	存活率 %
1	片段 1(H36B)	28/40	30
2	片段 2(H36B)	16/50	68
3	片段 3(H36B)	24/70	66
4	GBS67(H36B)	21/48	56
5	PBS	43/69	38

[0393] 因此,GBS67 2603 的片段 2 和 3 以及这些片段内的表位可替代全长 GBS672603 或全长 GBS67H36B 而用于免疫原性组合物。相似地,GBS67H36B 的片段 2 和 3 以及这些片段

内的表位可替代全长 GBS67 2603 或全长 GBS67H36B 而用于免疫原性组合物。

[0394] 材料和方法

[0395] 生物信息学

[0396] 无乳链球菌株系 2603V/R(V) 和 H36B(Ib) 的完整基因组序列可用登录号 AE009948 和 AAJS000000000 获得。逐对序列比对获自 ClustalW 算法。

[0397] 为了鉴定推定的结构,我们使用连接 NCBI-BLAST 数据库的 Pfam 程序。用 PsiPred (蛋白结构预测服务器;UCL 生物信息组) 软件进行二级结构预测。

[0398] 菌株和生长条件

[0399] 本研究中使用的 GBS 株系为 2603V/R(血清型 V)、515(Ia)、H36B(血清型 Ib) 和 5401(II)。细菌生长于 37 ° C 的托-休二氏培养基(THB; 迪菲克实验室(Difco Laboratories)) 或补充有 5% 绵羊血的胰蛋白酶大豆琼脂中。

[0400] 重组蛋白和抗血清的克隆、表达、纯化

[0401] GBS 菌株 2603 和 H36B 用作克隆序列的 DNA 来源,所述序列编码 GBS672603 和 H36B 等位基因变体的单一片段(片段 1、2 和 3)。通过革兰氏阳性菌标准方案用 NucleoSpin 组织试剂盒(MN 公司(Machery-Nagel))按照生产商说明分离基因组 DNA。通过在大肠杆菌 HK100 株系(246)中用 PIPE 克隆方法将对应于各结构域的基因克隆到 SpeedET 或 pET15-TEV 载体(N-末端 6xHIS 标签)中。使用的寡聚物列于表 6。检查 pET15-TEV 中的所得构建体以测序并随后转化到大肠杆菌 BL21(DE3) 中(诺瓦基公司(Novagen))。为了表达重组蛋白,对于 pET 克隆,培养物用 1mM IPTG 诱导后在 25 ° C 维持 5 小时,或对于 SpeedET 克隆用 0.2% 阿拉伯糖。所有重组蛋白用亲和色谱纯化。简单的说,通过离心收获细胞并在“裂解缓冲液”中裂解,所述缓冲液含溶于 PBS 的 10mM 咪唑、1mg/ml 溶菌酶、0.5mg/ml DNA 酶和 COMPLETE 抑制剂混合物(罗氏公司(Roche))。裂解物离心澄清并应用在含 10mM 咪唑的 PBS 中预平衡的 His 捕获 HP 柱上(阿莫山姆生物科学公司(Armesham Biosciences))。用含 250mM 咪唑的相同缓冲液进行蛋白洗脱,两次清洗步骤后用 20mM 和 50mM 咪唑缓冲液。纯组分的蛋白浓度用 BCA 试验(皮尔斯公司(PIERCE)) 评估。

[0402] 通过用前述纯的重组蛋白免疫 CD1 小鼠生产对各蛋白特异的抗血清 [247]。由 ELISA 监控所收集血清中的蛋白特异性免疫应答(总 Ig)。

[0403] 如前述报道 [2, 247] 生产对应于 2603 和 H36B 等位基因变体(分别为 TIGR 注释 SAG_1408 和 SAI_1512) 的全长重组 GBS67 蛋白。

[0404] 免疫印迹

[0405] 通过 4-12%NuPage Novex 预制凝胶(英杰公司(Invitrogen)) 分离 10ng 的各纯化蛋白并用 iBlot™ 干印迹系统(英杰公司) 将其电转到硝基纤维素膜上。在含 0.05% 吐温 20 和 10% 脱脂奶的 1X 磷酸盐缓冲盐水(PBS:140mM NaCl, 2.7mM KCl, 10mM Na2HPO4 和 1.8mM KH2PO4, pH 7.3) 中室温封闭 1 小时后,膜和 1:500 稀释的一抗室温(RT) 孵育 1 小时。用含 0.05% 吐温 20 的 PBS(PBST) 清洗三次后,所述膜用辣根过氧化物酶偶联的二抗(大科公司(Dako)) 孵育 1 小时。用 Opti-4CN 底物试剂盒(伯乐公司(Bio-Rad)) 观察阳性条带。

[0406] ELISA

[0407] 抗原特异型抗体应答用 ELISA 检测,每孔使用 100ng 的纯化重组抗原。通过比较

测试血清样品和参考血清样品的响应曲线,用参考线计算程序计算 IgG 抗体效价。所述参考血清样品是获自用纯化重组抗原免疫的小鼠的血清样品库,其中 150,000EU/mL 指定为任意效价。

[0408] FACS

[0409] 针对纯化的重组蛋白产生的小鼠血清用流式细胞仪在全细菌上分析以评价单一结构域的表面暴露。存在 0.08%(wt/vol) 多聚甲醛时固定指数期细菌细胞并 37° C 孵育 1 小时。然后固定的细菌用 PBS 清洗一次,在新生牛血清(西格玛公司(Sigma))中重悬并 25° C 孵育 20 分钟。所述细胞然后在预免疫或免疫血清中 4° C 孵育 1 小时,在稀释缓冲液(PBS, 20%新生牛血清, 0.1%BSA)中 1:200 稀释。细胞在 PBS-01%BSA 中清洗并用偶联 F(ab)₂ 山羊抗小鼠 IgG 的 R-藻红蛋白(杰克逊免疫研究实验室公司(Jackson ImmunoResearch Laboratories;Inc.))的 1:100 稀释液再 4° C 孵育 1 小时。清洗后,细胞在 PBS 中重悬并用 FACS Calibur 设备(BD 公司(Becton Dickinson),新泽西州富兰克林湖市)分析,使用 FlowJo 软件(树星公司(Tree Star),俄勒冈州艾士兰)。数据表示为用免疫血清与预免疫血清染色的细胞间的荧光差异。

[0410] 小鼠活性母体免疫模型

[0411] GBS 感染的母体免疫 / 新生幼仔攻击模型用于证实前述小鼠中产生的蛋白的保护效力 [247]。简单地说,CD-1 雌鼠(6-8 周龄)在第 1 天(CFA 中)、21 天和 35 天(IFA)用 PBS 或 20mg 重组蛋白免疫,最后一次免疫后饲养 3 天。出生 48 小时内,幼仔腹膜内注射一定剂量的不同 GBS 株系,该剂量计算产生 90% 致死率。攻击后监控 2 天幼仔的存活。用菲希尔精确检验进行统计分析。所有动物研究按照优秀健康研究所(Istituto Superiore di Sanità)(意大利)的指南进行。

[0412] 表 6 :用于克隆 GBS67 片段的引物

[0413]

引物	序列(5'-3')	扩增的基因
Fr 1-2603 正向	CTGTACTTCCAGGGCAATACCAATGTT TTAGGGGAA (SEQ ID NO:68)	编码 GBS67-2603 变体的 aa24-217 的片段
Fr 1-2603 反向	AATTAAGTCGCGTTATTTTCCACTGAC AGTAACTC (SEQ ID NO:69)	
Fr 2-2603 正向	CTGTACTTCCAGGGCACCATAGTAAAA CCAGTGGAC (SEQ ID NO:70)	编码 GBS67-2603 变体的 aa218-615aa 的片段
Fr 2-2603 反向	AATTAAGTCGCGTTATCCATTACCAAG CTGTAAATT (SEQ ID NO:71)	
Fr 3-2603 正向	CTGTACTTCCAGGGCCAAACATTACAG CCAAGTGAT (SEQ ID NO:72)	编码 GBS67-2603 变体的 aa616-866 片段
Fr 3-2603 反向	AATTAAGTCGCGTTATCCTTTCCACCT GTCATAGG (SEQ ID NO:73)	
Fr 1-H36B 正向	CTGTACTTCCAGGGCAATACCAATGTT TTAGGGGAA (SEQ ID NO:74)	编码 GBS67-H36B 变体的 aa24-217 片段
Fr 1-H36B 反向	AATTAAGTCGCGTTATTTACCGCTAAC AGTAACTC (SEQ ID NO:75)	
Fr 2-H36B 正向	CTGTACTTCCAGGGCTCCATAATAAAA ACTATAAATAAAG (SEQ ID NO:76)	编码 GBS67-H36B 变体的 aa218-610 片段
Fr 2-H36B 反向	AATTAAGTCGCGTTATCCGTTGCCAAG ATGTAAATT (SEQ ID NO:77)	
Fr 3-H36B 正向	CTGTACTTCCAGGGCCAAACATTGCAA CCAAGTGAT (SEQ ID NO:78)	编码 GBS67-H36B 变体的 aa611-861aa 片段
Fr 3-H36B 反向	AATTAAGTCGCGTTATCCTTTCCACCT GTCATC (SEQ ID NO:79)	

[0414] 参考文献

- [0415] [1]Rosini 等 Molecular Microbiology, 2006, 61 (1) :126-141
- [0416] [2]Margarit 等 Journal of Infectious Diseases, 2009, 199:108-115
- [0417] [3]Tettelin 等 Proc. Natl. Acad. Sci. U. S. A. 102 (39), 13950-13955 (2005)
- [0418] [4]Needleman 和 Wunsch (1970) J. Mol. Biol. 48, 443-453.
- [0419] [5]Rice 等. (2000) Trends Genet 16:276-277.
- [0420] [6]Bodanszky (1993) Principles of Peptide Synthesis (《肽合成原理》)

(ISBN:0387564314).

[0421] [7]Fields等.(1997)Meth Enzymol 289:Solid-Phase Peptide Synthesis.(《酶学方法 289: 固相肽合成》) ISBN:0121821900.

[0422] [8]Chan 和 White(2000)Fmoc Solid Phase Peptide Synthesis Fmoc (《固相肽合成》). ISBN:0199637245.

[0423] [9]Kullmann(1987)Enzymatic Peptide Synthesis (《酶 多 肽 合 成》). ISBN:0849368413.

[0424] [10]Ibba(1996)Biotechnol Genet Eng Rev 13:197-216.

[0425] [11]US 专利 5, 707, 829

[0426] [12]Vaccine Design(《疫苗设计》)(1995)Powell 和 Newman 编, ISBN:030644867X. 普莱努出版社(Plenum)

[0427] [13]W090/14837.

[0428] [14]W090/14837.

[0429] [15]Podda 和 Del Giudice(2003)Expert Rev Vaccines 2:197-203.

[0430] [16]Podda(2001)Vaccine 19:2673-2680.

[0431] [17]Vaccine Design:The Subunit and Adjuvant Approach(《疫苗设计: 亚基和佐剂方法》)(Powell 和 Newman 编)Plenum Press(普莱努出版社)1995(ISBN0-306-44867-X)。

[0432] [18]Vaccine Adjuvants:Preparation Methods and Research Protocols(《疫苗佐剂:制备方法 and 研究方案》)(Methods in Molecular Medicine series(《分子医学方法》)丛书的第 42 卷)。ISBN:1-59259-083-7. O' Hagan 编

[0433] [19]Allison 和 Byars(1992)Res Immunol 143:519-25.

[0434] [20]Hariharan 等.(1995)Cancer Res 55:3486-9.

[0435] [21]US 2007/014805.

[0436] [22]W095/11700.

[0437] [23]US 专利 6, 080, 725.

[0438] [24]W02006/113373.

[0439] [25]W02005/097181.

[0440] [26]US 5, 057, 540.

[0441] [27]W096/33739.

[0442] [28]EP A 0109942.

[0443] [29]W096/11711.

[0444] [30]W000/07621.

[0445] [31]Barr 等.(1998)Advanced Drug Delivery Reviews 32:247271.

[0446] [32]Sjolanderet 等.(1998)Advanced Drug Delivery Reviews 32:321338.

[0447] [33]Niikura 等.(2002)Virology 293:273-280.

[0448] [34]Lenz 等.(2001)J Immunol 166:5346-5355.

[0449] [35]Pinto 等.(2003)J Infect Dis 188:327-338.

[0450] [36]Gerber 等.(2001)J Virol 75:4752-4760.

- [0451] [37]W003/024480.
- [0452] [38]W003/024481.
- [0453] [39]Gluck 等. (2002)Vaccine 20:B10 B16.
- [0454] [40]EP A 0689454.
- [0455] [41]Johnson 等. (1999)Bioorg Med Chem Lett 9:2273-2278.
- [0456] [42]Evans 等. (2003)Expert Rev Vaccines 2:219-229.
- [0457] [43]Meraldi 等. (2003)Vaccine 21:2485-2491.
- [0458] [44]Pajak 等. (2003)Vaccine 21:836-842.
- [0459] [45]Kandimalla 等. (2003)Nucleic Acids Research 31:2393-2400.
- [0460] [46]W002/26757.
- [0461] [47]W099/62923.
- [0462] [48]Krieg(2003)Nature Medicine 9:831-835.
- [0463] [49]McCluskie 等. (2002)FEMS Immunology and Medical Microbiology32:179-185.
- [0464] [50]W098/40100.
- [0465] [51]US 6, 207, 646.
- [0466] [52]US 6, 239, 116.
- [0467] [53]US 6, 429, 199.
- [0468] [54]Kandimalla 等. (2003)Biochemical Society Transactions 31(第3部分):654-658.
- [0469] [55]Blackwell 等. (2003)J Immunol 170:4061-4068.
- [0470] [56]Krieg(2002)Trends Immunol 23:64-65.
- [0471] [57]W001/95935.
- [0472] [58]Kandimalla 等. (2003)BBRC 306:948-953.
- [0473] [59]Bhagat 等. (2003)BBRC 300:853-861.
- [0474] [60]W003/035836.
- [0475] [61]Schellack 等. (2006)Vaccine 24:5461-72.
- [0476] [62]W095/17211.
- [0477] [63]W098/42375.
- [0478] [64]Beignon 等. (2002)Infect Immun 70:3012-3019.
- [0479] [65]Pizza 等. (2001)Vaccine 19:2534-2541.
- [0480] [66]Pizza 等. (2000)Int J Med Microbiol 290:455-461.
- [0481] [67]Scharton-Kersten 等. (2000)Infect Immun 68:5306-5313.
- [0482] [68]Ryan 等. (1999)Infect Immun 67:6270-6280.
- [0483] [69]Partidos 等. (1999)Immunol Lett 67:209-216.
- [0484] [70]Peppoloni 等. (2003)Expert Rev Vaccines 2:285-293.
- [0485] [71]Pine 等. (2002)J Control Release 85:263-270.
- [0486] [72]Tebbey 等. (2000)Vaccine 18:2723-34.
- [0487] [73]Domenighini 等. (1995)Mol Microbiol 15:1165-1167.

- [0488] [74]W099/40936.
- [0489] [75]W099/44636.
- [0490] [76]Singh 等 (2001)J Cont Release 70:267-276.
- [0491] [77]W099/27960.
- [0492] [78]US 6,090,406.
- [0493] [79]US 5,916,588.
- [0494] [80]EP A 0626169.
- [0495] [81]W099/52549.
- [0496] [82]W001/21207.
- [0497] [83]W001/21152.
- [0498] [84]Andrianov 等. (1998)Biomaterials 19:109-115.
- [0499] [85]Payne 等. (1998)Adv Drug Delivery Review 31:185-196.
- [0500] [86]Stanley (2002)Clin Exp Dermatol 27:571-577.
- [0501] [87]Jones (2003)Curr Opin Investig Drugs 4:214-218.
- [0502] [88]W099/11241.
- [0503] [89]W094/00153.
- [0504] [90]W098/57659.
- [0505] [91] 欧洲专利申请 0835318, 0735898 和 0761231.
- [0506] [92]Donnelly 等. (1997)Annu Rev Immunol 15:617-648.
- [0507] [93]Strugnell 等. (1997)Immunol Cell Biol 75(4):364-369.
- [0508] [94]Cui (2005)Adv Genet 54:257-89.
- [0509] [95]Robinson 和 Torres (1997)Seminars in Immunol 9:271-283.
- [0510] [96]Brunham 等. (2000)J Infect Dis 181 Suppl 3:S538-43.
- [0511] [97]Svanholm 等. (2000)Scand J Immunol 51(4):345-53.
- [0512] [98]DNA Vaccination-Genetic Vaccination (《DNA 疫苗接种 - 基因疫苗接种》) (1998)Koprowski 等编. (ISBN 3540633928).
- [0513] [99]Gene Vaccination:Theory and Practice (《基因免疫 : 治疗和实践》) (1998) Raz 编 (ISBN 3540644288).
- [0514] [100]Findeis 等., Trends Biotechnol. (1993)11:202
- [0515] [101]Chiou 等. (1994)Gene Therapeutics:Methods And Applications Of Direct Gene Transfer. (《基因治疗 : 直接基因转移的方法和应用》) Wolff 编
- [0516] [102]Wu 等., J. Biol. Chem. (1988)263:621
- [0517] [103]Wu 等., J. Biol. Chem. (1994)269:542
- [0518] [104]Zenke 等., Proc. Natl. Acad. Sci. (USA) (1990)87:3655
- [0519] [105]Wu 等., J. Biol. Chem. (1991)266:338
- [0520] [106]Jolly, Cancer Gene Therapy (1994)1:51
- [0521] [107]Kimura, Human Gene Therapy (1994)5:845
- [0522] [108]Connelly, Human Gene Therapy (1995)1:185
- [0523] [109]Kaplitt, Nature Genetics (1994)6:148

- [0524] [110]WO 90/07936.
[0525] [111]WO 94/03622.
[0526] [112]WO 93/25698.
[0527] [113]WO 93/25234.
[0528] [114]US 专利 5, 219, 740.
[0529] [115]WO 93/11230.
[0530] [116]WO 93/10218.
[0531] [117]US 专利 4, 777, 127.
[0532] [118]GB 专利号 2, 200, 651.
[0533] [119]EP-A-0345242.
[0534] [120]WO 91/02805.
[0535] [121]WO 94/12649.
[0536] [122]WO 93/03769.
[0537] [123]WO 93/19191.
[0538] [124]WO 94/28938.
[0539] [125]WO 95/11984.
[0540] [126]WO 95/00655.
[0541] [127]Curiel, Hum. Gene Ther. (1992)3:147
[0542] [128]Wu, J. Biol. Chem. (1989)264:16985
[0543] [129]US 专利 5, 814, 482.
[0544] [130]WO 95/07994.
[0545] [131]WO 96/17072.
[0546] [132]WO 95/30763.
[0547] [133]WO 97/42338.
[0548] [134]WO 90/11092.
[0549] [135]US 专利 5, 580, 859
[0550] [136]US 专利 5, 422, 120
[0551] [137]WO 95/13796.
[0552] [138]WO 94/23697.
[0553] [139]WO 91/14445.
[0554] [140]EP 0524968.
[0555] [141]Philip, Mol. Cell Biol. (1994)14:2411
[0556] [142]Woffendin, Proc. Natl. Acad. Sci. (1994)91:11581
[0557] [143]US 专利 5, 206, 152.
[0558] [144]WO 92/11033.
[0559] [145]US 专利 5, 149, 655.
[0560] [146]Zwijnenburg 等. (2001)J Infect Dis 183:1143-6.
[0561] [147]Tettelin 等 Proc. Natl. Acad. Sci. U. S. A. 99 (19), 12391-12396 (2002)
[0562] [148]W009/101403

- [0563] [149]W02006/130328
- [0564] [150]Lewis 等. (2004)PNAS USA 101:11123-8.
- [0565] [151]W02006/050341.
- [0566] [152]Paoletti 等. (1992)J Clin Invest 89:203-9
- [0567] [153]W096/40795
- [0568] [154]Wessels 等. (1989)Infect Immun 57:1089-94.
- [0569] [155]W02006/082527.
- [0570] [156]Paoletti 等. (1990)J Biol Chem 265:18278-83.
- [0571] [157]Wessels 等. (1990)J Clin Invest 86:1428-33.
- [0572] [158]Paoletti 等. (1992)Infect Immun 60:4009-14.
- [0573] [159]Wessels 等. (1987)Proc Natl Acad Sci USA 84:9170-4.
- [0574] [160]Wang 等. (2003)Vaccine 21:1112-7.
- [0575] [161]Wessels 等. (1993)Infect Immun 61:4760-6
- [0576] [162]Wessels 等. (1995)J Infect Dis 171:879-84.
- [0577] [163]Baker 等. (2004)J Infect Dis 189:1103-12.
- [0578] [164]US 专利 4356170.
- [0579] [165]Paoletti 和 Kasper(2003)Expert Opin Biol Ther 3:975-84.
- [0580] [166]W02005/000346
- [0581] [167] 匿名 (2002 年 1 月)Research Disclosure, 453077.
- [0582] [168]Anderson(1983)Infect Immun 39(1):233-238.
- [0583] [169]Anderson 等. (1985)J Clin Invest 76(1):52-59.
- [0584] [170]EP-A-0372501.
- [0585] [171]EP-A-0378881.
- [0586] [172]EP-A 0427347.
- [0587] [173]W093/17712.
- [0588] [174]W094/03208.
- [0589] [175]W098/58668.
- [0590] [176]EP A 0471177.
- [0591] [177]W091/01146.
- [0592] [178]Falugi 等. (2001)Eur J Immunol 31:3816-3824.
- [0593] [179]Baraldo 等. (2004)Infect Immun 72(8):4884-7.
- [0594] [180]EP-A-0594610.
- [0595] [181]Ruan 等. (1990)J Immunol 145:3379-3384.
- [0596] [182]W000/56360.
- [0597] [183]W001/72337.
- [0598] [184]W000/61761.
- [0599] [185]W000/33882
- [0600] [186]W099/42130.
- [0601] [187]US 专利 4, 761, 283.

- [0602] [188]US 专利 4, 356, 170.
- [0603] [189]US 专利 4, 882, 317.
- [0604] [190]US 专利 4, 695, 624.
- [0605] [191]Mol. Immunol., 1985, 22, 907-919
- [0606] [192]EP-A-0208375.
- [0607] [193]Bethell G. S. 等., J. Biol. Chem., 1979, 254, 2572-4
- [0608] [194]Hearn M. T. W., J. Chromatogr., 1981, 218, 509-18
- [0609] [195]W000/10599.
- [0610] [196]Gever 等., Med. Microbiol. Immunol, 165:171-288 (1979).
- [0611] [197]US 专利 4, 057, 685.
- [0612] [198]US 专利 4, 673, 574; 4, 761, 283; 4, 808, 700.
- [0613] [199]US 专利 4, 459, 286.
- [0614] [200]US 专利 4, 965, 338.
- [0615] [201]US 专利 4, 663, 160.
- [0616] [202]W02007/000343.
- [0617] [203] Vaccines (《疫苗》). (Plotkin 和 Orenstein 编). 第 4 版, 2004, ISBN:0-7216-9688-0.
- [0618] [204]Rappuoli 等. (1991)TIBTECH 9:232-238.
- [0619] [205]Harper 等. (2004)Lancet 364(9447):1757-65.
- [0620] [206]Brandt 等. (2006)J Antimicrob Chemother. 58(6):1291-4. Epub 2006Oct 26
- [0621] [207]Winter 等., (1991)Nature 349:293-99
- [0622] [208]US 4, 816, 567.
- [0623] [209]Inbar 等., (1972)Proc. Natl. Acad. Sci. U. S. A. 69:2659-62.
- [0624] [210]Ehrlich 等., (1980)Biochem 19:4091-96.
- [0625] [211]Huston 等., (1988)Proc. Natl. Acad. Sci. U. S. A. 85:5897-83.
- [0626] [212]Pack 等., (1992)Biochem 31, 1579-84.
- [0627] [213]Cumber 等., (1992)J. Immunology 149B, 120-26.
- [0628] [214]Riechmann 等., (1988)Nature 332, 323-27.
- [0629] [215]Verhoeyan 等., (1988)Science 239, 1534-36.
- [0630] [216]GB 2, 276, 169.
- [0631] [217]Gennaro(2000)Remington:The Science and Practice of Pharmacy (《雷明顿:药学技术与实践》). 第 20 版, ISBN:0683306472.
- [0632] [218]Methods In Enzymology (《酶学方法》)(S. Colowick 和 N. Kaplan 编, 学术出版社公司(Academic Press, Inc.))
- [0633] [219]Handbook of Experimental Immunology (《实验免疫学手册》), 第 IIV 卷 (D. M. Weir 和 C. C. Blackwell 编, 1986, 布莱克威尔科学出版公司(Blackwell Scientific Publications))
- [0634] [220]Sambrook 等. (2001)Molecular Cloning:A Laboratory Manual (《分子

克隆:实验室手册》),第3版(冷泉港实验室出版社(Cold Spring Harbor Laboratory Press)).

[0635] [221]Handbook of Surface and Colloidal Chemistry(《表面和胶体化学手册》)(Birdi, K. S. 编, CRC 出版社(CRC Press), 1997)

[0636] [222]Ausubel 等.(编)(2002)Short protocols in molecular biology(《精编分子生物学实验指南》),第5版(Current Protocols).

[0637] [223]Molecular Biology Techniques:An Intensive Laboratory Course(《分子生物学技术:强化实验室课程》),(Ream 等., 编, 1998, 学术出版社(Academic Press))

[0638] [224]PCR(Introduction to Biotechniques Series)(《PCR(生物技术入门丛书)》),第2版.(Newton 和 Graham 编, 1997, 施普林格出版公司(Springer Verlag))

[0639] [225]Geysen 等.(1984)PNAS USA 81:3998-4002.

[0640] [226]Carter(1994)Methods Mol Biol 36:207-23.

[0641] [227]Jameson, BA 等. 1988, CABIOS 4(1):181-186.

[0642] [228]Raddrizzani 和 Hammer(2000)Brief Bioinform 1(2):179-89.

[0643] [229]Bublil 等.(2007)Proteins 68(1):294-304.

[0644] [230]De Lalla 等.(1999)J. Immunol. 163:1725-29.

[0645] [231]Kwok 等.(2001)Trends Immunol 22:583-88.

[0646] [232]Brusic 等.(1998)Bioinformatics 14(2):121-30

[0647] [233]Meister 等.(1995)Vaccine 13(6):581-91.

[0648] [234]Roberts 等.(1996)AIDS Res Hum Retroviruses 12(7):593-610.

[0649] [235]Maksyutov 和 Zagrebelnaya(1993)Comput Appl Biosci 9(3):291-7.

[0650] [236]Feller 和 de la Cruz(1991)Nature 349(6311):720-1.

[0651] [237]Hopp(1993)Peptide Research 6:183-190.

[0652] [238]Welling 等.(1985)FEBS Lett. 188:215-218.

[0653] [239]Davenport 等.(1995)Immunogenetics 42:392-297.

[0654] [240]Tsurui 和 Takahashi(2007)J Pharmacol Sci. 105(4):299-316.

[0655] [241]Tong 等.(2007)Brief Bioinform. 8(2):96-108.

[0656] [242]Schirle 等.(2001)J Immunol Methods. 257(1-2):1-16.

[0657] [243]Chen 等.(2007)Amino Acids 33(3):423-8.

[0658] [244]Current Protocols in Molecular Biology(《新编分子生物学实验指南》)(F. M. Ausubel 等., 编, 1987)附录 30

[0659] [245]Smith 和 Waterman(1981)Adv. Appl. Math. 2:482-489.

[0660] [246]Klock 等(2008), Proteins, 71:982-994

[0661] [247]Maione 等(2005), Science, 309:148-150

[0001]

序列表

SEQ ID NO:1 (GBS67 2603)

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAAGKGLVVKKTDDQNKPLSKATFVLKTTAHPES
KIEKVTAELTGEATFDNLI PGDYTLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNQEE
LDKQYPPTGIYEDTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLKRISEVGDLAHNSK
YKIELTVSGKTIVKPVKQKPLDVVFLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNR
VALVTYGSDI FDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTAPKAKWGS
TTNGLTPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTRSYAINNFKLGASYE
SQFEQMKNNGYLKNSNFLTDPKPEDIKNGGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYR
NGPVKEHGTPTKLYINSLKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGT FQKLKEEAFKLS DGEITEL
MRSFSSKPEYYTPIVTSADTSNNEILSKIQQQFETILTKENSIVNGTIEDPMGDKINLQLGNGQTLQPS
DYTLQNGDGSVMKDG IATGGPNNDGGILKGVKLEYIGNKLYVRGLNLGEGQKVTLT YDVKLDDSFISNK
FYDTNGRRTLNPKSEDPNTRDFPIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNKLLKLGATFEL
QEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPEDYQKITNKPILTFEVVKGSIKN
IIAVNKQISEYHEEGDKHLITNTHIPPKGIIPMTGGKILSFILIGGAMMSIAGGIYIWKRYKSSDMS
IKKD

SEQ ID NO:2 (GBS67 2603 片段1)

NTNVLGESTVPENGAAGKGLVVKKTDDQNKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLI PGDY
TLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNQEE LDKQYPPTGIYEDTKESYKLEHV
KGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLKRISEVGDLAHNSK YKIELTVSGK

SEQ ID NO:3 (GBS67 2603 片段2)

TIVKPVKQKPLDVVFLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNRVALVTYGSDI
FDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTAPKAKWGS TTNGLTPEQQ
KEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTRSYAINNFKLGASYESQFEQMKNNG
YLNKSNFLTDPKPEDIKNGGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVKEHGTP
TKLYINSLKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGT FQKLKEEAFKLS DGEITELMRSFSSKPEY
YTPIVTSADTSNNEILSKIQQQFETILTKENSIVNGTIEDPMGDKINLQLGNG

SEQ ID NO:4 (GBS67 2603 片段3)

QTLQPSDYTLQNGDGSVMKDG IATGGPNNDGGILKGVKLEYIGNKLYVRGLNLGEGQKVTLT YDVKLDD
SFISNKFYDTNGRRTLNPKSEDPNTRDFPIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNKLLK
GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPEDYQKITNKPILTFEVV
KGSIKNIIAVNKQISEYHEEGDKHLITNTHIPPKGIIPMTGGK

SEQ ID NO:5 (GBS67 H36B)

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAAGKGLVVKKTDDQNKPLSKATFVLKPTSHSES
KVEKVTTEVTGEATFDNLTPGDYTLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSDDKKSIIIEQRQEE
LDKQYPLTGAYEDTKESYNLEHVKN SIPNGKLEAKAVNPYSSEGEHIREIQEGTLKRISEVNDLDHNSK
YKIELTVSGKSI IKTINKDEPLDVVFLDNSNSMKNNGKNNKAKKAGEAVETIIKDV LGANVENRAALV
TYGSDIFDGRTVKVIKGFKEDPYGLETSTVQNTDYSYKFTNIAADI IKKIPKEAPEAKWGGTSLGL
TPEKKREYDLSKVGETFTMKAFMEADTLSSIQRSRKKIIVHLTDGVPTRSYAINSFVKGSTYANQFER
IKEKGYLDKNNYFITDDPEKIKNGGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVR
EHGTPTKLYINSLKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGT FQKLKEEAFELSDGEITELMNSFS
SKPEYYTPIVTSADVSNNEILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLHLGNGQTLQPSDYTLQ
GNDGSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLT YDVKLDDSFISNKFYDTN
GRTLNPKSEEPDTRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDKDNKLLKLGATFELQEFNE
DYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVVKGSIQNI IAVN
KQISEYHEEGDKHLITNTHIPPKGIIPMTGGKILSFILIGGAMMSIAGGIYIWKRRHKSSDASIEKD

SEQ ID NO:6 (GBS67 H36B 片段1)

NTNVLGESTVPENGAAGKGLVVKKTDDQNKPLSKATFVLKPTSHSES KVEKVTTEVTGEATFDNLTPGDY
TLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSDDKKSIIIEQRQEE LDKQYPLTGAYEDTKESYNLEHV
KN SIPNGKLEAKAVNPYSSEGEHIREIQEGTLKRISEVNDLDHNSK YKIELTVSGK

SEQ ID NO:7 (GBS67 H36B 片段2)

[0002]

SIKTKINKDEPLDVFVLDNSNSMKNNNGKNNKAKKAGEAVETI IKDVLGANVENRAALVTYGSDIFDGR
TVKVIKGFKEDPYGLETSTFTVQTNDSYKFKFTNIAADI IKKI PKEAPEAKWGGTSLGLTPEKKREYDL
SKVGETFTMKAFMEADTLSSSIQRKSRKI IVHLTDGVPTRSYAINSFVKGSTYANQFERIKEKGYLDKN
NYFITDDPEKIKGNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVREHGTPTKLYI
NSLKQKNYDIFNFGIDISGFRQVYNEDYKKNQDGTFOKLKEEAFELSDGEITELMNSFSSKPEYYTPIV
TSADVSNNEILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLHLGNG

SEQ ID NO:8 (GBS67 H36B 片段 3)

QTLQPSDYTLQGNDSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLYDVKLDD
SFISNKFYDTNGRDTLNPKEEPTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDDKNNKLLK
GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVV
KGSIQNI IAVNKQISEYHEEGDKHLITNTHIPPKGII PMTGGKG

SEQ ID NO:9 (GBS67 CJB111)

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAAGKLVVKKTDQNKPLSKATFVLKTTAHPES
KIEKVTAELTGEATFDNLI PGDYTLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNQEE
LDKQY PPTGIYEDTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREI PEGTLSKRISEVGDLAHNK
YKIELTVSGKTI VPKVVKQKPLDVFVLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNR
VALVTYGSDIFDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEII KRIPTAPKAKWGS
TTNGLTPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKI IVHVT DGVPTRSYAINNFKLGASYE
SQFEQMKNNGYLKNSNLLTDKPEDIKGNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYR
NGPVREHGTPTKLYINSLKQKNYDIFNFGIDISGFRQVYNEDYKKNQDGTFOKLKEEAFELSDGEITEL
MKSFSSKPEYYTPIVTSDDASNNEILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLQLGNGQTLQPS
DYTLQGNDSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLYDVKLDDSFISNK
FYDTNGRDTLNPKSEDPNTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDDKNNKLLK GATFEL
QEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVV KGSIQN
IIAVNKQISEYHEEGDKHLITNTHIPPKGII PMTGGKGILSFILIGGSMMSIAGGIYIWKRYKSSDIS
REKD*

SEQ ID NO:10 (GBS67 CJB111 片段 1)

NTNVLGESTVPENGAAGKLVVKKTDQNKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLI PGDY
TLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNQEE LDKQY PPTGIYEDTKESYKLEHV
KGSVPNGKSEAKAVNPYSSEGEHIREI PEGTLSKRISEVGDLAHNKYKIELTVSGK

SEQ ID NO:11 (GBS67 CJB111 片段 2)

TIVKPVKQKPLDVFVLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNRVALVTYGSDI
FDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEII KRIPTAPKAKWGSTTNGLTPEQQ
KEYYLSKVGETFTMKAFMEADDILSQVNRNSQKI IVHVT DGVPTRSYAINNFKLGASYESQFEQMKNNG
YLNKSNFLLTDKPEDIKGNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVREHGTP
TKLYINSLKQKNYDIFNFGIDISGFRQVYNEDYKKNQDGTFOKLKEEAFELSDGEITELMKSFSSKPEY
YTPIVTSDDASNNEILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLQLGNG

SEQ ID NO:12 (GBS67 CJB111 片段 3)

QTLQPSDYTLQGNDSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLYDVKLDD
SFISNKFYDTNGRDTLNPKSEDPNTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDDKNNKLLK
GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVV
KGSIQNI IAVNKQISEYHEEGDKHLITNTHIPPKGII PMTGGKG

SEQ ID NO:13 (GBS67 515)

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAAGKLVVKKTDQNKPLSKATFVLKTTAHPES
KIEKVTAELTGEATFDNLI PGDYTLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNQEE
LDKQY PPTGIYEDTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREI PEGTLSKRISEVGDLAHNK
YKIELTVSGKTI VPKVVKQKPLDVFVLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNR
VALVTYGSDIFDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEII KRIPTAPKAKWGS
TTNGLTPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKI IVHVT DGVPTRSYAINNFKLGASYE
SQFEQMKNNGYLKNSNLLTDKPDIDIKGNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYR
NGPVKEHGTPTKLYINSLKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGTFOKLKEEAFKLS DGEITEL
MRSFSSKPEYYTPIVTSADTSNNEILSKIQQQFETILTENSIVNGTIEDPMGDKINLQLGNGQILQPS

[0003]

DYTLQGNDSVMDGIATGGPNNDGGILKGVKLEYIGNKLYVRGLNLGEGQKVTLTVDKLDSDSFISNK
FYDTNGRRTLNPKSEDPNTRDFPIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNKLLKLGATFEL
QEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPEDYQKITNKPILTFEVVKGSIKN
IIAVNKQISEYHEEGDKHLITNTHIPPKGIIPKTGGKILSFILIGGAMMSIAGGIYIWKRYKSSDMS
IKKD*

SEQ ID NO:14 (GBS67 515 片段 1)

NTNVLGESTVPENGAAGKLVVKKTDQNKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLI PGDY
TLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPTGIYEDTKESYKLEHV
KGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLISKRISEVGDLAHNKYKIELTVSGK

SEQ ID NO:15 (GBS67 515 片段 2)

TIVKPVKQKPLDVVFLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNRVALVTYGS DI
FDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRI PTEAPKAKWGSTTNGLTPEQQ
KEYYLSKVGETFTMKA FMEADDILSQVNRNSQKIIVHVTDGVPTRSYAINNFKLGASYESQFEQMKNG
YLNKSNFLLTDKPD DIKNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVEHGTP
TKLYINSLKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGT FQKLKEEAFKLS DGEITELMRSFSSKPEY
YTPIVTSADTSNNEILSKIQQQFETILT KENSIVNGTIEDPMGDKINLQLGNG

SEQ ID NO:16 (GBS67 515 片段 3)

QILQPSDYTLQGNDSVMDGIATGGPNNDGGILKGVKLEYIGNKLYVRGLNLGEGQKVTLTVDKLDSD
SFISNKFYDTNGRRTLNPKSEDPNTRDFPIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNKLLK
GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPEDYQKITNKPILTFEVV
KGSIKNIIAVNKQISEYHEEGDKHLITNTHIPPKGIIPKTGGK

SEQ ID NO:17 (GBS67 NEM316)

MRKYQKFSKILTSLFCLSQIPLNTNVLGESTVPENGAAGKLVVKKTDQNKPLSKATFVLKTTAHPES
KIEKVTAELTGEATFDNLI PGDYTLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNH
LDKQYPPTGIYEDTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLISKRISEVGD
YKIELTVSGKTIIVKPVKQKPLDVVFLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNR
VALVTYGS DIFDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRI PTEAPRAK
WGSTTNGLTPEQQKQY YLSKVGETFTMKA FMEADDILSQVDRNSQKIIVHITDGVPTRSYAINNF
KLGASYESQFEQMKNGYLNKSNFLLTDKPEDIKNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLN
YPKGTIYRNGPVEHGTPTKLYINSLKQKNYDIFNFGIDISAFRQVYNEEYKKNQDGT FQKLKEE
AFELSDGEITELMKS FSSKPEY YTPIVTSSDASNNEILSKIQQQFEKVLTKENSIVNGTIEDPMG
DKINLQLGNGQTLQPSDYTLQGNDSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQ
KVTLTVDKLDSDSFISNKFYDTNGRRTLNPKSEDPNTRDFPIPKIRDVREYPTITIKNEKKLGEI
EFTKVDKDNKLLKLGATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVS
PKDYQKITNKPILTFEVVKGSIQNIIAVNKQISEYHEEGDKHLITNTHIPPKGIIPMTGGKILS
FILIGGSMMMSIAGGIYIWKRYKSSDISREKD

SEQ ID NO:18 (GBS67 NEM316 片段 1)

NTNVLGESTVPENGAAGKLVVKKTDQNKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLI PGDY
TLSEETAPEGYKKTNTQWQVKVESNGKTTIQNSGDKNSTIGQNH EELDKQYPPTGIYEDTKESYKLEHV
KGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLISKRISEVGDLAHNKYKIELTVSGK

SEQ ID NO:19 (GBS67 NEM316 片段 2)

TIVKPVKQKPLDVVFLDNSNSMNDGPNFQRHNKAKKAAEALGTAVKDI LGANS DNRVALVTYGS DI
FDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRI PTEAPRAKWGSTTNGLTPEQQ
KQY YLSKVGETFTMKA FMEADDILSQVDRNSQKIIVHVTDGVPTRSYAINNFKLGASYESQFEQMKNG
YLNKSNFLLTDKPEDIKNGESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVEHGTP
TKLYINSLKQKNYDIFNFGIDISAFRQVYNEEYKKNQDGT FQKLKEEAFELSDGEITELMKS FSSKPEY
YTPIVTSSDASNNEILSKIQQQFEKVLTKENSIVNGTIEDPMGDKINLQLGNG

SEQ ID NO:20 (GBS67 NEM316 片段 3)

QTLQPSDYTLQGNDSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLTVDKLDSD
SFISNKFYDTNGRRTLNPKSEDPNTRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDKDNKLLK

[0004]

GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVV
KGSIQNI IAVNKQISEYHEEGDKHLITNTHIPPKGIIPMTGGKG

SEQ ID NO:21 (GBS67 DK21)

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDQNKPLSKATFVLKPTSHSES
KVEKVTTEVTGEATFDNLT PGDYTLSEETAPEGYKKTQTQWQVKVESNGKTTIQNSDDKKSIEQRQEE
LDKQYPLTGAYEDTKESYNLEHVKN SIPNGKLEAKAVNPYSSEGEHIREIQEGTLSKR ISEVNDLDH NK
YKIELTVSGKSI IKTINKDEPLDVVFLDNSNSMKNNKAKKAGEAVETI IKDVLGANVENRAALV
TYGSDIFDGR TVKVIKGFKEDPYGLET SFTVQTN DYSYK KFTNIAADI IKKI PKEAPEAKWGGTSLGL
TPEKKREYDLSKVGETFTMKAFMEADTLLSSIQRKSRKI IVHLTDGVPTRSYAINS FVTGSTYANQFER
IKEKGYLDKNNYFITDDPEKIKGN GESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVR
EHGTPTKLYINSLKQKNYDIFNFGIDISGFRQVY NEDYKKNQDGT FQKLKEEAFELSDGEITELMNSFS
SKPEYYP IIVTSADVSNN EILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLQLGNGQTLQPSDYTLQ
GNDGSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLTYDVKLDDSFISNKFYDTN
GRTTLNPKSEEPD TLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKV DKNKLLKLGATFELQEFNE
DYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVVKGSIQNI IAVN
KQISEYHEEGDKHLITNTHIPPKGIIPMTGGKGILSFILIGGAMMSIAGGIYIWKHRHKS SDASIEKD*

SEQ ID NO:22 (GBS67 DK21 片段 1)

NTNVLGESTVPENGAKGKLVVKKTDQNKPLSKATFVLKPTSHSESKVEKVTTEVTGEATFDNLT PGDY
TLSEETAPEGYKKTQTQWQVKVESNGKTTIQNSDDKKSIEQRQEELDKQYPLTGAYEDTKESYNLEHV
KNSIPNGKLEAKAVNPYSSEGEHIREIQEGTLSKR ISEVNDLDH NKYKIELTVSGK

SEQ ID NO:23 (GBS67 DK21 片段 2)

SIIKTINKDEPLDVVFLDNSNSMKNNKAKKAGEAVETI IKDVLGANVENRAALV TYGSDIFDGR
TVKVIKGFKEDPYGLET SFTVQTN DYSYK KFTNIAADI IKKI PKEAPEAKWGGTSLGLTPEKKREYDL
SKVGETFTMKAFMEADTLLSSIQRKSRKI IVHLTDGVPTRSYAINS FVTGSTYANQFERI KEKGYLDKN
NYFITDDPEKIKGN GESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVR EHGTPTKLYI
NSLKQKNYDIFNFGIDISGFRQVY NEDYKKNQDGT FQKLKEEAFELSDGEITELMNSFSSKPEYYP IIV
TSADVSNN EILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLQLGNG

SEQ ID NO:24 (GBS67 DK21 片段 3)

QTLQPSDYTLQ GNDGSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLTYDVKLDD
SFI SNKFYDTN GRTTLNPKSEEPD TLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKV DKNKLLK
GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVV
KGSIQNI IAVNKQISEYHEEGDKHLITNTHIPPKGIIPMTGGKG

SEQ ID NO:25 (GBS67 CJB110)

MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDQNKPLSKATFVLKTTAHPES
KIEKVTA EVTGEATFDNLT PGDYTLSEETAPEGYKKTQTQWQVKVESNGKTTIQNSDDKKSIEQRQEE
LDKQYPLTGAYEDTKESYNLEHVKN SIPNGKLEAKAVNPYSSEGEHIREIQEGTLSKR ISEVNDLDH NK
YKIELTVSGKSI IKTINKDEPLDVVFLDNSNSMKNNKAKKAGEAVETI IKDVLGANVENRAALV
TYGSDIFDGR TVKVIKGFKEDPYHGLETSFTVQTN DYSYK KFTNIAADI IKKI PKEAPEAKWGGTSLGL
TPEKKREYDLSKVGETFTMKAFMEADTLLSSIQRKSRKI IVHLTDGVPTRSYAINS FVTGSTYANQFER
IKEKGYLDKNNYFITDDPEKIKGN GESYFLFPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVR
EHGTPTKLYINSLKQKNYDIFNFGIDISGFRQVY NEDYKKNQDGT FQKLKEEAFELSGGEITELMKSFS
SKPEYYP IIVTSADVSNN EILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLQLGNGQTLQPSDYTLQ
GNDGSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLTYDVKLDDSFISNKFYDTN
GRTTLNPKSEEPD TLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKV DKNKLLKLGATFELQEFNE
DYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVVKGSIQNI IAVN
KQISEYHEEGDKHLITNTHIPPKGIIPMTGGKGILSFILIGGAMMSIAGGIYIWKHKKSS DASIEKD*

SEQ ID NO:26 (GBS67 CJB110 片段 1)

NTNVLGESTVPENGAKGKLVVKKTDQNKPLSKATFVLKTTAHPESKIEKVTA EVTGEATFDNLT PGDY
TLSEETAPEGYKKTQTQWQVKVESNGKTTIQNSDDKKSIEQRQEELDKQYPLTGAYEDTKESYNLEHV
KNSIPNGKLEAKAVNPYSSEGEHIREIQEGTLSKR ISEVNDLDH NKYKIELTVSGK

SEQ ID NO:27 (GBS67 CJB110 片段 2)

[0005]

SIIKTINKDEPLDVVFLDNSNSMKNNKNNKAKKAGEAVETI IKDVLGANVENRAALVITYGSDIFDGR
TVKVIKGFKEPDYHGLETSFTVQTNDSYKFKFTNIAADI IKKIPKEAPEAKWGGTSLGLTPEKKREYDL
SKVGETFTMKAFMEADTLSSIQRKSRI IVHLTDGVPTRSYAINSFVTGSTYANQFERIKEKGYLDKN
NYFITDDPEKIKNGESYFLPPLDSYQTQII SGNLQKLHYLDLNLNYPKGTIYRNGPVRHGTPTKLYI
NSLKQKNYDIFNFGIDI SGFRQVYNEDYKKNQDGT FQKLKEEAFELSGGEITELMKSFSKPEYYTPIV
TSADVSNNEILSKIQQQFEKILTKENSIVNGTIEDPMGDKINLQLGNG

SEQ ID NO:28 (GBS67 CJB110 片段 3)

QTLQPSDYTLQGNDSIMKDSIATGGPNNDGGILKGVKLEYIKNKLYVRGLNLGEGQKVTLYDVKLDD
SFISNKFYDTNGRTTLNPKSEEPDTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDKDNKLLK
GATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPKDYQKITNKPILTFEVV
KSSIQNI IAVNKQISEYHEEGDKHLITNTHI PPKGII PMTGKKG

SEQ ID NO:29 (接头)

GSGS

SEQ ID NO:30 (接头)

GSGGGG

SEQ ID NO:31 (接头)

GSGSGGGG

SEQ ID NO:32 (IC 佐剂)

NCNCNCNCNCNCNCNCNCNCNCNCNCNC [各n为肌苷]

SEQ ID NO:33 (聚阳离子肽佐剂)

KLKLLLLLKLK

SEQ ID NO:34 (GBS80 2603)

MKLSKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTVNIYKLOADS YKSEITSNGG
IENKDG EVISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKCLTTVEAADAKVGTILEEGVSLPQKTN
AQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYKNNVVTDEPK
TDKDVKKLQDDAGYTI GEEFKWFLKSTIPANLGDYEKFEITDKFADGLTYKSVGKIKIGSKTLNRDEH
YTIDEPTVDNQNTLKITFKPEKFKEIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVL
GKAIENTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAEFDLLASDGTAVKWT
DALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKGLAYAVDANAEGTAVTYKLETKAPEGYVIPDK
EIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPSIPNTGGIGTAIFVAIGAAMFAVKGMRRTK
DN

SEQ ID NO:35 (GBS80 2603 无前导序列)

AEVSQERPAKTTVNIYKLOADS YKSEITSNGG IENKDG EVISNYAKLGDNVKGLQGVQFKRYKVKTDIS
VDELKCLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVP
FVLELPVANSTGTGFLSEINIYKNNVVTDEPKTDKDVKKLQDDAGYTI GEEFKWFLKSTIPANLGDYE
KFEITDKFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFKEIAELLKGMTLVK
NQDALDKATANTDDAAFLEIPVASTINEKAVL GKAIENTFELQYDHTPDKADNPKPSNPPRKPEVHTGG
KRFVKKDSTETQTLGGAEFDLLASDGTAVKWT DALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKG
LAYAVDANAEGTAVTYKLETKAPEGYVIPDK EIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS
IPNTGGIGTAIFVAIGAAMFAVKGMRRTKDN

SEQ ID NO:36 (GBS80 2603 无跨膜/胞质区域)

MKLSKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTVNIYKLOADS YKSEITSNGG
IENKDG EVISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKCLTTVEAADAKVGTILEEGVSLPQKTN
AQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYKNNVVTDEPK

[0006]

TDKDVKKLGQDDAGYTI GEEFKWFLKSTIPANLGDYEKFEITDKFADGLTYKSVGKIKIGSKTLNRDEH
 YTIDEPTVDNQNTLKITFKPEKFKEIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVL
 GKAIENFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAEFDLLASDGTAVKWT
 DALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKGLAYAVDANAEGTAVTYKPKETKAPEGYVIPDK
 EIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPSIPNTG

SEQ ID NO:37 (GBS80 2603

无跨膜/胞质区域和细胞壁锚定)

MKLSKKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVVSQERPAKTTVNIYKQLQADSYKSEITSNGG
 IENKDGVEVISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKLLTVEAADAKVGTILEEGVSLPQKTN
 AQLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVPFVLELPPVANSTGTGFLSEINIYPKNVVTDEPK
 TDKDVKKLGQDDAGYTI GEEFKWFLKSTIPANLGDYEKFEITDKFADGLTYKSVGKIKIGSKTLNRDEH
 YTIDEPTVDNQNTLKITFKPEKFKEIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVL
 GKAIENFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAEFDLLASDGTAVKWT
 DALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKGLAYAVDANAEGTAVTYKPKETKAPEGYVIPDK
 EIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS

SEQ ID NO:38 (GBS80 2603 无胞外结构域)

AEVSQERPAKTTVNIYKQLQADSYKSEITSNGGIENKDGVEVISNYAKLGDNVKGLQGVQFKRYKVKTDIS
 VDELKLLTVEAADAKVGTILEEGVSLPQKTNQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVP
 FVLELPPVANSTGTGFLSEINIYPKNVVTDEPKTDKDVKKLGQDDAGYTI GEEFKWFLKSTIPANLGDYE
 KFEITDKFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFKEIAELLKGMTLVK
 NQDALDKATANTDDAAFLEIPVASTINEKAVL GKAIENFELQYDHTPDKADNPKPSNPPRKPEVHTGG
 KRFVKKDSTETQTLGGAEFDLLASDGTAVKWT DALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKG
 LAYAVDANAEGTAVTYKPKETKAPEGYVIPDK EIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS

SEQ ID NO:39 (GBS80 2603的N-末端免疫原性片段

AEVSQERPAKTTVNIYKQLQADSYKSEITSNGGIENKDGVEVISNYAKLGDNVKGLQGVQFKRYKVKTDIS
 VDELKLLTVEAADAKVGTILEEGVSLPQKTNQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVP
 FVLELPPVANSTGTGFLSEINIYPKNVVTDEPKTDKDVKKLGQDDAGYTI GEEFKWFLKSTIPANLGDYE
 KFEITDKFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFKEIAELLK

SEQ ID NO:40 (GBS59 2603)

MKRINKYFAMFSALLLTLTSLLSVAPAFAD EATTNTVTLHKILQTESNLNKS NFPGTTGLNGKDYKGG A
 ISDLAGYFGEGSKEIEGAF FALALKEDEKSGKVQYVKAKEGNKLT PALINKDGTPEITVNI DEAVSGLTP
 EGD TGLVFN TKGLKGEFKIVEVKS KSTYNNNGSLLAASKAVPVNITLPLVNE DGVVADAHVY PKNTEEK
 PEIDKNFAKTNDL TALTDVNRLLTAGANYGN YARDKATATAEIGKVVPYEVKTKIHKGSKYENLVWTDI
 MSNGLTMGSTVSLKAGTTETF AKD TDYELS IDARGFTLKF TADGLGKLEKA AKTADIEFTLTYSATVN
 GQAI IDNPESNDIKLSYGNKPGKDL TELPVT PPSKGEVTVAKTWS DGIAPDGVNVVYTLKDKDKTVA SVS
 LTKTSKGTIDLNGIKFEVSGNFSGKFTGLENKSYMISERVSGYGSAINLENGKVTITNTKDS DNPTPL
 NPTEPKVETHGKKFVKTNEQGDR LAGAQFVVKNSAGKYLALKADQSE GQKTLAAKKIALDEAIAAYNKL
 SATDQKGEKGITAKELIKTKQADY DAAFIEARTAYEWITDKARAITYTSNDQGQFEVTGLADGTYNLEE
 T LAPAGFAKLAGNIK FVVNQGSYITGGNIDYVANSNQKDATR V ENKKVTIPQTGGIGTILFTIIGLSIM
 LGAVVIMKRRQSKEA

SEQ ID NO:41 (GBS59 515)

MKKINKYFAVFSALLLTVTSLFSVAPVFAEEAKTTDTVTLHKIVMPRTAFDGF TAGTKGKDNTDYVGKQ
 IEDLKTYFGSGEAKEIAGAYFAFKNEAGTKYITENGE EVDLDTTDAKGC AVLKGLTTDNGFKFNTSKL
 TGT YQIVELKEKSTYNN DGSILADSKAVPVKITLPLVNDNGVVKDAHVY PKNTEETK PQVDKNFADKELD
 YANNKDKGTVSASVGDVKKYHVGTKILKGS DYKLIWTD SMTKGLTFNNDIAVTLDGATLDATNYKLV
 ADDQGFRLVLTDKGLEAVAKAAKT KDVEIKITYSATLNGS AVVEVLETNDVKLDYGN NP TIE NEPKEGI
 PVDKKITVNTK TWAVDGNEVNKADETVDAVFTLQVKGDKGKWNVDSAKATAATSFKHTFENLDNAKTYRV
 IERVSGYAPEYVSFVNGVVTIKNNKDSNEPTPINPSEPKVVTYGRKFVKTNKDGERLAGATFLVKKDG
 KYLARKSGVATDAEKA AVDSTKSALDAAVKAYNDLTKEKQEQD GKSALATVSEKQKAYNDAFVKANYS
 YEWVEDKNAKNVVKLISNDKGQFEITGLTEGQYSLEETQAPTGYAKLSGDV SFVFNATSYSKGSAQDIE
 YTQGSKTKDAQQVINKKVTIPQTGGIGTIFFTIIGLSIMLGAVVIMKRRQSEEV

[0007]

SEQ ID NO:42 (GBS59 cjb111)

MKKINKCLTMFSTLLILLTSLFSVAPAFADDATTDTVTLHKIVMPQAAFDMFTEGTKGKNDSDYVVGKQI
NDLKS YFGSTDAKEIKGAFFVFKNETGTFKITENGKEVDTLLEAKDAEGGAVLSGLTKDNGFVENTAKLK
GIYQIVELKEKS NYDNNGSILADSKAVPVKITLPLVNNQGVVKAHIYPKNTETKPPQVDKNFADKDLDY
TDNRKDKGVVSATVGDKKEYIVGTKILKGS DYKLVWTD SMTKGLTFNNNVKTLDGEDFPVLNYKLV
DDQGFRALNATGLAAVAAA AKDKDVEIKITYSATVNGSTTVEI PETNDVKLDYGNNPTEESEPOEGTP
ANQEI KVIKDWAVDGTITDANVAVKAI FTLQEKQTDGTVVNVASHEATKPSRFEHTFTGLDNAKYRVV
ERVSGYTPPEYVSFKNGVVTIKNNKNSNDPTPINPSEPKVVTYGRKFVKTNQANTERLAGATFLVKKEGK
YLARKAGAATAEAKAAVKTA KLALDEAVKAYNDLTKEKQEQEGKTALATVDQKQKAYNDAFVKANYSY
EWVADKKADNVVKLISNAGGQFEITGLDKGTYGLEETQAPAGYATLSGDVNFVETATSYSKGATTDIAY
DKGSVKKDAQVQNKVVTIPQTGGIGTILFTIIGLSIMLGAVVIMKRRQSEEA

SEQ ID NO:43 (GBS59 h36b)

MKRINKYFAMFSALLLILTSLLSVAPVFAAEMGNITKTVTLHKIVQTS DN LAKPNFPINGLNGTKYMG
OKLTDISGYFGQGSKEIAGAFFAVMNSQTKYITESGTEVESIDAAGVLKGLTTENGITFNTANLKGT
QIVELLDKSNYKNGDKVLADSKAVPVKITLPLYNEEGIVDAEVYPKNTEEAPQIDKNFAKANKLLNDS
DNSAIAGGADYDKYQAEKAKATAEIGQEIPIYEVKTKIQKGSKYKNLAWVD TMSNGLTMGNTVNLEASSG
SFVEGTDYNVERDDRGFTLKFDTGLTKLQKEAETQAVEFTLYSATVNGAAIDDKPESNDIKLQYGNK
PGKVKVEIPVTPSNGEITVSKTWDKGS DLENANVVYTLKGGTAVASVSLTKTTPNGEINLNGIKFTV
TGAFAGKFSGLTDSKTYMISER IAGYGNITITGAGSAAITNTPDSDNPTPLNPTEPKVVTHGKKFVKT
STETERLQGAQFVVKDSAGKYLALKSSATI SAQTTAYTNAKTALDAKIAAYNKL SADDQKGTGETAKA
EIKTAQDAYNAAFIVARTAYEWWTNKEDANVVKVTSNADGQFEVSGLATGDYKLEETQAPAGYAKLAGD
VDFKVGNSKADDSGNI DYTASSNKKDAQRIENKVTIPQTGGIGTILFTIIGLSIMLGAVVIMKRRQ
SEEA

SEQ ID NO:44 (GBS59 CJB110)

MKKINKYFAVFSALLLTVTSLLSVAPAFAD EATNTVTLHKILQTESNLKNSNFP GTTGLNGDDYKGES
ISDLAEYFGSGSKEIDGAFFALALBEEKDGVVQYVKA KANDKLT PDLITKGT PATTTKVEEAVGGLTTG
TGIVFNTAGLKGNFKIIELKDKSTYNNNGSLLAASKAVPVKITLPLVSKDGVVKAHVYPKNTETKPEV
DKNFAKTNDLTALKDATLLKAGADYKNYSATKATVTA EIGKVIPIYEVKTKVLKGS KYEKLWTD TMSNG
LTMGDDVNLA VSGTTTTFIKDI DYTLSIDDRGFTLKFKATGLDKLEEA AKASDVEFTLYKATVNGQAI
IDNPEVNDIKLDYGNKPGTDLSEQPVT PEDGEVKVTKTWAAGANKADAKVVYTLKNATKQVVASVALTA
ADTKGTINLGKGMTFEITGAFSGTFKGLQNKAYTVSERVAGYTNAINVTGNAVAITNTPDSDNPTPLNP
TQPKVETHGKKFVKVGDADARLAGAQFVVKNSAGKFLALKE DAAVSGAQTELATAKTDLDNAIKAYNGL
TKAQQEGADGTS AKELINTKQSAYDAAF I KARTAYTWVDEKTKAITFTSNNQGGQFEVTVGLEVGSYKLEE
TLAPAGYAKLSGDI EFTVGHDSYTS GDIKYKTDASNNAQKVENKVTIPQTGGIGTILFTIIGLSIML
GAVVIMKRRQSEEA

SEQ ID NO:45 (GBS59 DK21)

MKKINKFFVAFSALLLILTSLLSVAPAFAE EERTTETVTLHKILQETNLKNSAFFPGTKGLDGT EYDGGK
AIDKLD SYFGNDSKDIGGAYFILANSKGEYIKANDKNLKP EFGNTPKTTLNISEAVGGLTEENAGIK
FETTLGRGDFQIIELKDKSTYNNGGAILADSKAVPVKITLPLINKDGVVKAHVYPKNTETKPPQIDKNF
ADKNLDYINNQDKGTISATVGDVKKYTVGTKILKGS DYKLVWTD SMTKGLTFNNDVTVTL DGANFEQ
SNYTLVADDQGFRLVLNATGLSKVAEAAKTKDVEIKINYSATVNGSTVVEKSENNDVKLDYGNNPTTEN
EPQTGNPVNKEITVRKTWAVD GNEVNKGDEKVDVFTLQVKDSDKVVNDSATATAATDFKYTFKNLDN
AKTYRVVERVSGYAPAYVS FVGGVVTIKNNKNSNDPTPINPSEPKVVTYGRKFVKTNQDGSERLAGATF
LVKNSQSQYLARKSGVATNEAHKAVTDAKVQLDEAVKAYNKLTKEQESODGKAALNLI DEKOTAYNEA
FAKANYSYEWVVDKNAANVVKLISNTAGKFEITGLNAGEYSLEETQAPTGYAKLSSDVSFKVNDTSYSE
GASNDIAYDKDSGKTD AQVVNKKVTIPQTGGIGTILFTIIGLSIMLGAVVIMKRRQSEEA

SEQ ID NO:46 (GBS1523 COH1)

MKKKMIQSLLVASLAFGMVAVS PVTPIAFAAETGTITVQDTQKGATYKAYKVFDAEIDNANVSDSNKDG
SYLIPQGKEAEYKASDFNSLFTTTNGGRTYVTKKDTASANEIATWAKSISANTTPVSTVTE SNNDGT
EVINVSQYGYVVSSTVNNGAVIMVTSVTPNAT IHEKNTDATWGDGGKTV DQKTYSVGDTVKYTTITYK
NAVNYHGTEKVYQYVIKDTMPSASVVDLNEGSYEVITD GSGNITTLTQGSEKATGKYNLLEENNNFTI
TIPWAATNTPGTQNGANDDFYKGIN TITVYTGVLKSGAKPGSADLPENTNIATINPNTSNDDPGQ
KVTVRDQGITIKKIDGSTKASLQGAIFVLKNATGQFLNFNDTNNVEWGTEANATEYTTGADGIITITGL

[0008]

KEGTYYLVEKKAPLGYNLLDNSQKVI LGDGATDTTNSDNLLVNPTVENNKGTELPSTGGIGTTIFYIIG
AILVIGAGIVLVARRRLRS

SEQ ID NO:47 (GBS1523 COH1 无信号序列区域)

AETGTTVQDTQKGATYKAYKVFDAEIDNANVSDSNKDGASYLIPQGKEAEYKASTDFNSLFTTTT
NGGRTYVTKKDTASANEIATWAKSISANTTPVSTVTVESNNDGTEVINVSQYGYYYVSSTVNNGAVI
MVTSVTPNATIHEKNTDATWGDGGGKTVDQKTYSVGDTVKYTITYKNAVNYHGTEKVIYQVIKDTM
PSASVVDLNEGSYEVTITDGSGNITTLTQGSEKATGKYNLLEENNNFTITI PWAATNTPGTQNG
ANDDFYKGIN TITVYTYGVLKSGAKPGSADLPENTNIATINPNTSNDDPGQKVTVRDQGITIKKI
DGSTKASLQGAIFVLKNATGQFLNFNDTNNVEWGTEANATEYTTGADGIITITGLKEGTYYLVEKK
APLGYNLLDNSQKVI LGDGATDTTNSDNLLVNPTVENNKGTE

SEQ ID NO:48 (GBS1523 COH1 位置41突变)

AETGTTVQDTKKGATYKAYKVFDAEIDNANVSDSNKDGASYLIPQGKEAEYKASTDFNSLFTTTT
NGGRTYVTKKDTASANEIATWAKSISANTTPVSTVTVESNNDGTEVINVSQYGYYYVSSTVNNGAVI
MVTSVTPNATIHEKNTDATWGDGGGKTVDQKTYSVGDTVKYTITYKNAVNYHGTEKVIYQVIKDTM
PSASVVDLNEGSYEVTITDGSGNITTLTQGSEKATGKYNLLEENNNFTITI PWAATNTPGTQNG
ANDDFYKGIN TITVYTYGVLKSGAKPGSADLPENTNIATINPNTSNDDPGQKVTVRDQGITIKKI
DGSTKASLQGAIFVLKNATGQFLNFNDTNNVEWGTEANATEYTTGADGIITITGLKEGTYYLVEKK
APLGYNLLDNSQKVI LGDGATDTTNSDNLLVNPTVENNKGTE

SEQ ID NO:49 (GBS80-GBS1523 杂合物)

MASAE TGTITVQDTQKGATYKAYKVFDAEIDNANVSDSNKDGASYLIPQGKEAEYKASTDFNSLFT
TTTNGGRTYVTKKDTASANEIATWAKSISANTTPVSTVTVESNNDGTEVINVSQYGYYYVSSTVNNG
AVIMVTSVTPNATIHEKNTDATWGDGGGKTVDQKTYSVGDTVKYTITYKNAVNYHGTEKVIYQVIK
DTMPSASVVDLNEGSYEVTITDGSGNITTLTQGSEKATGKYNLLEENNNFTITI PWAATNTPGTQNG
QNGANDDFYKGIN TITVYTYGVLKSGAKPGSADLPENTNIATINPNTSNDDPGQKVTVRDQGITI
KKIDGSTKASLQGAIFVLKNATGQFLNFNDTNNVEWGTEANATEYTTGADGIITITGLKEGTYYLV
EKKAPLGYNLLDNSQKVI LGDGATDTTNSDNLLVNPTVENNKGTEGGGGSGGGGSGGGGSELAEVS
QERPAKTTVNIYKLQADSYKSEITSNGGIENKDG EVISNYAKLGDNVKGLQGVQFKRYKVKTDISV
DELKCLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYA
VPFVLELPVANSTGTGFLSEINIYPKNVVTDEPKTKDKVKKLQDDAGYTIGEEFKWFLKSTIPAN
LGDYEKFEITDKFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFKIEAEL
LKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIENFELQYDHTPDKADNPKPSN
PPRKPEVHTGGKRFVKKDSTETQTLGGAEFDLLASDGTAVKWTDALIKANTNKNYIAGEAVTGQPI
KLKSHTDGTFEIKGLAYAVDANAEGTAVTYKPKETKAPEGYVIPDKEIEFTVSQTSYNTKPTDITV
DSADATPDTIKNNKRPS

SEQ ID NO:50 (GBS80-GBS1523 杂合物)

MASAE TGTITVQDTQKGATYKAYKVFDAEIDNANVSDSNKDGASYLIPQGKEAEYKASTDFNSLFT
TTTNGGRTYVTKKDTASANEIATWAKSISANTTPVSTVTVESNNDGTEVINVSQYGYYYVSSTVNNG
AVIMVTSVTPNATIHEKNTDATWGDGGGKTVDQKTYSVGDTVKYTITYKNAVNYHGTEKVIYQVIK
DTMPSASVVDLNEGSYEVTITDGSGNITTLTQGSEKATGKYNLLEENNNFTITI PWAATNTPGTQNG
QNGANDDFYKGIN TITVYTYGVLKSGAKPGSADLPENTNIATINPNTSNDDPGQKVTVRDQGITI
KKIDGSTKASLQGAIFVLKNATGQFLNFNDTNNVEWGTEANATEYTTGADGIITITGLKEGTYYLV
EKKAPLGYNLLDNSQKVI LGDGATDTTNSDNLLVNPTVENNKGTEGGGGGELAEVSQERPAKTTV
NIYKLQADSYKSEITSNGGIENKDG EVISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKCLTTV
EAADAKVGTILEEGVSLPQKTNAQGLVVDALDSKSNVRYLYVEDLKNSPSNITKAYAVPFVLELPV
ANSTGTGFLSEINIYPKNVVTDEPKTKDKVKKLQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEI
TDKFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFKIEAELLKGM TLVKN
QDALDKATANTDDAAFLEIPVASTINEKAVLGKAIENFELQYDHTPDKADNPKPSNPPRKPEVHT
GGKRFVKKDSTETQTLGGAEFDLLASDGTAVKWTDALIKANTNKNYIAGEAVTGQPIKLKSHTDGT
FEIKGLAYAVDANAEGTAVTYKPKETKAPEGYVIPDKEIEFTVSQTSYNTKPTDITVDSADATPDT
IKNNKRPS

SEQ ID NO:51 (GBS80-GBS1523 杂合物)

MASAE TGTITVQDTKKGATYKAYKVFDAEIDNANVSDSNKDGASYLIPQGKEAEYKASTDFNSLFTTTT
NGGRTYVTKKDTASANEIATWAKSISANTTPVSTVTVESNNDGTEVINVSQYGYYYVSSTVNNGAVIMVT

[0009]

SVTPNATIHEKNTDATWGDGGGKTVDQKTYSVGDTVKYTITYKNAVNYHGTEKQYQYVIKDTMPSASV
 DLNEGSYEVTITDGSNIITLTQGSEKATGKYNLLENNFTITIPWAATNTPGTQNGANDDFYKGI
 INTITVYTYGVLKSGAKPGSADLPENTNIATINPNTSNDPPGQKVTVRDQGITIKKIDGSTKASLQGA
 FVLKNTAQGLNFNDTNNVEWGTEANATEYTTGADGITITGLKEGTYLVEKKAPLGYNLLDNSQKVI
 LGDGATDTTNSDNLVNPTVENNKGTEGGGGSGGGGGGGSELAEVSQERPAKTTVNIYKLQADSYSK
 EITSNGGIENKDGVEISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKKLTTVEAADAKVGTILEEGV
 SLPQKTNAOGLVVDALDSKSNVRYLYVEDLKNSPSNIKAYAVPFVLELPPVANSTGTGFLSEINIYKPN
 VVTDEPKTDKDVKKLQDDAGYTI GEEFKWFLKSTIPANLGDYEKFEITDKFADGLTYKSVGKIKIGSK
 TLNRDEHYTIDEPTVDNQNTLKITFKPEKFKIEAELLKGMTLVKNQDALDKATANTDDAAFLEIPVAST
 INEKAVLGKAIENTFELQYDHTPDKADNPKPSNPPRKPVEVHTGGKRFVKKDSTETQTLGGAEFDLLASD
 GTAVKWTDALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKGLAYAVDANAEGTAVTYKLETKAPE
 GYVIPDKEIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS

SEQ ID NO:52 (GBS80-GBS1523 杂合物)

MASAEVTITVQDTKKGATYKAYKVFDAEIDNANVSDSNKDGASYLIPQGKEAEYKASTDFNSLFTTTT
 NGGRTYVTKKDTASANEIATWAKSISANTT PVSTVTESSNDGTEVINVSQYGYVVSSTVNNGAVIMVT
 SVTPNATIHEKNTDATWGDGGGKTVDQKTYSVGDTVKYTITYKNAVNYHGTEKQYQYVIKDTMPSASV
 DLNEGSYEVTITDGSNIITLTQGSEKATGKYNLLENNFTITIPWAATNTPGTQNGANDDFYKGI
 NTITVYTYGVLKSGAKPGSADLPENTNIATINPNTSNDPPGQKVTVRDQGITIKKIDGSTKASLQGA
 FVLKNTAQGLNFNDTNNVEWGTEANATEYTTGADGITITGLKEGTYLVEKKAPLGYNLLDNSQKVI
 LGDGATDTTNSDNLVNPTVENNKGTEGGGGSGGGGGGGSELAEVSQERPAKTTVNIYKLQADSYSK
 EITSNGGIENKDGVEISNYAKLGDNVKGLQGVQFKRYKVKTDISVDELKKLTTVEAADAKVGTILEEGV
 SLPQKTNAOGLVVDALDSKSNVRYLYVEDLKNSPSNIKAYAVPFVLELPPVANSTGTGFLSEINIYKPN
 VVTDEPKTDKDVKKLQDDAGYTI GEEFKWFLKSTIPANLGDYEKFEITDKFADGLTYKSVGKIKIGSK
 TLNRDEHYTIDEPTVDNQNTLKITFKPEKFKIEAELLKGMTLVKNQDALDKATANTDDAAFLEIPVAST
 INEKAVLGKAIENTFELQYDHTPDKADNPKPSNPPRKPVEVHTGGKRFVKKDSTETQTLGGAEFDLLASD
 GTAVKWTDALIKANTNKNYIAGEAVTGQPIKLSHTDGTFEIKGLAYAVDANAEGTAVTYKLETKAPE
 GYVIPDKEIEFTVSQTSYNTKPTDITVDSADATPDTIKNNKRPS

SEQ ID NO:53 (GBS104 2603)

MKKRQKIWRGLSVTLILLSQIPFGILVQGETQDTNQAALGKIVKKTGDNATPLGKATFVLKNDNDKSET
 SHETVEGSGEATFENIKPGDYTLREETAPIGYKKTDKTWKVKVADNGATIEGMDADKAEKRKEVLNAQ
 YPKSAIYEDTKENYPLVNVEGSKVGEQYKALNPNINGKDRREIAEGWLSKKITGVNDLKNKYKIELTV
 EGKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVALVYASTI
 FDGTEATVSKGVADQNGKALNDSVSWDYHKTFTTATTHNYSYLNLTNDANEVNIKSRIPKEAEHINGD
 RTLYQFGATFTQKALMKANEILETQSSNARKKLI FHVTDGCVPTMSYAINFNPIYISTSYQNQFNSFLNKI
 PDRSGILQEDFIINGDDYQIVKGDGESFKLFSRDKVPEVTGGTQAAYRVPQNQLSVMSNEGYAINSGYI
 YLYWRDYNWVYFPDPKTKKVSATKQIKTHGEPTTYENGNIRPKGYDIFTVIGVNGDPGATPLEAEKF
 MQSISSKTENYTNVDDTNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLV
 GNDGSSQLKNGVALGGPNSDGGILKDVTVYDKTSQTIKINHLNLGSGQKVVLTYDVRKLDNYISNKFYN
 TNNRTLSPKSEKEPNTIRDFPIPKIRDVREFFVLTISNQKKMGEVEFIKVNKDKHSESLGAKFQLQI
 EKDFSGYKQFVPEGSVDTTKNDGKIYFKALQDGNKYLYEISSPDGYIEVKTKPVVTFITQNGEVTNLKA
 DPNANKNQIGYLEGNGKHLITNTPKRPPGVFPKTTGGIGTIVYILVGSFMILTICSFRKQL

SEQ ID NO:54 (GBS1524)

MLKKCQTFIIESLKKKKHPKEWKIIMWSLMILTTFLTTYFLILPAITVEETKTDVVGITLENKNSSQVT
 SSTSSSQSSVEQSKPQTPASSVTETSSSEEAAYREPLMFRGADYTVTVTLTKEAKIPKNADLKVTELK
 DNSATFKDYKKKALTEVAKQDSEIKNFKLYDITIE SNGKEAEPQAPVKVEVNYDKPLEASDENLKVVHF
 KDDGQTEVLKSKDQTAETKNTSSDVAFKTDSFSIYAIVQEDNTEVPRLTYPHFQNNNDGTDYDFLTASGMQV
 HHQIIKDGESLGEVGIPTIKAGEHFNGWYTYDPTTGKYGDPVKFGEPIVTVTETKEICVRPFMSKVATVT
 LYDDSAKGSILERYQVPLDSSGNGTADLSSFVSPPTSTLLFVGSKTQNGAPLSESEIQALPVSSDIS
 LYPVFKESYGVFNTEGDLSTGVYIAPRRVLTGQAPASTIKPNDPTRPGYTFAGWYTAASGGAAFDNQV
 LTKDITLYAHWSPAQTTYTINYWQQSATDNKNATDAQTYEYAGQVTRSGLSLSNQTLTQQDINDKLP
 GFKVNNRTRTETSMIKDDGSSVVNVYDRKLI TIKFAKYGGYSLPEYYSYNWSSDADTYTGLYGTTLA
 ANGYQWKTGAWGYLANVGNQVGTYGMSYLGELFIPNDTVSDVILKFPKGNIVQTYRFFKQGLDGTYS
 LADTGGGAGADEFTTEKYLGFNVKYYQRLYPDNYLFDQYASQTSAGVKVPI SDEYYDRYGAYHKDYL
 LVVWYERNYSYIKYLDPLDNTELPNFPVKDVLVEQNLSSYAPDTTTVQPKPSRPGYVWDGKWKDQAQT
 QVDFDNTMPPHDVKVYAGWQKVYRNVNIDPNCGRLSKTDYLYDLHYGDRI PDYTDITRDYIQDPSGT
 YYYKYDSRDKDPDSTKDAYTTDTLSNVDTTKYKYVKDAYKLVGWYVNPDSIRPYNFSGAVTQDI
 NLRAIWRKAGDYHIYSNDAVGTGKPALDASGQQLQTSNEPTDPDSYDDGSHSALLRPTMPDGYRFR

[0010]

GWVYNGKIYNPYDSIDIDAHLADANKNITIKPVIIPVGDIKLEDTSIKYNGNGGTRVENGNNVVTQVETP
 RMELNSTTIPENQYFTRTGYNLIGWHHDKDLADTGRVEFTAGQSIGIDNNPDATNTLYAVWQPKEYTV
 RVSKTVVGLDEDKTKDFLNPSETLQQENFPLRDGQTKFEKVPYGTSSISIDEQAYDEFKVSSEITEKNL
 ATGEADKTYDATGLQSLTVSGDVIDISFTNTRIKQKQVRLQKVVNENDNNFLAGAVFDIYESDANGNKASH
 PMYSGLVTNDKGLLLVDANNYLSLPVKGYYLTETKAPPGYLLPKNDISVLVISTGVTFEQNGNATPIK
 ENLVDGSTVYTFKITNSKGTLPSTGGIGTHIYILVGLALALPSGLILYRKKI*

SEQ ID NO:55 (GBS3 2603)

MNNNEKKVKYFLRKTAYGLASMSAAFAVCSGIVHADTSSGISASIPHKKQVNLGAVTLKNLISKYRGND
 KAIAILLSRVNDNFNRASQDTLPQLINSTEAEIRNILYQGQIGKQNKPSVTTTHAKVSDQELGKQSRRSQD
 IIKSLGFLSSDQKDIILVKSISSSKDSQLILKFVTQATQLNNAESTKAKQMAQNDAVALIKNISPEVLEEY
 KEKIQRASTKSQVDEFVAEAKKVVNSNKETLVNQANGKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNE
 GKLNI TAAMNALNSIKQAAQEVAQKNLQKQYAKKIERISSKGLALS KKAKEIYEKHK SILPTPGYYADS
 VGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKMLDEVKLLKELQDLTRGTKEDKKPDVKPEAKPEAKPD
 VKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPDVKIVSLIVMLSAGLITIVLKHKKK

SEQ ID NO:56 (GBS3 2603 无信号序列区域)

TSSGISASIPHKKQVNLGAVTLKNLISKYRGNDKAIAILLSRVNDNFNRASQDTLPQLINSTEAEIRNIL
 YQGQIGKQNKPSVTTTHAKVSDQELGKQSRRSQDI IKS LGFLSSDQKDIILVKSISSSKDSQLILKFVTQA
 TQLNNAESTKAKQMAQNDAVALIKNISPEVLEEYKEKIQRASTKSQVDEFVAEAKKVVNSNKETLVNQAN
 GKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNEGKLNI TAAMNALNSIKQAAQEVAQKNLQKQYAKKIE
 RISSKGLALS KKAKEIYEKHK SILPTPGYYADSVGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKMLDEV
 KLLKELQDLTRGTKEDKKPDVKPE

SEQ ID NO:57 (GBS3 2603 卷曲螺旋和富含脯氨酸区段)

TSSGISASIPHKKQVNLGAVTLKNLISKYRGNDKAIAILLSRVNDNFNRASQDTLPQLINSTEAEIRNIL
 YQGQIGKQNKPSVTTTHAKVSDQELGKQSRRSQDI IKS LGFLSSDQKDIILVKSISSSKDSQLILKFVTQA
 TQLNNAESTKAKQMAQNDAVALIKNISPEVLEEYKEKIQRASTKSQVDEFVAEAKKVVNSNKETLVNQAN
 GKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNEGKLNI TAAMNALNSIKQAAQEVAQKNLQKQYAKKIE
 RISSKGLALS KKAKEIYEKHK SILPTPGYYADSVGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKMLDEV
 KLLKELQDLTRGTKEDKKPDVKPEAKPEAKLEAKPEAKPATKKS VNTSGNLAAKKAIENKKYSK

SEQ ID NO:58 (GBS3 2603 信号序列和卷曲螺旋)

MNNNEKKVKYFLRKTAYGLASMSAAFAVCSGIVHADTSSGISASIPHKKQVNLGAVTLKNLISKYRGND
 KAIAILLSRVNDNFNRASQDTLPQLINSTEAEIRNILYQGQIGKQNKPSVTTTHAKVSDQELGKQSRRSQD
 IIKSLGFLSSDQKDIILVKSISSSKDSQLILKFVTQATQLNNAESTKAKQMAQNDAVALIKNISPEVLEEY
 KEKIQRASTKSQVDEFVAEAKKVVNSNKETLVNQANGKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNE
 GKLNI TAAMNALNSIKQAAQEVAQKNLQKQYAKKIERISSKGLALS KKAKEIYEKHK SILPTPGYYADS
 VGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKMLDEVKLLKELQDLTRGTKEDKKPDVKPEAKPEAKPD
 VKPEAKPDVKPEAKPDVKPEAK

SEQ ID NO:59 (GBS3 2603 信号序列区段)

TSSGISASIPHKKQVNLGAVTLKNLISKYRGNDKAIAILLSRVNDNFNRASQDTLPQLINSTEAEIRNIL
 YQGQIGKQNKPSVTTTHAKVSDQELGKQSRRSQDI IKS LGFLSSDQKDIILVKSISSSKDSQLILKFVTQA
 TQLNNAESTKAKQMAQNDAVALIKNISPEVLEEYKEKIQRASTKSQVDEFVAEAKKVVNSNKETLVNQAN
 GKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNEGKLNI TAAMNALNSIKQAAQEVAQKNLQKQYAKKIE
 RISSKGLALS KKAKEIYEKHK SILPTPGYYADSVGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKMLDEV
 KLLKELQDLTRGTKEDKK

SEQ ID NO:60 (GBS3 2603 信号序列、卷曲螺旋和富含脯氨酸区段)

[0011]

MNNNEKKVKYFLRKTAYGLASMSAAFVCSGIVHADTSSGISASIPHKQVNLGAVTLKNLISKYRGND
KAIAILLSRVNDFNRASQDTPQLINSTEAEIRNILYQGGIGKQNKPSVTHAKVSDQELGKQRRSQD
I I KSLGFLSSDQKDILVKSISSSKDSQLILKFVQTATQLNNAESTKAKQMAQNDVALIKNISPEVLEEY
KEKIQRASTKSQVDEFVAEAKKVNSNKETLVNQANGKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNE
GKLNITAMNALNSIKQAAQEVAKNLQKQYAKKIERISSKGLALSCKAKEIYEKHKSI LPTPGYYADS
VGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKMLDEVKLLKELQDLTRGTKEDKKPDVKPEAKPEAKPD
VKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPDVKPEAKPEAKPEAKSEAKPEAK
LEAKPEAKPATKKSVENTSGNLAACKAIENKKYSKK

SEQ ID NO:61 (GBS3 515)

MNNNEKKVKYFLRKTAYGLASMSAAFVCSGIVHADTSSGISASIPHKQVNLGAVTLKNLISKYRGND
KAIAILLSRVNDFNRASQDTPQLINSTEAEIRNILYQGGIGKQNKPSVTHAKVSDQELGKQRRSQD
I I KSLGFLSSDQKDILVKSISSSKDSQLILKFVQTATQLNNAESTKAKQMAQNDVALIKNISPEVLEEY
KEKIQRASTKSQVDEFVAEAKKVNSNKETLVNQANGKKQEI AKLENLSNDEMLRYNTAIDNVVKQYNE
GKLNITAMNALNSIKQAAQEVAKNLQKQYAKKIERISSKGLALSCKAKEIYEKHKSI LPTPGYYADS
VGTYLNRFRDKQTFGNRSVWTGQSGLDEAKKNA

SEQ ID NO:62 (GBS3 cjb111)

MNNNEKKVKYFLRKTAYGLASMSAAFVCSGIVHADTSSGISDSIPHKQVNLGAVTLKNLISKYRGND
KAIAILLSRVDDFNRSQDTPQLINSTEAEINNTLPQGR I I KQSI PVVRLKVERLGSGAIKAESINNI
KAESINKIQGKSTNTIKAESINKIKVESINTIKAESINKIQAKPINTIKAESINTIKAESIHKIKPQSI
KSTSATHVKVSDQELAKQRRSQDI I KSLGFLSSDQKDILVKSISSSKDSQLILKFVQTATQLNNAEST
KAKHMAQNDVASIKNISLEVL EYKEKIQRASTKSQVDELVAEAKKVNSNKETLVNQANGKKQEI AKL
ENLSNDEMLRYNTAIDNVVKQYNEGKLNITDAMNALNSIKQAAQEVAKNLQKQYAKKIERI SLKGLAL
SKKAKEIYEKHKSI LPTPGYYADSVGTYLNRFRDKRTFGNRSVWTGQSGLDEAKKNA

SEQ ID NO:63 (GBS3 coh1)

MNNNEKKVKYFLRKTAYGLASMSAAFVCSGIVNTPTVSADSPDTLKVEKLGKLDVKSVHELTPISIP
NELKGAKEQALSSII SHPNITNSEVDKLASDYSFRINTSNDVNDVKRLLEFYNVAVARKQLDNTSADYR
SKIDNISTTGLAIALEAKEIYEANKSILPHRYKDSVGTYVNSFEERRSPGKFNIWNGQEGFNAAQKLE
DVKKLLELQNLTKNNKPNIQVPKQAPTEAAKPALESPEALTRLTWYNQAKDLLKDDQVKDYVDILSV
QKAVDQAYDHVEEGKFITTDQANQLANKLRDALQSLELKDKKVAKPVAKGTYDVKYVDTEGKEVAKSRH
FEGEEGAAFVTSAKEVAGYKLVRTGAVSNVFTAGAQVRTYVYEKVKPEVKPDVKPEAKPEAKPEVKPD
VKPEAKPEAKPEVKS DVKPEAKPEAKPEAKPEVKPDVKPEAKPEAKPATKKSVENTSGNLVAKKAIENKK
YSKKLPSTGEAASPLLAIVSLIVMLSAGLITIVLKHKKK

SEQ ID NO:64 (SAN1485 coh1)

MLKKQFGNFGKESRKRVRVCMRKSGKHVKSVMTQIGYVILSRFSGKEKSSKVQTTSEDLRKTASIL
TAVAALGAVVGGTTDTTSVSAEETPTATELTGNEKTATAETVVVAPEVKT VNSDSSSHSTSESQSMST
STLQSTSASLSASESLMDSTSASLSESSSLSEYSSLSSSESVSASESVQSSEAAATTARVQPRAMRVV
SSASDMETLPAALISGEGDVTTVQGDVTDKQLNDIKLSGGVQAKAGVINMDKSESMHMSLKFIDSV
NRGDTFEIKLSDNIDTNGASNYSIVEPIKSPTEGYATGIYDSQKKSIVYSFTDFAASKNNINGILDIP
LWPDDTTVQNTKEDVLFVSKIKDQEATIKETVKYDPPVRIDFAGGVSVDSRITNIDVVGKMTYISQIN
VDGKSLYNYNGLYTRIYNYSKESTADLKNSTIKIYKTTSDNIVESMVQDYSSMEDVTSKFANSYPEKGW
YDIYWGQFIASNETYVIVVETPFTNAVTLNNTLSDYNNNGVEHNHTYSSSESGYSDVNAQERKILSELV
SSSESVSSESVSNSESI STSESVSNSESI SSSSESVSSESISTSESVSTSESISSESVSSESVSSES
ESISSESVSNSESI SSSSESVSNSESI SSSSESVSSESISNSESI SSSSESVSTSESISSESVSNSESI
SSSESVSSESVSSTSESISSESVSNSESI STSESVSTSESISSESVSSESISSESVSNSESI SSSSESV
ESVSSSESVSNSESI SSSSESVSNSESI STSESVSTSESISSESVSNSESI SSSSESVSNSESI SSSSESV
SNSESI SSSSESVSNSESI SSSSESVSSESISSTSESVSTSESISSESVSNSESI SSSSESVSNSESI
ESISSESVSNSESI SSSSESVSSESISSESVSSESVSNSESI SSSSESVSNSESI SSSSESVSSESIS
SSSESVSNSESI SSSSESVSSESISSESISSESISSESISSESISSESISSESISSESISSESISSESIS
HTEIKSDSESQHEVKHQVLPETGDNSASALGLLGAGLLL GATKSRKKKKD

SEQ ID NO:65 (GBS147 2603)

MDKHHSKKAILKLTTLITTSILLMHSNQVNAEEQELKNQEQSPVIANVAQQPSPSVTNTVEKTSVTAAS
ASNTAKEMGDTSVKNDKTEDELLEELSKNLDTSNLGADLEEEYPSKPETTNNKESNVVTNASTAIAQKV

[0012]

PSAYEEVKPESKSSLAFLDTSKITKLOAITQRGKGNVVAIIDTGFDINHDI FRLDSPKDDKHSFKTKTE
 FEELKAKHNI TYGKWVNDKIVFAHNYANNTETVADIAAAMKGYGSEAKNISHGTHVAGIFVGNKRPA
 INGLLEGAAPNAQVLLMRI PDKIDSDKFEAYAKAITDAVNLGAKTINMSIGKTADSLIALNDKVKLA
 LKLASEKGVAVVVAAGNEGAFGMDYSKPLSTNPDYGTVNSPAISEDTL SVASYESLKTISEVVETIEG
 KLVKLP I VTSKPFDKGKAYDVVYANYGAKKDFEGKDFKGI ALIERGGGLDFMTKI THATNAGVVGIVI
 FNDQEKRGNF LI PYREL PVGI I SKVDGERIKNTSSQLTFNQSF EVVDSQGGNRMLEQSSWGVTAEGAIK
 PDVTASGF E IYSSTYNNQYQ TMSGTSMASPHVAGLMTMLQSHLAEKYKGMNLD SKKLELSKNILMSSA
 TALYSEEDKAFYS PRQQGAGVVD AEA KAIQAQY I TGN D GKAKINLKRMGDKFDITVTI HKLVEGVKELY
 YQANVATEQV NKGK FALKPQALLD TNWQKVI LRDKETQVRF TIDASQFSQKLKEQMANGY FLEGFVRFK
 EAKDSNQELMSIPFVGFNGDFANLQALE TPIYKTL SKGSFYKPN DTHK DQLEYNESAPFESNNTAL
 LTQSASWGYVDYVKN GGELELAPESPKRI ILGT FENKVEDKTI HLLERDAANNPYFAISP NKDGNRDEI
 TPQATFLRNVKDI SAQVLDQNGNVIWQSKVLP SYRKNFHN NPKQSDGHYRMDALQWSGLDKDGKVVADG
 FYTYRLRYTPVAEGANSQESDFKVQVSTKSNLPSRAQFDETNRTL SLAMPKESYVPT YRLQLVLSHV
 VKDEEYGD ETSYHYFHIDQEGKVTLPKTVKIGESEVAVDPKAL TLVVEDKAGNFATVKLSDLLNKAVVS
 EKENAIVISNSFKYFDNLKKEPMFISKKEKVVNKNLEE I ILVKPQTTVTTQSLSK EITKSGNEKVL TST
 NNNSSRVAKIISPKNHGD SVNHTLPSTSDRATNGLFVGT LALLSLLLYLKP KKTKNNSK

SEQ ID NO: 66 (GBS328 2603)

MKKKIILKSSVLGLVAGTSIMFSSVFADQVGVQVIGVND F HGALDNTGTANMPDGKVANAGTAAQLDAY
 MDDAQKDFKQTNPNGESIRVQAGDMVGAS PANSGLLQDEPTVKNFNAMNVEYGT LGNHEFDEGLAEYNR
 IVTGKAPAPDSNINNITKSY PHEAAKQEI VVANVIDKVNKQI PYNWKPYAIKNI PVNNSVNVGFIGIV
 TKDIPNLVLRK NYEQYEF LDEAETIVKYAKELQAKNVKAI VVLAHV PATSKNDIAEGEAAEMMKVNQL
 FPENSVDIVFAGH NHQYTNGLVGKTRIVQALSQ GKAYADVRGVLDTDTQDFIETPSAKVIAVAPGKKTG
 SADIQAIVDQANTIVKQVTEAKIGTAEVSVMI TRSVDQDNVSPVGS LITEAQLAIARKSWPDI DFAMTN
 NNGIRADLLIKPDGTTIWGAAQAVQ PFGNI LQVVEITGRDLYKALNEOYDQKONFFLOIAGLR YTYTDN
 KEGGEETPFKVVKAYKSN GEEINPDAKYKLVINDFLFGGGDFASFRNAKLLGAINPDTEVFMAYITDL
 EKAGKVSVPNNKPKIYVTMKMVNETITQNDGTHS I IKKLYLDRQGNIVAQEIVSDTLNQT KSKSTKIN
 PVTTIHKKQLHQFTAINPMRNYGKPSNSTTVKSKQLPKNSEYQSF LMSVFGVGLIGIALNTKKKHKM

SEQ ID NO: 67 (GBS84 2603)

MKGQKI IALAGLVLSCHFALTACHTQEHKNSHHI KTKQVAKKKANKKKVSVKESHKKRKG VAGVDFPTD
 DGFLLT KDSKILSHTDSGIVVAHGNSHFIFYS DLKGSKFSYLI PNANTKT NKNQAVRNFKAGAVAVNT
 LNDGYVFN PADIVSEDANGYVVRHGDHFHYI PKASLSQKQVQASRAVSR LGNQSN SHYRVNS SKIAGL
 HYPTSNGLFNGRGIGTTP TGI LVEHHNLHFISFADLRKGGWGS IADRYQPQKKADSKKQSPSSKKP
 RTENTLPKDIKDKLAYLARELHLDISRIRVLKTLN GEIGFEYPHDDH THVIMAKDIDL SKPIPNPHDD
 EDH HKGHHHDES DHKHEEHEHTSKNLSDE DQKLIYLAEKLG LNPNQIEVLT SEDGSIIFKYPHDDHS
 HTIASKDIEIGKPIPDGHH DSHAKDKVGMATLKQIGFDD EIIQDILHADAPT PFP SNETNPEKMRQWL
 ATVTKINIGQRTNPFQRFGLSLMPNIEVLGIGFT PINDMTPVLQFKK LKQLWMTNTGITDYSFLDKMPL
 LEGLDISQNGIKDLSFLTKYKQLSLIAAANNGITSLKPLAELPNLQFLVLSHNNISDLT PLSNLT KLOE
 LYLDHNNVKNLSALS GKKDLKVL DLSNNSADLSTLKTTSLETLLN ETNTSNLSFLKQNPKVSNL TIN
 NAKLASLDGIEESDEIVKVEAEGNQIKSLVLKNKQGS LKFLNVTNNQLTSLEGVNNYTSLETLSVSKNK
 LESLDIKTPNKTVTNLDFSHNNVPTS QLKLEKNIPEAVAKNFP AVVEGSMVGNGLAEKAAMASKEDK
 QVSDNTNHQNKTEKSAQANADSKKENPKTHDEHHDHEETDHAHVGH HHH

SEQ ID NO: 68 (引物)

CTGTACTTCCAGGGCAATACCAATGTTT TAGGGGAA

SEQ ID NO: 69 (引物)

AATTAAGTCGCGTTATTTCCACTGACAGTTAACTC

SEQ ID NO: 70 (引物)

CTGTACTTCCAGGGCACCATAGTAAAACCACTGGAC

SEQ ID NO: 71 (引物)

AATTAAGTCGCGTTATCCATTACCAAGCTGTAAATF

SEQ ID NO: 72 (引物)

CTGTACTTCCAGGGCCAAACAT TACAGCCAAGTGAT

[0013]

SEQ ID NO:73 (引物)

AATTAAGTCGCGTTATCCTTTCCACCTGTCATAGG

SEQ ID NO:74 (引物)

CTGTACTIONCCAGGGCAATACCAATGTTTTAGGGGAA

SEQ ID NO:75 (引物)

AATTAAGTCGCGTTATTTACCGCTAACAGTTAACTC

SEQ ID NO:76 (引物)

CTGTACTIONCCAGGGCTCCATAATAAAAACTATAAATAAAG

SEQ ID NO:77 (引物)

AATTAAGTCGCGTTATCCGTTGCCAAGATGTAAATT

SEQ ID NO:78 (引物)

CTGTACTIONCCAGGGCCAAACATTGCAACCAAGTGAT

SEQ ID NO:79 (引物)

AATTAAGTCGCGTTATCCTTTCCACCTGTCATC

		Fr. 1	Fr. 2	Fr. 3
		1		50
SAG_1408 2603	(1)	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDQ		
SAI_1512 H36B	(1)	MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDQ		
		51		100
SAG_1408 2603	(51)	NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP		
SAI_1512 H36B	(51)	NKPLSKATFVLKPTSHSESKVEKVTEVTGEATFDNLTPGDYTLSEETAP		
		101		150
SAG_1408 2603	(101)	EGYKKTQQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPPTGIYE		
SAI_1512 H36B	(101)	EGYKKTQQTWQVKVESNGKTTIQNSDDKKSIIEQRQEELDKQYPLTGAYE		
		151		200
SAG_1408 2603	(151)	DTKESYKLEHVKGSYPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV		
SAI_1512 H36B	(151)	DTKESYNLEHVKNSIPNGKLEAKAVNPYSSEGEHIREIQEGTLSKRISEV		
		201		250
SAG_1408 2603	(201)	GDLAHNKYKIELTVSGKTIIVKPVDKQKPLDVVFLDNSNSMNDGPNFQR		
SAI_1512 H36B	(201)	NDLDHNKYKIELTVSGKSIIKTINKDEPLDVVFLDNSNSMKNNG---K		
		251		300
SAG_1408 2603	(251)	HNKAKRAAEALGTAVKDILGANSNDRVALVTYGSDFDGRSVDVVKGFKE		
SAI_1512 H36B	(247)	NNKAKKAGEAVETIIKDV LGANVENRAALVTYGSDFDGRTVKVIKGFKE		
		301		350
SAG_1408 2603	(301)	DDKYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTNGL		
SAI_1512 H36B	(297)	DP-YYGLETSFTVQTNDSYKFTNIAADIKKIPKEAPEAKWGTSGLGL		
		351		400
SAG_1408 2603	(351)	TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVT DGVPTR		
SAI_1512 H36B	(346)	TPEKKREYDLSKVGETFTMKAFMEADTLSSIQRKSRIIVHVT DGVPTR		
		401		450
SAG_1408 2603	(401)	SYAINNFKLGASYESQFEQMKNNGYLNKSNFLLTDKPEDIKNGESYFLF		
SAI_1512 H36B	(396)	SYAINSFVKGSTYANQFERIKEKGYLDKNNYFITDDPEKIKNGESYFLF		
		451		500
SAG_1408 2603	(451)	PLDSYQTQIIISGNLQKLHYLDLNLNYPKGTIYRNGPVKEHGTPTKLYINS		
SAI_1512 H36B	(446)	PLDSYQTQIIISGNLQKLHYLDLNLNYPKGTIYRNGPVREHGTPTKLYINS		
		501		550
SAG_1408 2603	(501)	LKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGT FQKLKEEAFKLSDGEIT		
SAI_1512 H36B	(496)	LKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGT FQKLKEEAFELSDGEIT		

图 1

		551		600
SAG_1408 2603	(551)	ELMRSFSSKPEYYTPIVTSADTSNNEILSKIQQQFETILTKENSIVNGTI		
SAI_1512 H36B	(546)	ELMNSFSSKPEYYTPIVTSADVSNNEILSKIQQQFEKILTKENSIVNGTI		
		601		650
SAG_1408 2603	(601)	EDPMGDKINLQLGNGDTLQPSDYTLQNDGSVMKDG IATGGPNNDGGILK		
SAI_1512 H36B	(596)	EDPMGDKINLHLGNGDTLQPSDYTLQNDGSIMKDSIATGGPNNDGGILK		
		651		700
SAG_1408 2603	(651)	GVKLEYIGNKLYVRGLNLCEGQKVTLYDVKLD DSFISNKFYDTNGRITL		
SAI_1512 H36B	(646)	GVKLEYIKNKLYVRGLNLCEGQKVTLYDVKLD DSFISNKFYDTNGRITL		
		701		750
SAG_1408 2603	(701)	NPKSEDPNTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNK L		
SAI_1512 H36B	(696)	NPKSEPDTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFTKVDKDNK L		
		751		800
SAG_1408 2603	(751)	LLKGATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLI		
SAI_1512 H36B	(746)	LLKGATFELQEFNEDYKLYLPIKNNNSKVVTGENGKISYKDLKDGKYQLI		
		801		850
SAG_1408 2603	(801)	EAVSPEDYQKITNKPILTFEVVKGSIKNI IAVNKQISEYHEEGDKHLITN		
SAI_1512 H36B	(796)	EAVSPKDYQKITNKPILTFEVVKGSIQNI IAVNKQISEYHEEGDKHLITN		
		851		900
SAG_1408 2603	(851)	THIPKGIIPMTGGKILS FILIGGAMMSIAGGIYIWKRYKKSSDMSIKK		
SAI_1512 H36B	(846)	THIPKGIIPMTGGKILS FILIGGAMMSIAGGIYIWKRRHKSSDASIEK		
		901		
SAG_1408 2603	(901)	D-		
SAI_1512 H36B	(896)	D-		

图 1(续)

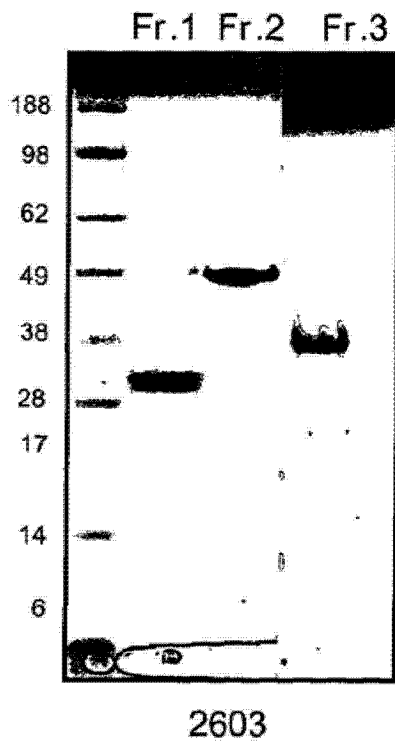


图 2A

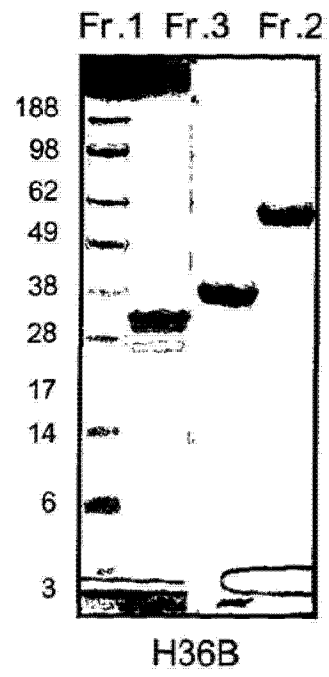


图 2B

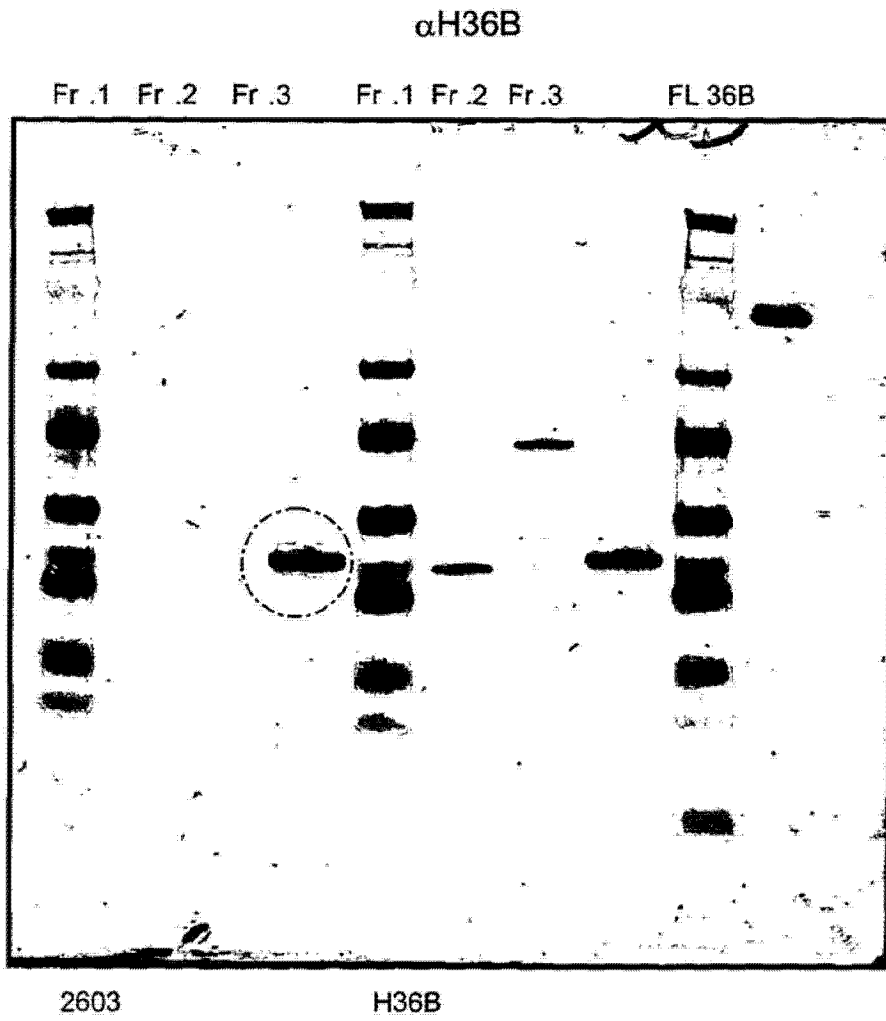
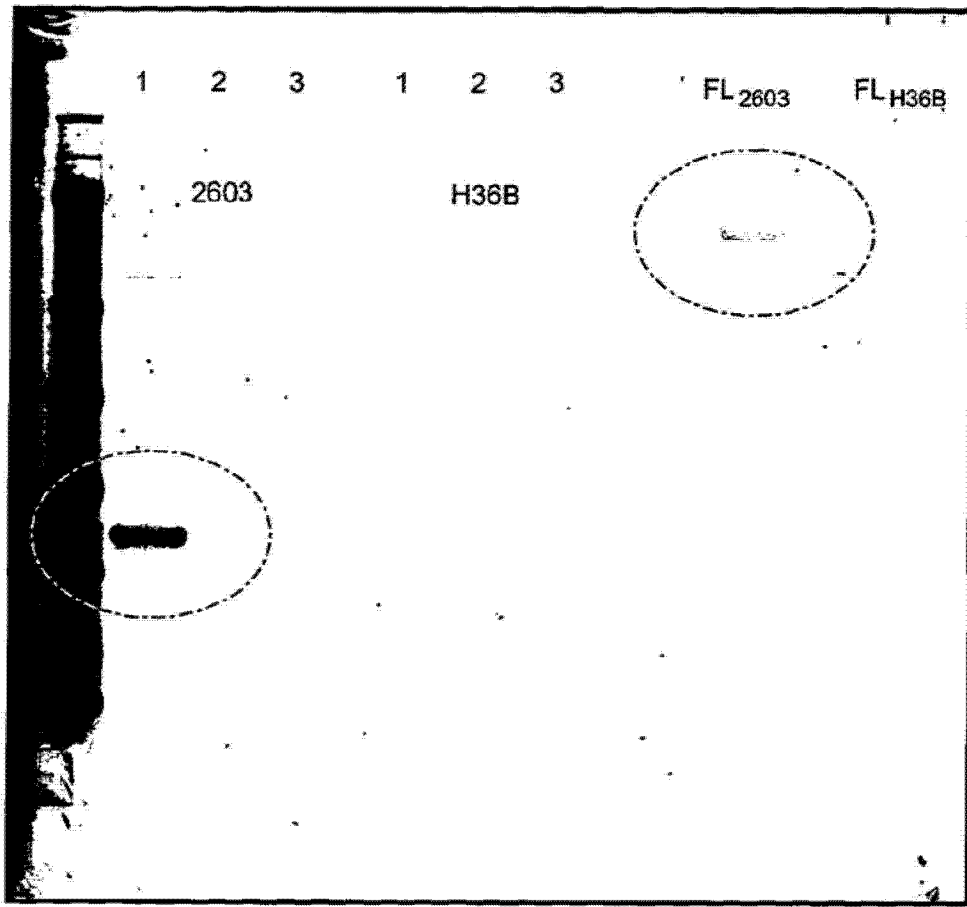
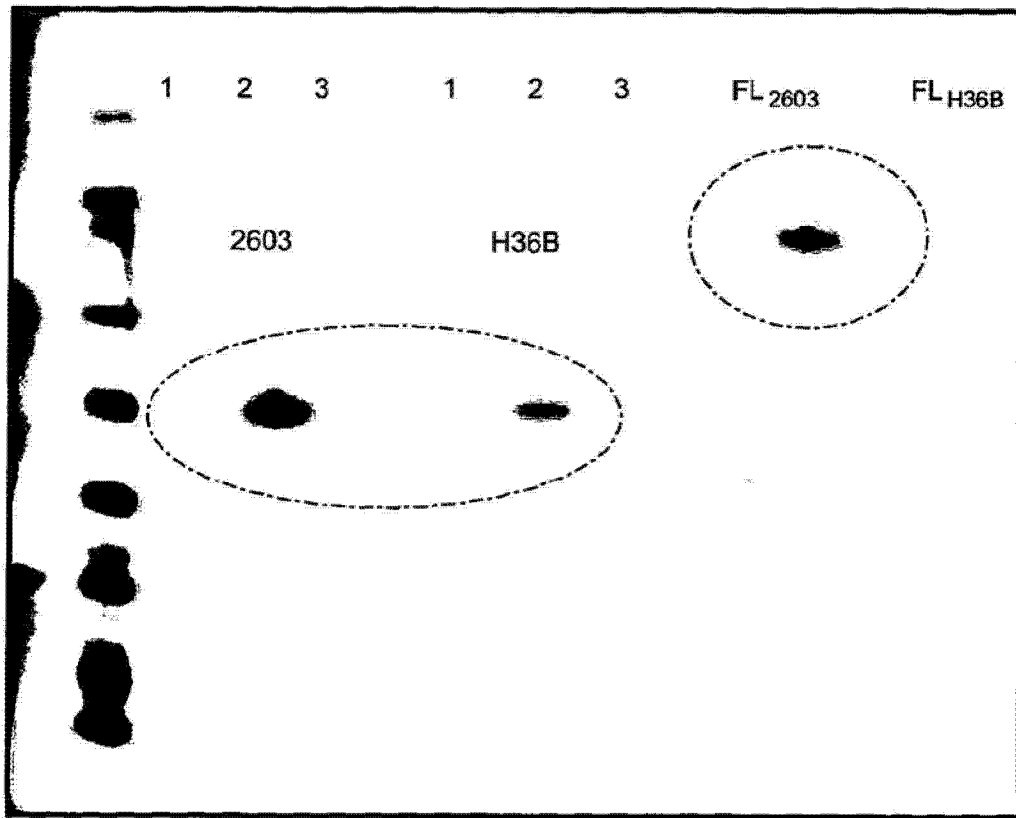


图 3B



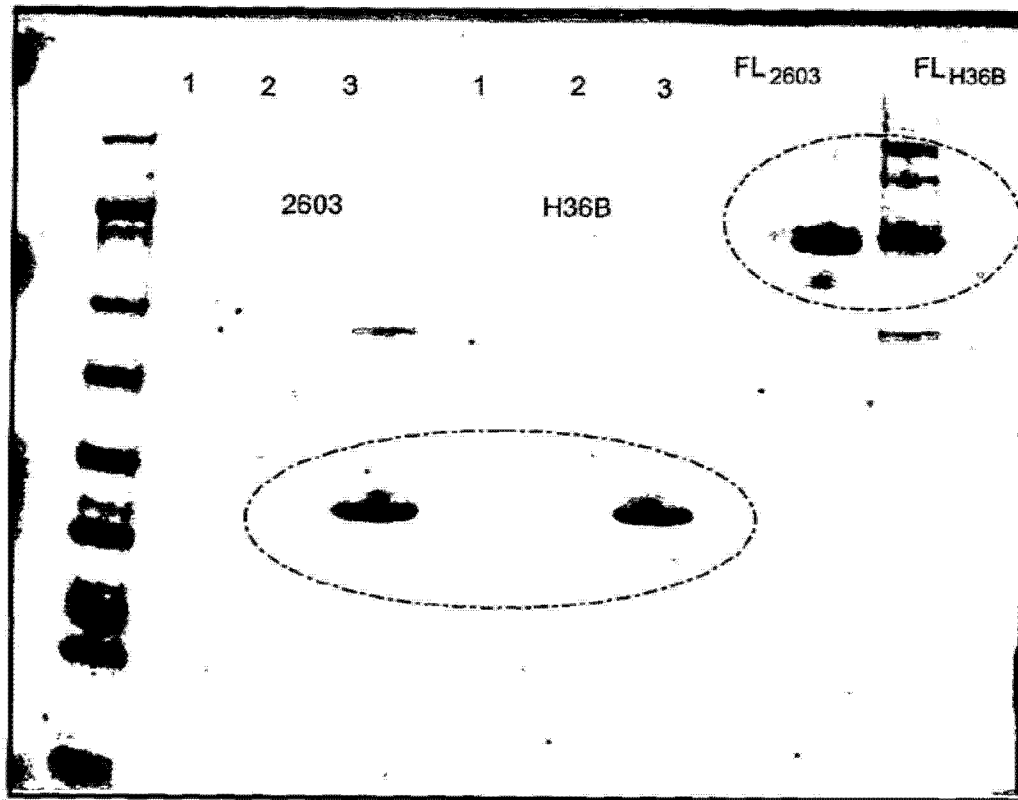
α 片段 1 (2603)

图 4



α 片段 2 (2603)

图 5A



α 片段 3 (2603)

图 5B