

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 02822187.7

[51] Int. Cl.

C07K 14/22 (2006.01)

C12N 15/62 (2006.01)

A61K 38/16 (2006.01)

[45] 授权公告日 2008年5月28日

[11] 授权公告号 CN 100390196C

[22] 申请日 2002.9.6 [21] 申请号 02822187.7

[30] 优先权

[32] 2001.9.6 [33] GB [31] 0121591.2

[86] 国际申请 PCT/IB2002/003904 2002.9.6

[87] 国际公布 WO2003/020756 英 2003.3.13

[85] 进入国家阶段日期 2004.5.8

[73] 专利权人 启龙有限公司

地址 意大利锡耶纳

[72] 发明人 M·皮扎

[56] 参考文献

WO9957280A2 1999.11.11

WO0023595A1 2000.4.27

审查员 唐 慧

[74] 专利代理机构 上海专利商标事务所有限公司
代理人 周承泽

权利要求书 3 页 说明书 32 页 附图 7 页

[54] 发明名称

奈瑟球菌蛋白质的杂合和串联表达

[57] 摘要

连接两种或多种奈瑟球菌蛋白质从而使它们被翻译成一条多肽链。提供分子式为 $\text{NH}_2 - \text{A} - [- \text{X} - \text{L} -]_n - \text{B} - \text{COOH}$ 的杂合蛋白质，其中 X 是一种氨基酸序列，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 是大于 1 的整数。蛋白质中各 n - X - 部分的序列与其它 - X - 部分具有序列相同性，这种蛋白质是一种串联蛋白质。

1. 一种为以下分子式的杂合蛋白质： $\text{NH}_2\text{-A-X}_1\text{-L-X}_2\text{-L-B-COOH}$ ，
其特征在于，其中 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，所述序列是指导蛋白质运输的前导序列或有利于克隆或纯化的短肽序列，B 是任选的 C 末端氨基酸序列，所述序列是指导蛋白质运输的序列或有利于克隆或纯化的短肽序列，和
 X_1 是 936 氨基酸序列， X_2 是 741 氨基酸序列。
2. 如权利要求 1 所述的杂合蛋白质，其中 X_1 来自菌株 2996。
3. 如权利要求 1 所述的杂合蛋白质，其中 X_1 是 W099/57280 中的 SEQ ID 2884。
4. 如权利要求 1 所述的杂合蛋白质，其中 X_2 来自菌株 MC58。
5. 如权利要求 1 所述的杂合蛋白质，其中 X_2 是 W099/57280 中的 SEQ ID 2536。
6. 如权利要求 1 所述的杂合蛋白质，其中，L 具有 20 个或较少的氨基酸。
7. 如权利要求 1 所述的杂合蛋白质，其中，L 是一种聚甘氨酸接头。
8. 如权利要求 1-7 任一所述的杂合蛋白质，其中 L 的氨基酸序列是 GSGGGG。
9. 如权利要求 1 所述的杂合蛋白质，其中，A 具有 40 个或较少的氨基酸。
10. 如权利要求 1 所述的杂合蛋白质，其中，B 具有 40 个或较少的氨基酸。
11. 如权利要求 1 所述的杂合蛋白质，其特征在于，所述蛋白质的氨基酸序列是 SEQ ID 30。
12. 一种核酸，其中，所述核酸编码权利要求 1-11 的所述蛋白质。
13. 一种组合物，其中，所述组合物包含权利要求 1 所述的蛋白质或权利要求 11 所述的核酸编码的蛋白质，还包括一种或多种蛋白质，选自：
 - (1) 287；
 - (2) 741；
 - (3) ORF46.1；
 - (4) 961；
 - (5) $\text{NH}_2\text{-A-X}_1\text{-L-X}_2\text{-L-B-COOH}$ ，其中 $\text{X}_1=287$ ， $\text{X}_2=953$ ；
 - (6) $\text{NH}_2\text{-A-X}_1\text{-L-X}_2\text{-L-B-COOH}$ ，其中 $\text{X}_1=287$ ， $\text{X}_2=919$ ；
 - (7) $\text{NH}_2\text{-A-X}_1\text{-L-X}_2\text{-L-B-COOH}$ ，其中 $\text{X}_1=287$ ， $\text{X}_2=961$ ；
 - (8) $\text{NH}_2\text{-A-X}_1\text{-L-X}_2\text{-L-B-COOH}$ ，其中 $\text{X}_1=287$ ， $\text{X}_2=741$ 。
14. 如权利要求 13 所述的组合物，其中，所述组合物还包含蛋白质(4)和(5)。

15. 如权利要求 14 所述的组合物，其中，蛋白质(4)是 SEQ ID 31，蛋白质(5)是 SEQ ID 28 或 SEQ ID 29。

16. 如权利要求 13-15 任一所述的组合物，其特征在于，该组合物还包含：

- 脑膜炎奈瑟球菌蛋白质抗原；
- 脑膜炎奈瑟球菌外膜泡囊（OMV）制品；
- 脑膜炎奈瑟球菌糖抗原；
- 肺炎链球菌糖抗原；
- 甲肝、乙肝或丙肝病毒抗原；
- 百日咳杆菌抗原；
- 白喉抗原；
- 破伤风抗原；
- 幽门螺杆菌蛋白质抗原
- 流感嗜血杆菌 (*Haemophilus influenzae*) 糖抗原；
- 淋病奈瑟球菌抗原；
- 肺炎衣原体抗原；
- 沙眼衣原体抗原；
- Porphyromonas gingivalis* 的抗原；
- polio 抗原；
- 狂犬病毒抗原；
- 麻疹、流行性腮腺炎和/或风疹病毒抗原；
- 流感病毒抗原；
- 卡定性莫拉菌 (*Moraxella catarrhalis*) 抗原；
- 无乳链球菌 (*Streptococcus agalactiae*) 抗原；
- 化脓性链球菌 (*Streptococcus pyogenes*) 抗原；
- 金黄色葡萄球菌抗原。

17. 如权利要求 13 所述的组合物，其中，所述组合物还包括药学上可接受的载体。

18. 如权利要求 17 所述的组合物，其中，所述组合物可用作药物。

19. 权利要求 17 所述的组合物的用途，用于制备治疗奈瑟氏菌导致的疾病的药物。

20. 如权利要求 19 所述的用途，其中，所述疾病是细菌性脑膜炎。

21. 如权利要求 13 所述的组合物的用途, 其特征在于, 用于制备针对奈瑟氏菌的疫苗。

奈瑟球菌蛋白质的杂合和串联表达

所有本文所引用文献全部纳入本文供参考。

技术领域

本发明属于蛋白质表达领域。具体涉及奈瑟球菌(*Neisseria*) (如淋病奈瑟球菌(*N. gonorrhoeae*)或优选脑膜炎奈瑟球菌(*N. meningitidis*))的蛋白质表达。

背景技术

参考文献1和2揭示了参考文献3到6所述奈瑟球菌蛋白质表达的一种另和改进的方法。一个这种方法是生成“杂合性”蛋白质，此蛋白中两种或多种奈瑟球菌蛋白质表达为一条多肽链。此方法有两个优点。首先，自身表达可能不稳定或弱的蛋白质可通过加入适当的伴侣来协助克服其问题。第二，简化工业生产，因为仅需采用一次表达和纯化即可产生两种单独有用的蛋白质。

本发明的一个目标是进一步提供另一种和改进的奈瑟球菌蛋白质表达方法。

发明概述

杂合蛋白质

因此发明提供同时表达两种或多种(如3、4、5、6或更多)奈瑟球菌蛋白质的方法，其中连接所述两种或多种蛋白质从而使它们被翻译成一条多肽链。一般，本发明的杂合蛋白质可表示为公式 $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中X是一条氨基酸序列，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是大于1的整数。

n的值在2和x之间，x的值通常是3、4、5、6、7、8、9或10。n优选是2、3或4；更优选是2或3；最优选n=2。

-X-部分

本发明有两种主要的杂合蛋白质组。这两组不互相排斥。

在第一组中，各-X-部分是：

(a) orf1、orf4、orf25、orf40、orf46.1、orf83、NMB1343、230、233、

287、292、594、687、736、741、907、919、936、953、961 或 983 氨基酸序列；

(b) 与(a)的氨基酸序列有序列相同性的氨基酸序列；或

(c) 含(a)的氨基酸序列片段的氨基酸序列

优选(a)的亚组是：orf46.1、230、287、741、919、936、953、961 和 983。

更优选的(a)的亚组是：orf46.1、287、741 和 961。图 3 显示优选的杂合蛋白质。

在第二组中，该杂合蛋白质包括第一个-X-部分(-X_a-)和第二个-X-部分(-X_b-)。-X_a-部分具有下列氨基酸序列之一：

(d) 参考文献 3 所述的 446 个偶数序列编号 SEQ IDs(即 2、4、6...890、892)；

(e) 参考文献 4 所述的 45 个偶数 SEQ IDs(即 2、4、6...88、90)；

(f) 参考文献 5 所述的 1674 个偶数 SEQ IDs 2-3020、偶数 SEQ IDs 3040-3114 和所有的 SEQ IDs 3115-3241；

(g) 参考文献 7 的 2160 个氨基酸序列，NMB0001 到 NMB2160；

(h) 参考文献 1 或 2 所述的氨基酸序列。

-X_b-部分与-X_a-的关系如下：(i)-X_b-与-X_a-具有序列相同性和/或(j)-X_b-包括-X_a-的片段。

此第二种杂合性蛋白质例子包括的蛋白质中有两个或多个相同的-X-部分，或它们是同一蛋白质的变体，如同一蛋白质的两种多态，可表示为-X_a-X_b-；三种多态型，可表示为-X_a-X_b-X_c等。

-X_a-X_b-部分可以从 N-末端到 C-末端的任一顺序。

-X_a-部分优选是 orf1、orf4、orf25、orf40、orf46.1、orf83、NMB1343、230、233、287、292、594、687、736、741、907、919、936、953、961 或 983 氨基酸序列。-X_a-部分更优选是 orf46.1、230、287、741、919、936、953、961 和 983。-X_a-部分最优选是：orf46.1、287、741 或 961 氨基酸序列。

这些蛋白质中，各 n-X-部分与其它各-X-部分具有序列相同性，此蛋白质称为“串联蛋白质”。优选 n=2 的串联蛋白质。

(b)和(i)中的‘序列相同性’程度优选大于 50%(如 60%、70%、80%、90%、95%、99%或更高直至 100%)。这包括突变体、同源物、直向同源物、等位基因变体等(如参见参考文献 8)。相同性优选通过 MPSRCH 程序(Oxford Molecular)中应用的 Smith-Waterman 同源性检索算法来确定，使用参数空格开放罚分(gap open penalty)=12 和空格延伸罚分(gap extension penalty)=1 的相关空格检索。通常，将两个蛋白质间 50%或更高的相同性视为功能相等的指征。

(c)和(j)中所指‘片段’应包括(a)、(d)、(e)、(f)、(g)或(h)的氨基酸序列的至少 m 个连续氨基酸,取决于具体序列。m 是 7 或更大(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100、150、200 或更大)。

这些片段优选包含(a)、(d)、(e)、(f)、(g)或(h)氨基酸序列的抗原表位。优选的片段见参考文献 9 和 10 所述。

优选的(c)和(j)片段是 C-和/或 N-末端的截短物(如 Δ 1-287、 Δ 2-287 等)。

优选的(b)、(c)和(j)序列删除了聚甘氨酸序列。已发现这样有助于表达[参考文献 2]。聚甘氨酸序列可表示为(Gly)_g, 其中 $g \geq 3$ (如 4、5、6、7、8、9 或更大)。如果-X-部分包含其野生型中的聚甘氨酸序列,在本发明的杂合蛋白质中优选删除此序列。这可以通过破坏或去除(Gly)_g, 如通过缺失(如 CGGGGS \rightarrow CGGGS、CGGS、CGS 或 CS)、通过取代(如 CGGGGS \rightarrow CGXGGS、CGXXGS、CGXGXS 等)和/或通过插入(如 CGGGGS \rightarrow CGGXGGS、CGXGGGS 等)而实现。优选缺失(Gly)_g, 缺失直至包含聚甘氨酸序列的蛋白质 N-末端部分(如缺失 SEQ ID 1 中的残基 1-32)本文称为‘ Δ G’。删除聚甘氨酸序列具体用于蛋白质 287、741、983 和 Tbp2(Δ G287、 Δ G 741、 Δ G 983 和 Δ G Tbp2-参考文献 1 和 2)。

优选的(c)和(j)片段删除了整个蛋白质结构域。这具体用于蛋白质 961、287 和 ORF46。一旦蛋白质在概念上分成结构域,(c)和(j)片段可删除一个或多个这些结构域(如 287B、287C、287BC、ORF46₁₋₄₃₃、ORF46₁₃₃₋₆₀₈、961c-参考文献 2; 本文图 4 和 5)。

287 蛋白质在概念上分成三个结构域,称为 A、B 和 C(参见参考文献 2 的图 5)。结构域 B 具有 IgA 蛋白酶的排列,结构域 C 具有转铁蛋白-结合蛋白的排列,结构域 A 显示与数据库序列无强排列。287 的多态型排列见参考文献 8 所述。

ORF46 在概念上分成两个结构域:种和血清型之间非常保守的第一结构域(氨基酸 1-433: ORF46.1)和不很保守的第二结构域(氨基酸 433-608)。优选缺失该第二结构域,留下 ORF46.1。ORF46 的多态型排列见参考文献 8 所述。

961 蛋白质在概念上分成几个结构域(图 4)。

如果-X-部分含有野生型的前导肽序列,可将它包含在本发明的杂合蛋白质中或删除。当删除此前导肽时,这是(c)和(j)中氨基酸序列的优选例子。在一个实施方案中,缺失此前导肽,除了位于杂合蛋白质 N-末端的-X-部分,即保留 X₁ 的前导肽,但出去删去 X₂...X_n 的前导肽。这相当于缺失所有前导肽并采用 X₁ 的前导肽作为部分-A-。

当 $n=2$ 时, $-X$ -部分的优选配对是: $\Delta G287$ 和 230; $\Delta G287$ 和 936; $\Delta G287$ 和 741; 961c 和 287; 961c 和 230; 961c 和 936; 961cL 和 287; 961cL 和 230; 961cL 和 936; ORF46.1 和 936; ORF46.1 和 230; 230 和 961; 230 和 741; 936 和 961; 936 和 741。当 $n=2$ 时, 串联蛋白质的 $-X$ -部分优选配对是: $\Delta G741$ 和 741; $\Delta G287$ 和 287。更具体的是, 当 $n=2$ 时优选下列 X_1 和 X_2 的组合:

X_1	X_2
$\Delta G287$	230
$\Delta G287$	936
$\Delta G287$	741
$\Delta G287$	961
$\Delta G287$	ORF46.1
$\Delta G287$	919
$\Delta G287$	953
961c	287
961c	230
961c	936
961c	741
961c	983
961c	$\Delta G983$
961c	ORF46.1
961	ORF46.1
961cL	287
961cL	230
961cL	936
ORF46.1	936
ORF46.1	230
ORF46.1	741
ORF46.1	$\Delta G741$
ORF46.1	983
ORF46.1	$\Delta G983$
230	961
230	741
230	$\Delta G741$
936	961
936	741
936	$\Delta G741$
$\Delta G741$	741
ORF46.1	983
$\Delta G741$	ORF46.1

X_1	X_2
230	$\Delta G287$
936	$\Delta G287$
741	$\Delta G287$
961	$\Delta G287$
ORF46.1	$\Delta G287$
919	$\Delta G287$
953	$\Delta G287$
287	961c
230	961c
936	961c
741	961c
983	961c
$\Delta G983$	961c
ORF46.1	961c
ORF46.1	961
287	961cL
230	961cL
936	961cL
936	ORF46.1
230	ORF46.1
741	ORF46.1
$\Delta G741$	ORF46.1
983	ORF46.1
$\Delta G983$	ORF46.1
961	230
741	230
$\Delta G741$	230
961	936
741	936
$\Delta G741$	936
$\Delta G287$	287
983	ORF46.1
ORF46.1	$\Delta G741$

$\Delta G741$	983
$\Delta G741$	961
$\Delta G741$	961c
$\Delta G983$	ORF46.1
$\Delta G983$	961
$\Delta G983$	961c

983	$\Delta G741$
961	$\Delta G741$
961c	$\Delta G741$
ORF46.1	$\Delta G983$
961	$\Delta G983$
961c	$\Delta G983$

其中 287 以全长形式使用, 优选 287 位于杂合蛋白质的 C-末端; 如果 287 要用于 N-末端, 优选使用 287 的 ΔG 形式。类似的, 当 741 以全长形式使用时, 优选位于杂合蛋白质的 C-末端; 如果它要用于 N-末端, 优选使用 741 的 ΔG 形式。

-L-部分

对于[-X-L-]的各 n 实例, 接头氨基酸序列-L-可以存在或没有。例如, 当 n=2 时, 杂合物可以是 $NH_2-X_1-L_1-X_2-L_2-COOH$ 、 $NH_2-X_1-X_2-COOH$ 、 $NH_2-X_1-L_1-X_2-COOH$ 、 $NH_2-X_1-X_2-L_2-COOH$ 等。

接头氨基酸序列-L-通常较短(如 20 个或更少氨基酸即 19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、1)。例子包括能促进克隆的短肽序列、聚甘氨酸接头(即 Gly_n , 其中 n=2、3、4、5、6、7、8、9、10 或更大)和组氨酸尾(即 His_n , 其中 n=3、4、5、6、7、8、9、10 或更大)。其它适合的接头氨基酸序列是本领域技术人员懂得的。有用的接头是 GSGGGG(SEQ ID 27), 其中 Gly-Ser 二肽形成自 *Bam*HI 限制性酶切位点从而有助于克隆和操作, Gly_4 四肽是典型的聚甘氨酸接头。

如果 X_{n+1} 是 ΔG 蛋白质且 L_n 是甘氨酸接头, 这相当于不是 ΔG 蛋白质且没有 L_n 的 X_{n+1} 。

-A-部分

-A-是任选的 N-末端氨基酸序列。它通常较短(如 40 个或更少氨基酸即 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)。例子包括指导蛋白质运输的前导序列, 或者有利于克隆或纯化的短肽序列(如含有组氨酸尾即 His_n , 其中 n=3、4、5、6、7、8、9、10 或更大)。其它适合的 N-末端氨基酸序列是本领域技术人员懂得的。如果 X_1 缺乏自身的 N-末端甲硫氨酸, -A-可以是甲硫氨酸残基。

-B-部分

-B-是任选的 C-末端氨基酸序列。它通常较短(如 40 个或更少氨基酸即 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)。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(如包含组氨酸尾即 His_n,其中 n=3、4、5、6、7、8、9、10 或更大)。或提高蛋白质稳定性的序列。其它适当的 C-末端氨基酸序列是本领域技术人员懂得的。

蛋白质的多态型形式

本发明可使用任何脑膜炎奈瑟球菌菌株的氨基酸序列。因此涉及的具体蛋白质(如‘287’或‘ORF46.1’)包括任何菌株的蛋白质。菌株间的序列变化包括在(b)、(c)、(i)和(j)中。

脑膜炎奈瑟球菌血清群 B 的参考序列包括:

蛋白质	参考文献	蛋白质	参考文献
orf1	参考文献 3, SEQ ID 650	orf4	参考文献 3, SEQ ID 218
orf25	参考文献 3, SEQ ID 684	orf40	参考文献 4, SEQ ID 4
orf46	参考文献 6, SEQ ID 1049	orf83	参考文献 3, SEQ ID 314
NMB1343	参考文献 7, NMB 1343	230	参考文献 5, SEQ ID 830
233	参考文献 5, SEQ ID 860	287	参考文献 5, SEQ ID 3104
292	参考文献 5, SEQ ID 1220	594	参考文献 5, SEQ ID 1862
687	参考文献 5, SEQ ID 2282	736	参考文献 5, SEQ ID 2506
741	参考文献 5, SEQ ID 2536	907	参考文献 5, SEQ ID 2732
919	参考文献 5, SEQ ID 3070	936	参考文献 5, SEQ ID 2884
953	参考文献 5, SEQ ID 2918	961	参考文献 5, SEQ ID 940
983	参考文献 7, NMB 1969		

参考文献 8 揭示了蛋白质 ORF4、ORF40、ORF46、225、235、287、519、726、919 和 953 的多态型形式。961 的多态形式见参考文献 11 和 12 所述。任何这些多态形式可用于本发明。

本文所列序列包括已鉴定的蛋白质 741(SEQ IDs 1-22)和 NMB 1343(SEQ IDs 23-24)的多态形式。

血清群和菌株

本发明的优选蛋白质包括-X-部分, 其具有脑膜炎奈瑟球菌血清群 B 中发现的

氨基酸序列。在本发明的单个蛋白质中，各个-X-部分可以来自一种或多种菌株。例如当 $n=2$ 时， X_2 可以来自与 X_1 相同或不同的菌株。当 $n=3$ 时，菌株可以是(i) $X_1=X_2=X_3$ (ii) $X_1=X_2 \neq X_3$ (iii) $X_1 \neq X_2 = X_3$ (iv) $X_1 \neq X_2 \neq X_3$ 或(v) $X_1 = X_3 \neq X_2$ 等。

在血清群 B 中，优选的-X-部分来自菌株 2996、MC58、95N477 或 394/98。菌株 95N477 有时在本文中称为 ‘ET37’，这是其电泳形式。菌株 394/98 有时在本文中称为 ‘nz’，因为它是新西兰菌株。

当使用 287 的一种形式时，优选来自菌株 2996 或来自菌株 394/98。

当使用 741 的一种形式时，优选来自血清群 B 的菌株 MC58、2996、394/98 或 95N477，或者血清群 C 的菌株 90/18311。

当使用 961 的一种形式时，优选来自菌株 2996。

菌株以下标表示，如 741_{MC58} 是菌株 MC58 的蛋白质 741。除非另有说明，本文所提及的蛋白质(如没有下标)来自脑膜炎奈瑟球菌菌株 2996，它可作为 ‘参考’ 菌株。然而要理解的是本发明一般不受菌株的限制。如上所述，一般涉及的蛋白质(如 ‘287’ 或 ‘919’ 等)包括任何菌株的蛋白质。通常与 2996 有 90%或更高的序列相同性(如 91%、92%、93%、94%、95%、96%、97%、98%、99%或更高)。

蛋白质 961 的结构域为基础的表达

参考文献 1 和 2 揭示了蛋白质如何在概念上分成结构域且蛋白质如何以这些结构域的基础进行操作。本发明将此方法扩展应用于蛋白质 961(也称为 ‘NadA’ [11、12])。

在脑膜炎奈瑟球菌血清群 B 的菌株 2996 中，NadA 有 405 个氨基酸。此蛋白质在概念上分成下列九个结构域(图 4)：

结构域名称	氨基酸	结构域名称	氨基酸
961-1 ‘L’	1-23	961-6	269-286
961-2	24-87	961-7	287-330
961-3	88-143	961-8	331-350
961-4	144-180	961-9	351-405
961-5	181-268		

此信息可用于定位其它 961 形式中的相同结构域

已用各种方法从菌株 2996 中的 961 缺失掉这些结构域(图 5)。961 的优选片段删除了些几个结构域如下的一个或多个：

- 961-2 到 961-5('961a')
- 961-6 到 961-9('961b')
- 961-1 到 961-8('961cL')
- 961-2 到 961-8('961c')
- 961-2 到 961-6 和结构域 961-7 的氨基酸 287-325('961d')
- 961-2 到 961-8 和结构域 961-9 的氨基酸 351-383('961 Δ 1')
- 961-1 到 961-8 和结构域 961-9 的氨基酸 351-383('961 Δ 1L')
- 961-1 到 961-7 和结构域 961-8 的氨基酸 331-343('961cL- Δ aro')
- 961-1 到 961-6 和结构域 961-7 的氨基酸 287-315('961cL- Δ cc')
- 961-1 到 961-5('961aL')
- 961-1 到 961-4('961aL- Δ 1')
- 961-1 到 961-3('961aL- Δ 2')
- 961-1 到 961-2('961aL- Δ 3')

这 13 个片段(和其一端或两端缺少 1、2、3、4 或 5 个氨基酸的亚片段)是优选的(c)和(j)片段,但它们也可自身表达,即不是本发明的杂合蛋白质形式。因此本发明提供含这些片段之一的蛋白质,此蛋白质不是全长 961 且不是参考文献 1 或 2 具体描述的蛋白质。此蛋白质可以是融合蛋白(如 GST-融合或 His-尾融合)序列

本发明还提供具有 SEQ IDs 1-24 的氨基酸序列的蛋白质。也提供与这些蛋白质具有序列相同性的蛋白质和核酸。如上所述,‘序列相同性’程度优选大于 50%(如 60%、70%、80%、90%、95%、99%或更高)。

本发明还提供编码这些蛋白质的核酸。

此外,本发明提供优选在‘高严谨性’条件下(如 65°C 在 0.1xSSC、0.5%SDS 溶液中)可与此核酸杂交的核酸,

发明也提供编码本发明蛋白质的核酸。

还应理解的是本发明提供的核酸包含与上述序列互补的序列(如用于反义或探测目的)。

本发明的核酸当然可以许多方法制备(如通过化学合成,从基因组或 cDNA 文库,从生物体本身制备等),并可采用各种形式(如单链、双链、载体、探针等)。

此外,术语“核酸”包括 DNA 和 RNA 以及它们的类似物,如含修饰主链的类似物和肽核酸(PNA)等。

混合物

本发明还提供一种组合物，该组合物包含两种或多种(即 2、3、4、5、6 或 7)以下蛋白质：

- (1) 287
- (2) 741
- (3) ORF46.1
- (4) 961
- (5) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 $n=2$ ， $X_1=287$ ， $X_2=953$
- (6) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 $n=2$ ， $X_1=287$ ， $X_2=919$
- (7) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 $n=2$ ， $X_1=287$ ， $X_2=961$

此混合物可包含一种或两种下列蛋白质，与(1)到(7)的两种或多种或仅与(1)到(7)中的一种蛋白质组合：

- (8) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 $n=2$ ， $X_1=287$ ， $X_2=741$
- (9) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 $n=2$ ， $X_1=936$ ， $X_2=741$

当蛋白质 287 和 741 包括在此混合物中时(即在蛋白质 1、2、5、6、7 或 8 中)，它们可能是‘ ΔG ’形式。当包括蛋白质 961 时，优选‘961c’形式，其中没有 N-末端前导序列和 C-末端膜锚定序列。

优选的混合物包含下列三种蛋白质：

- (1) 961c，优选 961c₂₉₉₆(如本文的 SEQ ID 31)；
- (2) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 n 为 2， $-X_1-$ 是 $\Delta\text{G}287$ (优选 $\Delta\text{G}287_{\text{N2}}$)， $-X_2-$ 是缺乏前导肽的 953(优选 953₂₉₉₆)， $-L_1-$ 是 GSGGGG 且 $-A-$ 包括 N-末端甲硫氨酸(如 $-A-$ 是 M 或 MA)(如本文的 SEQ IDs 28 和 29)；
- (3) $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$ ，其中 $n=2$ ， $X_1=936$ (优选 936₂₉₉₆)， $X_2=\Delta\text{G}741$ (优选 $\Delta\text{G}741_{\text{MC58}}$)， $L_1=\text{GSGGGG}$ (如本文的 SEQ ID 30)。

此混合物也可包含脑膜炎奈瑟球菌的外膜囊泡。

异源宿主

当本发明的蛋白质在奈瑟球菌中表达时，本发明优选使用异源宿主。异源宿主可以是原核生物(如细菌)或真核生物细胞。优选大肠杆菌，但其它适当的宿主，包括枯草芽孢杆菌(*Bacillus subtilis*)、霍乱弧菌(*Vibrio cholerae*)、伤寒沙门杆菌(*Salmonella typhi*)、鼠伤寒沙门杆菌(*Salmonella typhimurium*)、乳酸胺奈瑟菌(*Neissria lactamica*)、脊髓灰质炎奈瑟菌(*Neissria cinerea*)、分枝

杆菌(*Mycobacteria*) (如结核分枝杆菌)、酵母等。

载体等

本发明提供(a)编码上述蛋白质的核酸(b)含有这些核酸序列的载体(c)含所述载体的宿主细胞(d)含本发明蛋白质或核酸的组合物,也许其适合用作免疫原性组合物(如疫苗)或用作诊断试剂(e)这些组合物作为药物(如疫苗)或作为诊断试剂的用途(f)这些组合物在生产(1)治疗或预防奈瑟球菌感染的药物(2)检测奈瑟球菌或抗奈瑟球菌抗体存在的诊断试剂,和/或(3)提高抗奈瑟球菌抗体的制剂的用途,以及(g)治疗病人的方法,包括给予病人治疗有效量的这些组合物。

实施本发明通常包括以下基本步骤:获得编码第一种蛋白质的第一种核酸;获得编码第二种蛋白质的第二种核酸;连接第一种和第二种核酸。可将所得核酸插入表达载体中或可作为表达载体的一部分。

为改进溶解性,纯化杂合蛋白质时可包括本文所述重折叠技术。

免疫原性组合物和药物

本发明的组合物优选是免疫原性组合物,更优选疫苗组合物。该组合物的pH值优选在6和7之间。此pH值可通过采用缓冲液来维持该组合物可以是无菌的。

本发明的疫苗可以是预防性的(即预防感染)或治疗性的(即治疗感染),但一般是预防性的。

本发明也提供本发明的组合物作为药物用途。药物优选能提高哺乳动物免疫应答(即它是免疫原性组合物),更优选是疫苗。

本发明也提供本发明组合物在生产能提高哺乳动物免疫应答的药物中的用途。药物优选是疫苗。

本发明也提供提高哺乳动物免疫应答的方法,该方法包括给予有效量的本发明组合物的步骤。免疫应答优选是保护性的。此方法可提高加强的免疫应答。

哺乳动物优选是人。当此疫苗用于预防时,优选的人是儿童(如初学走路的孩子或婴儿);当疫苗用于预防时,优选的人是成年人。针对儿童的疫苗也可给予成人,以例如评估安全性、剂量、免疫原性等。

这些用途和方法优选用于预防和/或治疗奈瑟球菌引起的疾病(如脑膜炎、败血症、淋病等)。优选预防和/或治疗细菌脑膜炎。

此组合物的其它成分

除上述成分外,本发明的组合物通常包含一种或多种‘药学上可接受的远载体’,包括自身不诱导接受此组合物的个体产生有害抗体的任何远载体。适当的远

载体一般是大的代谢缓慢的大分子、如蛋白质、多糖、聚乳酸、聚乙醇酸、聚合氨基酸、氨基酸共聚物、海藻糖(WO 00/56365)和脂类凝聚物(如油滴或脂质体)。这种远载体是本领域技术人员熟知的。此疫苗也可包含稀释液,如水、盐水、甘油等。此外,辅助物质如润湿剂或乳化剂、pH 缓冲物质等可存在。药学上可接受赋形剂的完整论述可获自《Remington 药物科学》(Remington's Pharmaceutical Sciences)。

用作疫苗的免疫原性组合物包括免疫有效量的抗原以及任何其它需要的上述成分。‘免疫有效量’指将此剂量以单剂量或作为系列剂量的一部分给予个体,能有效地治疗或预防。此剂量的变化取决于待治疗个体的健康和身体情况、年龄、待治疗个体的分类(如非人灵长类动物、灵长类动物等)、个体免疫系统合成抗体的能力、所需保护的程 度、疫苗的配方、治疗医生对医疗情况的评估和其它相关因素。预期此剂量有相对宽的范围,可通过常规试验确定。剂量治疗可以是单剂量方案或多剂量方案(如包括强化剂量)。此疫苗可结合其它免疫调节剂施用。

疫苗可结合其它免疫调节剂施用。

除(或取代)铝盐外,该组合物可包含其它佐剂。能提高该组合物效果的优选佐剂包括但不限于:(1)水包油乳剂(含有或没有其它特定免疫刺激剂如胞壁酰肽(见下)或细菌细胞壁成分),例如(a)MF59™(WO90/14837;参考文献中的第10章),含5%鲨烯、0.5%吐温80和0.5%Span85(任选含MTP-PE),用微流化剂制成亚微米颗粒、(b)SAF,含10%鲨烯、0.4%吐温80和5% pluronic-嵌段聚合物L121和thr-MDP,微流化成亚微米乳剂或旋涡混合产生较大颗粒大小的乳剂、(c)Ribi™佐剂系统(RAS)(Ribi Immunochem, Hamilton, MT),含2%鲨烯、0.2%吐温80和一种或多种细菌细胞壁成分,选自单磷酸脂质A(MPL)、海藻糖 dimycolate(TDM)和细胞壁骨架(CWS),优选MPL+ CWS(Detox™);(2)皂角苷佐剂,如可使用QS21或Stimulon™(Cambridge Bioscience, Worcester, MA)或它们所产生的颗粒如ISCOMs(免疫刺激复合物),ISCOMs可避免采用其它去污剂,如WO 00/07621;(3)弗氏完全佐剂(CFA)和弗氏不完全佐剂(IFA);(4)细胞因子,如白介素(例如IL-1、IL-2、IL-4、IL-5、IL-6、IL-7、IL-12(WO 99/44636)等)、干扰素(例如γ干扰)、巨噬细胞集落刺激因子(M-CSF)、肿瘤坏死因子(TNF)等;(5)单磷酸脂质A(MPL)或3-O-脱酰基MPL(3d MPL),如GB-2220221、EP-A-0689454;(6)3d MPL与例如QS21和/或水包油乳剂的组合,如EP-A-0835318、EP-A-0735898、EP-A-0761231;(7)含CpG基序的寡核苷酸[Krieg *Vaccine* 2000, 19, 618-622; Krieg *Curr Opin Mol*

Ther 2001 3:15-24; Roman 等, *Nat. Med*, 1997, 3, 849-854; Weiner 等, *PNAS USA*, 1997, 94, 10833-10837; Davis 等, *J. Immunol.*, 1998, 160, 870-876; Chu 等, *J. Exp. Med.*, 1997, 186, 1623-1631; Lipford 等, *Eur. J. Immunol.*, 1997, 27, 2340-2344; Moldoveanu 等, *Vaccine*, 1998, 16, 1216-1224; Krieg 等, *Nature*, 1995, 374, 546-549; Klinman 等, *PNAS USA*, 1996, 93, 2879-2883; Ballas 等, *J. Immunol.*, 1996, 157, 1840-45; Cowdery 等, *J. Immunol*, 1996, 156, 4570-75; Halpern 等 *Cell Immunol.*, 1996, 167, 72-78; Yamamoto 等, *Jpn. J. Cancer Res.*, 1988, 79, 866-873; Stacey 等, *J. Immunol.*, 1996, 157, 2116-2122; Messina 等, *J. Immunol.*, 1991, 147, 1759-1764; Yi 等, *J. Immunol.*, 1996, 157, 4918-4925; Yi 等, *J. Immunol.*, 1996, 157, 5394-5402; Yi 等, *J. Immunol.*, 1998, 160, 4755-4761; Yi 等, *J. Immunol.*, 1998, 160, 5898-5906; 国际专利申请 W096/02555、W098/16247、W098/18810、W098/40100、W098/55495、W098/37919 和 W098/52581]即含至少一个 CG 二核苷酸, 5-甲基胞嘧啶可任选用于取代胞嘧啶; (8) 聚氧乙烯醚或聚氧乙烯酯如 W099/52549; (9) 结合辛氧基(octoxynol)的聚氧乙烯山梨糖酯类表面活性剂(如 WO 01/21207)或结合至少一种其它非离子表面活性剂如辛氧基(octoxynol)聚氧乙烯烷基醚或酯的表面活性剂(如 WO 01/21152); (10) 免疫刺激寡核苷酸(如 CpG 寡核苷酸)和皂角苷如 WO 00/62800; (11) 免疫刺激剂和金属盐颗粒如 WO 00/23105; (12) 皂角苷和水包油乳剂如 W099/11241; (13) 皂角苷(如 QS21)+3dMPL+IL-12(任选类固醇)如 W098/57659; (14) 作为免疫刺激剂能提高此组合物效果的其它物质。

胞壁酰肽包括 N-乙酰-胞壁酰-L-苏氨酸-D-异谷氨酰胺(thr-MDP)、N-乙酰-nor胞壁酰-L-丙氨酸-D-异谷氨酰胺(nor-MDP)、N-乙酰胞壁酰-L-丙氨酸-D-异谷氨酸-L-丙氨酸-2(1'2'-二棕榈酰-sn-甘油-3-羟基磷酸氧基(phosphoryloxy))-乙醇胺 MTP-PE)等。

其它抗原

可包含在本发明组合物中的其它抗原包括:

-脑膜炎奈瑟球菌血清群 B 的外膜泡囊(OMV)制品, 如见参考文献 13、14、15、16、17 等所述。

-脑膜炎奈瑟球菌血清群 A、C、W135 和/或 Y 的糖抗原, 如参考文献 18 中所述的血清群 C 的寡糖[同样参见参考文献 19]或参考文献 20 的寡糖。

-肺炎链球菌的糖抗原

-幽门螺杆菌的蛋白质抗原,如 CagA[如 24]、VacA[如 24]、NAP[如 25]、HopX[如 26]、HopY[如 26]和/或尿素酶。

-甲肝病毒抗原,如灭活病毒[如 27、28]。

-乙肝病毒抗原,如表面抗原和/或核心抗原[如 28、29]。

-丙肝病毒抗原[如 30]。

-百日咳杆菌(*Bordetella pertussis*)抗原,如百日咳杆菌的百日咳全毒素(PT)和丝状血凝素(FHA),还任选与百日核毒素(pertactin)和/或凝集素 2 和 3[如参考文献 31 和 32]组合使用。

-白喉抗原,例如白喉类毒素[如参考文献 33 的第 3 章],如 CRM₁₉₇ 突变体[如 34]。

-破伤风抗原,如破伤风类毒素[参考文献 33 的第 4 章]。

-流感嗜血杆菌 B(*Haemophilus influenzae*)的糖抗原[如 19]。

-淋病奈瑟球菌抗原[如 3、4、5]。

-肺炎衣原体(*Chlamydia pneumoniae*)抗原[如 3、4、5]。

-沙眼衣原体(*Chlamydia trachomatis*)抗原[如 42]。

-牙龈卟啉单胞菌(*Porphyromonas gingivalis*)的抗原[如 43]。

-脊髓灰质炎病毒(polio)的抗原[如 44、45],例如 IPV 或 OPV。

-狂犬病毒抗原[如 46],例如冻干的灭活病毒[如 47、RabAvert™]。

-麻疹、流行性腮腺炎和/或风疹抗原[如参考文献 33 的第 9、10 和 11 章]。

-流感病毒抗原[如参考文献 33 的第 19 章],例如血凝素和/或神经氨酸酶表面蛋白。

-卡它性莫拉菌(*Moraxella catarrhalis*)抗原[如 48]。

-无乳链球菌(*Streptococcus agalactiae*) (B 组链球菌)蛋白质抗原[如 49、50]。

-无乳链球菌糖抗原。

-化脓性链球菌(*Streptococcus pyogenes*) (A 组链球菌)抗原[如 50、51、52]。

-金黄色葡萄球菌(*Staphylococcus aureus*)抗原[如 53]。

该组合物可包括一种或多种这些其它抗原。

当采用糖类或碳水化合物抗原时,优选将其偶联于载体蛋白以提高免疫原性[如参考文献 54 到 63]。优选的载体蛋白是细菌毒素或类毒素,如白喉或破伤风类毒素。尤其优选 CRM₁₉₇ 白喉类毒素。其它适当的载体蛋白包括脑膜炎奈瑟球菌外膜

蛋白[如参考文献 64]、合成肽[如 65、66]、热激蛋白[如 67]、百日咳蛋白[如 68、69]、流感嗜血杆菌的蛋白质 D[如 70]、艰难梭菌(*C. difficile*)的毒素 A 或 B[如 71]等。当混合物包含血清群 A 和 C 的荚膜糖类时, MenA 糖:MenC 糖的比例(w/w) 优选大于 1(如 2:1、3:1、4:1、5:1、10:1 或更高)。可将脑膜炎奈瑟球菌不同血清群的糖类偶联于相同或不同的载体蛋白。

可利用任何适当的偶联反应,必要时可采用任何合适的接头。

必要时可使毒性蛋白质抗原脱毒(如通过化学和/或遗传方法[32]使百日毒素脱毒)。

当该组合物包含白喉抗原时,也优选包含破伤风抗原和百日咳抗原。类似的,当包含破伤风抗原时,也优选包含白喉和百日咳抗原。类似的,当包含百日咳抗原时,也优选包含白喉和破伤风抗原。

抗原优选与铝盐(如磷酸铝、氢氧化铝、羟基磷酸铝、正氢氧化铝、正磷酸铝、硫酸铝)混合铝盐可采用任何适当的形式(如凝胶,晶体,无定形等)。

此组合物中的抗原通常以各至少 $1\mu\text{g/ml}$ 的浓度存在。一般,任何给定抗原的浓度应足够引起抗此抗原的免疫应答。

采用本发明组合物中的蛋白质抗原的另一种方法,可采用编码此抗原的核酸(如参考文献 72 到 80)。因此本发明组合物的蛋白质成分可用编码此蛋白质的核酸取代(优选 DNA,如以质粒形式)。

定义

术语“包括”指“包含”以及“由...组成”,如“包括”X 的组合物可只由 X 组成或包含一些另外成分如 X+Y。

涉及数值 x 的术语“约”指例如 $x \pm 10\%$ 。

附图简述

图 1 显示蛋白质 741 的 23 种序列的排列。这些是 SEQ IDs1 到 22 加上 MC58 的序列。

图 2 显示淋菌(顶部; SEQ ID 25)和脑膜炎球菌(底部; SEQ ID 26)的 NMB1343 序列的排列。

图 3 显示本发明的杂合和串联蛋白质。

图 4 显示 961₂₉₉₆ 中的 9 个结构域,图 5 显示如何操作这些结构域。

实施本发明的模式

杂合蛋白质-X1= Δ G287

除了参考文献 1 和 2 所述外, 构建了七个 N-末端含有菌株 2996 的 Δ G287 的杂合蛋白质, 还制备了八个 287 串联蛋白质(见下)。

#	n	X ₁	L ₁	X ₂	L ₂
1	2	Δ G287	-	230	(His) ₆
2	2		-	936	(His) ₆
3	2		-	741 _{MC58}	(His) ₆
4	2		-	741 _{ET37}	(His) ₆
5	2		-	741 _{90/18311}	(His) ₆
6	2		-	741 _{95N177}	(His) ₆
7	2	Δ G287 _{nz}	-	741 _{MC58}	(His) ₆

这些蛋白质用弗氏完全佐剂(FCA)或 3mg/ml 明矾作为佐剂用于免疫小鼠。所得血清用杀菌试验测试抗不同奈瑟球菌菌株抗体。采用蛋白质#3 测得的滴定度如下:

菌株(血清群)	2996 ^(B)	MC58 ^(B)	NGH38 ^(B)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)
氢氧化铝	8192	32768	8192	>2048	16384	8192
FCA	16384	262144	8192	>2048	>32768	8192

在用蛋白质#3 和氢氧化铝的进一步实验中, 特定了抗 287 和抗 741 抗体, 各 ELISA 滴度超过 984150, BCA 滴度如下:

2996 ^(B)	MC58 ^(B)	NGH38 ^(b)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)	BZ133 ^(C)
8000	65000	4000	4000	32000	8000	16000

用参考文献 1 和 2 所述蛋白质免疫后测定抗同源菌株抗体获得的结果如下:

n	X ₁	L ₁	X ₂	L ₂	杀菌滴度		ELISA	
					FCA	明矾	FCA	明矾
2	△G287 _{394/98}	-	961	(His) ₆	-	32768	-	>109350
			919		32768	4096	4718	3678
			953		>32768	>16384	1900	6936
			741		16384	2048	232	862
2	△G287 ₂₉₉₆	-	961	(His) ₆	65536	32768	108627	>109350
			919		128000	32000	11851	2581
			953		65536	-	3834	-
			741		16384	8192	315	4645

杂合蛋白质-X1=961c 或 961cL

除了参考文献 1 和 2 所述外,构建了八个在 N-末端含有 961c 或 961cL(即 961c+前导肽)的杂合蛋白质:

#	n	X ₁	L ₁	X ₂	L ₂
1	2	961c	-	287	-
2	2		-	287	(His) ₆
3	2		-	230	(His) ₆
4	2		-	936	(His) ₆
5	2	961cL	-	287	-
6	2		-	287	(His) ₆
7	2		-	230	(His) ₆
8	2		-	936	(His) ₆

这些蛋白质用弗氏完全佐剂(FCA)或 3.3mg/ml 明矾作为佐剂用于免疫小鼠。所得血清用杀菌试验测试抗不同奈瑟球菌菌株抗体。采用蛋白质#8 测得的滴定度如下:

菌株(血清群)	2996 ^(B)	MC58 ^(B)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)
氢氧化铝	8192	8192	512	1024	<16
FCA	65536	16384	>2048	>2048	8192

用 961c-741[参考文献 1 和 2]免疫后获得的滴度如下:

菌株(血清群)	2996 ^(B)	MC58 ^(B)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)	BZ133 ^(C)
氢氧化铝	65536	32768	4096	>32768	16384	>2048
FCA	>16384	262144	4096	>16384	-	>2048

这些结果可通过将 961c-741 与 ORF46.1 或 Δ G287-919 混合来改进。

用参考文献 1 和 2 所述蛋白质免疫后测定抗同源菌株抗体获得的结果如下:

n	X ₁	L ₁	X ₂	L ₂	杀菌滴度		ELISA	
					FCA	明矾	FCA	明矾
2	961c	-	ORF46.1	(His) ₆	32768	1024	>109350	>109350
			741		>16384	8192	>109350	>109350
			936		>32768	8192	>109350	>109350

杂合蛋白质-X1= ORF46.1

除了参考文献 1 和 2 所述外, 构建了两个在 N-末端含有 ORF46.1 的杂合蛋白质:

#	n	X ₁	L ₁	X ₂	L ₂
1	2	ORF46.1	-	936	(His) ₆
2	2		-	230	(His) ₆

这些蛋白质用弗氏完全佐剂(FCA)或 3mg/ml 明矾作佐剂用于免疫小鼠。所得血清用杀菌试验和 ELISA 测试抗同源菌株抗体。

用参考文献 1 和 2 所述蛋白质免疫后获得的结果如下:

n	X ₁	L ₁	X ₂	L ₂	杀菌滴度		ELISA	
					FCA	明矾	FCA	明矾
2	ORF46.1	-	961	(His) ₆	8192	8192	21558	>109350
		-	961c	(His) ₆	8192	128	9020	76545

杂合蛋白质-X1=230

除了参考文献 1 和 2 所述外, 构建了四个在 N-末端含有 230 的杂合蛋白质:

#	n	X ₁	L ₁	X ₂	L ₂
1	2	230	-	ORF46.1	(His) ₆
2	2		-	961	(His) ₆
3	2		-	961c	(His) ₆
4	2		-	741 _{MC58}	(His) ₆

杂合蛋白质-X1=936

除了参考文献 1 和 2 所述外, 构建了七个在 N-末端含有 936 的杂合蛋白质:

#	n	X ₁	L ₁	X ₂	L ₂
1	2	936	-	ORF46.1	(His) ₆
2	2		-	961	(His) ₆
3	2		-	741 _{ET37}	(His) ₆
4	2		-	741 _{MC58}	(His) ₆
5	2		-	741 _{90/18311}	(His) ₆
6	2		-	741 _{95N177}	(His) ₆
7	2		-	741	(His) ₆

这些蛋白质用弗氏完全佐剂(FCA)或 3mg/ml 明矾作佐剂用于免疫小鼠。所得血清用杀菌试验测试抗不同的奈瑟球菌菌株抗体。采用蛋白质#2 测得的滴定度如下:

菌株(血清群)	2996 ^(B)	MC58 ^(B)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)
氢氧化铝	16348	32768	1024	2048	<16
FCA	65536	65536	>2048	8192	2048(36%)

采用蛋白质#4 测得的滴定度如下:

菌株(血清群)	2996 ^(b)	MC58 ^(B)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)
氢氧化铝	256	>262144	>2048	32768	8192
FCA	1024	>262144	>2048	>32768	>32768

采用蛋白质#7 测得的滴定度如下:

菌株(血清群)	2996 ^(B)	MC58 ^(B)	394/98 ^(B)	44/76 ^(B)	F6124 ^(A)	BZ133 ^(C)
氢氧化铝	256	13000	16000	32000	8000	16000

用参考文献 1 和 2 所述蛋白质免疫后测定抗同源菌株抗体获得的结果如下：

n	X ₁	L ₁	X ₂	L ₂	杀菌滴度		ELISA	
					FCA	明矾	FCA	明矾
2	936	-	741	(His) ₆	1024	256	1466	5715
			936		>32768	>32768	>109350	>109350

杂合蛋白质的混合物

小鼠用氢氧化铝为佐剂的三种蛋白质免疫，单种或三种蛋白质的组合：(1)287_{NZ-953}；(2)936-741；和(3)961c。混合物能诱导抗不同菌株的高杀菌滴度：

	2996 ^(B)	MC58 ^(B)	NGH38	394/98 ^(B)	H44/76 ^(B)	F6124 ^(A)	BZ133 ^(C)	C11 ^(C)
(1)	32000	16000	130000	16000	32000	8000	16000	8000
(2)	256	131000	128	16000	32000	8000	16000	<4
(3)	32000	8000	-	-	-	8000	-	32000
混合物	32000	32000	65000	16000	260000	65000	>65000	8000
(X)	4000	4000	1000	1000	>4000	1000	4000	n. c.

‘-’ 表示此菌株不含 Nad A 基因

(X)是蛋白质 287 和外膜泡囊的组合，用于比较

观察各个小鼠，该混合物诱导了高而恒定的杀菌滴度：

#	1	2	3	4	5	6	7	8	9	10
2996	32768	16384	65536	32768	32768	65536	65536	32768	65536	8192
MC58	65536	32768	65536	65536	65536	8192	65536	32768	32768	65536
394/98	65536	4096	16384	4096	8192	4096	32768	16384	8192	16384

串联蛋白质

本发明的杂合蛋白质可表示为公式 NH₂-[-X-L-]_n-COOH。其中-X-的所有 n 实例是相同的碱性蛋白质(相同或者不同菌株或种类的同种蛋白质)，此蛋白质称为“串联”蛋白质。

十二种具体串联蛋白质是：

#	n	X ₁	L ₁	X ₂	L ₂
1	2	△G741 _{MC58}	-	741 _{MC58}	(His) ₆
2	2	△G287 ₂₉₉₆	(Gly) ₆	△G287 _{394/98}	(His) ₆
3	2	△G287 ₂₉₉₆	-(Gly) ₆	△G287 ₂₉₉₆	(His) ₆
4	2	△G287 _{394/98}	-(Gly) ₆	△G287 _{394/98}	(His) ₆
5	2	△G287 _{394/98}	-(Gly) ₆	△G287 ₂₉₉₆	(His) ₆
6	2	△G287 ₂₉₉₆	-(Gly) ₆	△G287 _{394/98}	-
7	2	△G287 ₂₉₉₆	-(Gly) ₆	△G287 ₂₉₉₆	-
8	2	△G287 _{394/98}	-(Gly) ₆	△G287 _{394/98}	-
9	2	△G287 _{394/98}	-(Gly) ₆	△G287 ₂₉₉₆	-
10	2	△G741 _{MC58}	-	741 _{394/98}	(His) ₆
11	2	△G741 _{MC58}	-	741 _{90/18311}	(His) ₆
12	2	△G741 _{MC58}	-	741 _{95N477}	(His) ₆

蛋白质#1 到#5 都已在大肠杆菌中以可溶形式表达。表达水平在 0.24 和 0.50 mg 蛋白质/每升培养物之间。纯化这些串联蛋白质并与作为佐剂的磷酸铝混合。串联蛋白质#2、#4 和#5 易被磷酸铝吸附；串联蛋白质#1 和#3 吸附不完全。

等位基因变体-741

发现了 741 的 22 种多态性序列 (SEQ IDs 1 到 22)。这些序列和 MC58 序列排列于图 1 中。

等位基因变体-NMB1343

对不同血清群脑膜炎球菌的 42 种菌株使用 PCR，在 24/42 菌株中发现编码 NMB1343 蛋白的基因但在 18/42 菌株中无 (表 1)。测序 10 个 NMB1343+ 菌株的 NMB1343 基因 (表 1，第 3 列)。所有 10 个菌株中此核酸序列 (和氨基酸序列 SEQ ID 23; GenBank AAF41718) 相同。

NMB1343 也在 2 种淋病奈瑟球菌菌株 (F62 和 SN4) 中检测到。淋球菌的氨基酸序列是 SEQ ID 24。脑膜炎球菌序列的排列是：


```

      . . . .10 . . . .20 . . . .30 . . . .40 . . . .50
Ng    1: INNLWEISSYLNROISCQNDEQNISGLDLPNGNIFAVAIRVDRPMDGLAT: 50
Nm    1: ~~~~~GNFLYRGISCQDEQNISGLDLPNGNIFAVAIRVDRPMDGLAT: 45

      . . . .60 . . . .70 . . . .80 . . . .90 . . . .100
Ng   51: HGPSVKNAVYAHQIETDLYDGGYISTTITDKEIANKFATSSGIEYGNIANI:100
Nm   46: HGPSVKNAVYAHQIETDGLYDSYISTTIDKEIANKFATSSGIEYGNIANI: 95

      . . . .110 . . . .120 . . . .130 . . . .140 . . . .150
Ng  101: NRDLFSGQYSLFEVEVEHPENPDERSVTIRAEDGGIPEEVIAELIEIN:150
Nm   96: NRDLFSGQYSLFEVEVEHPENPHEREVTIRAEDGGIPEEVIAELIEIN:145

```

对应的核苷酸序列的排列示于图 2 中。这显示淋球菌序列在 NMB1343 基因 5' 区域有一个 4 聚体插入，导致移码和随后失去 5' 甲硫氨酸残基。

结构域缺失-961

961 不存在于脑膜炎奈瑟球菌血清群 a 的基因组序列[81]中，尽管周围区域在血清群 A 和 B 间保守(>90%)。参考文献 11 和 12 揭示了 961 的多态型形式。发现此基因存在于 91%的血清群 B 菌株中，属于高毒性谱系 ET-5、ET-37 和 A4 簇，但该基因在所有测试的谱系 3 菌株中缺乏。即使不属于高毒性谱系，大部分测试的血清群 C 菌株为阳性。对于含血清型 2a 和 2b 的血清群 B 菌株也同样如此。对于血清群 A，属于亚群 III 的一菌株为阳性，而另两种属于亚群 IV-1 的菌株为阴性。淋病奈瑟菌中和乳酰胺奈瑟菌与脊髓灰质炎奈瑟球菌的共生种类中缺乏 961。

图 4 和 5 显示蛋白质 961 中的结构域。

当蛋白质 961 的锚定区域(结构域 9)缺失并在大肠杆菌中表达时，此蛋白质被输出到周质，分泌于培养物的上清中。

为进一步研究，在结构特征基础上(就 961c Δ aro 突变体而言缺失芳香族残基，和其它情况中缺失卷曲螺旋区)构建了缺失 961 C-末端区域的突变体(961cL- Δ aro、961cL Δ cc、961aL、961aL- Δ 1、961aL- Δ 2、961aL- Δ 3)。分析这些突变体表达和分泌入周质和培养物上清中的情况。在所有这些缺失突变体中，蛋白质大量产生、存在于周质组分中且释放到培养物上清中。

Δ G287-杂交菌株的杀菌活性

克隆了 5 个不同脑膜炎奈瑟球菌血清群 B 菌株的 287 并对其进行操作以去掉 N-末端至聚甘氨酸区域末端并引入 C-末端组氨酸尾。这产生了 5 种 Δ G287 蛋白质。这些蛋白质用 FCA 作佐剂用于产生小鼠免疫血清，随后测试了血清对所有 5 种血清群 B 菌株以及血清群 A 和 C 菌株的杀菌活性。杀菌滴度如下：

蛋白质 菌株	测试血清抗诸菌株的杀菌活性						
	2996	BZ232	MC58	1000	394/98	F6124	BZ133

2996	16000	128	4096	4096	1024	8000	16000
BZ232	>8000	256	2048	8000	2048	16000	8000
MC58	>8000	64	>8000	8000	2048	8000	8000
1000	>8000	64	4096	8000	1024	16000	16000
394/98	>16000	128	16000	>2048	>16000	-	-

*抗同源菌株的滴度以粗体表示

重折叠

为改进一些杂合蛋白质的可溶蛋白质水平, 采纳了参考文献 2 所述的另外重折叠方案。

包含体(IBs)的分离如下:

1. 细胞 (5mg 湿重) 用 ultraturrax (1000rpm) 4°C 在 25ml 0.1M Tris-HCl pH7.1、1mM EDTA 中匀浆。
2. 每克细胞加入 1.5ml 溶菌酶, 用 ultraturrax 简单混合, 4°C 孵育 30 分钟。
3. 用超声波或高压匀浆(French press)破裂细胞。
4. 为消化 DNA, 加入 MgCl₂ 到 3mM 终浓度和加 DNA 酶到 10 μg/ml 终浓度, 25°C 孵育 30 分钟。
5. 加入 0.5 体积的 60mM EDTA、6% Triton X-100、1.5M NaCl pH7 到溶液中, 4°C 孵育 30 分钟。
6. 4°C 31000g (20,000rpm) 离心 10 分钟沉降包含体。
7. 用 ultraturrax 重悬浮沉淀于 40ml 0.1M tris-HCl pH7、20mM EDTA 中。
8. 重复离心步骤 6。
9. 可使用包含体沉淀或 -20°C 冻存。

杂合蛋白质如下在大肠杆菌中表达:

蛋白质	培养物体积 (升)	烧瓶体积 (升)	温度 (°C)	最终的 OD ₆₀₀	产生的包 含体(w/w)
ORF46.1-961-His	1	2	37	1.51	33.2%
ORF46.1-961c-His	1	2	37	1.6	28.3%
961c- ORF46.1His	1	2	37	1.18	23.5%
orf46.1-741 His	5	5	37	12.42	35.2%

溶解沉淀、重折叠、超滤、透析, 随后纯化得到蛋白质:

ORF46.1-961-His IBs 溶解如下: 将 IB 蛋白质重悬浮于 4ml 6M 盐酸胍、1mM EDTA pH8.5 缓冲液中到 1mg/ml 的蛋白质终浓度。为重折叠此蛋白质, 取 2ml 溶解的蛋白质在 400ml 重折叠缓冲液(0.1M Tris-HCl、1M L-精氨酸、2mM EDTA pH8.2) 中稀释并 15°C 孵育 1 小时, 得到 5 μ g/ml 浓度的蛋白质。随后, 另加入 2ml 溶解的蛋白质并在相同温度再孵育 1 小时, 得到 10 μ g/ml 终浓度的蛋白质。该物质用 300ml Amicon 超滤装置(8400)超滤, 在 30kDa 截留值的 Amicon 膜(YM30)上施加 3 巴压力, 产生 130ml 的终体积。超滤后物质用 12-14kDa 截留值的再生纤维素管状(Cellusep-Step bio)膜以 10L 0.1M Tris-HCl pH8.2 缓冲液透析 24 小时。进行第二次透析, 以 10L 300mM NaCl、50mM 磷酸钠 pH8.0 缓冲液透析 24 小时。透析后物质在 Beckman 离心机转子 JA25.5 中以 22000rpm 4°C 离心 45 分钟。离心后分离的上清用于进行 His-尾纯化。

orf46.1-961c His IBs 溶解如下: 将 IB 蛋白质重悬浮于 4ml 6M 盐酸胍、1mM EDTA pH8.5 缓冲液中到 1mg/ml 的蛋白质终浓度。为重折叠此蛋白质, 取 2ml 溶解的蛋白质在 400ml 重折叠缓冲液(0.5M Tris-HCl、1M L-精氨酸、2mM EDTA pH8.2) 中稀释并 15°C 孵育 1 小时, 得到 5 μ g/ml 浓度的蛋白质。随后, 另加入 2ml 溶解的蛋白质并在相同温度下再孵育 1 小时, 得到 10 μ g/ml 终浓度的蛋白质。该物质用 300ml Amicon 超滤装置(8400)超滤, 在 30kDa 截留值的 Amicon 膜(YM30)上施加 3 巴压力, 产生 150ml 的终体积。超滤后物质用 12-14kDa 截留值的再生纤维素管状(Cellusep-Step bio)膜以 10L 0.1M Tris-HCl pH8.2 缓冲液透析 24 小时。进行第二次透析, 以 10L 300mM NaCl、50mM 磷酸钠 pH8.0 缓冲液透析 24 小时。透析后物质在 Beckman 离心机转子 JA25.5 中以 22000rpm 4°C 离心 45 分钟。离心后分离的上清用于进行 His-尾纯化。

961c- orf46.1His IBs 溶解如下: 将 IB 蛋白质重悬浮于 4ml 6M 盐酸胍、1mM EDTA pH8.5 缓冲液中到 1mg/ml 的蛋白质终浓度。为重折叠此蛋白质和 2ml 溶解的蛋白质在 400ml 重折叠缓冲液(0.1M Tris-HCl、0.5M L-精氨酸、2mM EDTA pH8.2) 中稀释并 15°C 孵育 1 小时, 得到 5 μ g/ml 浓度的蛋白质。随后, 另加入 2ml 溶解的蛋白质并在相同温度下再孵育 1 小时, 得到 10 μ g/ml 终浓度的蛋白质。该物质用 300ml Amicon 超滤装置(8400)超滤, 在 30kDa 截留值的 Amicon 膜(YM30)上施加 3 巴压力, 产生 150ml 的终体积。超滤后物质用 12-14kDa 截留值的再生纤维素管状(Cellusep-Step bio)膜以 10L 0.1M Tris-HCl pH8.2 缓冲液透析 24 小时。进行第二次透析, 以 10L 300mM NaCl、50mM 磷酸钠 pH8.0 缓冲液透析 24 小时。透

析后物质在 Beckman 离心机转子 JA25.5 中以 22000rpm 4℃离心 45 分钟。离心后分离的上清用于进行 His-尾纯化。

orf46.1-741 His IBs 溶解如下：将 IB 蛋白质重悬浮于 4ml 6M 盐酸胍、1mM EDTA pH8.5 缓冲液中到 10mg/ml 的蛋白质终浓度。为重折叠此蛋白质，取 2ml 溶解的蛋白质在 400ml 重折叠缓冲液(0.5M Tris-HCl、0.7M L-精氨酸、2mM EDTA pH7.2)中稀释并 15℃孵育 1 小时，得到 50 μg/ml 浓度的蛋白质。随后，另加入 2ml 溶解的蛋白质并在相同温度再孵育 1 小时，得到 100 μg/ml 终浓度的蛋白质。该物质用 300ml Amicon 超滤装置(8400)超滤，在 30kDa 截留值的 Amicon 膜(YM30)上施加 3 巴压力，产生 120ml 的终体积。超滤后物质用 12-14kDa 截留值的再生纤维素管状(Cellusep-Step bio)膜以 10L 0.1M Tris-HCl pH8.2 缓冲液透析 24 小时。进行第二次透析，以 10L 300mM NaCl、50mM 磷酸钠 pH8.0 缓冲液透析 24 小时。透析后物质在 Beckman 离心机转子 JA25.5 中以 22000rpm 4℃离心 45 分钟。离心后分离的上清用于进行 His-尾纯化。

与参考文献 2 所述纯化的蛋白质相比，杀菌试验滴定度如下：

蛋白质	参考文献 2		重折叠后		
	CFA	氢氧化铝	氢氧化铝	MF59	磷酸铝
ORF46.1-961-His	8192	8192	32768	-	-
ORF46.1-961c-His	8192	128	<64	8192	-
961c- ORF46.1His	32768	1024	16384	-	-
orf46.1-741 His	<4	16	<4	256	-

当没有 His-尾表达时(‘ORF46.1K’)，将类似过程用于 ORF46.1 以从 IBs 纯化此蛋白质：

蛋白质	培养物体积(升)	烧瓶体积(升)	温度(℃)	最终的 OD ₆₀₀	产生的包含体(w/w)
ORF46.1K	5	5	37	13.7	29.4

将 IB 蛋白质重悬浮于 4ml 6M 盐酸胍、1mM EDTA pH8.5 缓冲液中到 10mg/ml 的蛋白质终浓度。为重折叠，取 2ml 溶解的蛋白质在 400ml 重折叠缓冲液(0.5M Tris-HCl、0.7M L-精氨酸、2mM EDTA pH7.2)中稀释并 15℃孵育 1 小时，得到 50 μg/ml 浓度的蛋白质。随后，另加入 2ml 溶解的蛋白质并在相同温度下再孵育 1 小时，得到 100 μg/ml 终浓度的蛋白质。该物质用 300ml Amicon 超滤装置(8400)超滤，在 30kDa 截留值的 Amicon 膜(YM30)上施加 3 巴压力，产生 120ml 的终体

积。超滤后物质用 12-14kDa 截留值的再生纤维素管状(Cellusep-Step bio)膜以 10L 50mM 磷酸钠、2mM EDTA pH7.2 缓冲液透析 12 小时。以 10L 相同缓冲液进行第二次 24 小时透析。透析后物质在 Beckman 离心机转子 JA25.5 中以 22000rpm 4℃ 离心 45 分钟。离心后分离的上清用于阳离子交换层析。纯化在 AKTA 探测器层析系统(Amersham-Pharmacia Biotech)上进行, 采用 5ml HiTrap SP 琼脂糖 HP 柱 (Amersham-Pharmacia Biotech)。应用流速为 1.5ml/每分钟。尾析柱用 35ml 50mM 磷酸钠缓冲液 pH7.2 洗涤。用 50mM 磷酸钠缓冲液 pH7.2 进行线性梯度(0-1M NaCl)洗脱。蛋白质洗脱在 92mM 和 320mM NaCl 的两个峰中。收集含各个峰的组分并分别命名为合并液 1 和合并液 2。

与参考文献 2 所述纯化的蛋白质相比, 用氢氧化铝作佐剂时杀菌试验滴度从 <4 提到到 1024。此重折叠蛋白质用磷酸铝作佐剂的滴度为 2048。ELISA 滴度如下:

蛋白质	铝佐剂	Elisa (M7)	SBA (2996)
Orf46. 1k (合并液 1)	氢氧化铝 3. 3mg/ml	1212	512
	磷酸铝 0. 6mg/ml	154	1024
Orf46. 1k (合并液 2)	氢氧化铝 3. 3mg/ml	1085	1024
	磷酸铝 0. 6mg/ml	250	1024

应理解本发明仅通过实施例作了描述, 可进行修改且同时使这些修改保持在本发明的范围和思路内。

表 1

菌株	1343	序列	菌株分类
72/00	+		ET5 B:15:P1. 7, 13, 13a
30/00	+		ET5 B:15:P1. 7, 16
39/99	+		ET5 C:15:P1. 7, 16
95330	+		ET5 B:4:P1. 15
M4102	+		ET5 nd
MC58 (21)	+	+	ET5 B:15:P1. 7, 16b
BZ169 (7)	+	+	ET5 B:NT:P1. 16
BZ83 (19)	+		ET5 B:15:-. -
CU385	+	+	ET5 B:4:P1. 15
2201731	+		ET5 NG:4:P1. 15
64/96	+	+	ET5 NG:15:P1. 7, 16 (载体)

2201731	+		ET5 B:4:P1.15(载体)
ISS1071	+		nd B:15:P1.7,16(ET5?)
BZ198(2)	+	+	lin.3 B:8:P1.1
980-2543	+	+	lin.3 B:NT:P1.4
16060	+	+	其它 B:4:P1.14(载体)
394-98	+		nd B:4:P1.4(lin 3?)
ISS1106	+		nd B:4:P1.4(lin 3?)
BZ133(10)	+	+	sub I B:NT:-.-
S3446	+	+	nd B:14:P1.23,14
ISS1001	+	+	nd B:14:P1.13
2411751	+		其它 NG:21:P1.16(载体)
1712741	+		其它 NG:15:- (载体)
66/96	+		其它 B:17:P1.15 (载体)
961-5945	-		A4
96217	-		A4
312294	-		A4
90/18311(24)	-		ET37
93/4286(25)	-		ET37
M986	-		ET37
1000(5)	-		其它
NGE28(13)	-		其它载体
NGH38(14)	-		其它载体
BZ232(18)	-		其它
F6124(23)	-		sub III A:-.-
C11	-		C:-
NMB	-		nd
8047	-		nd
ISS759	-		nd C:2b:P1.2
ISS1113	-		nd C:2:P1.5
65/96	-		nd 4:P1.14
2996(96)	-		nd B:2b:P1.5,2

参考文献(在此并入以供参考)

- 1 - 国际专利申请 WO01/64920.
- 2 - 国际专利申请 WO01/64922.
- 3 - 国际专利申请 WO99/24578.
- 4 - 国际专利申请 WO99/36544.
- 5 - 国际专利申请 WO99/57280.
- 6 - 国际专利申请 WO00/22430.
- 7 - Tettelin 等, (2000) *Science* 287:1809-1815.
- 8 - 国际专利申请 WO00/66741.
- 9 - 国际专利申请 WO00/71574.
- 10 - 国际专利申请 WO01/04316.
- 11 - 国际专利申请 PCT/IB02/03396.
- 12 - Comanducci 等, (2002) *J Exp Med* 195:1445-1454.
- 13 - *Vaccine Design: subunit & adjuvant approach* (1995) Powell & Newman (ISBN: 030644867X).
- 14 - 国际专利申请 WO01/52885.
- 15 - Bjune 等, (1991) *Lancet* 338(8775):1093-1096.
- 16 - Fukasawa 等, (1999) *Vaccine* 17:2951-2958.
- 17 - Rosenqvist 等, (1998) *Dev. Biol. Stand.* 92:323-333.
- 18 - Costantino 等, (1992) *Vaccine* 10:691-698.
- 19 - Costantino 等, (1999) *Vaccine* 17:1251-1263.
- 20 - 国际专利申请 PCT/IB02/03191.
- 21 - Watson (2000) *Pediatr Infect Dis J* 19:331-332.
- 22 - Rubin (2000) *Pediatr Clin North Am* 47:269-285, v.
- 23 - Jedrzejewski (2001) *Microbiol Mol Biol Rev* 65:187-207.
- 24 - 国际专利申请 WO93/18150.
- 25 - 国际专利申请 WO99/53310.
- 26 - 国际专利申请 WO98/04702.
- 27 - Bell (2000) *Pediatr Infect Dis J* 19:1187-1188.
- 28 - Iwarson (1995) *APMIS* 103:321-326.
- 29 - Gerlich 等, (1990) *Vaccine* 8 Suppl:S63-68 & 79-80.
- 30 - Hsu 等, (1999) *Clin Liver Dis* 3:901-915.
- 31 - Gustafsson 等, (1996) *N. Engl. J. Med.* 334:349-355.
- 32 - Rappuoli 等, (1991) *TIBTECH* 9:232-238.
- 33 - *Vaccines* (1988) 编, Plotkin & Mortimer. ISBN 0-7216-1946-0.
- 34 - Del Giudice 等, (1998) *Molecular Aspects of Medicine* 19:1-70.
- 35 - 国际专利申请 WO02/02606.
- 36 - Kalman 等, (1999) *Nature Genetics* 21:385-389.
- 37 - Read 等, (2000) *Nucleic Acids Res* 28:1397-406.
- 38 - Shirai 等, (2000) *J. Infect. Dis.* 181(Suppl 3):S524-S527.
- 39 - 国际专利申请 WO99/27105.
- 40 - 国际专利申请 WO00/27994.
- 41 - 国际专利申请 WO00/37494.
- 42 - 国际专利申请 WO99/28475.
- 43 - Ross 等, (2001) *Vaccine* 19:4135-4142.
- 44 - Sutter 等, (2000) *Pediatr Clin North Am* 47:287-308.
- 45 - Zimmerman & Spann (1999) *Am Fam Physician* 59:113-118, 125-126.

- 46 – Dreesen (1997) *Vaccine* 15 Suppl:S2-6.
- 47 – *MMWR Morb Mortal Wkly Rep* 1998 Jan 16;47(1):12, 19.
- 48 – McMichael (2000) *Vaccine* 19 Suppl 1:S101-107.
- 49 – Schuchat (1999) *Lancet* 353(9146):51-6.
- 50 – WO02/34771.
- 51 – Dale (1999) *Infect Dis Clin North Am* 13:227-43, viii.
- 52 – Ferretti 等, (2001) *PNAS USA* 98: 4658-4663.
- 53 – Kuroda 等, (2001) *Lancet* 357(9264):1225-1240; 也可见1218-1219页.
- 54 – Ramsay 等, (2001) *Lancet* 357(9251):195-196.
- 55 – Lindberg (1999) *Vaccine* 17 Suppl 2:S28-36.
- 56 – Buttery & Moxon (2000) *J R Coll Physicians Lond* 34:163-168.
- 57 – Ahmad & Chapnick (1999) *Infect Dis Clin North Am* 13:113-133, vii.
- 58 – Goldblatt (1998) *J. Med. Microbiol.* 47:563-567.
- 59 – 欧洲专利0 477 508.
- 60 – 美国专利5,306,492.
- 61 – 国际专利申请WO98/42721.
- 62 – *Conjugate Vaccines* (Cruse 等编,) ISBN 3805549326, 特别卷,10:48-114.
- 63 – Hermanson (1996) *Bioconjugate Techniques* ISBN: 0123423368 或 012342335X.
- 64 – 欧洲专利申请0372501.
- 65 – 欧洲专利申请 0378881.
- 66 – 欧洲专利申请0427347.
- 67 – 国际专利申请WO93/17712.
- 68 – 国际专利申请 WO98/58668.
- 69 – 欧洲专利申请0471177.
- 70 – 国际专利申请WO00/56360.
- 71 – 国际专利申请WO00/61761.
- 72 – Robinson & Torres (1997) *Seminars in Immunology* 9:271-283.
- 73 – Donnelly 等, (1997) *Annu Rev Immunol* 15:617-648.
- 74 – Scott-Taylor & Dalgleish (2000) *Expert Opin Investig Drugs* 9:471-480.
- 75 – Apostolopoulos & Plebanski (2000) *Curr Opin Mol Ther* 2:441-447.
- 76 – Ilan (1999) *Curr Opin Mol Ther* 1:116-120.
- 77 – Dubensky 等, (2000) *Mol Med* 6:723-732.
- 78 – Robinson & Pertmer (2000) *Adv Virus Res* 55:1-74.
- 79 – Donnelly 等, (2000) *Am J Respir Crit Care Med* 162(4 Pt 2):S190-193.
- 80 – Davis (1999) *Mt. Sinai J. Med.* 66:84-90.
- 81 – Parkhill 等, (2000) *Nature* 404:502-506.

序列表

SEQ ID 1 - 741 来自菌株1000

5 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTPLDHKDKSLQSLTLDQSVRKNEK
LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQITLASGEFQIYKQNHSAVVALQIEKIN
NPKIDSLINQRSFLVSLGGGEHTAFNQLPDGKAEYHGKAFSSDDPNRGLHYSIDFTKKQGYGRIEHLKT
PEQVELASAELKADEKSHAVILGDTRYGGEKGTYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ

*SEQ ID 2 - 741*来自菌株2201731（提前终止密码子），通过可靠序列

10 MTRSKPVNRTAFCCLSLTTALILTACSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLTLDQSVRKNEK
LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQVYKQSHSALTAFQTEQIQ
DSEHSGKMAKRQFRIGDIAGEHTSFDKLEGGGRATYRGTAFGSDDAGGKLYTIDFAAQQNGKIEHLK
SPELNVDLAAADIKPDGKRHAVISGSVLYNQAEGSYSLGIFGGKA

SEQ ID 3 - 741 来自菌株90/18311（不完全）

15 GLADALTAPLDHKDKSLQSLTLDQSVRKNEK LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEV
DGQLITLESGEFQIYKQDHSAVVALQIEKIN NPKIDSLINQRSFLVSLGGGEHTAFNQLPSGKAEYHGK
AFSSDDPNRGLHYSIDFTKKQGYGRIEHLKTPEQVELASAELKADEKSHAVILGDTRYGGEKGTYHLA
LFGDRAQEIAGSATVKIREKVHET

SEQ ID 4 - 741 来自菌株L93/4286（不完全）

20 VAADIGAGLADALTAPLDHKDKGLOSMLDOSVRKNEK LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFD
FIRQIEVDGQITLASGEFQIYKQNHSAVVALQIEKIN NPKIDSLINQRSFLVSLGGGEHTAFNQLPDG
KAEYHGKAFSSDDPNRGLHYSIDFTKKQGYGRIEHLKTPEQVELASAELKADEKSHAVILGDTRYGGE
KGTYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ

SEQ ID 5 - 741 来自菌株 2996

25 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN
NPKIDSLINQRSFLVSLGGGEHTAFNQLPDGKAEYHGKAFSSDDAGGKLYTIDFAAQQHGKIEHLKT
PEQVELAAAELKADEKSHAVILGDTRYGSEEKGTYHLALFGDRAQEIAGSATV KIGEKVHEIGIAGKQ

SEQ ID 6 - 741 来自菌株 30/00

30 KDKGLQSLTLDQSVRKNEK LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEF
QVYKQSHSALTAFQTEQIQDSEHSGKMAKRQFRIGDIAGEHTSFDKLEGGGRATYRGTAFGSDDAGGKLT
TYTIDFAAQQNGKIEHLKSPENVDLAAADIKPDGKRHAVISGSVLYNQAEGSYSLGIFGGKAQEVAG
SAEVTVNGIRHIGLAAKQ

SEQ ID 7 - 741 来自菌株 312294

35 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN
NPKIDSLINQRSFLVSLGGGEHTAFNQLPDGKAEYHGKAFSSDDAGGKLYTIDFAAQQHGKIEHLKT
PEQVELAAAELKADEKSHAVILGDTRYGSEEKGTYHLALFGDRAQEIAGSATV KIGEKVHEIGIAGKQ

SEQ ID 8 - 741 来自菌株 39/99（不完全）

40 DKGLQSLTLDQSVRKNEK LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQ
VYKQSHSALTAFQTEQIQDSEHSGKMAKRQFRIGDIAGEHTSFDKLEGGGRATYRGTAFGSDDAGGKLT
YTIDFAAQQNGKIEHLKSPENVDLAAADIKPDGKRHAVISGSVLYNQAEGSYSLGIFGGKAQEVAGS
AEVTVNGIRHIGLAAKQ

SEQ ID 9 - 741 来自菌株 5/99

45 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLTLDQSVRKNEK
LKLAAQGAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN
NPKIDSLINQRSFLVSLGGGEHTAFNQLPSGKAEYHGKAFSSDDPNRGLHYSIDFTKKQGYGRIEHLKT
PEQVELASAELKADEKSHAVILGDTRYGGEKGTYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ

SEQ ID 10 - 741 来自菌株 67/00

MTRSKPVNRTAFCCFSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTEQEQ
 DPEHSGKMAKRRFKIGDIAGEHTSFDKLPKDVMTYRGTAFGSDDAGGKLYTIDFAAQQGHGKIEHLK
 5 SPELNVELATAYIKPDEKHHAVISGSVLYNQDEKGSYSLGIFGGQAQEVAGSAEVETANGIHHIGLAAKQ

SEQ ID 11 - 741 来自菌株 BZ169

LQSLTLDQSVRKNEK LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQVYK
 QSHSALTALQTEQIQDSEHSGKMAKRRQFRIGDIAGEHTSFDKLPPEGGRATYRGTAFGSDDAGGKLYTID
 DFAAQQNGKIEHLK SPELNVDLAAADIKPDGKRHAVISGSVLYNQAEKGSYSLGIFGGQAQEVAGSAEV
 10 KTVNGIRHIGLAAKQ

SEQ ID 12 - 741 来自菌株 72/00

LQSLTLDQSVRKNEK LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQVYK
 QSHSALTALQTEQIQDSEHSGKMAKRRQFRIGDIAGEHTSFDKLPPEGGRATYRGTAFGSDDAGGKLYTID
 DFAAQQNGKIEHLK SPELNVDLAAADIKPDGKRHAVISGSVLYNQAEKGSYSLGIFGGQAQEVAGSAEV
 15 KTVNGIRHIGLAAKQ

SEQ ID 13 - 741 来自菌株 93/114

MTRSKPVNRTAFCCFSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTEQEQ
 DPEHSGKMAKRRFKIGDIAGEHTSFDKLPKDVMTYRGTAFGSDDAGGKLYTIDFAAQQGHGKIEHLK
 20 SPELNVELATAYIKPDEKHHAVISGSVLYNQDEKGSYSLGIFGGQAQEVAGSAEVETANGIHHIGLAAKQ

SEQ ID 14 - 741 来自菌株 95N477

MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN
 NPDKIDSLINQRSFLVSLGGGEHTAFNQLPSGKAIEYHGKAFSSDDPNGLRHLHYSIDFTKKQGYGRIEHLK
 25 PEQNVELASAELKADEKSHAVILGDTRYGGEEKGTYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ

SEQ ID 15 - 741 来自菌株 96217

MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN
 NPDKIDSLINQRSFLVSLGGGEHTAFNQLPDGKAIEYHGKAFSSDDAGGKLYTIDFAAQQGHGKIEHLK
 30 PEQNVELAAAEKKADEKSHAVILGDTRYGSEEEKGTYHLALFGDRAQEIAGSATVKIGEKVHEIGIAGKQ

SEQ ID 16 - 741 来自菌株 BZ133

MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTEQVQ
 DSEHSGKMAKRRQFRIGDIAGEHTSFDKLPPEGGRATYRGTAFGSDDASGKLYTIDFAAQQGHGKIEHLK
 35 SPELNVDLAAADIKPDKRRHAVISGSVLYNQAEKGSYSLGIFGGQAQEVAGSAEVETANGIRHIGLAAKQ

SEQ ID 17 - 741 来自菌株 BZ232

MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQITITLASEGFQIYKQNHSAVVALQIEKIN
 NPDKIDSLINQRSFLVSLGGGEHTAFNQLPDGKAIEYHGKAFSSDDPNGLRHLHYSIDFTKKQGYGRIEHLK
 40 PEQNVELASAELKADEKSHAVILGDTRYGGEEKGTYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ

SEQ ID 18 - 741 来自菌株 C11

MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN
 NPDKIDSLINQRSFLVSLGGGEHTAFNQLPSGKAIEYHGKAFSSDDPNGLRHLHYSIDFTKKQGYGRIEHLK
 45 PEQNVELASAELKADEKSHAVILGDTRYGGEEKGTYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ

SEQ ID 19 - 741 来自菌株 M1090

MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
 LKLAQAQAEKTYGNGDSLNTGKLNKDKVSRFDFIRQIEVDGQLITLESGEFQIYKQDHSAVVALQIEKIN

NPKIDSLINQRSFLVSLGGEHTAFNQLPSGKAEYHGKAFSSDDAGGKLYTIDFAAKQGHGKIEHLKT
PEQNVELASAEKKADEKSHAVILGDTRYGGEKGYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ
SEQ ID 20 - 741 来自菌株 M1096

5 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTPLDHKDKSLQSLTLDQSVRKNEK
LKLAQAQGAEKTYGNGDSLNTGKLNKDKVSRDFIRQIEVDGQTITLASGEFQIYKQNHSAVVALQIEKIN
NPKIDSLINQRSFLVSLGGEHTAFNQLPDGKAEYHGKAFSSDDPNRGLHYSIDFTKKQGYGRIEHLKT
PEQNVELASAEKKADEKSHAVILGDTRYGGEKGYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ
SEQ ID 21 - 741 来自菌株 M198/172

10 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
LKLAQAQGAEKTYGNGDSLNTGKLNKDKVSRDFIRQIEVDGQLITLESGEFQVYKQSHSALTALQTEBQVQ
DSEHSGKMVAKRQFRIGDIAGEHTSFDKLPEGGRATYRGTAFGSDDASGKLYTIDFAAKQGHGKIEHLK
SPELNVDLAASDIKPKDKRHAVISGSVLYNQAEKGSYSLGIFGGQAQEVAGSAEVETANGIRHIGLAAKQ
SEQ ID 22 - 741 来自菌株 NGH38

15 MTRSKPVNRTAFCCLSLTAALILTACSSGGGGVAADIGAGLADALTAPLDHKDKSLQSLTLDQSVRKNEK
LKLAQAQGAEKTYGNGDSLNTGKLNKDKVSRDFIRQIEVDGQTITLASGEFQIYKQNHSAVVALQIEKIN
NPKIDSLINQRSFLVSLGGEHTAFNQLPDGKAEYHGKAFSSDDPNRGLHYSIDFTKKQGYGRIEHLKT
PEQNVELASAEKKADEKSHAVILGDTRYGGEKGYHLALFGDRAQEIAGSATVKIREKVHEIGIAGKQ
SEQ ID 23 - NMB1343 来自十种脑膜炎球菌菌株

20 MGNELYRGI SCQDEQNNGQLKPKGNKAEVAIRYDGFYDYGKATHGSPVKNVYAHQIETGLYDGCYIS
TTTTKEIAKKFATSSGIENGYIYVLRDLFGQYSIFEYEVEHPENPNEKEVTIRAEDCGCIPPEVIAKE
LIEIN

SEQ ID 24 - NMB1343 来自淋球菌

25 INNLEWISYLYRGI SCQDEQNNGQLKPKGNKAEVAIRYDGFYDYGKATHGSPVKNVYAHQIETDLYD
GCYISTTTTKEIAKKFATSSGIENGYIYVLRDLFGQYSIFEYEVEHPENPDEKEVTIRAEDCGCIPPEV
IIAKELIEIN

SEQ ID 25 - NMB1343 核酸序列 (淋球菌)

30 TCCGCCGCATTACCTTATAAAAATAAAACATCCCTCTCAAGCAGTCTGATAATGTTTGGATTGCTTGAGAT
TGATGAGTGATGGTGTAAATTCAAACCTTTAAATTAATAACTTATGGGAAATTTCTTATATAGAGG
CATTAGTTGCCAACAAAGATGAGCAAAATAATGGACAGTTAAACCTAAAGGTAATAAAGCTGAAGTTGCA
ATTCGTTATGATGGTAAGTTTAAATATGATGGTAAAGCTACACATGGTCCAAAGTGTGAAGAATGCAGTTT
ACGCCCATCAAATTGAAACAGATCTATATGACGGATGTTATATATCTACGACAACAGACAAGGAAATTC
CAAGAAATTTGCAACAAGCTCCGGCATCGAAAATGGCTATATATATGTTTAAATAGAGATTTGTTGGT
CAATATCTATTTTTGAATATGAGGTTGAACATCCAGAAAACCCAGATGAGAAGGAAGTAACAATCAGAG
CTGAAGATTGTGGCTGTATTCCTGAAGAAGTGATTATTGCTAAAGAGTTGATAGAAATTAACATAAGTTGA
35 AAGGTCAATATAATGGCTTTAGTTGAATTGAAAGTGCCCGACATTTGGCGGACACGAAAATGTAGATATTA
TCGC

SEQ ID 26 - NMB1343 核酸序列 (脑膜炎球菌)

40 TCCGCCGCATTACCTTATAAAAATAAAACATCCCTCTCAAGCAGTCTGATAATGTTTGGATTGCTTGAGAT
TGATGAGTAATGGTGTAAATTCAAACCTTTAAATTAATAACTTATGGGAAATTTCTTATATAGAGGCATT
AGTTGCCAACAAAGATGAGCAAAATAATGGACAGTTAAACCTAAAGGTAATAAAGCTGAAGTTGCAATTC
GTTATGATGGTAAGTTTAAATATGATGGTAAAGCTACACATGGTCCAAAGTGTGAAGAATGCAGTTTACGC
CCATCAAATTGAAACAGGCTATATGACGGATGTTATATATCTACGACAACAGACAAGGAAATGCAAG
AAATTTGCAACAAGTTCCGGCATCGAAAATGGCTATATATATGTTTAAATAGGGATTTGTTGGTCAAT
ATTTCTATTTTTGAATATGAGGTTGAACATCCAGAAAACCCAAATGAGAAGGAAGTAACAATCAGAGCTGA
45 AGATTGTGGCTGTATTCCTGAAGAAGTGATTATTGCTAAAGAGTTGATAGAAATTAACATAAGTTGAAAGG
TCAATATAATGGCTTTAGTTGAATTGAAAGTGCCCGACATTTGGCGGACACGAAAATGTAGATATTAATTCGC

SEQ ID 27 - 接头

GSGGGG

SEQ ID 28 - 优选 AG287-953 杂交体

MASPDVKSADTLISKPAAPVVAEKETEVEDAPQAGSQGQAPSTQGSQDMAAVSAENTGNNGAATTDKPK
 NEDEGPQNDMPQNSAESANQTGNNQPADSSDSAPASNPAANGGNSFRVLDLANGVLDGSPQNTLTHC
 KGDSCNGDNLLDEEAPSKSEFENLNERIEKYKKGKSDKFTNLVAVQANGTNKYVIYKDKSASSS
 5 SARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNRYRLTYGAEKLPGGSYALR
 VQGEPAKGEMLAGTAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDGIIDSGDDLHMGTQKFKA
 DGNGFKGTWTENGCGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGFGVFAGKKEQDGS~~GGGG~~GATYKVDEYH
 ANARFAIDHFNTSTNVGGFYGLTGSVEFDQAKRDGKIDITIPVANLQSGSQHF~~TDHL~~KSADIFDAAQYPD
 IRFVSTKFNFGKKLVSVDGNLTMHGKTAPVKLKA~~EKF~~NQYQSPMAKTEVCGGDFSTTIDRTKWGV~~DYLV~~
 10 NVGMTKSVRIDIQIEAAKQ

SEQ ID 29 - AG287_{NZ}-953 杂交体

MASPDVKSADTLISKPAAPVVSEKETEAKEDAPQAGSQGQAPSAQGGQDMAAVSEENTGNNGAAATDKPK
 NEDEGAQNMPQNAADTDSLTPNHTPASNMPAGNMENQAPDAGESEQPANQPD~~MANTADGMQDD~~PSAGG
 ENAGNTAAQGTNQAENNTAGSQNPASSTNPSATNSGGDFGR~~TNVGNSVVIDGSPQNTLTHC~~KGDS~~CSG~~
 15 NNFLDEEVQLKSEFEKLSADKISNYKKGKNDGKNDK~~FVGLVADSVQMG~~INQYIIFYKPKP~~TSFARFR~~
 RSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNRYRLTYGAEKLP~~GGSYALRVQGE~~
 SKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGIIDSGDGLHMGTQKFKA
 IDGNGFKGTWTENGCGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGFGVFAGKKEQDGS~~GGGG~~GATYKVDEYHANARF
 AIDHFNTSTNVGGFYGLTGSVEFDQAKRDGKIDITIPVANLQSGSQHF~~TDHL~~KSADIFDAAQYPD
 20 IRFVSTKFNFGKKLVSVDGNLTMHGKTAPVKLKA~~EKF~~NQYQSPMAKTEVCGGDFSTTIDRTKWGV~~DYLV~~
 NVGMTKSVRIDIQIEAAKQ

SEQ ID 30 - 936-AG741 杂交体

MKPKPHTVRTLIAAIFSLALSGCVSAVIGSAAVGA~~SAVDRRTTGAQTD~~DDNVMALRIETTARSYL~~RQNNQ~~
 TKGYTPQISVVGYNRHL~~LL~~LGGVATEGEKQFV~~QGIARSEQAAEGVYNYITVASL~~PRTAGDIAGDTWNTSK
 25 VRATLLGISPATQARVKIVTYGNVTVMGILTPEEQAQITQKVSTTVGVQKVITLYQNYV~~QRGSSGGG~~VVA
 ADIGAGLADALTAPLDHKDKGLQSLTL~~DQSVRKNEK~~LKLAQAQGAEKTYGNGDSLNTGK~~LKNDKVS~~RFD~~FI~~
 RQIEVDGQLITLESGEFQVYKQSHSALTAFQTEQIQDSEHSGK~~MAKRQFRIGDIAGEHTSF~~DKLPEGGR
 ATYRGTAFGSDDAGGKLT~~YTFIDFAAKQ~~GNGKIEHLKSP~~ELNVDLAAADIKPDGKR~~HAVISGSVLYNQA~~EK~~
 GSYSLGIFGGKAQEVAGSAEVKTVNGIRHIGLAAKQ

30 SEQ ID 31 - 961c

MATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLK~~KVV~~TN
 LTKTVNENKQNVDAKVAAESEIEKLT~~TKLADTDAALADTDAALDAT~~TNALNKLGENIT~~TTFAET~~KTNI~~V~~
 KIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADEAVKTANEAKQTAEETKQNVDAKVAAETAAGKAE
 AAAGTANTAADKAEVA~~AKVTDIKADIATNKDNI~~AKKANSADVY~~TREES~~DSK~~FVRIDGLNATTE~~KL~~DTRL~~
 35 ASAEKSIADHD~~TRLNGLDKTV~~SDLRKE~~TRQGLAEQAAL~~SGLFQ~~PYN~~VG

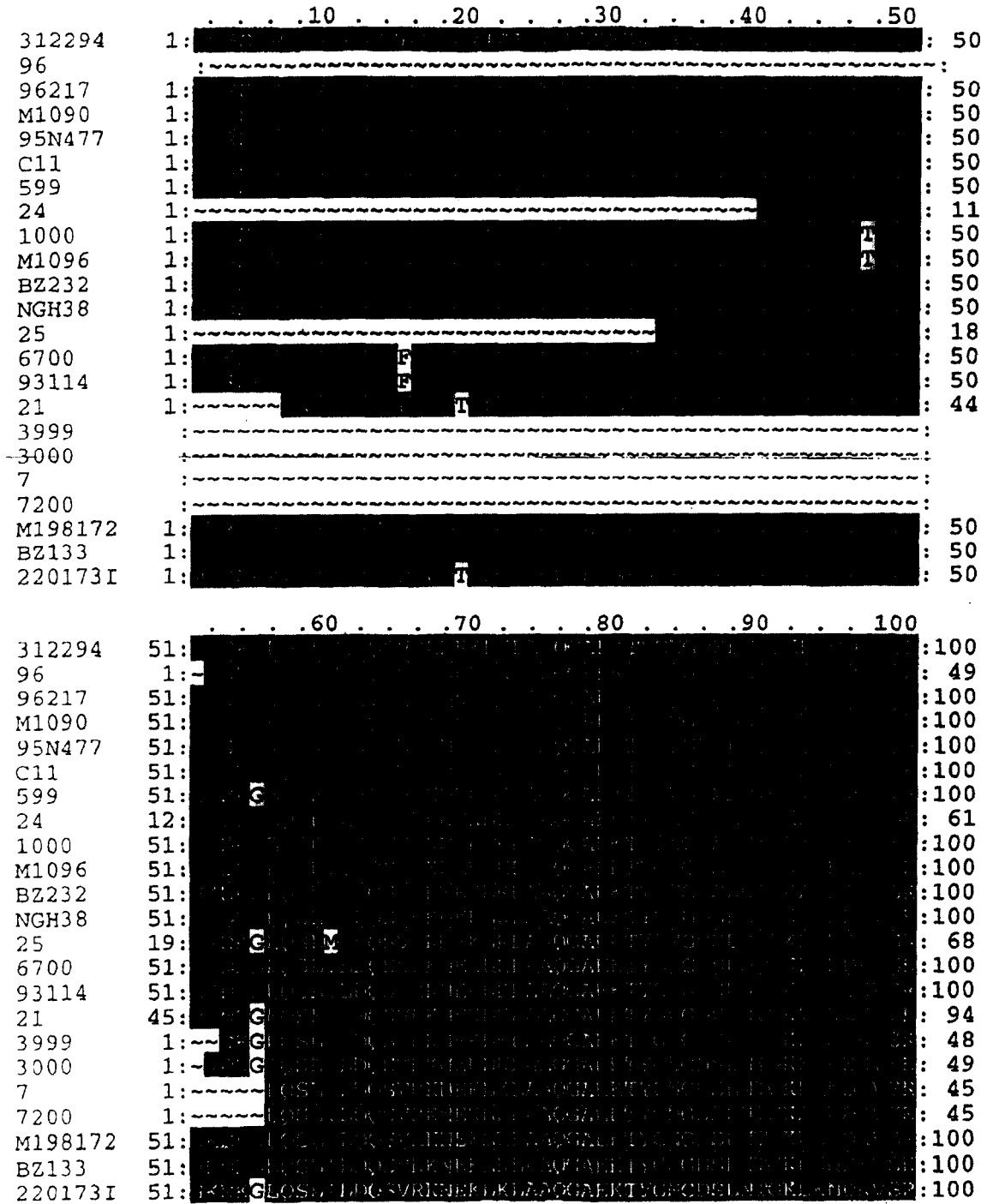


图 1

```

      . . . 110 . . . 120 . . . 130 . . . 140 . . . 150
312294 101: [REDACTED] I K N D K I D S E I N : 150
96      50: [REDACTED] I K N D K I D S E I N : 99
96217  101: [REDACTED] I K N D K I D S E I N : 150
M1090  101: [REDACTED] I K N D K I D S E I N : 150
95N477 101: [REDACTED] I K N D K I D S E I N : 150
C11     101: [REDACTED] I K N D K I D S E I N : 150
599     101: [REDACTED] I K N D K I D S E I N : 150
24      62: [REDACTED] I K N D K I D S E I N : 111
1000    101: [REDACTED] I K N D K I D S E I N : 150
M1096   101: [REDACTED] I K N D K I D S E I N : 150
BZ232   101: [REDACTED] I K N D K I D S E I N : 150
NGH38   101: [REDACTED] I K N D K I D S E I N : 150
25      69: [REDACTED] I K N D K I D S E I N : 118
6700    101: [REDACTED] S E I T F T Q O D S E H S G K M V A : 150
93114   101: [REDACTED] S E I T F T Q O D S E H S G K M V A : 150
21      95: [REDACTED] S E I T F T Q O D S E H S G K M V A : 144
3999    49: [REDACTED] S E I T F T Q O D S E H S G K M V A : 98
3000    50: [REDACTED] S E I T F T Q O D S E H S G K M V A : 99
7        46: [REDACTED] S E I T F T Q O D S E H S G K M V A : 95
7200    46: [REDACTED] S E I T F T Q O D S E H S G K M V A : 95
M198172 101: [REDACTED] S E I T F T Q O D S E H S G K M V A : 150
BZ133   101: [REDACTED] S E I T F T Q O D S E H S G K M V A : 150
220173I 101: [REDACTED] S E I T F T Q O D S E H S G K M V A : 150

```

```

      . . . 160 . . . 170 . . . 180 . . . 190 . . . 200
312294 151: Q S L V S G L G [REDACTED] A N Q [REDACTED] E I K K S [REDACTED] G [REDACTED] : 199
96      100: Q S L V S G L G [REDACTED] A N Q [REDACTED] E I K K S [REDACTED] G [REDACTED] : 148
96217  151: Q S L V S G L G [REDACTED] A N Q [REDACTED] E I K K S [REDACTED] G [REDACTED] : 199
M1090  151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
95N477 151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
C11     151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
599     151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
24      112: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 160
1000    151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
M1096   151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
BZ232   151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
NGH38   151: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 199
25      119: Q S L V S G L G [REDACTED] A N Q [REDACTED] S [REDACTED] E I K K S [REDACTED] P N R H S [REDACTED] T K : 167
6700    151: K R K I G D I A [REDACTED] S D K [REDACTED] K D V M [REDACTED] T R S T [REDACTED] G [REDACTED] : 200
93114   151: K R K I G D I A [REDACTED] S D K [REDACTED] K D V M [REDACTED] T R S T [REDACTED] G [REDACTED] : 200
21      145: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 194
3999    99: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 148
3000    100: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 149
7        96: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 145
7200    96: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 145
M198172 151: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 200
BZ133   151: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 200
220173I 151: K Q R I G D I A [REDACTED] S D K [REDACTED] G R T R S T [REDACTED] G [REDACTED] : 200

```

图 1续

```

      . . . 210 . . . 220 . . . 230 . . . 240 . . . 250
312294 200: [REDACTED] A A A S L DTR G S [REDACTED] A :249
96      149: [REDACTED] A A A S L DTR G S [REDACTED] A :198
96217  200: [REDACTED] A A A S L DTR G S [REDACTED] A :249
M1090  200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
95N477 200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
C11     200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
599     200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
24      161: [REDACTED] S A A S L DTR G G [REDACTED] A :210
1000    200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
M1096   200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
BZ232   200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
NGH38   200: [REDACTED] S A A S L DTR G G [REDACTED] A :249
25      168: [REDACTED] S A A S L DTR G G [REDACTED] A :217
6700    201: [REDACTED] T Y F H S SVL N Q D S S G :250
93114   201: [REDACTED] T Y F H S SVL N Q D S S G :250
21      195: [REDACTED] D A D P G R S SVL N Q A S S G :244
3999    149: [REDACTED] D A D P G R S SVL N Q A S S G :198
3000    150: [REDACTED] D A D P G R S SVL N Q A S S G :199
7        146: [REDACTED] D A D P G R S SVL N Q A S S G :195
7200    146: [REDACTED] D A D P G R S SVL N Q A S S G :195
M198172 201: [REDACTED] D A D P K R S SVL N Q A S S G :250
BZ133   201: [REDACTED] D A D P K R S SVL N Q A S S G :250
220173I 201: [REDACTED] D A D P G R S SVL N Q A S S G :250

```

```

      . . . 260 . . . 270 . . . 280
312294 250: [REDACTED] T I G E K V E I G :279
96      199: [REDACTED] T I G E K V E I G :212
96217  250: [REDACTED] T I G E K V E I G :279
M1090  250: [REDACTED] T I G E K V E I G :279
95N477 250: [REDACTED] T I G E K V E I G :279
C11     250: [REDACTED] T I G E K V E I G :279
599     250: [REDACTED] T I G E K V E I G :279
24      211: [REDACTED] T I G E K V E T ~ ~ ~ ~ ~ :234
1000    250: [REDACTED] T I G E K V E I G :279
M1096   250: [REDACTED] T I G E K V E I G :279
BZ232   250: [REDACTED] T I G E K V E I G :279
NGH38   250: [REDACTED] T I G E K V E I G :279
25      218: [REDACTED] T I G E K V E I G :247
6700    251: [REDACTED] E T A N G I R H L A :280
93114   251: [REDACTED] E T A N G I R H L A :280
21      245: [REDACTED] E T V N G I R H L A :274
3999    199: [REDACTED] E T V N G I R H L A :228
3000    200: [REDACTED] E T V N G I R H L A :229
7        196: [REDACTED] E T V N G I R H L A :225
7200    196: [REDACTED] E T V N G I R H L A :225
M198172 251: [REDACTED] E T A N G I R H L A :280
BZ133   251: [REDACTED] E T A N G I R H L A :280
220173I 251: [REDACTED] E T A N G I R H L A :256

```

图 1续

```

      10      20      30      40      50      60      70
TCCGCCGATTACCTTATAAAATAAAACATCCCTCTCAAGCAGTCTGATAATGTTTGGATTGCTTGAGATTGATGAG
.....
TCCGCCGATTACCTTATAAAATAAAACATCCCTCTCAAGCAGTCTGATAATGTTTGGATTGCTTGAGATTGATGAG

      80      90      100     110     120     130     140     150
TGATGGTGTAAATTCAAACCTTAAATTAATAACTTATGGGAAATTTCTTATTTATATAGAGGCATTAGTTGCCAAC
.....
TAATGGTGTAAATTCAAACCTTAAATTAATAACTTATGGGAAATTTCTTA----TATAGAGGCATTAGTTGCCAAC

      160     170     180     190     200     210     220     230
AAGATGAGCAAAATAATGGACAGTTAAACCTAAAGGTAATAAAGCTGAAGTTGCAATTCGTTATGATGGTAAGTTT
.....
AAGATGAGCAAAATAATGGACAGTTAAACCTAAAGGTAATAAAGCTGAAGTTGCAATTCGTTATGATGGTAAGTTT

      240     250     260     270     280     290     300
AAATATGATGGTAAAGCTACACATGGTCCAAGTGTGAAGAATGCAGTTTACGCCCATCAAATTGAAACAGATCTATA
.....
AAATATGATGGTAAAGCTACACATGGTCCAAGTGTGAAGAATGCAGTTTACGCCCATCAAATTGAAACAGATCTATA

      310     320     330     340     350     360     370     380
TGACGGATGTTATATATCTACGACAACAGACAAGGAAATTGCCAAGAAATTTGCAACAAGTCCGGCATCGAAAATG
.....
TGACGGATGTTATATATCTACGACAACAGACAAGGAAATTGCCAAGAAATTTGCAACAAGTCCGGCATCGAAAATG

      390     400     410     420     430     440     450     460
GCTATATATATGTTTTAAATAGAGATTGTTTGGTCAATATCTATTTTTGAATATGAGGTTGAACATCCAGAAAAC
.....
GCTATATATATGTTTTAAATAGGGATTGTTTGGTCAATATCTATTTTTGAATATGAGGTTGAACATCCAGAAAAC

      470     480     490     500     510     520     530
CCAGATGAGAAGGAAGTAACAATCAGAGCTGAAGATTGTGGCTGTATTCCTGAAGAAGTGATTATTGCTAAAGAGTT
.....
CCAAATGAGAAGGAAGTAACAATCAGAGCTGAAGATTGTGGCTGTATTCCTGAAGAAGTGATTATTGCTAAAGAGTT

      540     550     560     570     580     590     600     610
GATAGAAATTAACATAAGTTGAAAGGTCAATATAATGGCTTTAGTTGAATTGAAAGTGCCCGACATTGGCGGACACGA
.....
GATAGAAATTAACATAAGTTGAAAGGTCAATATAATGGCTTTAGTTGAATTGAAAGTGCCCGACATTGGCGGACACGA

      620     630
AAATGTAGATATTATCGC
.....
AAATGTAGATATTATCGC

```

图 2

Δ G287-919-His			
Δ G287-Orf46.1-His	961c-741 _{MC58} -His		Orf46.1-287-His
Δ G287-953-His	961c-983-His		Orf46.1-919-His
Δ G287-961-His	961c-Orf46.1-His		Orf46.1-741 _{MC58} -His
Δ G287-230-His	961cL-741 _{MC58}		Orf46.1-961-His
Δ G287-936-His	961cL-287		Orf46.1-961c-His
Δ G287-287-His	961c-230-His		Orf46.1-983-His
Δ G287-287 _{nz} -His	961c-936-His		Orf46.1-936-His
Δ G287-741 _{MC58} -His			Orf46.1-230-His
Δ G287-741 _{ET37} -His			
Δ G287 _{nz} -919-His	Δ G741 _{MC58} -961c-His		230-741_{MC58}-His
Δ G287 _{nz} -953-His	Δ G741 _{MC58} -961-His		230-Orf46.1-His
Δ G287 _{nz} -961-His	Δ G741 _{MC58} -983-His		230-961-His
Δ G287_{nz}-287-His	Δ G741 _{MC58} -Orf46.1-His		230-961c-His
Δ G287 _{nz} -287 _{nz} -His	Δ G741_{MC58}-741_{MC58}-His		936-741_{MC58}-His
Δ G287 _{nz} -741 _{MC58} -His	Δ G741_{MC58}-741_{ET37}-His		936-Orf46.1-His
			936-961-His
			936-741_{ET37}-His
Δ G287-919-Orf46.1-His			Δ G983-741 _{MC58} -His
Δ G287-Orf46.1-919-His	919-287		Δ G983-961c-His
919-287-Orf46-His	953-287		Δ G983-961-His
Orf46.1-287-919-His	919-Orf46.1-His		Δ G983-Orf46.1-His



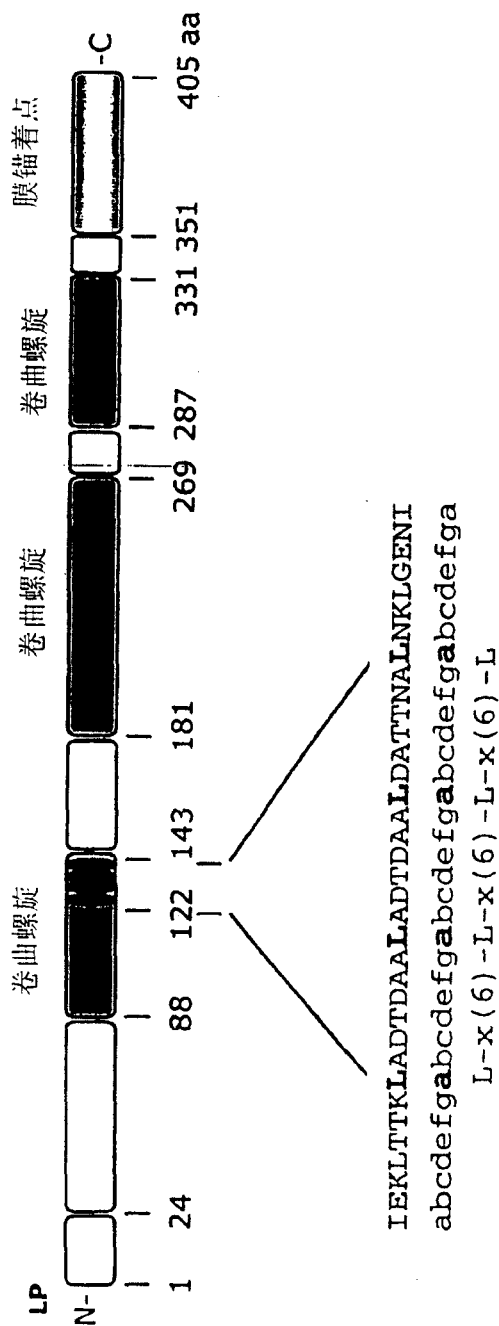


图 4

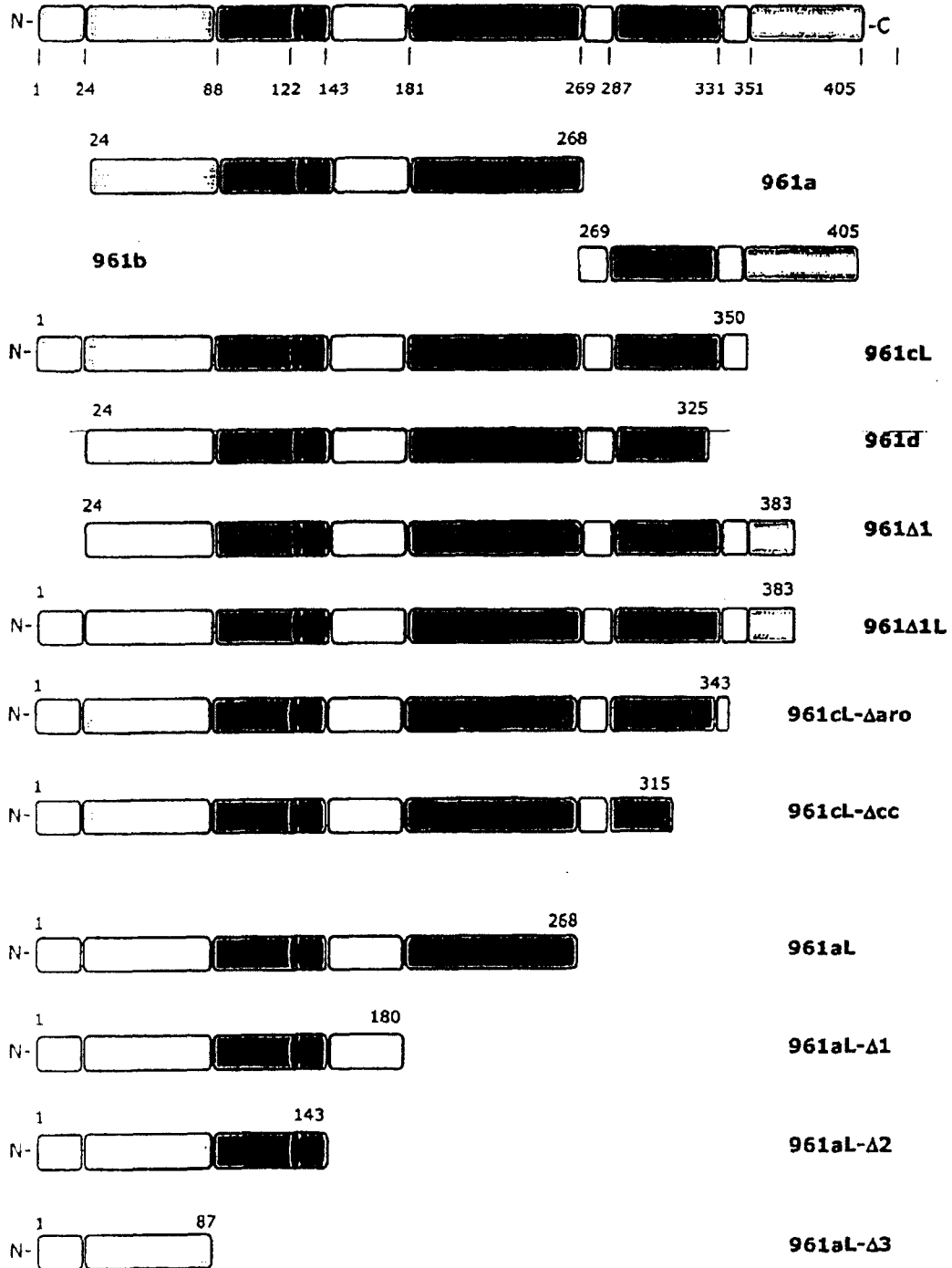


图 5