

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

A61K 39/09

A61K 39/40 C07K 16/12

C07K 14/315



# [12] 发明专利说明书

[21] ZL 专利号 99810854.5

[45] 授权公告日 2005 年 3 月 9 日

[11] 授权公告号 CN 1191851C

[22] 申请日 1999.4.7 [21] 申请号 99810854.5

[30] 优先权

[32] 1998. 4. 7 [33] US [31] 60/080,878

[32] 1998. 4. 7 [33] US [31] 09/056,019

[86] 国际申请 PCT/US1999/007669 1999.4.7

[87] 国际公布 WO1999/051188 英 1999.10.14

[85] 进入国家阶段日期 2000.12.7

[71] 专利权人 圣朱德儿童研究医院

地址 美国田纳西州

共同专利权人 医学免疫公司

[72] 发明人 E·I·托曼恩 H·R·马舒尔

T·M·维泽曼 L·S·约翰逊

S·科尼格

审查员 孙大龙

[74] 专利代理机构 中国国际贸易促进委员会专利  
商标事务所

代理人 李 瑛

权利要求书 7 页 说明书 131 页 附图 8 页

[54] 发明名称 包含胆碱结合蛋白 AN - 末端截取物的氨基酸的多肽、由该多肽衍生的疫苗及其应用

供药物组合物、疫苗和用其进行诊断和治疗的方法。

[57] 摘要

本发明提供一种包含胆碱结合蛋白 AN - 末端截取物的氨基酸序列的分离多肽，其中的氨基酸是 SEQ ID NOS: 1、3 - 7 或 9 - 11 所述的，也包括其片段、突变体、变体、类似物或衍生物。而且，本发明提供一种包含胆碱结合蛋白 AN - 末端截取物的氨基酸序列的分离多肽及其制备方法，其中的氨基酸是 SEQ ID NO 24 所述的，该多肽保留其天然的三级结构。本发明提供一种包含胆碱结合蛋白 AN - 末端截取物的氨基酸序列的分离多肽，其中多肽具有凝集素活性而不结合胆碱。本发明提供一种分离的免疫原性多肽，含有胆碱结合蛋白 AN - 末端截取物的氨基酸序列。本发明提供一种编码多肽的分离的核酸，所述多肽包含胆碱结合蛋白 AN - 末端截取物的氨基酸序列。最后，本发明提

ISSN 1008-4274

1. 分离的多肽，其含有 SEQ ID NO: 1、3、或 24 所示的氨基酸序列，其中所述多肽不与胆碱结合。
2. 权利要求 1 的分离的多肽，其中所述氨基酸序列包含至多 475 个氨基酸。
3. 权利要求 1 的分离的多肽，其中所述氨基酸序列包含至多 460 个氨基酸。
4. 分离的多肽，其含有 SEQ ID NO: 1、3、或 24 所示的氨基酸序列的变体，其中所述变体包含 SEQ ID NO: 1、3、或 24 的至少 1-57 个氨基酸替代，不与胆碱结合，并具有免疫原性。
5. 分离的多肽，其含有 SEQ ID NO: 1、3、或 24 所示的氨基酸序列的类似物或衍生物，其中所述多肽不与胆碱结合，并具有免疫原性。
6. 分离的多肽，其含有 SEQ ID NO: 5 所示的氨基酸序列，其中所述多肽不与胆碱结合。
7. 权利要求 6 的分离的多肽，其中所述氨基酸序列包含至多 475 个氨基酸。
8. 权利要求 6 的分离的多肽，其中所述氨基酸序列包含至多 460 个氨基酸。
9. 分离的多肽，其含有 SEQ ID NO: 5 所示的氨基酸序列的变体，其中所述变体包含 SEQ ID NO: 5 的至少 1-57 个氨基酸替代，包含至多 398 个氨基酸，不与胆碱结合，并具有免疫原性。
10. 分离的多肽，其含有 SEQ ID NO: 5 所示的氨基酸序列的类似物或衍生物，其中所述多肽不与胆碱结合，并具有免疫原性。
11. 分离的多肽，其含有 SEQ ID NO: 4 或 22 所示的氨基酸序列，其中所述多肽不与胆碱结合。
12. 权利要求 11 的分离的多肽，其中所述氨基酸序列包含至多 475 个氨基酸。
13. 权利要求 11 的分离的多肽，其中所述氨基酸序列包含至多 460

个氨基酸。

14. 分离的多肽，其含有 SEQ ID NO: 4 或 22 所示的氨基酸序列的变体，其中所述变体包含 SEQ ID NO: 4 或 22 的至少 1-57 个氨基酸替代，包含至多 398 个氨基酸，不与胆碱结合，并具有免疫原性。

15. 分离的多肽，其含有 SEQ ID NO: 4 或 22 所示的氨基酸序列的类似物或衍生物，其中所述多肽不与胆碱结合并具有免疫原性。

16. 权利要求 1、2、3、4、或 5 的分离的多肽，其中所述多肽与抗体相互作用，所述抗体能够与全长 CbpA 多肽相互作用。

17. 权利要求 6、7、8、9、或 10 的分离的多肽，其中所述多肽与抗体相互作用，所述抗体能够与全长 CbpA 多肽相互作用。

18. 权利要求 11、12、13、14、或 15 的分离的多肽，其中所述多肽与抗体相互作用，所述抗体能够与全长 CbpA 多肽相互作用。

19. 分离的多肽，其含有 SEQ ID NO: 24 的片段，其中所述片段含有 SEQ ID NO: 24 的至少 138 个连续氨基酸，其中所述多肽不与胆碱结合。

20. 分离的多肽，由 SEQ ID NO: 1、3、4、5、22、或 24 所示的氨基酸序列组成。

21. 分离的多肽，其含有 SEQ ID NO: 7 或 9 所示的氨基酸序列，其中所述多肽不与胆碱结合，并含有至多 376 个氨基酸。

22. 分离的多肽，其含有 SEQ ID NO: 7 或 9 所示的氨基酸序列的变体，其中所述变体包含 SEQ ID NO: 7 或 9 的至少 1 个至 57 个氨基酸替代，不与胆碱结合，具有免疫原性，并含有至多 376 个氨基酸。

23. 分离的多肽，其含有 SEQ ID NO: 7 或 9 所示氨基酸序列的类似物或衍生物，其中所述多肽不与胆碱结合，含有至多 376 个氨基酸，并具有免疫原性。

24. 分离的多肽，其含有 SEQ ID NO: 10、11、或 23 所示的氨基酸序列，其中所述多肽不与胆碱结合，并含有至多 328 个氨基酸。

25. 分离的多肽，其含有 SEQ ID NO: 10 或 23 所示的氨基酸序列的变体，其中所述变体包含 SEQ ID NO: 10 或 23 的至少 1 个至 57 个氨基酸替代，不与胆碱结合，具有免疫原性，并含有至多 147 个氨基酸。

26. 分离的多肽，其含有 SEQ ID NO: 10 或 23 所示的氨基酸序列的类似物或衍生物，其中所述多肽不与胆碱结合，含有至多 328 个氨基酸，并具有免疫原性。

27. 分离的多肽，其含有 SEQ ID NO: 11 所示的氨基酸序列的变体，其中所述变体包含 SEQ ID NO: 11 的至少 1 个至 57 个氨基酸替代，不与胆碱结合，具有免疫原性，并含有至多 328 个氨基酸。

28. 分离的多肽，其含有 SEQ ID NO: 11 所示的氨基酸序列的类似物或衍生物，其中所述多肽不与胆碱结合，含有至多 328 个氨基酸，并具有免疫原性。

29. 分离的多肽，由 SEQ ID NO: 7、9、10、11、或 23 所示的氨基酸序列组成。

30. 权利要求 21、22、23、24、25、26、27、或 28 的分离的多肽，其中所述多肽与抗体相互作用，所述抗体能够与全长 CbpA 多肽相互作用。

31. 权利要求 1、2、3、6、7、8、11、12、13、21、或 24 的分离的多肽，其中所述多肽具有免疫原性。

32. 权利要求 4、9、14、或 22 的分离的多肽，其中所述氨基酸替代包含保守的氨基酸替代。

33. 权利要求 1-15 或 19-29 任一项的分离的多肽，其中所述多肽具有凝集素活性。

34. 权利要求 1-5 任一项的分离的多肽，其中该多肽通过用羟胺裂解全长胆碱结合蛋白 A 制得，其中羟胺在第 475 位氨基酸处裂解胆碱结合蛋白 A，由此产生胆碱结合蛋白 A 的 N-末端截取物。

35. 权利要求 5、10、15、23、或 26 的分离的多肽，其中该类似物包含 N-末端甲硫氨酸或 N-末端聚组氨酸。

36. 药物组合物，其含有与包含胆碱结合蛋白 A N-末端截取物的氨基酸序列的多肽结合的分离的抗体，其中所述分离的抗体不与胆碱结合区域结合，并且所述分离的抗体基本上没有结合胆碱结合区域的抗体，其中所述氨基酸序列包含 SEQ ID NO: 7、9、10、11、或 23。

37. 药物组合物，其含有与包含胆碱结合蛋白 A N-末端截取物的氨

氨基酸序列的多肽结合的分离的抗体，其中所述分离的抗体不与胆碱结合区域结合，并且所述分离的抗体基本上没有结合胆碱结合区域的抗体，其中所述氨基酸序列包含 SEQ ID NO: 1、3、4、5、22、或 24。

38. 权利要求 36 或 37 的药物组合物，其中该抗体是单克隆抗体。

39. 权利要求 36 或 37 的药物组合物，其中该抗体是多克隆抗体。

40. 权利要求 36 或 37 的药物组合物，其中该抗体是嵌合(双特异性)抗体。

41. 药物组合物，其含有权利要求 1-15 或 19-29 任一项的多肽和可药用载体或稀释剂。

42. 药物组合物的用途，用于生产诱导暴露于或感染了肺炎球菌个体的免疫应答的药剂，其中所述药物组合物包含具有不与胆碱结合的胆碱蛋白 AN-末端截取物的氨基酸序列的多肽，其中所述氨基酸序列包含权利要求 1-29 任一项的多肽。

43. 权利要求 42 的用途，其中诱导保护性的免疫应答。

44. 药物组合物的用途，用于生产预防个体被肺炎球菌感染的药剂，其中所述药物组合物包含具有不与胆碱结合的胆碱结合蛋白 AN-末端截取物的氨基酸序列的多肽，其中所述氨基酸序列包含权利要求 1-29 任一项的多肽。

45. 权利要求 44 的用途，其中所述药物组合物被递送到呼吸道或鼻咽中。

46. 药物组合物的用途，用于生产预防被肺炎球菌感染的药剂，其中所述组合物包含与含有胆碱结合蛋白 AN-末端截取物的氨基酸序列的多肽结合的分离的抗体和可药用载体或稀释剂，其中所述分离的抗体不与胆碱结合区域结合，并且所述分离的抗体基本上没有结合胆碱结合区域的抗体，其中所述抗体能够结合权利要求 1-29 任一项的多肽。

47. 权利要求 46 的用途，其中所述药物组合物被递送到呼吸道或鼻咽中。

48. 疫苗，其包含多肽和可药用佐剂或载体，其中所述多肽包含不与胆碱结合的胆碱结合蛋白 AN-末端截取物的氨基酸序列，其中所述多肽

包含权利要求 1-29 任一项的氨基酸序列。

49. 疫苗，其包含多肽和可药用佐剂或载体，其中所述多肽包含不与胆碱结合的胆碱结合蛋白 A N-末端截取物的氨基酸序列，其中所述多肽包含如图 2 所示胆碱结合蛋白 A N-末端截取物的氨基酸序列。

50. 分离的核酸分子，其包含编码 SEQ ID NO: 1、3 或 24 所示多肽的核苷酸序列，其中所述多肽不与胆碱结合。

51. 权利要求 50 的分离的核酸分子，其中所述核苷酸序列包含 SEQ ID NO: 12 或 14。

52. 分离的核酸分子，其包含编码含有 SEQ ID NO: 1、3 或 24 的变体的多肽的核苷酸序列，其中所述变体包含 SEQ ID NO: 1、3 或 24 的至少 1-57 个氨基酸替代，不与胆碱结合并且具有免疫原性。

53. 分离的核酸分子，其包含编码 SEQ ID NO: 22 或 4 中所示的多肽的核苷酸序列，其中所述多肽不与胆碱结合。

54. 权利要求 53 的分离的核酸分子，其中所述核苷酸序列包含 SEQ ID NO: 15。

55. 分离的核酸分子，其包含编码含有 SEQ ID NO: 22 或 SEQ ID NO: 4 之变体的多肽的核苷酸序列，其中所述变体包含至少 1-57 个氨基酸替代，其中所述多肽不与胆碱结合并具有免疫原性。

56. 分离的核酸分子，其包含编码 SEQ ID NO: 5 所示多肽的核苷酸序列，其中所述多肽不与胆碱结合。

57. 分离的核酸分子，其包含编码含有 SEQ ID NO: 5 之变体的多肽的核苷酸序列，其中所述变体包含 SEQ ID NO: 5 的至少 1-57 个氨基酸替代，不与胆碱结合，有免疫原性，并且包含至多 398 个氨基酸。

58. 权利要求 56 的分离的核酸分子，其中所述核苷酸序列包含 SEQ ID NO: 16。

59. 分离的核酸分子，其包含编码具有 SEQ ID NO: 24 的至少 138 个连续氨基酸的多肽的核苷酸序列，其中所述多肽不与胆碱结合。

60. 分离的核酸分子，其包含具有 SEQ ID NO: 12 的至少 318 个连续核苷酸的核苷酸序列，其中所述核苷酸序列编码不与胆碱结合的多肽。

61. 分离的核酸分子, 由 SEQ ID NO: 12、14、15 或 16 组成。
62. 分离的核酸分子, 其包含编码 SEQ ID NO: 7 或 9 所示多肽的核苷酸序列, 其中所述多肽包含至多 376 个氨基酸, 并且不与胆碱结合。
63. 分离的核酸分子, 其包含编码含有 SEQ ID NO: 7 或 9 之变体的多肽的核苷酸序列, 其中所述变体包含 SEQ ID NO: 7 或 9 的至少 1-57 个氨基酸替代, 包含至多 376 个氨基酸, 不与胆碱结合, 并且有免疫原性。
64. 分离的核酸分子, 其包含编码 SEQ ID NO: 10、11、或 23 所示多肽的核苷酸序列, 其中所述多肽包含至多 328 个氨基酸并且不与胆碱结合。
65. 分离的核酸分子, 其包含编码含有 SEQ ID NO: 10 或 23 之变体的多肽的核苷酸序列, 其中所述变体包含 SEQ ID NO: 10 或 23 的至少 1-57 个氨基酸替代, 包含至多 147 个氨基酸, 不与胆碱结合, 并且有免疫原性。
66. 分离的核酸分子, 其包含编码含有 SEQ ID NO: 11 之变体的多肽的核苷酸序列, 其中所述变体包含 SEQ ID NO: 11 的至少 1-57 个氨基酸替代, 包含至多 328 个氨基酸, 不与胆碱结合, 并且有免疫原性。
67. 分离的核酸分子, 其包含编码 SEQ ID NO: 17 或 19 所示的核苷酸序列, 其中所述核苷酸序列编码不与胆碱结合的多肽。
68. 分离的核酸分子, 由 SEQ ID NO: 17、19、20、或 21 组成。
69. 分离的核酸分子, 其包含 SEQ ID NO: 20 或 21 所示核苷酸序列, 其中所述核苷酸序列编码具有至多 328 个氨基酸的 N-末端胆碱结合多肽, 并且所述多肽不与胆碱结合。
70. 权利要求 50-69 任一项的分离的核酸分子, 其中所述核酸分子进一步包含启动子。
71. 权利要求 50-69 任一项的分离的核酸分子, 其中所述核酸分子是 DNA。
72. 权利要求 71 的分离的核酸分子, 其中所述核酸分子是 cDNA。
73. 权利要求 71 的分离的核酸分子, 其中所述核酸分子是基因组 DNA。
74. 权利要求 50-69 任一项的分离的核酸分子, 其中所述核酸分子是

RNA。

75. 权利要求 50-69 任一项的分离的核酸分子，其中所述核酸分子与启动子有效地连接。

76. 权利要求 75 的分离的核酸分子，其中所述启动子是 RNA 转录的启动子。

77. 载体，其含有权利要求 50-69 任一项的核酸分子。

78. 权利要求 77 的载体，其中所述载体进一步包含启动子。

79. 权利要求 77 的载体，其中所述载体是质粒、粘粒、酵母人工染色体 (YAC)、噬菌体或真核生物病毒 DNA。

80. 用于产生多肽的宿主载体系统，其包含适当非人类宿主细胞中的权利要求 77 的载体。

81. 权利要求 77 的宿主载体系统，其中所述适当非人类宿主细胞包含原核或真核细胞。

82. 含有权利要求 77 的载体的非人类细胞。

83. 获得纯化形式多肽的方法，包括：

(a) 将权利要求 77 的载体导入适当宿主细胞；

(b) 培养获得的宿主细胞以产生该多肽；

(c) 回收步骤(b)中产生的多肽；和

(d) 纯化步骤(c)中回收的多肽。

84. 疫苗，其包含编码多肽的分离的核酸分子和可药用佐剂或载体，其中所述多肽包含不与胆碱结合的胆碱结合蛋白 A N-末端截取物的氨基酸序列，其中所述核酸分子包含权利要求 50-69 任一项的核苷酸序列。

85. 疫苗，其包含权利要求 76 的载体和可药用佐剂或载体。

86. 权利要求 48 或 84 任一项的疫苗的用途，用于生产治疗感染了或暴露于肺炎球菌的个体的药剂。



包含胆碱结合蛋白 A N-末端截取物的氨基酸的多肽、  
由该多肽衍生的疫苗及其应用

发明领域

本发明普遍涉及胆碱结合蛋白 A N-末端截取物的多肽。本发明也涉及对细菌感染，特别是肺炎球菌感染提供保护作用或激发产生保护性抗体的疫苗和用于诊断和被动免疫治疗的对抗这种多肽的抗体和拮抗剂。该多肽和/或编码该多肽的核酸也可以用作肺炎球菌细菌附着的竞争性抑制剂。最后，本发明涉及利用该多肽进行治疗的治疗方法。

发明背景

肺炎链球菌是革兰氏阳性菌，它是引起侵袭性感染如败血症、脑膜炎、中耳炎、和大叶性肺炎的主要原因 (Tuomanen 等人, NEJM 332: 1280-1284, 1995)。肺炎球菌强烈地与上下呼吸道细胞结合。象大多数细菌一样，通过与真核生物的碳水化合物结合的细菌表面蛋白质的呈递以类似凝集素的方式来完成肺炎球菌与人类细胞的粘着(Cundell, D.&Tuomanen, E. (1994) *Microb Pathog* 17:361-374)。肺炎球菌与非炎性上皮结合可看作是无症状运输过程。已经有人提出，转变为侵袭性疾病与局部产生炎性因子有关，所述因子通过激活人类细胞改变可在人类细胞上获得的受体的数目和类型(Cundell,D.等人 (1995) *Nature* 377:435-438)。在这种新的环境中肺炎球菌便有机会取得优势并且利用了这些未调整受体之一，血小板激活因子(PAF)受体(Cundell,D.等人 (1995) *Nature* 377:435-438)。在 PAF 受体出现的几分钟内，肺炎球菌经历粘着性和侵袭性增强的波动。例如通过可溶性受体类似物来抑制细菌与激活细胞的结合可在动物模型中阻断疾病的进展(Idanpaan-Heikkila, I.等人(1997) *J. Infect. Dis.*, 176:704-712)。在这点上特别有效的是包含有或没有附加唾液酸的乳糖-N-新四糖的可溶性碳水化合物，它在体外预防肺炎球菌附着在人类细胞上并且在体内预防移生于

肺中。

#### 胆碱结合蛋白：候选结构附着基因

肺炎球菌产生一类能够与细菌表面结合的表面蛋白，这种结合是通过与细胞壁磷壁酸或脂磷壁质非共价结合的。肺炎链球菌表面覆盖着一类 CBPs(胆碱结合蛋白)，这类蛋白与磷脂酰胆碱非共价结合。CbpA 是一种 75kD 表面暴露的胆碱结合蛋白，显示嵌合结构。具有独特的 N-末端区域，即富含脯氨酸的区域，其后是由 10 个负责结合胆碱的重复区组成的 C-端区域。

CbpA 是一种复合糖的粘附素(配体)，复合糖上包含存在于真核细胞表面上的受体。具有 CbpA 缺陷的突变在鼻咽移生的未成年大鼠模型中显示降低的毒力。这种结合涉及修饰磷壁酸的胆碱决定簇并且通过该类蛋白质各个成员中的特征胆碱结合区域来介导。Lopez 等人在他们的自溶酶研究中发现了这种胆碱结合区域并且对其特性进行了充分地确定(Ronda 等人(1987)*Eur. J. Biochem*, 164:621-624)。含有该区域的其它蛋白包括肺炎球菌噬菌体的自溶素和保护性抗原，即肺炎球菌表面蛋白 A(PspA) (Ronda,C.等人(1987)*Eur. J. Biochem*, 164:621-624 和 McDaniel, L. S., 等人(1992)*Microb Pathog*, 13:261-269)。CbpA 不能移生于鼻咽区域，该区域是与 CbpA 的共同具有的其它类型成员(C 端)并且其与人细胞结合的活性是由其特有的 N-末端区域产生的。由于移生的过程和疾病的发展取决于作为第一步骤的肺炎球菌附着于人细胞上，通过交叉反应性抗体或者通过与模拟该区的肽的竞争性抑制作用阻断 N-末端区域的功能对于阻断疾病来说可能是关键性的。在 PCT 国际申请 PCT/US97/07198 中讨论了抗肺炎球菌疫苗的胆碱结合蛋白并且将该 PCT 申请的全部内容引入本文供参考。目前抗肺炎链球菌的疫苗使用该细菌 23 个最普通血清型的荚膜的纯化碳水化合物，但该疫苗仅仅有 50% 的保护作用(Shapiro 等人, *NJEM* 325:1453, 1991)并且在 2 岁以下不具有免疫原性。而且，治疗性多肽在被多抗性生物体感染的病例中会提供治疗的选择性。所以，本发明通过提供保护性疫苗满足了这一长期的需要。

### 发明概述

本发明提供一种分离的多肽，该多肽包含胆碱结合蛋白 A N-末端截取物的氨基酸序列。所述多肽包含 SEQ ID NOS 1, 3-7 或 9-11 中所述的氨基酸序列，包括其片段、突变体、变体、类似物或衍生物。而且，本发明提供一种分离的多肽，该多肽包含具有 SEQ ID NO 24 氨基酸的胆碱结合蛋白 A N-末端截取物的氨基酸序列，其中该多肽显示其三级结构并且也提供制备该多肽的方法。分离的多肽适用于使动物和人

对细菌感染，优选对肺炎球菌感染产生免疫作用。

本发明也涉及具有凝集素活性但没有胆碱结合活性的胆碱结合蛋白 A 的 N-末端截取物。而且，本发明提供一种免疫原性胆碱结合蛋白 A 的 N-末端截取物或其片段。

本发明也涉及分离的核酸，如重组 DNA 分子或克隆的基因或其简并变体、突变体、类似物或其片段，其编码分离的多肽或者其竞争性抑制多肽的活性。优选地，分离的核酸(包括简并物、变体、突变体、类似物或其片段)具有 SEQ ID NOS: 12, 14-17, 19-22 或 23 所述的序列。在本发明进一步的实施方案中，可以将如此确定的重组 DNA 分子或克隆基因的完整 DNA 序列有效地连接到表达控制序列上，该表达控制序列可以被引入到合适的宿主中。因此，本发明还涉及用克隆的基因或重组 DNA 分子转化的单细胞宿主，所述克隆基因或重组 DNA 分子包含编码本发明多肽的 DNA 序列，特别是由上文所述序列确定的 DNA 序列或其片段。

对抗分离多肽的抗体包括天然产生的和重组方法制备的抗体。这些抗体可包括通过已知的遗传技术制备的多克隆和单克隆抗体，以及双特异性的(嵌合的)抗体，和包括适于诊断使用及能够调节细菌粘附包括但不限于起竞争剂作用的其它功能的抗体。

本发明的再一目的是提供一种治疗哺乳动物来控制细菌或其亚单位的量或活性以便处理或预防侵入、自生或特发病态所产生的不利后果的方法。本发明提供在治疗方法中使用的药物组合物，其中包含或基于分离的多肽、其亚单位或其结合配对物。

最后，本发明提供药物组合物、疫苗及使用它们的诊断和治疗方法。

### 附图简述

图 1.胆碱结合蛋白 A(CbpA)和重组截取物 R1(如图 2 所示从 CbpA 的 N-末端开始约氨基酸 16 到氨基酸 321)和 R2(如图 2 所示从 CbpA 的 N-末端开始约氨基酸 16 到氨基酸 444) 的图解。区域 A 是如图 2 所示从 CbpA 氨基酸序列的 N-末端开始约氨基酸 153 到氨基酸 321; 区域 B 是如图 2 所示从 CbpA 氨基酸序列的 N-末端开始约氨基酸 270 到氨基酸 326; 并且区域 C 是如图 2 所示从 CbpA 氨基酸序列的 N-末端开始约氨基酸 327 到氨基酸 433。

图 2A-B.比较 CbpA N-末端区域的核酸和氨基酸序列各种血清型的同源性。

图 3.重组 R1 和 R2 的表达和纯化。

图 4.在小鼠中被动保护的结果。对抗重组 R2 的免疫血清保护小鼠免受致死性肺炎链球菌的激发。

图 5.抗-R2 抗体对粘附到 LNT-HSA 包被平板上的 R6x 的滴定。

图 6.具有阻断肺炎球菌与 LNT-HSA 包被平板粘附活性的抗-CbpA 和吸附的抗- CbpA 抗体的滴定。

图 7.在小鼠中的主动保护结果。对抗重组 R1 的免疫血清保护小鼠免受致死性肺炎链球菌的激发(激发物: 560cfu 血清型 6B)。

### 发明详述

本发明涉及一种分离的多肽,该多肽包含胆碱结合蛋白 A N-末端截取物的氨基酸序列。该多肽适用于使动物对肺炎球菌感染产生免疫作用。这些多肽或其肽片段,当用合适的佐剂配制时,用于对抗肺炎球菌和对抗具有交叉反应性蛋白的其它细菌的疫苗中。

本发明提供一种分离的多肽,该多肽包含胆碱结合蛋白 A N-末端截取物的氨基酸序列。在一个实施方案中,该多肽具有 SEQ ID NO 1, 3-5, 7 或 9-11 所述的任一氨基酸序列,包括其片段、突变体、变体、类似物或衍生物。在另一实施方案中,该多肽具有氨基酸 KXXE(SEQ ID NO 6)。

本发明提供一种分离的多肽,该多肽包含图 2 所述的胆碱结合蛋白

A N-末端截取物的氨基酸序列。在一个实施方案中，该多肽具有图 2 所述保守区的氨基酸序列。例如，保守区包括但不限于氨基酸序列 158-210; 158-172; 300-321; 331-339; 355-365; 367-374; 379-389; 409-427 和 430-447。图 2 描述本发明 CbpA N-末端区域的核酸和氨基酸序列各种血清型的同源性。

另外，本发明提供一种分离的多肽，该多肽包含具有 SEQ ID NO 24 所述氨基酸的胆碱结合蛋白 A N-末端截取物的氨基酸序列，其中所述多肽显示其三级结构。在一个实施方案中，所述多肽是其类似物、片段、突变体或变体。所述变体如图 2 所示。本发明也提供一种分离的多肽，该多肽包含具有图 2 所述血清型 4 的约 16 位到约 474 位氨基酸或如图 2 所述血清型 4 相应氨基酸的胆碱结合蛋白 A N-末端截取物的氨基酸序列，其中所述多肽显示其三级结构。在一个实施方案中，三级结构与存在于天然蛋白质中的三级结构相对应。

制备所述多肽的方法的例子如下：用羧胺裂解全长的胆碱结合蛋白 A，其中所述羧胺裂解胆碱结合蛋白 A 血清型 R6x 和血清型 4 的 475 位处的氨基酸天冬酰胺(N)或者图 2 所示不同血清型中血清型 R6x 或血清型 4 的相应氨基酸，由此产生胆碱结合蛋白 A 的 N-末端截取物。产生截断的胆碱结合蛋白 A 或其片段并且保持天然三级结构(即，全长胆碱结合蛋白 A 的三级结构)的其它方法是有所描述的并且对本领域技术人员来说是已知的。由于所述多肽保持其三级结构，分离的多肽适用于作为免疫原使动物和人对细菌感染，优选肺炎球菌产生免疫作用。

含有胆碱结合蛋白 A(CbpA)血清型 4 型氨基酸序列的多肽如下：

```

ENEGATQVPTSSNRANESQAEQGEQPKKLDSEKDKARKEVEEYVKKIVGESY
AKSTKKRHTITVALVNELNNIKNEYLNKIVESTSESQIQILMMESRSKVDEAV
SKFEKDSSTSSSSSDSSTKPEASDTAKPNKPTPEGKVAEAKKKVEEAEKKAKD
QKEEDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKANEPKDEQKIKQAE
AEVESKQAEATRLKKIKTDREEAEAEAKRRADAKEQKPKGRAKRGVPGEL
ATPKKENDAKSSDSSVGEETLPSPLKPEKKVAEAEKKVVEEAKKKAEDQKE
EDRRNYPTNTYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKVKQAKAE

```

VESKKAETRLEKIKTDRKKAEEEEAKRKA AEEDKVKEKP AEQPQPAPAPKAE  
 KPAPAPKPEN (SEQ ID NO 24).

“多肽 R2”意指含有胆碱结合蛋白 A(CbpA)血清型 4 型(见图 1)的 N-末端截取物的氨基酸序列的多肽，其具有下列序列：

ENEGATQVPTSSNRANESQAEQGEQPKKLDSE RDKARKEVEEYVKKIVGESY  
 AKSTKKRHTITVALVNELNNIKNEYLNKIVESTSESQLQILMMESRSKVDEAV  
 SKFEKDSSSSSSSDSSTKPEASDTAKPNKPTEPG EKVAEAKKKVEEAEKKAKD  
 QKEEDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKANEP RDEQKIKQAE  
 AEVESKQAEATRLKKIKTDREEAE EEAARRADAKEQGKPKGRAKRGVPGEL  
 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAEAEKKVVEEAKKKAEDQKE  
 EDRRNYPTNTYKTLELEIAESDVEVKKAELELVKEEAK EPRNEEKVKQAKAE  
 VESKKAETRLEKIKTDRKKAEEEEAKRKA AEEDKVKEKPA (SEQ ID NO 1).

编码胆碱结合蛋白 A(CbpA)血清型 4 型的 N-末端截取物的多肽 R2 的 DNA 序列：

GAGAACGAGGGAGCTACCCAAGTACCCACTTCTTCTAATAGGGCAAATGA  
 AAGTCAGGCAGAACAAGGAGAACAACCTAAAAAACTCGATTCAGAACGA  
 GATAAGGCAAGGAAAGAGGTCGAGGAATATGTAAAAAAAATAGTGGGTG  
 AGAGCTATGCAAAATCAACTAAAAGCGACATACAATTACTGTAGCTCTA  
 GTTAACGAGTTGAACAACATTAAGAACGAGTATTTGAATAAAATAGTTGA  
 ATCAACCTCAGAAAGCCAACACTACAGATACTGATGATGGAGAGTCGATCAA  
 AAGTAGATGAAGCTGTGTCTAAGTTTGAAAAGGACTCATCTTCTTCGTCAA  
 GTTCAGACTCTTCCACTAAACCGGAAGCTTCAGATACAGCGAAGCCAAAC  
 AAGCCGACAGAACCAGGAGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTG  
 AAGAAGCTGAGAAAAAAGCCAAGGATCAAAAAGAAGAAGATCGTCGTAA  
 CTACCCAACCACTTACTTACAAAACGCTTGA ACTTGAAATTGCTGAGTCCG  
 ATGTGGAAGTTAAAAAAGCGGAGCTTGA ACTAGTAAAAGTGAAAGCTAA  
 CGAACCTCGAGACGAGCAAAAAATTAAGCAAGCAGAAGCGGAAGTTGAG

AGTAAACAAGCTGAGGCTACAAGGTTAAAAAAATCAAGACAGATCGTG  
 AAGAAGCAGAAGAAGAAGCTAAACGAAGAGCAGATGCTAAAGAGCAAG  
 GTAAACCAAAGGGCGGGCAAACGAGGAGTTCCTGGAGAGCTAGCAAC  
 ACCTGATAAAAAAGAAAATGATGCGAAGTCTTCAGATTCTAGCGTAGGTG  
 AAGAACTCTTCCAAGCCATCCCTGAAACCAGAAAAAAGGTAGCAGA  
 AGCTGAGAAGAAGGTTGAAGAAGCTAAGAAAAAGCCGAGGATCAAAAA  
 GAAGAAGATCGCCGTAACCTACCAACCAATACTTACAAAACGCTTGAAC  
 TGAAATTGCTGAGTCCGATGTGGAAGTTAAAAAGCGGAGCTTGAAC  
 TAAAAGAGGAAGCTAAGGAACCTCGAAACGAGGAAAAAGTTAAGCAAGC  
 AAAAGCGGAAGTTGAGAGTAAAAAGCTGAGGCTACAAGGTTAGAAAA  
 ATCAAGACAGATCGTAAAAAGCAGAAGAAGAAGCTAAACGAAAAGCAG  
 CAGAAGAAGATAAAGTTAAAGAAAAACCAGCTG (SEQ ID NO 12).

#### **CbpA 血清型 4 的氨基酸序列:**

ENEGATQVPTSSNRANESQAEQGEQPKKLDSEKDKARKEVEEYVKKIVGESY  
 AKSTKKRHTITVALVNELNNIKNEYLNKIVESTSESQIQILMMESRSKVDEAV  
 SKFEKDSSTSSSSSSSSTKPEASDTAKPNKPTPEGKVAEAKKKVEEAEEKAKD  
 QKEEDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKANEPKDEQKIKQAE  
 AEVESKQAEATRLKKIKTDREEAEEEAKRRADAKEQGKPKGRAKRGVPGEL  
 ATPDKKENDAKSSDSSVGEETLPSPSLKPEKKVAEAEKKVVEAKKKAEDQKE  
 EDRRNYPTNTYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKVKQAKAE  
 VESKKAETRLEKIKTDRKKAEEEAKRKAEEEDKVKEKPAEQQPAPAPKAE  
 KPAPAPKPENPAEQPKAEKPADQQAEEYARRSEEEYNRLTQQQPPKTEKPA  
 QPSTPKTGWKQENGMWYFYNTDGSMATGWLQNNGSWYYLNSNGAMATG  
 WLQNNGSWYYLNANGSMATGWLQNNGSWYYLNANGSMATGWLQYNGS  
 WYYLNANGSMATGWLQYNGSWYYLNANGDMATGWVKDGDWYYLEAS  
 GAMKASQWFKVSDKWYVNGSGALAVNITVDGYGVNANGWVN. (SEQ ID  
 NO 2)

#### **编码 CbpA 血清型 4 的氨基酸序列的 DNA 序列:**

GAGAACGAGGGAGCTACCCAAGTACCCACTTCTTCTAATAGGGCAAATGA  
AAGTCAGGCAGAACAAGGAGAACAACCTAAAAAACTCGATTCAGAACGA  
GATAAGGCAAGGAAAGAGGGTCGAGGAATATGTAAAAAAAATAGTGGGTG  
AGAGCTATGCAAAATCAACTAAAAAGCGACATACAATTACTGTAGCTCTA  
GTTAACGAGTTGAACAACATTAAGAACGAGTATTTGAATAAAAATAGTTGA  
ATCAACCTCAGAAAGCCAACACTACAGATACTGATGATGGAGAGTCGATCAA  
AAGTAGATGAAGCTGTGTCTAAGTTTGAAAAGGACTCATCTTCTTCGTCAA  
GTTTCAGACTCTTCCACTAAACCGGAAGCTTCAGATACAGCGAAGCCAAAC  
AAGCCGACAGAACCAGGAGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTG  
AAGAAGCTGAGAAAAAAGCCAAGGATCAAAAAGAAGAAGATCGTCGTAA  
CTACCCAACCATTACTTACAAAACGCTTGAACCTGAAATTGCTGAGTCCG  
ATGTGGAAGTTAAAAAGCGGAGCTTGAACCTAGTAAAAGTGAAAGCTAA  
CGAACCTCGAGACGAGCAAAAAATTAAGCAAGCAGAAGCGGAAGTTGAG  
AGTAAACAAGCTGAGGCTACAAGGTTAAAAAAAATCAAGACAGATCGTG  
AAGAAGCAGAAGAAGAAGCTAAACGAAGAGCAGATGCTAAAGAGCAAG  
GTAAACCAAAGGGGCGGGCAAACGAGGAGTTCCTGGAGAGCTAGCAAC  
ACCTGATAAAAAAGAAAATGATGCGAAGTCTTCAGATTCTAGCGTAGGTG  
AAGAACTCTTCCAAGCCCATCCCTGAAACCAGAAAAAAGGTAGCAGA  
AGCTGAGAAGAAGGTTGAAGAAGCTAAGAAAAAAGCCGAGGATCAAAA  
GAAGAAGATCGCCGTAACCTACCCAACCAATACTTACAAAACGCTTGAAC  
TGAAATTGCTGAGTCCGATGTGGAAGTTAAAAAAGCGGAGCTTGAAC  
GTAAAAGAGGAAGCTAAGGAACCTCGAAACGAGGAAAAAGTTAAGCAAG  
CAAAAGCGGAAGTTGAGAGTAAAAAAGCTGAGGCTACAAGGTTAGAAAA  
AATCAAGACAGATCGTAAAAAAGCAGAAGAAGAAGCTAAACGAAAAGCA  
GCAGAAGAAGATAAAGTTAAAGAAAAACCAGCTGAACAACCACAACCAG  
CGCCGGCTCCAAAAGCAGAAAAACCAGCTCCAGCTCCAAAACCAGAGAA  
TCCAGCTGAACAACCAAAGCAGAAAAACCAGCTGATCAACAAGCTGAA  
GAAGACTATGCTCGTAGATCAGAAGAAGAATATAATCGCTTGAACCAACA  
GCAACCGCCAAAAACTGAAAAACCAGCACAACCATCTACTCCAAAAACA  
GGCTGGAAACAAGAAAACGGTATGTGGTACTTCTACAATACTGATGGTTC



AATGGCGACAGGATGGCTCCAAAACAAAGGCTCAAGGTAcTACcTCAACAG  
CAATGGCGCTATGGCGACAGGATGGCTCCAAAACAATGGTTCATGGTACT  
ATCTAAACGCTAATGGTTCATGGCAACAGGATGGCTCCAAAACAATGGT  
TCATGGTACTACCTAAACGCTAATGGTTCATGGCGACAGGATGGCTCCA  
ATACAATGGCTCATGGTACTACCTAAACGCTAATGGTTCATGGCGACAG  
GATGGCTCCAATACAATGGCTCATGGTACTACCTAAACGCTAATGGTGAT  
ATGGCGACAGGTTGGGTGAAAGATGGAGATACCTGGTACTATCTTGAAGC  
ATCAGGTGCTATGAAAGCAAGCCAATGGTTCAAAGTATCAGATAAATGGT  
ACTATGTCAATGGCTCAGGTGCCCTTGCAAGTCAACACAACCTGTAGATGGC  
TATGGAGTCAATGCCAATGGTGAATGGGTAAACTAA (SEQ ID NO 13).

“多肽 R1”意指含有胆碱结合蛋白 A(CbpA)血清型 4 型的 N-末端截取物的 16 位到 321 位氨基酸序列的多肽，其具有下列序列：

ENEGATQVPTSSNRANESQAEQGEQPKKLDSEKDKARKEVEEYVKKIVGESY  
AKSTKKRHTITVALVNELNLIKNEYLNKIVESTSESQIQILMMESRSKVDEAV  
SKFEKDS SSSSSSSSSTKPEASDTAKPNKPTEPGKVAEAKKKVEEAEEKKAKD  
QKEEDRRNYPTITYKTLELEIAESDVEVKKALELVKVKANEPREDEQKIKQAE  
AEVESKQAEATRLKKJKTDREEAE EEAARRADAKEQ GKPKGRAKRGVPGEL  
ATPDKKENDAKSSDSSVGEETL (SEQ ID NO 3).

编码多肽 R1 的 DNA 序列为：

GAGAACGAGGGAGCTACCCAAGTACCCACTTCTTCTAATAGGGCAAATGA  
AAGTCAGGCAGAACAAGGAGAACAACCTAAAAAACTCGATTCAGAACGA  
GATAAGGCAAGGAAAGAGGTTCGAGGAATATGTAAAAAAAATAGTGGGTG  
AGAGCTATGCAAAAATCAACTAAAAAGCGACATACAATTACTGTAGCTCTA  
GTTAACGAGTTGAACAACATTAAGAACGAGTATTTGAATAAAAATAGTTGA  
ATCAACCTCAGAAAGCCAACACTACAGATACTGATGATGGAGAGTCGATCAA  
AAGTAGATGAAGCTGTGTCTAAGTTTGAAAAGGACTCATCTTCTTCGTCAA  
GTTTCAGACTCTTCCACTAAACCGGAAGCTTCAGATACAGCGAAGCCAAAC  
AAGCCGACAGAACCAGGAGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTG  
AAGAAGCTGAGAAAAAAGCCAAGGATCAAAAAGAAGAAGATCGTCGTAA  
CTACCCAACCATTACTTACAAAACGCTTGAAGTTGAAATTGCTGAGTCCG  
ATGTGGAAGTTAAAAAAGCGGAGCTTGAAGTAGTAAAAGTGAAAGCTAA  
CGAACCTCGAGACGAGCAAAAAATTAAGCAAGCAGAAGCGGAAGTTGAG  
AGTAAACAAGCTGAGGCTACAAGGTTAAAAAAAATCAAGACAGATCGTG  
AAGAAGCAGAAGAAGAAGCTAAACGAAGAGCAGATGCTAAAGAGCAAG  
GTAAACCAAGGGGCGGGCAAACGAGGAGTTCCTGGAGAGCTAGCAAC  
ACCTGATAAAAAAGAAAATGATGCGAAGTCTTCAGATTCTAGCGTAGGTG  
AAGAAACTCTTC (SEQ ID NO 14).

“多肽 C/R2”意指在 R2 内包含重复区 C 的多肽，其中重复区 C 具有胆碱结合蛋白 A(CbpA)血清型 4 型的 N-末端截取物的 327 位到 433 位氨基酸序列，其具有下列序列：

KPEKKVAEAEKKVVEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAESDVEVK  
KAELELVKEEAKEPRNEEKVKQAKAEVESKKAETRLEKIKTDRKKAEEEEAK  
RKA (SEQ ID NO 4).

### 多肽 C/R2 的 DNA 序列

AAACCAGAAAAAAGGTAGCAGAAGCTGAGAAGAAGGTTGAAGAAGCTA  
AGAAAAAAGCCGAGGATCAAAAAGAAGAAGATCGCCGTA ACTACCCAAC  
CAATACTTACAAAACGCTTGA ACTTGAAATTGCTGAGTCCGATGTGGAAG  
TTAAAAAAGCGGAGCTTGA ACTAGTAAAAGAGGAAGCTAAGGAACCTCG  
AAACGAGGAAAAAGTTAAGCAAGCAAAAAGCGGAAGTTGAGAGTAAAAAA  
GCTGAGGCTACAAGGTTAGAAAAAATCAAGACAGATCGTAAAAAAGCAG  
AAGAAGAAGCTAAACGAAAAGCA (SEQ ID NO 15)

“多肽 A/R2”意指在 R2 内包含重复区 A 的多肽，其中重复区 A 具有胆碱结合蛋白 A(CbpA)血清型 4 型的 N-末端截取物的 153 位到 269 位氨基酸序列，其具有下列序列：

TEPGEKVAEAKKKVEEAEKKAKDQKEEDRRNYPTITYKTLELEIAESDVEVK  
KAELELVKVKANEP RDEQKIKQAEAEVESKQAEATRLKIKTDREEAE EEEAK

RRADA (SEQ ID NO 5). 如图 1 所示，多肽 R2 的 A 区与 R1 中的 A 区是相同的。

编码多肽 A/R2 的 DNA 序列为：

ACAGAACCAGGAGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTGAAGAA  
GCTGAGAAAAAAGCCAAGGATCAAAAAGAAGAAGATCGTCGTA ACTACC

CAACCATTACTTACAAAACGCTTGAACCTGAAATTGCTGAGTCCGATGTG  
 GAAGTTAAAAAAGCGGAGCTTGAACCTAGTAAAAGTGAAAGCTAACGAAC  
 CTCGAGACGAGCAAAAAATTAAGCAAGCAGAAGCGGAAGTTGAGAGTAA  
 ACAAGCTGAGGCTACAAGGTTAAAAAAAATCAAGACAGATCGTGAAGAA  
 GCAGAAGAAGAAGCTAAACGAAGAGCAGATGCT (SEQ ID NO 16).

可以改变或修饰一个或多个氨基酸残基的同一性或位置以包括变体, 例如缺失物, 包含比蛋白质全部特定残基少的残基, 取代物, 其中一个或多个特定残基被其它残基取代和加成物, 其中一个或多个氨基酸残基被加入到所述多肽的末端或中间部分(见图 2)。这些分子包括: 掺入对于由所选非哺乳动物宿主表达来说“优选的”密码子; 提供限制性核酸内切酶的切割位点; 并且提供促进易于表达载体构建的其它起始、末端或中间的 DNA 序列。具体来说, 下列是血清型 4 中氨基酸取代的例子, 其包括但不限于: 154 位的 E 被 K 取代; 155 位的 P 被 L 取代; 156 位的 G 被 E 取代; 157 位的 E 被 K 取代; 181 位的 K 被 E 取代; 182 位的 D 被 A 取代; 187 位的 R 被 Y、H 或 L 取代; 194 位的 I 被 N 取代; 200 位的 E 被 D 取代; 202 位的 E 被 D 取代; 209 位的 E 被 K 取代; 212 位的 K 被 E 取代; 218 位的 V 被 L 取代; 220 位的 V 被 K 或 E 取代; 221 位的 K 被 E 取代; 223 位的 N 被 D 或 K 取代; 225 位的 P 被 S、T 或 R 取代; 227 位的 D 被 N 取代; 228 位的 E 被 K 取代; 229 位的 Q 被 E、G 或 D 取代; 230 位的 K 被 T 取代; 232 位的 K 被 N 取代; 235 位的 E 被 K 取代; 236 位的 A 被 E 取代; 237 位的 E 被 K 取代; 240 位的 S 被 N 取代; 241 位的 K 被 E 取代; 242 位的 Q 被 K 取代; 249 位的 K 被 E 取代; 250 位的 K 被 N 取代; 257 位的 E 被 Q 或 K 取代; 263 位的 A 被 L 取代; 264 位的 K 被 E 取代; 265 位的 R 被 N 取代; 266 位的 R 被 I 取代; 267 位的 A 被 K 或 V 取代; 258 位的 D 被 T 取代; 269 位的 A 被 D 取代; 291 位的 A 被 T、V、P、G 或 X 取代; 294 位的 G 被 G、A 或 E 取代; 295 位的 V 被 D 或 A 取代; 295 位的 P 被 L 或 F 取代; 299 位的 L 被 P 或 Q 取代; 328 位的 P 被 S 取代; 329 位的 E 被 G 取代; 340 位的 E 被 A 取代; 343 位的 K 被 E 或 D 取代; 347 位的 E 被 K 取代; 349 位的 D 被 A 取

代；354位的R被H取代；366位的E被D取代；375位的E被K取代；378位的K被E取代；390位的E被G取代；391位的P被S取代；393位的N被D取代；397位的V被I取代；和408位的K被Q取代。

“多肽R2血清型-R6x”是指包含胆碱结合蛋白A(CbpA)血清型R6x的N-末端截取物的16位到444位氨基酸序列的多肽，其具有下列序列：

ENEGSTQAATSSNMAKTEHRKAAKQVVDEYIEKMLREIQLDRRKHTQNVAL  
 NIKLSAIKTKYLRELVLEEKSKDELPSKAKLDAAFEKFKKDTLKPGEKVA  
 EAKKKVEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAEFDVKVKEAELELVK  
 EEAKESRNEGTIKQAKEKVESKKAETRLENIKTDRKKAEEEAKRKADAKLK  
 EANVATSDQGKPKGRAKRGVPGELATPDKKENDAKSSDSSVGEETLPSSSLK  
 SGKKVAEAEKKVVEAEKKAKDQKEEDRRNYPTNTYKTLDLEIAESDVKVKE  
 AELELVKEEAKEPRDEEKIKQAKAKVESKKAETRLENIKTDRKKAEEEAKR  
 KAAEEDKVKEKPA (SEQ ID NO 7)

#### 编码多肽R2血清型R6x的DNA序列：

GAAAACGAAGGAAGTACCCAAGCAGCCACTTCTTCTAATATGGCAAAGAC  
 AGAACATAGGAAAGCTGCTAAACAAGTCGTCGATGAATATATAGAAAAA  
 ATGTTGAGGGAGATTCAACTAGATAGAAGAAAACATACCCAAAATGTGCGC  
 CTAAACATAAAGTTGAGCGCAATTA AACGAAGTATTTGCGTGAATTAA  
 ATGTTTTAGAAAGAGAAGTCGAAAGATGAGTTGCCGTCAGAAATAAAAGCA  
 AAGTTAGACGCAGCTTTTGAGAAGTTTAAAAAAGATACATTGAAACCAGG  
 AGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTGAAGAAGCTAAGAAAAAA  
 GCCGAGGATCAAAAAGAAGAAGATCGTCGTA ACTACCCAACCAATACTTA  
 CAAAACGCTTGAACTTGAAATTGCTGAGTTCGATGTGAAAGTTAAAGAAG  
 CGGAGCTTGA ACTAGTAAAAGAGGAAGCTAAAGAACTCGAAACGAGGGC  
 ACAATTAAGCAAGCAAAAGAGAAAGTTGAGAGTAAAAAAGCTGAGGCTA  
 CAAGGTTAGAAAACA CAAGACAGACGTAAAAAAGCAGAAGAAGAAGCT  
 AAACGAAAAGCAGATGCTAAGTTGAAGGAAGCTAATGTAGCGACTTCAG

A:CAAGGTAAACCAAAGGGGCGGGCAAACGAGGAGTTCCTGGAGAGCTA  
 GCAACACCTGATAAAAAAGAAAATGATGCGAAGTCTTCAGATTCTAGCGT  
 AGGTGAAGAACTCTTCCAAGCTCATCCCTGAAATCAGGAAAAAAGGTAG  
 CAGAAGCTGAGAAGAAGGTTGAAGAAGCTGAGAAAAAAGCCAAGGATCA  
 AAAAGAAGAAGATCGCCGTAACCTACCCAACCAATACTTACAAAACGCTTG  
 ACCTTGAAATTGCTGAGTCCGATGTGAAAGTTAAAGAAGCGGAGCTTGAA  
 CTAGTAAAAGAGGAAGCTAAGGAACCTCGAGACGAGGAAAAAATTAAGC  
 AAGCAAAAGCGAAAGTTGAGAGTAAAAAAGCTGAGGCTACAAGGTTAGA  
 AAACATCAAGACAGATCGTAAAAAAGCAGAAGAAGAAGCTAAACGAAAA  
 GCAGCAGAAGAAGATAAAGTTAAAGAAAAACCAGCTG (SEQ ID NO 17)

**CbpA 血清型 R6x 的氨基酸序列:**

ENEGSTQAATSSNMAKTEHRKAAKQVVDEYIEKMLREIQLDRRKHTQNVAL  
 NIKLSAIKTKYLRELVLEEKSKDELPSBIKAKLDAAFEKFKKDTLKPGEKVA  
 EAKKKVEEAKKKAEDQKEEDRRNYPTNTYKTLLELEIAEFDVKVKEAELELVK  
 EEAKESRNEGTIKQAKEKVESKKAETRLENIKTDRKKAEEEAKRKADAKLK  
 EANVATSDQGPKGRAKRGVPGELATPDKKENDAKSSDSSVGEETLPSSSLK  
 SGKKVAEAEKKVEEAEKKAKDQKEEDRRNYPTNTYKTLDELEIAESDVKVKE  
 AELELVKEEAKEPRDEEKIKQAKAKVESKKAETRLENIKTDRKKAEEEAKR  
 KAAEEDKVKEKPAEQPQAPATQPEKPAKPEKPAEQPKAEKTDQQAEDY  
 ARRSEEEYNRLTQQQPPKTEKPAQPSTPKTGWKQENGMWYFYNTDGSMT  
 GWLQNNGSWYYLNANGAMATGWLQNNGSWYYLNANGSMATGWLQNNG  
 SWYYLNANGAMATGWLQYNGSWYYLNSNGAMATGWLQYNGSWYYLNA  
 NGDMATGWLQNNGSWYYLNANGDMATGWLQYNGSWYYLNANGDMATG  
 WVKDGDWYYLEASGAMKASQWFKVSDKWYYVNGSGALAVNTTVDGYG  
 VNANGWVN (SEQ ID NO 8).

**编码 CbpA 血清型 R6x 氨基酸序列的 DNA 序列:**

GAAAACGAAGGAAGTACCCAAGCAGCCACTTCTTCTAATATGGCAAAGAC  
 AGAACATAGGAAAGCTGCTAAACAAGTCGTCGATGAATATATAGAAAAA  
 ATGTTGAGGGAGATTCAACTAGATAGAAGAAAACATACCCAAAATGTCGC

CTTAAACATAAAGTTGAGCGCAATTA AACGAAGTATTTGCGTGAATTAA  
ATGTTTTAGAAAGAGAAGTCGAAAGATGAGTTGCCGTCAGAAATAAAAGCA  
AAGTTAGACGCAGCTTTTGAGAAGTTTAAAAAAGATACATTGAAACCAGG  
AGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTGAAGAAGCTAAGAAAAAA  
GCCGAGGATCAAAAAGAAGAAGATCGTCGTA ACTACCCAACCAATACTTA  
CAAAACGCTTGA ACTTGAAATTGCTGAGTTCGATGTGAAAGTTAAAGAAG  
CGGAGCTTGA ACTAGTAAAAGAGGAAGCTAAAGAACTCGAAACGAGGGC  
ACAATTAAGCAAGCAAAAAGAGAAAGTTGAGAGTAAAAAAGCTGAGGCTA  
CAAGGTTAGAAAACA CAAGACAGACGTAAAAAAGCAGAAGAAGAAGCT  
AAACGAAAAGCAGATGCTAAGTTGAAGGAAGCTAATGTAGCGACTCAGA  
CAAGGTAAACCAAAGGGGCGGGCAAAACGAGGAGTTCCTGGAGAGCTAG  
CAACACCTGATAAAAAAGAAAATGATGCGAAGTCTTCAGATTCTAGCGTA  
GGTGAAGAACTCTTCCAAGCTCATCCCTGAAATCAGGAAAAAAGGTAGC  
AGAAGCTGAGAAGAAGGTTGAAGAAGCTGAGAAAAAAGCCAAGGATCAA  
AAAGAAGAAGATCGCCGTA ACTACCCAACCAATACTTACAAAACGCTTGA  
CCTTGAAATTGCTGAGTCCGATGTGAAAGTTAAAGAAGCGGAGCTTGAAC  
TAGTAAAAGAGGAAGCTAAGGAACCTCGAGACGAGGAAAAAATTAAGCA  
AGCAAAAGCGAAAGTTGAGAGTAAAAAAGCTGAGGCTACAAGGTTAGAA  
AACATCAAGACAGATCGTAAAAAAGCAGAAGAAGAAGCTAAACGAAAAG  
CAGCAGAAGAAGATAAAGTTAAAGAAAAACCAGCTGAACAACCACAACC  
AGCGCCGGCTACTCAACCAGAAAAACCAGCTCCAAAACCAGAGAAGCCA  
GCTGAACAACCAAAAGCAGAAAAACAGATGATCAACAAGCTGAAGAAG  
ACTATGCTCGTAGATCAGAAGAAGAATATAATCGCTTGACTCAACAGCAA  
CCGCCAAAACCTGAAAAACCAGCACAACCATCTACTCCAAAACAGGCT  
GGAAACAAGAAAACGGTATGTGGTACTTCTACAATACTGATGGTTCAATG  
GCAACAGGATGGCTCCAAAACAACGGTTCATGGTACTATCTAAACGCTAA  
TGGTGCTATGGCGACAGGATGGCTCCAAAACAATGGTTCATGGTACTATC  
TAAACGCTAATGGTTCAATGGCAACAGGATGGCTCCAAAACAATGGTTCA  
TGGTACTACCTAAACGCTAATGGTGCTATGGCGACAGGATGGCTCCAATA  
CAATGGTTCATGGTACTACCTAAACAGCAATGGCGCTATGGCGACAGGAT  
GGCTCCAATACAATGGCTCATGGTACTACCTCAACGCTAATGGTGATATG

GCGACAGGATGGCTCCAAAACAACGGTTCATGGTACTACCTCAACGCTAA  
 TGGTGATATGGCGACAGGATGGCTCCAATACAACGGTTCATGGTATTACC  
 TCAACGCTAATGGTGATATGGCGACAGGTTGGGTGAAAGATGGAGATACC  
 TGGTACTATCTTGAAGCATCAGGTGCTATGAAAGCAAGCCAATGGTTCAA  
 AGTATCAGATAAATGGTACTATGTCAATGGCTCAGGTGCCCTTGCAGTCA  
 ACACA ACTGTAGATGGCTATGGAGTCAATGCCAATGGTGAATGGGTAAAC  
 TAA (SEQ ID NO 18).

“多肽 R1 血清型 R6x”是指包含胆碱结合蛋白 A(CbpA)血清型 R6x 的 N-末端截取物的 16 位到 321 位氨基酸序列的多肽，其具有下列序列：

ENEGSTQAATSSNMAKTEHRKAAKQVVDEYIEKMLREIQLDRRKHTQNVAL  
 NIKLSAIKTKYLRELVLEEKSKDELPSSEIKAKLDAAFEKFKKDTLKPGEKVA  
 EAKKKVEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAEFDVKVKEAELELVK  
 EEAKESRNEGTIKQAKEKVESKKAETRLENIKTDRKKAEEEEAKRKADAKLK  
 EANVATSDQGPKGRAKRGVPGELATPDKKENDAKSSDSSVGEETL (SEQ ID  
 NO 9).

编码多肽 R1 的 DNA 序列为：

GAAAACGAAGGAAGTACCCAAGCAGCCACTTCTTCTAATATGGCAAAGAC  
 AGAACATAGGAAAGCTGCTAAACAAGTCGTCGATGAATATATAGAAAAA  
 ATGTTGAGGGAGATTCAACTAGATAGAAGAAAACATACCCAAAATGTTCGC  
 CTTAAACATAAAGTTGAGCGCAATTAACGAAGTATTTGCGTGAATTAA  
 ATGTTTTAGAAAGAGAAGTCGAAAGATGAGTTGCCGTCAGAAATAAAAGCA  
 AAGTTAGACGCAGCTTTTGAGAAGTTTAAAAAAGATACATTGAAACCAGG  
 AGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTGAAGAAGCTAAGAAAAAA  
 GCCGAGGATCAAAAAGAAGAAGATCGTCGTA ACTACCCAACCAACTTA  
 CAAAACGCTTGA ACTTGA AATTGCTGAGTTCGATGTGAAAGTTAAGAAG  
 CGGAGCTTGA ACTAGTAAAAGAGGAAGCTAAAGAATCTCGAAACGAGGG  
 CACAATTAAGCAAGCAAAAAGAGAAAGTTGAGAGTAAAAAAGCTGAGGCT  
 ACAAGGTTAGAAAACA CAAGACAGATCGTAAAAAAGCAGAAGAAGAAG



CTAAACGAAAAGCAGATGCTAAGTTGAAGGAAGCTAATGTAGCGACTTCA  
GATCAAGGTAAACCAAAGGGGCGGGCAAAACGAGGAGTTCCTGGAGAGC  
TAGCAACACCTGATAAAAAAGAAAATGATGCGAAGTCTTCAGATTCTAGC  
GTAGGTGAAGAACTCTTC (SEQ ID NO 19).

“多肽 C/R2 血清型 R6x”是指在 R2 内包含重复区 C 的多肽(参见图 2), 其中所述重复区 C 具有胆碱结合蛋白 A(CbpA)血清型 R6x 的 N-末端截取物的 327 位到 433 位的氨基酸序列, 其具有下列序列:

KSGKKVAEAEKKVEEAEKKAKDQKEEDRRNYPTNTYKTLDEIAESDVKVK  
EALELVKEEAKEPRDEEKIKQAKAKVESKKAETRLENIKTDRKKAEEEEAK  
RKA (SEQ ID NO 10)

#### 多肽 C/R2 血清型 R6x 的 DNA 序列:

AAATCAGGAAAAAAGGTAGCAGAAGCTGAGAAGAAGGTTGAAGAAGCTG  
AGAAAAAAGCCAAGGATCAAAAAGAAGAAGATCGCCGTAACCTACCCAAC  
CAATACTTACAAAACGCTTGACCTTGAAATTGCTGAGTCCGATGTGAAAG  
TTAAAGAAGCGGAGCTTGAAGTAGTAAAAGAGGAAGCTAAGGAACCTCG  
AGACGAGGAAAAAATTAAGCAAGCAAAAAGCGAAAAGTTGAGAGTAAAAAA  
GCTGAGGCTACAAGGTTAGAAAACATCAAGACAGATCGTAAAAAAGCAG  
AAGAAGAAGCTAAACGAAAAGCA (SEQ ID NO 20).

“多肽 A/R2 血清型 R6x”是指在 R2 内包含重复区 A 的多肽(参见图 2), 其中所述重复区 A 具有胆碱结合蛋白 A(CbpA)血清型 R6x 的 N-末端截取物的 155 位到 265 位的氨基酸序列, 其具有下列序列:

PGEKVAEAKKKVEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAEFDVKVKE  
AELELVKEEAKESRNEGTIKQAKEKVESKKAETRLENIKTDRKKAEEEEAKR  
KADA (SEQ ID NO 11)

#### 编码多肽 A/R2 血清型 R6x 的 DNA 序列:

CCAGGAGAAAAGGTAGCAGAAGCTAAGAAGAAGGTTGAAGAAGCTAAGA  
 AAAAAGCCGAGGATCAAAAAGAAGAAGATCGTCGTAACCTACCCAACCAA  
 TACTTACAAAACGCTTGAACCTGAAATTGCTGAGTTCGATGTGAAAGTTA  
 AAGAAGCGGAGCTTGAACCTAGTAAAAGAGGAAGCTAAAGAAICTCGAAAC  
 GAGGGCACAATTAAGCAAGCAAAAAGAGAAAGTTGAGAGTAAAAAAGCTG  
 AGGCTACAAGGTTAGAAAACAiCAAGACAGATCGTAAAAAAGCAGAAGA  
 AGAAGCTAAACGAAAAGCAGATGCT (SEQ ID NO 21).

本发明涉及一种分离的多肽，其中该分离的多肽由如 SEQ ID NOS 22 或 23 所述的氨基酸序列所组成，包括其片段、突变体、变体、或类似物，或衍生物。

SPSLKPEKKVAEAEKKEVEAKKKAEDQKEEDRRNYPTNTYKTLELEIAESDV  
 EVKKAELVKEEAKEPRNEEKVKQAKAEVESKKAETRLEKIKTDRKKAEE  
 EAKRKAEEEDKVKEKPA (SEQ ID NO 22; 血清型 4; 323-434 位) 或  
 PSSSLKSGKKVAEAEKKEVEAKKAKDQKEEDRRNYPTNTYKTLDLEIAESD  
 VKVKEAELELVKEEAKEPRDEEKIKQAKAKVESKKAETRLENIKTDRKKAEE  
 EEAKRKAEEEDKVKEKRA (SEQ ID NO 23; 血清型 R6x; 322-434 位).

“多肽 B/R2”是指包含图 2 所述的胆碱结合蛋白 A(CbpA)血清型 4 型的 N-末端截取物的 270 位到 326 位氨基酸序列的多肽。“多肽 B/R2 血清型-R6x”是指包含图 2 所述的胆碱结合蛋白 A(CbpA)血清型 R6x 的 N-末端截取物的 264 位到 326 位氨基酸序列的多肽。本发明研究具有图 1 所述的 A、B、C、A+B、B+C、A+C 区氨基酸序列的多肽。

此外，本发明提供一种分离的多肽，该多肽包含胆碱结合蛋白 A 的 N-末端截取物的氨基酸序列，其中所述多肽具有氨基酸 KXXE(SEQ ID NO 6)。

本发明涉及包含胆碱结合蛋白 A N-末端截取物的氨基酸序列的多肽，其中所述氨基酸序列阐述于图 2 中。在一个实施方案中，所述多肽具有一个氨基酸序列，该序列是图 2 所述的保守区。例如，保守区包括但不限于氨基酸序列 158-172; 300-321; 331-339; 355-365; 367-374; 379-389; 409-427; 和 430-447。图 2 阐述了本发明所研究的

各种血清型 CbpA 的 N 末端区的核酸和氨基酸序列的同源性。

本发明提供一种分离的多肽，该多肽包含从胆碱结合蛋白 A 的 N-末端截取的氨基酸序列，其中所述多肽具有凝集素活性并且不与胆碱结合。在一个实施方案中，所述多肽具有 SEQ ID NO 1、3-5、7 或 9-11 中所述的任一氨基酸序列，包括其片段、突变体、变体、类似物，或衍生物。

本文所使用的“具有凝集素活性的多肽”是指一种以非共价键与碳水化合物结合的多肽、肽或蛋白质。本文所定义的“粘附素”是指细菌与人类细胞的非共价键结合或足够稳定以经得起洗涤的分泌。本文所定义的“结合 LNnT”是指结合乳糖-N-新四糖包裹的底物多于结合白蛋白-对照。

本发明提供一种分离的免疫原性多肽，该多肽包含从胆碱结合蛋白 A 的 N 末端截取的氨基酸序列。本发明预测所述免疫原性多肽具有 SEQ ID NOS 1、3-7 或 9-11 中所述的任一氨基酸序列，包括其片段、突变体、变体、类似物，或衍生物。本发明提供了一种分离的多肽，该多肽包含图 2 所述从胆碱结合蛋白 A 的 N 末端截取的氨基酸序列。在一个实施方案中，所述多肽具有图 2 所述保守区的氨基酸序列。

本发明涉及包含上述氨基酸序列的多肽的类似物。所述类似物多肽可能具有不连接或连接到包含所述氨基酸序列的多肽的 N 或 COOH 末端的 N-末端蛋氨酸或 N-末端聚组氨酸。

在另一实施方案中，本发明研究了由所述多肽的蛋白水解消化产物产生的多肽的肽片段。在另一实施方案中，所述多肽衍生物上连有一个或多个化学部分。在另一实施方案中，所述化学部分是水溶性聚合物。在另一实施方案中，所述化学部分是聚乙二醇。在另一实施方案中，所述化学部分是单、二、三或 tetrapegylated。在另一实施方案中，所述化学部分是 N-末端 monopegylated。

聚乙二醇(PEG)连接到化合物上是特别有用的，因为 PEG 对哺乳动物具有很低的毒性(Carpenter 等人,1971)。例如，美国批准利用腺苷脱氨酶的 PEG 加合物来治疗人类严重的合并免疫缺陷综合症。与 PEG 结合获得的第二点好处是有效地降低异种化合物的免疫原性和抗原性。例如，人类蛋白质的 PEG 加合物可用于治疗其它哺乳动物物种的

疾病并且没有诱发严重免疫反应的危险。本发明化合物可以以微囊封装装置递送以便于减少或预防对抗所述化合物或对抗可产生所述化合物的细胞的宿主免疫反应。本发明化合物也可以以包裹在膜中的微囊化形式如脂质体递送。

已经描述了许多适用于与蛋白质直接反应的 PEG 活性形式。用于与蛋白质氨基反应的 PEG 试剂包括羧酸的活性酯或碳酸酯衍生物，特别是其中离去基团为 N-羧基琥珀酰亚胺、对-硝基苯酚、咪唑或 1-羧基-2-硝基苯-4-磺酸酯的化合物。包含马来酰亚氨基或卤代乙酰基的 PEG 衍生物是适用于蛋白质游离巯基改性的试剂。同样地，包含氨基胍或酰基胍的 PEG 试剂可用于与醛反应，所述醛是由蛋白质中碳水化合物基团的高碘酸盐氧化产生的。

在一个实施方案中，本文所描述的多肽的氨基酸残基优选以“L”同分异构形式存在。在另一实施方案中，以“D”同分异构形式存在的残基可替代任何 L-氨基酸残基，只要所述多肽能够保留所需要的凝集素活性的功能特性即可。NH<sub>2</sub> 是指在多肽的氨基端存在的游离氨基。COOH 是指在多肽的羧基端存在的游离羧基。本文所使用的缩写与标准的多肽命名一致，*J.Biol.Chem.*,243:3552-59(1969)。

应该注意，本文所有氨基酸残基序列都是由式子表示的，其中所述式子的左和右取向都是在氨基端到羧基端的常规方向上。此外，应该注意，在氨基酸残基序列开始或结尾的破折号表示与另外的一个或多个氨基酸残基序列连接的肽键。

利用公知的固相、液相或多肽缩合技术或其组合制备的合成多肽可包括天然和非天然氨基酸。用于肽合成的氨基酸可以是标准的 Boc(N<sup>α</sup>-氨基保护的 N<sup>α</sup>-叔丁氧基羰基)氨基酸树脂，通过 Merrifield 的最初固相方法中标准的脱保护、中和、偶联和洗涤方法(1963, *J.Am.Chem.Soc.*85:2149-2154)合成，或者是由 Carpino 和 Han(1972, *J.Org.Chem.* 37:3403-3406)首先描述的碱不稳定的 N<sup>α</sup>-氨基保护的 9-芴基甲氧基羰基(Fmoc)氨基酸。因此，本发明多肽可包含 D-氨基酸、D-氨基酸和 L-氨基酸的组合和各种用于传达特殊性质的“设计者”氨基酸(例如β-甲基氨基酸、C<sub>α</sub>-甲基氨基酸和 N<sub>α</sub>-甲基氨基酸等)。合成氨基酸包括用于合成赖氨酸的鸟氨酸，用于合成苯丙氨酸的氟苯丙氨酸

和用于合成亮氨酸或异亮氨酸的正亮氨酸。另外，通过在特定偶联步骤指定特定氨基酸，可产生 $\alpha$ -螺旋、 $\beta$ -转角、 $\beta$ 折叠、 $\gamma$ -转角和环状肽。

一方面，本发明所述肽在C-末端可包含特殊的氨基酸，它结合 $\text{CO}_2\text{H}$ 或 $\text{CONH}_2$ 侧链以模拟游离的甘氨酸或甘氨酸-酰胺基团。考虑该特殊残基的另一方法是具有由接头组成的或与珠粒键合的侧链的D或L氨基酸类似物。在一个实施方案中，假游离的C-末端残基可为D或L光学构型；在另一实施方案中，可以使用D和L异构体的外消旋混合物。

在另外的实施方案中，焦谷氨酸可作为肽的N-末端残基包含在内。尽管焦谷氨酸经不起通过Edman降解所得序列的检验，但是通过将取代仅限制到所给出的具有N-末端焦谷氨酸的珠粒上的50%的肽，珠粒上仍将保留足够的非焦谷氨酸肽用于测序。普通技术人员会容易地意识到该技术可用于任何在N-末端掺入耐受Edman降解的残基的肽的测序。下文详细描述了表征表现出所需活性的各个肽的其它方法。在50%的肽中存在特定的N-末端基团时，含有被保护的N-末端基团如焦谷氨酸的肽的比活性通过比较完全(100%)被保护的肽与未被保护的(0%)肽的活性将会容易地得到证明。

另外，预计本发明制备的肽具有更好定义的结构特性，并且使用肽模拟物和肽模拟键如酯键来制备具有新特性的肽。在另一实施方案中，可以产生有还原肽键(即， $\text{R}_1\text{-CH}_2\text{-NH-R}_2$ ，其中 $\text{R}_1$ 和 $\text{R}_2$ 是氨基酸残基或序列)的肽。可以作为二肽亚单位引入还原肽键。这样的分子将会耐受肽键水解，如，蛋白酶活性。这类肽将会提供具有独特功能和活性的配体，如，由于对代谢分解或蛋白酶活性具有耐受性而延长了体内的半衰期。而且，在某些系统中受约束的肽显示增强的功能活性是公知的(Hruby, 1982, *Life Sciences* 31: 189-199; Hruby等人, 1990, *Biochem J.* 268:249-262); 本发明提供一种生产受约束肽的方法，该肽在所有其它位置掺入随机序列。

可以合成受约束的、环状的或刚性化的肽，条件是在肽序列的至少两个位置插入氨基酸或氨基酸类似物以提供能够交联以在处理约束、环化或刚性化该肽形成交联物的化学官能团。当掺入回转诱导氨基酸时将有利于环化。能够使肽交联的氨基酸的例子是形成二硫化物的半胱氨酸、形成内酯或乳糖酶的天冬氨酸和螯合剂如 $\gamma$ -羧基-谷氨酸

(Gla)(Bachem)以螯合过渡金属并形成交联物。通过改良 Zee-Cheng 和 Olson(1980, *Biophys, Biochem. Res. Commun.* 94:1128-1132)所述的合成方法可以制备保护的 $\gamma$ -羧基-谷氨酸。可通过如氧化半胱氨酸残基以形成二硫化物或加入金属离子以形成螯合物来处理肽序列中含有至少两个能够交联的氨基酸的肽,使肽交联并形成受约束的、环状的或刚性的肽。

本发明提供系统制备交联物的策略。例如,如果将4个半胱氨酸残基掺入肽序列中,可使用不同的保护基(Hiskey, 1981, in *The Peptides: Analysis, Synthesis, Biology*, Vol. 3, Gross and Meienhofer, eds., Academic Press: New York, 137-167 页; Ponsanti 等人, 1990, *Tetrahedron* 46:8255-8266)。第一对半胱氨酸可被脱保护并氧化,然后将第二组脱保护并氧化。在该方法中,可形成特定组的二硫化的交联物。另外,可掺入一对半胱氨酸和一对核对的氨基酸类似物以便使交联物具有不同的化学特性。

可以将下列非经典氨基酸结合到肽中以便引入特定构象的基序: 1,2,3,4-四氢异喹啉-3-羧酸酯(Kazmierski 等人, 1991, *J. Am. Chem. Soc.* 113:2275-2283); (2S,3S)-甲基-苯丙氨酸、(2S, 3R)-甲基-苯丙氨酸、(2R,3S)-甲基-苯丙氨酸和(2R, 3R)-甲基-苯丙氨酸(Kazmierski 和 Hruby, 1991, *Tetrahedron Lett.*); 2-氨基四氢萘-2-羧酸(Landis, 1989, Ph.D. Thesis, University of Arizona); 羟基-1,2,3,4-四氢异喹啉-3-羧酸酯(Miyake 等人, 1989, *J. Takeda Res. Labs.* 43:53-76);  $\beta$ -吡啶(D 和 L)(Kazmierski, 1988, Ph.D. Thesis, University of Arizona); HIC(组胺酸异喹啉羧酸)(Zechel 等人, 1991, *Int. J. Pep. Protein Res.* 43); 和 HIC(组胺酸环脲)(Dharanipragada)。

下列氨基酸类似物和肽模拟物可被引入肽中以诱导或有利于形成特异的二级结构: LL-Acp(LL-3-氨基-2-丙烯酮-6-羧酸), 一种诱导 $\beta$ -转角的二肽类似物(Kemp 等人, 1985, *J. Org. Chem.* 50:5834-5838); 诱导 $\beta$ -折叠的类似物(Kemp 等人, 1988, *Tetrahedron Lett.* 29:5081-5082); 诱导 $\beta$ -转角的类似物(Kemp 等人, 1988, *Tetrahedron Lett.* 29:5057-5060); 诱导 $\alpha$ -螺旋的类似物(Kemp 等人, 1988, *Tetrahedron Lett.* 29:4935-4938); 诱导 $\gamma$ -转角的类似物(Kemp 等人, 1989, *J. Org.*

Chem. 54:109-115)和由下列参考文献提供的类似物: Nagai 和 Sato, 1985, *Tetrahedron Lett.* 26:647-650; DiMaio 等人, 1989, *J. Chem. Soc. Perkin Trans.* 1687 页; 以及 Gly-Ala 转角类似物(Kahn 等人, 1989, *Tetrahedron Lett.* 30:2317); 酰胺键同配物(Jones 等人, 1988, *Tetrahedron Lett.* 29:3853-3856); 四唑(Zabrocki 等人, 1988, *J. Am. Chem. Soc.* 110:5875-5880); DTC(Samanen 等人, 1990, *Int. J. Protein Pep. Res.* 35:501-509); 和 Olson 等人, 1990, *J. Am. Chem. Sci.* 112:323-333 和 Garvey 等人, 1990, *J. Org. Chem.* 56:436 中教导的类似物。在 Kahn 的美国专利 5440013(出版于 1995 年 8 月 8 日)中描述了 $\beta$ -转角的和 $\beta$ -凸起的构象限制模拟物。

本发明还提供本发明多肽或肽的修饰或衍生方法。肽的修饰对普通技术人员来说是公知的, 并且包括磷酸化、羧甲基化和酰化。修饰可通过化学或酶方式来进行。在另一方面, 可以制备糖基化或脂肪酰化的肽衍生物。在本领域中, 糖基化或脂肪酰化的肽的制备方法是公知的。也可以制备脂肪酰基肽衍生物。例如但不限于, 游离氨基(N-末端的或赖氨酰基上的)可被酰化, 如十四烷酰化。在另一实施方案中, 可以将包含结构式为 $-(CH_2)_nCH_3$ 的脂族侧链的氨基酸引入肽中。适于本发明使用的这种和其它肽-脂肪酸缀合物公开于英国专利 GB 8809162.4、国际专利申请 PCT/AU89/00166 和上文中的参考文献 5 中。

在编码这种多肽的核酸中可以进行突变使特定的密码子改变为编码不同氨基酸的密码子。这种突变通常通过改变尽可能最少的核苷酸来进行。这种类型的取代突变可被进行而使所得蛋白质中的氨基酸以非保守的方式(即, 由属于具有特定大小或特性的氨基酸类别的氨基酸的密码子改变为属于另一类别的氨基酸所特有的密码子)或以保守方式(即, 由属于具有特定大小或特性的氨基酸类别的氨基酸的密码子改变为属于相同类别的氨基酸所特有的密码子)改变。这种保守改变通常使所得蛋白质的结构和功能上发生较少的变化。非保守改变更有可能改变所得蛋白质的结构、活性或功能。应当认为本发明包括含有不明显改变所得蛋白质活性或结合特性的保守改变的序列。序列中氨基酸的取代可选自该类氨基酸的其它成员。例如, 非极性(疏水性)氨基酸包括丙氨酸、亮氨酸、异亮氨酸、缬氨酸、脯氨酸、苯丙氨酸、色氨酸

和蛋氨酸。含有芳香环结构的氨基酸是苯丙氨酸、色氨酸和酪氨酸。极性中性氨基酸包括甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺。带正电荷的(碱性)氨基酸包括精氨酸、赖氨酸和组氨酸。带负电荷的(酸性)氨基酸包括天冬氨酸和谷氨酸。不希望这种变化影响通过聚丙烯酰胺凝胶电泳确定的表现分子量或等电点。

特别优选的取代为：

-Lys 取代 Arg 和 Arg 取代 Lys, 这样的取代可保持正电荷；

-Glu 取代 Asp 和 Asp 取代 Glu, 这样的取代可保持负电荷；

-Ser 取代 Thr, 这样的取代可保持游离的-OH； 和

-Gln 取代 Asn, 这样的取代可保持游离的 NH<sub>2</sub>。

合成的 DNA 序列允许方便构建表达类似物或“突变蛋白”的基因。Noren 等人, *Science*, 244:182-188(1989年4月)描述了将非天然氨基酸定点引入蛋白质中的一般方法。该方法可用于产生基于非天然氨基酸的类似物。

在本发明中, 可以使用本领域常规分子生物学、微生物学和重组 DNA 技术。在文献中充分说明了这些技术。例如, 参见 Sambrook 等人, “*Molecular Cloning: A Laboratory Manual*” (1989); “*Current Protocols in Molecular Biology*” I-III 卷[Ausubel, R. M., ed.(1994)]; “*Cell Biology: A Laboratory Handbook*” I-III 卷[J. E. Celis, ed. (1994)]; “*Current Protocols in Immunology*” I-III 卷[Coligan, J. E., ed.(1994)]; “*Oligonucleotide Synthesis*” (M.J. Gait ed. 1984); “*Nucleic Acid Hybridization*” [B.D. Hames & S.J. Higgins eds.(1985)]; “*Transcription And Translation*” [B.D. Hames & S.J. Higgins eds.(1984)]; “*Animal Cell Culture*” [R.I. Freshney, ed.(1986)]; “*Immobilized Cells And Enzymes*” [IRL Press, (1986)]; B. Perbal, “*A Practical Guide To Molecular Cloning*” (1984)。

在另外的实施方案中, 焦谷氨酸可作为肽的 N-末端残基包含在内。尽管焦谷氨酸经不起通过 Edman 降解所得序列的检验, 但是通过将取代仅限制到所给出的具有 N-末端焦谷氨酸的珠粒上的 50% 的肽, 珠粒上仍将保留足够的非焦谷氨酸的肽。普通技术人员会容易地意识到该技术可用于任何在 N-末端掺入耐受 Edman 降解的残基的肽的测序。



下文详细描述了表征表现出所需活性的各个肽的其它方法。在 50% 肽中存在特定的 N-末端基团时,含有被保护的 N-末端基团如焦谷氨酸的肽的具体活性会通过比较完全(100%)被保护的肽与未被保护(0%)的肽的活性将会容易地得到证明。

用于衍生化的化学部分:适于衍生化的化学部分可选自水溶性聚合物中。所选择的聚合物应当是水溶性的以便其连接的成分不会沉淀在水性环境如,生理环境中。优选地,对于成品制剂的治疗使用来说,聚合物将是可药用的。本领域技术人员能够根据诸如聚合物/成分缀合物是否将在治疗上使用之类的考虑来选择所需的聚合物,如果用于治疗,则要考虑所需剂量、循环时间、对蛋白水解作用的耐受性以及其它因素。对于本发明的一种或多种成分来说,使用本文提供的测定方法是可以确定的。

例如,水溶性聚合物可选自聚乙二醇、乙二醇/丙二醇的共聚物、羧甲基纤维素、葡聚糖、聚乙烯醇、聚乙烯吡咯烷酮、聚-1,3-二氧戊烷、聚-1,3,6-三氧杂环己烷、乙烯/马来酸酐共聚物、聚氨基酸(均聚物或随机共聚物)和葡聚糖或聚(n-乙烯基吡咯烷酮)聚乙二醇、丙二醇均聚物、聚环氧丙烷/环氧乙烷共聚物、聚氧乙烯化多元醇和聚乙烯醇。聚乙二醇丙醛由于其在水中的稳定性可能在生产中具有优势。

聚合物可以具有任意的分子量,并且可以是分枝的或不分枝的。就聚乙二醇来说,为了易于操作和制备,优选的分子量在约 2kDa 到约 100kDa 之间(术语“约”表示在聚乙二醇的制备中一些分子比所述的分子量更重而另一些则更轻)。可使用其它大小的聚乙二醇,这取决于所需的治疗全貌(如,所需的持续释放期,对生物活性的影响,如果有的话,操作的难易度,抗原性的有无或大小和聚乙二醇对治疗性蛋白或类似物的其它已知影响)。

如此连接的聚合物分子的数量是可变的,并且本领域技术人员能够确定其对功能的影响。可以对相同或不同的化学部分(例如,聚合物,如不同重量的聚乙二醇)进行单衍生化,或者可以将衍生作用进行二、三、四或某些组合。聚合物分子与一种或多种成分分子的比例是可变的,它们在反应混合物中的浓度也是可变的。一般来说,最佳比(从反应的效率来说,没有过量的未反应的一种或多种成分和聚合物)将通过

各种因素来确定,如所需的衍生化程度(如,单、二、三等)、所选聚合物的分子量、聚合物是分枝的还是不分枝的以及反应条件。

聚乙二醇分子(或其它化学部分)应当连接到一种或多种成分上,所述成分被认为对蛋白质的功能性或抗原性区域有影响。对本领域技术人员来说,有许多可以使用的连接方法,如参见引入本文供参考的 EP 0401384(PEG 偶合到 G-CSF 上),也可参见 Malik 等人,1992, Exp. Hematol. 20:1028-1035(报告了用 tresyl chloride 使 GM-CSF pegylation)。例如,聚乙二醇可以经氨基酸残基通过反应基团如游离氨基或羧基来共价结合。反应基团是那些可以与活化的聚乙二醇分子结合的基团。具有游离氨基的氨基酸残基包括赖氨酸残基和 N-末端的氨基酸残基;具有游离羧基的氨基酸残基包括天冬氨酸残基、谷氨酸残基和 C-端的氨基酸残基。巯基也可用作连接聚乙二醇分子的反应性基团。出于治疗目的,连接在氨基上是优选的,如连接在 N-末端或赖氨酸基团上。

本发明提供一种分离的核酸,它编码有含有胆碱结合蛋白 A N-末端截取物的氨基酸序列的多肽。本发明提供一种分离的核酸,它编码有如图 2 所述的胆碱结合蛋白 A N-末端截取物的氨基酸序列的多肽。在一个实施方案中,核酸是 SEQ ID NOS 12, 14-17 和 19-21 所述的,包括其片段、突变体、变体、类似物或衍生物。核酸是 DNA、cDNA、基因组 DNA、RNA。另外,分离的核酸可有效地连接到 RNA 转录的启动子上。期待使用该核酸来竞争性抑制凝集素活性。

“载体”是一种复制子,如质粒、噬菌体或粘粒。可与另一 DNA 节段连接以便复制所连接的节段。

“DNA”指以其单股形式或双股螺旋存在的脱氧核糖核苷酸(腺嘌呤、鸟嘌呤、胸腺嘧啶或胞嘧啶)的聚合形式。该术语仅指分子的一级和二级结构,并且不将它限制到任何特定的三级结构。因此,该术语包括双股 DNA,特别是在线性 DNA 分子(如限制片段)、病毒、质粒和染色体中发现的。在特定双股 DNA 分子结构的讨论中,本文可按照仅在沿 DNA 的非转录股(即,具有与 mRNA 同源序列的股)的 5'到 3'方向上给出序列的正常惯例来描述序列。

当表达控制序列控制和调节 DNA 序列的转录和翻译时,将 DNA 序

列“有效地连接”到表达控制序列上。术语“有效地连接”包括在被表达的 DNA 序列前面具有合适的起始信号(如, ATG)和维持正确的读框以允许在表达控制序列的控制下表达 DNA 序列并产生由该 DNA 序列编码的所需产物。如果需要插入重组 DNA 分子中的基因不含有合适的起始信号, 则可以将这样的起始信号在该基因前面插入。

另外, 本发明也提供包含上述核酸分子的载体。启动子可以是或者等同于细菌、酵母、昆虫或哺乳动物的启动子。另外, 载体可以是质粒、粘粒、酵母人造染色体(YAC)、噬菌体或真核病毒 DNA。

可以使用可用于表达蛋白的本领域已知的其它为数众多的载体骨架。这些载体包括但不限于: 腺病毒(AV)、腺伴随病毒(AAV)、猿猴病毒 40(SV40)、巨细胞病毒(CMV)、小鼠乳腺肿瘤病毒(MMTV)、莫洛尼鼠类白血病病毒、DNA 递送系统, 即脂质体和表达质粒递送系统。此外, 一类载体包含由病毒衍生的 DNA 元件, 病毒如牛乳头瘤病毒、多形瘤病毒、杆状病毒、逆转录病毒或 Semliki Forest 病毒。这些载体可以商购或者由所述序列通过本领域公知的方法来装配。

本发明也提供生产多肽的宿主载体系统, 它包含合适宿主细胞的载体。合适的宿主细胞包括但不限于原核细胞或真核细胞, 如细菌细胞(包括革兰氏阳性细胞)、酵母细胞、真菌细胞、昆虫细胞和动物细胞。许多哺乳动物的细胞都可用作宿主, 包括但不限于小鼠成纤维细胞 NIH 3T3、CHO 细胞、HeLa 细胞、Ltk 细胞、Cos 细胞等。

在表达本发明的 DNA 序列中, 可以采用各种各样的宿主/表达载体的组合。例如, 有用的表达载体可以由染色体、非染色体和合成 DNA 序列的片段组成。合适的载体包括 SV40 和已知的细菌质粒, 如大肠杆菌质粒 col E1, pCR1, pBR322, pMB9 及其衍生物, 质粒如 RP4 的衍生物; 噬菌体 DNAs, 如噬菌体  $\lambda$  的各种衍生物, 如 NM989, 和其它噬菌体 DNA, 如 M13 和丝状单股噬菌体 DNA; 酵母质粒如  $2\mu$  质粒或其衍生物; 真核细胞中可以使用的载体, 如在昆虫或哺乳动物细胞中可以使用的载体; 由质粒和噬菌体 DNA 的组合衍生的载体, 如已被修饰以利用噬菌体 DNA 或其它表达控制序列的质粒等。

可在这些载体中使用各种各样的表达控制序列 - 控制与其有效连接的 DNA 序列的表达的序列 - 来表达本发明的 DNA 序列。这些可以

使用的表达控制序列包括如 SV40、CMV、牛痘、多形瘤或腺病毒的早期或晚期启动子、*lac* 系统、*trp* 系统、*TAC* 系统、*TRC* 系统、*LTR* 系统、噬菌体 $\lambda$ 的主要操纵子和启动子区域、fd 外壳蛋白的控制区、3-磷酸甘油酸激酶或其它糖酵解酶的启动子、酸性磷酸酶(如, Pho5)的启动子、酵母 $\alpha$ -交配因子的启动子和其它已知的控制原核细胞或真核细胞或其病毒基因表达的序列及其各种组合。

各种各样的单细胞宿主细胞也可以用于表达本发明的 DNA 序列。这些宿主可包括公知的真核和原核宿主, 如大肠杆菌、假单胞菌属、芽孢杆菌属、链霉菌属、真菌如酵母的菌株和动物细胞如 CHO, R1.1, B-W 和 L-M 细胞、非洲绿猴肾细胞(如, COS1、COS7、BSC1、BSC40 和 BMT10)、昆虫细胞(如, SF9)和组织培养中的人细胞和植物细胞。

应该理解对于表达本发明的 DNA 序列来说不是所有的载体、表达控制序列和宿主都能发挥同样好的功能。即使表达系统相同所有宿主也不能发挥同样好的功能。然而, 在不脱离本发明范围的情况下, 本领域技术人员不需要过多的实验就能够选择合适的载体、表达控制序列和宿主来完成所需的表达。例如, 在选择载体时, 必须考虑宿主, 因为载体必须在宿主中发挥作用。也应考虑载体的拷贝数目、控制拷贝数目的能力和载体所编码的任何其它蛋白的表达, 如抗生素标记物。

在选择表达控制序列时, 通常将考虑各种因素。例如, 这些因素包括系统的相对强度、其控制能力及其与被表达的特定 DNA 序列或基因的相容性, 特别是考虑潜在的二级结构。将通过考虑下列因素来选择合适的单细胞宿主: 如, 它们与所选载体的相容性、其分泌特性、其正确折叠蛋白的能力、其发酵的要求以及对被表达的 DNA 序列所编码的产物的宿主的毒性和表达产物纯化的难易。

本发明进一步提供一种制备多肽的方法, 它包括使上述宿主载体系统在适宜的允许多肽产生的条件下生长并回收所产生的多肽。

本发明进一步提供一种能够特异性识别分离的多肽或与其结合的抗体。所述抗体可以是单克隆抗体或多克隆抗体。此外, 所述抗体可以用可检测的标记物标记, 所述标记物是放射性的、可比色的、荧光的或发光的标记物。所述标记抗体可以是单克隆抗体或多克隆抗体。

在一个实施方案中，所述标记抗体是纯化的标记抗体。标记抗体的方法是本领域公知的。

术语“抗体”的实例包括天然存在的和非天然存在的抗体。具体地说，术语“抗体”包括多克隆和单克隆抗体及其片段。此外，术语“抗体”包括嵌合抗体和全合成抗体及其片段。所述抗体包括但不限于多克隆、单克隆、嵌合、单链、Fab片段和Fab表达文库。

可以使用本领域已知的各种方法来制备多肽或其衍生物或类似物的多克隆抗体(例如，参见**抗体—实验室手册**, Harlow and Lane, eds., Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 1988)。为了制备抗体，可通过给各种宿主动物注射截取的CbpA或其衍生物(例如片段或融合蛋白)使其被免疫，所述宿主动物包括但不限于兔子、小鼠、大鼠、绵羊、山羊等。在一个实施方案中，可将多肽缀合到免疫原性载体，例如牛血清白蛋白(BSA)或钥孔蠔血蓝蛋白(KLH)上。依赖于宿主的种类，各种佐剂可用于增强免疫反应。

为了制备单克隆抗体或其片段、其类似物或其衍生物，可使用由任何在培养物中的传代细胞系产生抗体分子的技术(例如，参见**抗体—实验室手册**, Harlow and Lane, eds., Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York, 1988)。这些技术包括但不限于最初由Kohler和Milstein开发的杂交瘤技术(1975, *Nature* 256: 495-497)以及trioma技术、人B-细胞杂交瘤技术(Kozbor等人, 1983, *今日免疫学* 4:72)和制备人单克隆抗体的EBV-杂交瘤技术(Cole等人, 1985, *单克隆抗体和癌症治疗*, Alan R. Liss, Inc., pp. 77-96)。在本发明另一实施方案中，可在无菌动物中，利用最新技术(PCT/US90/02545)生产单克隆抗体。按照本发明，人抗体可使用并且可通过使用人杂交瘤(Cote等人, 1983, *Proc. Natl. Acad. Sci. U. S. A.* 80:2026-2023)或通过在外用EBV病毒转化人B细胞(Cole等人, 1985, *单克隆抗体和癌症治疗*, Alan R. Liss, Inc., pp. 77-96)获得。事实上，按照本发明，可使用所开发的通过将多肽具有特异性的小鼠抗体分子的基因与具有适宜的生物活性的人抗体分子的基因拼接起来制备“嵌合抗体”的技术(Morrison等人, 1984, *J. Bacteriol.* 159:870; Neuberger等人, 1984, *Nature* 312:604-608; Takeda等人, 1985, *Nature* 314:452-454); 所述抗体包含在本发明范

国内。所述人或人源化嵌合抗体优选用于治疗人类疾病(如下所述),因为人或人源化抗体比异种抗体诱导免疫反应,特别是其本身的过敏反应的可能性要小得多。本发明另一实施方案是利用构建 Fab 表达文库中所述的技术(Huse 等人, 1989, Science 246:1275-1281)来快速和方便地鉴定具有所需的对多肽的特异性的单克隆 Fab 片段、或其衍生物或类似物。

包含抗体分子独特型的抗体片段可通过已知技术来产生。例如,这些片段包括但不限于:可通过抗体分子的胃蛋白酶消化产生的  $F(ab')_2$  片段;可通过还原  $F(ab')_2$  片段的二硫桥产生的 Fab' 片段和可通过用木瓜蛋白酶和还原剂处理抗体分子产生的 Fab 片段。

在抗体的产生中,可通过本领域已知的技术来筛选所需抗体,如,放射免疫测定、ELISA(酶联免疫吸附测定)、“夹心”免疫测定、免疫放射测定、凝胶扩散沉积素反应、免疫扩散测定、原位免疫测定(例如使用胶体金、酶或放射性同位素标记)、蛋白质印迹、沉淀反应、凝集测定(如,凝胶凝集测定、血细胞凝集测定)、补体结合测定、免疫荧光测定、蛋白质 A 测定和免疫电泳测定等。在一个实施方案中,通过检测原发抗体上的标记来检测抗体结合。在另一个实施方案中,通过检测继发抗体或试剂与原发抗体的结合来检测原发抗体。在另一实施方案中,标记继发抗体。本领域中已知有多种免疫测定中检测结合的方法并且在本发明的范围之内。

为了在体外进行检测,可用标记物标记抗体,标记物如酶、荧光团、发色团、放射性同位素、染料、胶体金、乳胶粒和化学发光剂。另外,为了在体内进行检测,可用如放射性同位素(优选铊或碘);磁共振位移试剂(如钆和锰);或不透射线的试剂来标记抗体。

这些研究中最常用的标记物是放射性元素、酶、当暴露于紫外光下时有荧光的化学药品等。许多荧光材料都是已知的并且可用作标记物。它们包括如荧光素、若丹明、金胺、德克萨斯红、AMCA 蓝和荧光黄。具体的检测材料是在山羊中制备的并经异硫氰酸酯与荧光素缀合的抗兔抗体。多肽也可用放射性元素或用酶标记。可通过目前可利用的计数方法来检测放射性标记物。优选的同位素可选自  $^3\text{H}$ 、 $^{14}\text{C}$ 、 $^{32}\text{P}$ 、 $^{35}\text{S}$ 、 $^{36}\text{Cl}$ 、 $^{51}\text{Cr}$ 、 $^{57}\text{Co}$ 、 $^{58}\text{Co}$ 、 $^{59}\text{Fe}$ 、 $^{90}\text{Y}$ 、 $^{125}\text{I}$ 、 $^{131}\text{I}$  和  $^{186}\text{Re}$ 。

同样也可以使用酶标记物并且可通过任何目前采用的比色、分光光度、荧光分光光度、电流或气流技术来检测。通过与桥接分子如碳化二亚胺、异硫氰酸酯、戊二醛等反应，将酶与所选颗粒缀合。在这些方法中可以使用的许多酶都是已知的并且可以使用。优选的是过氧化物酶、 $\beta$ -葡糖醛酸酶、 $\beta$ -D-葡糖苷酶、 $\beta$ -D-半乳糖苷酶、脲酶、葡萄糖氧化酶加过氧化物酶和碱性磷酸酶。通过举例的方式提及美国专利 3654090、3850752 和 4016043 中公开了候补的标记材料和方法。

在本发明的另一实施方案中，可以制备适于医学专家使用的商购试验试剂盒以确定预定的结合活性或者预定的与被怀疑的靶细胞结合活性能力的存在或缺乏。根据以上所讨论的试验技术，一类试剂盒至少包含标记的多肽或其结合配对物，例如其特异性抗体和用法说明，当然，这取决于所选方法，如“竞争性的”、“夹心的”、“DASP”等。该试剂盒也可以包含外围试剂如缓冲剂、稳定剂等。

因此，可以制备证明有利于预定的细菌结合活性的细胞的存在或能力的试验试剂盒，它包含：

(a)预定量的至少一种标记的免疫化学反应性成分，该成分是通过将本发明的多肽或其特异性结合配对物直接或间接连接到可检测标记物上获得的；

(b)其它试剂；和

(c)使用所述试剂盒的说明。

本发明提供拮抗剂或阻断剂，它包括但不限于：肽片段、模拟物、核酸分子、核酶、多肽、小分子、碳水化合物分子、单糖、寡糖或抗体。本发明也包括竞争性阻断或抑制肺炎球菌的药剂。本发明提供一种药剂，它包含无机化合物、核酸分子、寡核苷酸、有机化合物、肽、肽模拟化合物或抑制该多肽的蛋白质。

本发明提供一种疫苗，它包含具有 SEQ ID NOS: 1, 3-7, 9-11, 22 和 23 中所述氨基酸序列的多肽和可药用佐剂或载体。所述多肽可包含图 2 所述的胆碱结合蛋白 A N-末端截取物的氨基酸序列。本发明提供一种疫苗，它包含多肽和可药用佐剂或载体，所述多肽具有包含图 2 所述保留区的氨基酸序列。例如，保留区包括但不限于氨基酸序列 158-172; 300-321; 331-339; 355-365; 367-374; 379-389; 409-427 和 430-447。

本发明提供一种疫苗，它包含编码该多肽的分离的核酸和可药用佐剂或载体。

可通过用免疫原量的多肽或其衍生物或片段和佐剂进行免疫(接种疫苗)来诱导抗革兰氏阳性菌，特别是肺炎球菌的自动免疫，其中所述多肽或其抗原衍生物或片段是所述疫苗的抗原组分。

可在与制备疫苗的佐剂的混合物中制备本发明多肽或其衍生物或片段。优选地，用作所述疫苗抗原组分的本发明多肽的衍生物或其片段是粘附素。更优选地，用作所述疫苗抗原组分的多肽或肽衍生物或其片段是与所有或很多革兰氏阳性菌种的菌株共有的或与密切相关的细菌菌种共有的抗原。最优选地，所述疫苗的抗原组分是粘附素，它是一种共同抗原。

可通过本领域已知的方法，例如转染、电穿孔、微量注射、转导、细胞融合、DEAE葡聚糖、磷酸钙沉淀、脂转染(溶菌体融合)、基因枪的应用或DNA载体转运蛋白，将包含本发明以核酸为主药的疫苗的载体引入需要的宿主(例如参见Wu等人,1992, J. Biol. Chem. 267:963-967; Wu和Wu,1988, J. Biol. Chem. 263:14621-14624; Hartmut等人,1990年3月15日申请的加拿大专利申请2,012,311)。

所述疫苗可通过任何非肠道途径给药，所述非肠道途径包括但不限于肌肉内、腹膜内、静脉内等。优选地，因为理想的疫苗接种结果是阐明与抗原并因此与病原生物体的免疫反应，所以直接或通过寻靶或选择病毒载体间接给予到淋巴组织，例如淋巴结或脾是理想的。因为免疫细胞在不断地复制，所以它们是以逆转录病毒载体为主药的核酸疫苗的理想靶细胞，因为逆转录病毒需要复制细胞。

可通过将抗本发明多肽的抗血清、多克隆抗体或中和单克隆抗体给予怀疑受到革兰氏阳性菌，优选链球菌，并且更优选肺炎球菌感染的受治疗动物而使该动物获得被动免疫。尽管被动免疫不会带来长期的保护作用，但它可能是治疗尚未接种疫苗的受治疗者细菌感染的有价值的工具。对于革兰氏阳性抗生素耐药菌株的治疗来说，被动免疫是特别重要的，因为没有其它治疗可以利用。优选地，所给予的用于被动免疫治疗的抗体是自体抗体。例如，如果所述受治疗者是人，为了使针对抗体的免疫反应的可能性降低到最小，优选地，所述抗体是由



人体获得的或者是“人源化的”。本发明主动性或被动性疫苗或给予粘附素可用于保护受治疗动物使其免受革兰氏阳性菌，优选链球菌，并且更优选肺炎球菌的感染。

本发明提供一种药物组合物，它包含一定量本文所描述的多肽和可药用载体或稀释剂。

例如，这种用于阻止肺炎球菌附着到粘膜表面的药物组合物可包括抗凝集素区和/或可溶性过剩凝集素区蛋白质的抗体。通过任一机理阻断附着都能阻断感染的最初步骤，由此减少移生作用。这又转过来减少人与人之间的传播并预防症状性疾病的发展。

本发明提供一种诱导已经接触到或感染了肺炎球菌的受治疗者的免疫反应的方法，它包括给予所述受治疗者一定量的药物组合物，由此诱导免疫反应。

本发明提供一种预防受治疗者被肺炎球菌感染的方法，它包括给予所述受治疗者一定量的阻止肺炎球菌附着有效量的药物组合物，由此预防肺炎球菌感染。

本发明提供一种预防受治疗者被肺炎球菌感染的方法，它包括给予所述受治疗者一定量预防包含抗体和可药用载体或稀释剂的药物组合物，由此预防肺炎球菌感染。

本发明提供一种治疗感染了或接触到肺炎球菌的受治疗者的方法，它包括给予所述受治疗者治疗有效量本发明的疫苗，由此治疗所述受治疗者。

本发明提供一种抑制已经接触到或感染了肺炎球菌的受治疗者宿主细胞移生的方法，它包括给予所述受治疗者一定量的包含由 SEQ ID NOS 1, 3-5, 7 或 9-11 中所述氨基酸序列组成的多肽的药物组合物，由此诱导免疫反应。所述阻断移生作用的治疗性肽是由呼吸粘膜递送的。包含由图 2 中所述氨基酸序列组成的多肽的药物组合物。

本文所使用的“药物组合物”是指通过例如阻止肺炎球菌移生提供治疗效果或益处的、治疗有效量的本发明多肽产品以及适宜的稀释剂、防腐剂、增溶剂、乳化剂、辅剂和/或载体。本文所使用的“治疗有效量”是指提供对所述疾病治疗效果和给药方案的量。所述组合物是液体或冻干或用其它方法干燥的制剂并且包括由各种缓冲剂内含物

(例如 Tris-HCl、醋酸盐、磷酸盐)、pH 和离子浓度组成的稀释剂、添加剂如阻止表面吸收的白蛋白或明胶、去污剂(例如吐温 20、吐温 80、Pluronic F68、胆汁酸盐)、增溶剂(例如甘油、聚乙二醇)、抗氧化剂(例如抗坏血酸、焦亚硫酸钠)、防腐剂(例如、硫柳汞、苯甲醇、对羟基苯甲酸酯类)、填充物质或张力调节剂(例如乳糖、甘露糖醇)、聚合物如聚乙二醇共价连接到蛋白质上、与金属离子络合或将所述物质掺入聚合物如聚乳酸、聚乙醇酸、水凝胶等的特定制剂中或加到其上、或者加到脂质体、微乳液、微团、单层或多层囊泡、红细胞血影或原生质球上。所述组合物将影响治疗剂的物理状态、溶解度、稳定性、体内释放率和体内清除率。组合物的选择将依赖于具有治疗活性的蛋白质的物理或化学性质。例如,由膜结合型活性物衍生的产物可能需要含去污剂的制剂。控释或缓释组合物包括在亲脂性贮库(例如脂肪酸、蜡、油)中的制剂。本发明也包括由聚合物(例如泊咯沙姆或 poloxamines)包被的微粒组合物和偶联到针对组织特异性受体、配体或抗原的抗体上或偶联到组织特异性受体的配体上的活性物。本发明组合物的其它实施方案包括微粒形式的保护性包衣、蛋白酶抑制剂或渗透增强剂用于各种途径包括非肠道、肺、鼻和口服给药。

此外,本文所使用的“可药用载体”是本领域技术人员公知的并且包括但不限于 0.01-0.1M 并且优选 0.05M 的磷酸盐缓冲剂或 0.8% 的盐水。因此,所述可药用载体可以是水或非水溶液、悬浮液和乳液。非水溶剂的实例是丙二醇、聚乙二醇、植物油如橄榄油和可注射有机酯如油酸乙酯。水性载体包括水、醇/水溶液、悬浮液或乳液,包括盐水和缓冲介质。肠胃外载体包括氯化钠溶液、林格氏葡萄糖、葡萄糖和氯化钠、乳酸化林格氏溶液或固定油。静脉内载体包括液体和营养补充剂、电解质补充剂如那些以林格氏葡萄糖为主药的电解质补充剂等。也可以存在防腐剂和其它添加剂如抗微生物剂、抗氧化剂、对照剂、惰性气体等。

术语“佐剂”是指增强对抗原的免疫反应的化合物或混合物。佐剂可充当缓慢释放抗原的组织库并且也可以充当非特异性增强免疫反应的淋巴系统激活剂(Hood 等人, *Immunology, Second Ed.*, 1984, Benjamin/Cummings: Menlo Park, California, p. 348)。通常,在不存在

佐剂的情况下，单独用抗原进行初级激发将不能引起体液或细胞免疫反应。佐剂包括但不限于完全弗氏佐剂、不完全弗氏佐剂、皂甙、无机凝胶如氢氧化铝、表面活性物质如溶血卵磷脂、普卢兰尼克多元醇、聚阴离子、肽、油或烃乳液、钥孔蛾血蓝蛋白、二硝基苯酚和可能有用的人佐剂如 BCG(卡介苗)和小棒杆菌(*Corynebacterium parvum*)。优选地，所述佐剂是可药用的。

控释或缓释组合物包括在亲脂性贮库(例如脂肪酸、蜡、油)中的制剂。本发明也包括由聚合物(例如泊咯沙姆或 poloxamines)包被的微粒组合物和偶联到针对组织特异性受体、配体或抗原的抗体上或偶联到组织特异性受体的配体上的化合物。本发明组合物的其它实施方案包括微粒形式的保护性包衣、蛋白酶抑制剂或渗透增强剂用于各种途径包括非肠道、肺、鼻和口服给药。

当给药时，化合物通常迅速从粘膜表面或循环中清除并且因此可引起相对短效的药理学活性。所以，要求经常给予相对大剂量的生物活性化合物来维持疗效。已知通过共价连接水溶性聚合物如聚乙二醇、聚乙二醇和聚丙二醇的共聚物、羧甲基纤维素钠、葡聚糖、聚乙烯醇、聚乙烯吡咯烷酮或聚脯氨酸修饰的化合物在静脉注射后明显显示比相应的未修饰化合物更长的半衰期(Abuchowski 等人,1981; Newmark 等人,1982;和 Katre 等人,1987)。该修饰也可以增加所述化合物在水溶液中的溶解度、消除聚集作用、增强化合物的物理和化学稳定性并且大大降低化合物的免疫原性和反应性。结果，与未修饰化合物相比，可通过给予更少次数或更低剂量的所述聚合物-化合物加合物获得所需要的体内生物学活性。

剂量：足够的量包括但不限于大约为 1 $\mu$ g/kg-1000mg/kg。所述量可以为 10mg/kg。组合物的可药用形式包括可药用载体。

如上所述，本发明提供包含药物组合物的治疗组合物，其中所述药物组合物包含载体、疫苗、多肽、核酸和抗体、抗-抗体和与肺炎球菌竞争病原活性如附着到宿主细胞上的药剂。

包含活性组分的治疗组合物的制备是本领域公知的。典型地，所述组合物可制备成递送到鼻咽部的多肽气雾剂或可注射的液体溶液或悬浮液，然而，也可以制备成适用于在注射前溶于或悬浮于液体中的固体

形式。该制剂也可以被乳化。通常，将活性治疗成分与可药用并且可与活性成分相容的赋形剂混合。例如适宜的赋形剂是水、盐水、葡萄糖、甘油、乙醇等及其组合。此外，如果需要，所述组合物可包含少量增强活性成分有效性的辅助物质如润湿剂或乳化剂、pH缓冲剂。

可将活性组分配制成以中和的可药用盐形式存在的治疗组合物。可药用盐包括酸加成盐(用多肽或抗体分子的游离氨基形成的)，它是与无机酸如盐酸或磷酸，或有机酸如乙酸、草酸、酒石酸、苦杏仁酸等形成的。由游离羧基形成的盐也可以衍生于无机碱如钠、钾、铵、钙或铁的氢氧化物或有机碱如异丙胺、三甲胺、2-乙基氨基乙醇、组氨酸、普鲁卡因等。

当组合物中至少大约 75% (重量) 的所述蛋白质、DNA、载体(依赖于 A 和 B 所属种类的范畴)是“A”时，包含“A”(其中“A”是单一的蛋白质、DNA 分子、载体等)的组合物基本上不含“B”(其中“B”包含一个或多个污染蛋白质、DNA 分子、载体等)。优选地，在组合物中，“A”包含至少大约 90% (重量) 的 A+B，最优选地，至少大约 99% (重量)。

本文所使用的术语“治疗有效量”是指足够减少至少大约 15%，优选至少大约 50%，更优选至少大约 90% 并且最优选预防宿主的活性、功能和反应方面的临床上明显不足的量。另外，治疗有效量足够引起宿主临床上明显病症的改善。在本发明的上下文中，宿主反应方面的不足可由持续或传播细菌感染来证明。宿主临床上明显病症的改善包括细菌载荷的减少、细菌从移生的宿主细胞中的清除、与感染有关的发烧或炎症的减少或者与细菌感染有关的任何症状的减少。

按照本发明，可通过非肠道、透过粘膜，例如通过口、鼻、肺或直肠，或者透皮引入本发明治疗组合物的成分。优选地，给药途径是非肠道的，例如通过静脉内注射并且也包括但不限制于动脉内、肌肉内、真皮内、皮下、腹膜内、心室内和颅内给药。可优选通过口或肺递送来激活粘膜免疫；因为肺炎球菌通常移生在鼻咽和肺粘膜中，所以粘膜免疫可能是特别有效的预防性治疗。当在本发明治疗组合物中使用，术语“单位剂量”是指适宜作为单次剂量用于人的物理上分立的单位，各单位包含为产生所需要治疗效果而计算的预定量的活性物质

和所需要的稀释剂，即载体。

在另一实施方案中，所述活性化合物可在泡囊特别是脂质体中递送(参见 Langer, 科学 249:1527-1533(1990); Treat 等人, 脂质体在传染病和癌症的治疗中, Lopez-Berestein 和 Fidler(编者), Liss, 纽约, pp. 353-365(1989); Lopez-Berestein, 出处同上, pp. 317-327; 参见上文)。

在另一实施方案中，所述治疗化合物可在控释系统中递送。例如，可利用静脉输注、可植入的渗透泵、透皮贴片、脂质体或其它给药方式给予所述多肽。在一个实施方案中，可以使用泵(参见 Langer, 如上所述; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201(1987); Buchwald 等人, 外科 88:507(1980); Saudek 等人, N. Engl. J. Med. 321:574(1989))。在另一实施方案中，可使用高分子材料(参见控释的医学应用, Langer 和 Wise(编者), CRC 出版社, Boca Raton, Florida(1974); 控释药物的生物利用度、药品设计和性能, Smolen 和 Ball(编者), Wiley, New York (1984); Ranger and Peppas, J. Macromol. Sci. Rev. Macromol. Chem. 23:61(1983); 也参见 Levy 等人, 科学 228:190(1985); During 等人, Ann. Neurol. 25:351(1989); Howard 等人, J. Neurosurg. 71:105(1989))。在另一实施方案中，可将控释系统放置在治疗靶，即脑的附近，因此仅需要小部分所述系统剂量(例如参见 Goodson, 控释的医学应用, 如上所述, vol. 2, pp.115-138(1984))。优选地，将控释装置引入治疗对象不适宜免疫激活部位或肿瘤的附近。在 Langer 的评论中讨论了其它控释系统(科学 249:1527-1533(1990))。

优选地，以给予上述活性组分作为细菌感染有效治疗方案的治疗对象是人，但是也可以的任何动物。因此，本领域的普通技术人员都能迅速领会，本发明所述方法和药物组合物特别适于给予任何动物，特别是哺乳动物并且包括但不限于家养动物如猫或犬，农场动物如但不限制于牛、马、山羊、绵羊和猪，野生动物(不管是野外的还是在动物园中的)，研究动物如小鼠、大鼠、兔子、山羊、绵羊、猪、狗、猫等，即供兽医使用。

在本发明治疗方法和组合物中，提供治疗有效剂量的活性组分。治疗有效剂量可由普通医学工作者根据病人的特点(年龄、体重、性别、状况、并发症，其它疾病等)确定，这是本领域技术人员公知的方法。

此外，随着进行其它常规研究，将出现有关治疗各种病人各种疾病适宜的剂量水平的更具体的资料，并且根据治疗情况、被治疗者的年龄和一般健康状况，普通技术人员能够确定适宜的给药剂量。通常，对于静脉内注射或输注来说，剂量可以低于腹膜内、肌肉内或其它给药途径。给药计划表可依赖于循环半衰期和所使用的制剂而改变。将治疗有效量的所述组合物以与剂量制剂相容的方式给予。所需要给予的活性组分的精确量依赖于医师的判断并且是每个个体所特有的。然而，适宜的剂量可大约为每天每公斤体重 0.1-20，优选大约为 0.5-10 并且更优选 1-几微克活性组分并且依赖于给药途径。适宜的初次给药和加强注射的方案也是可改变的，但是其特征是在初次给药后，间隔一小时或多小时，通过随后注射或其它给药途径给予多次量。或者，可以考虑连续静脉内输注的方法来足以维持血中 10 纳摩尔-10 微摩尔的浓度。

与其它化合物一起给予：就细菌感染的治疗而言，可以与一种或多种用于治疗细菌感染的药物组合物一起给予本发明活性组分，所述药物组合物包括但不限于(1)抗生素；(2)细菌粘附素的可溶性碳水化合物抑制剂；(3)细菌粘附素的其它小分子抑制剂；(4)细菌代谢、转运或转化的抑制剂；(5)细菌溶胞的刺激物或(6)针对其它细菌抗原上的抗菌抗体或疫苗。其它有效的活性组分包括抗炎剂如甾体或非甾体抗炎药物。可同时给药(例如本发明活性组分和抗生素混合物的给予)，或者可依次给药。

因此，在具体的实施方案中，所述治疗组合物可进一步包括治疗有效量的活性组分和一种或多种下列活性组分：抗生素、甾族化合物等。制剂的实例如下：

### 制剂

#### 静脉内制剂 I

<u>组分</u>	<u>mg/ml</u>
头孢噻肟	250.0
多肽	10.0
葡萄糖 USP	45.0
亚硫酸氢钠 USP	3.2

乙二胺四乙酸二钠 USP	0.1
注射用水适量加至	1.0ml

**静脉内制剂 II**

<u>组分</u>	<u>mg/ml</u>
氨苄青霉素	250.0
多肽	10.0
亚硫酸氢钠 USP	3.2
乙二胺四乙酸二钠 USP	0.1
注射用水适量加至	1.0ml

**静脉内制剂 III**

<u>组分</u>	<u>mg/ml</u>
庆大霉素(硫酸盐)	40.0
多肽	10.0
亚硫酸氢钠 USP	3.2
乙二胺四乙酸二钠 USP	0.1
注射用水适量加至	1.0ml

**静脉内制剂 IV**

<u>组分</u>	<u>mg/ml</u>
多肽	10.0
葡萄糖 USP	45.0
亚硫酸氢钠 USP	3.2
乙二胺四乙酸二钠 USP	0.1
注射用水适量加至	1.0ml

**静脉内制剂 V**

<u>组分</u>	<u>mg/ml</u>
多肽拮抗剂	5.0
亚硫酸氢钠 USP	3.2

乙二胺四乙酸二钠 USP	0.1
注射用水适量加至	1.0ml

因此，在要求减少或抑制由细菌介导的细菌与宿主细胞、或其抗体、或其配体或该配体的抗体结合引起的感染的特定实例中，引入多肽来阻断细菌与宿主细胞的相互作用。

本文也考虑通过肺递送本发明的具有凝集素活性并且发挥本发明粘附素抑制剂(或其衍生物)作用的多肽。将粘附素抑制剂(或其衍生物)递送到哺乳动物的肺中，在肺中，它能干扰细菌，即链球菌并且优选肺炎球菌与宿主细胞结合。在本领域发现了其它关于通过肺递送的蛋白质制剂的报道[Adjei 等人, 药物研究, 7: 565-569 (1990); Adjei 等人, *International Journal of Pharmaceutics*, 63:135-144 (1990)(醋酸亮丙瑞林); Braquet 等人, *Journal of Cardiovascular Pharmacology*, 13(suppl. 5): 143-146(1989)(内皮素-1); Hubbard 等人, *Annals of Internal Medicine*, Vol. III, pp.206-212(1989)( $\alpha$ 1-抗胰蛋白酶); Smith 等人, *J. Clin. Invest.* 84:1145-1146 (1989) ( $\alpha$ 1-蛋白酶); Oswein 等人, “蛋白质的气雾化作用”, *Proceedings of Symposium on Respiratory Drug Delivery II*, Keystone, Colorado, March, (1990)(重组人生长激素); Debs 等人, *J. Immunol.* 140:3482-3488(1988)(干扰素- $\gamma$ 和肿瘤坏死因子 $\alpha$ ); Platz 等人, 美国专利 5,284,656(粒细胞集落刺激因子)]. 在 1995 年 9 月 19 日颁布给 Wong 等人的美国专利 5,451,569 中描述了用于通过肺递送药物的方法和组合物。

所有这样的装置均要求使用适于分配粘附素抑制剂(或其衍生物)的制剂。典型地，各制剂对于所使用的装置类型是特异性的并且除了治疗中所使用的通常的稀释剂、佐剂和/或载体外，可包括应用适宜的推进剂材料。也可以考虑使用脂质体、微囊或微球、包合配合物或其它类型的载体。依赖于化学修饰的类型或所使用装置的类型，也可以将通过化学方法修饰的粘附素抑制剂制备在不同的制剂中。

典型地，适用于利用喷射式或超声波式喷雾器来应用的制剂将包含以每 ml 溶液大约 0.1-25mg 生物活性粘附素抑制剂的浓度溶解在水中的粘附素抑制剂(或其衍生物)。所述制剂也可以包括缓冲剂和单糖(例如用于稳定粘附素抑制剂和调节渗透压)。所述喷雾器制剂也可以包含



表面活性剂，用于减少或预防形成气雾剂时由溶液雾化引起的粘附素抑制剂在表面聚集。

利用计量剂量的吸入器装置来应用的制剂通常包含细碎的粉剂，其中所述粉剂包含在表面活性剂的帮助下悬浮在推进剂中的粘附素抑制剂(或其衍生物)。所述推进剂可以是用于该目的的任何常规物质如氯氟烃、氢氯氟烃、氢氟烃或烃，包括三氯氟甲烷、二氯二氟甲烷、二氯四氟乙醇和 1,1,1,2-四氟乙烷或其组合。适宜的表面活性剂包括脱水山梨糖醇三油酸酯和大豆卵磷脂。油酸也可以用作表面活性剂。

液体气雾剂制剂在生理上可接受的稀释剂中包含粘附素抑制剂和分散剂。本发明干粉气雾剂制剂由细碎的固体形式的粘附素抑制剂和分散剂组成。对于液体或干粉气雾剂制剂来说，必须将所述制剂气雾化。即，必须将所述制剂破碎成液体或固体微粒以确保气雾化的剂量真正到达鼻通道或肺粘膜。本文所使用的术语“气雾剂微粒”用于描述适用于鼻或肺给药，即能够到达粘膜的液体或固体微粒。其它考虑如递送装置的结构、制剂中的其它组分和微粒特点是重要的。通过肺给药的这些特征是本领域公知的并且制剂的使用、雾化方法和递送装置的结构至多要求由本领域普通技术人员进行常规实验。在特定的实施方案中，所述微粒的中数动态直径将为 5 微米或更小以确保所述药物微粒到达肺泡[Wearley, L.L., Crit. Rev. in Ther. Drug Carrier Systems 8:333(1991)]。

气雾剂递送系统如加压计量剂量的吸入器和干粉吸入器公开于 Newman, S.P., 气雾剂和肺, Clarke, S.W.和 Davia, D. (编者), pp. 197-22 中并且可与本发明联合使用。

在另一实施方案中，如下文所详细讨论的，本发明的气雾剂制剂除了粘附素抑制剂外，可包括其它治疗学或药理学活性组分，如包括但不限于抗生素、甾体化合物、非甾体抗炎药等。

液体气雾剂制剂：本发明提供用于治疗由细菌，例如链球菌，特别是肺炎球菌感染的患者的气雾剂制剂和剂量形式。通常所述剂量形式在可药用稀释剂中包含粘附素抑制剂。可药用稀释剂包括但不限于无菌水、盐水、缓冲的盐水、葡萄糖溶液等。在特定实施方案中，可以在本发明或本发明的药物制剂中使用的稀释剂是磷酸盐缓冲的盐水

或通常 pH 在 7.0-8.0 范围的缓冲的盐水溶液或水。

本发明的液体气雾剂制剂可不包含或包含可药用甾体、稀释剂、增溶剂或乳化剂、表面活性剂和赋形剂。所述制剂可包括载体。所述载体是可溶解在循环系统中并且是生理上可接受的大分子，其中生理可接受是指本领域技术人员允许将所述载体作为治疗方案的一部分注射给患者，优选地，所述载体在循环系统中是相对稳定的，具有适宜的血浆清除半衰期。所述大分子包括但不限于大豆卵磷脂、油酸和脱水山梨糖醇三油酸酯，优选脱水山梨糖醇三油酸酯。

本实施方案的制剂也可以包括用于维持 pH、稳定溶液或用于调节渗透压的其它药剂。所述药剂的实例包括但不限于盐如氯化钠或氯化钾和碳水化合物如葡萄糖、半乳糖或甘露糖等。

本发明进一步考虑包含粘附素抑制剂和另一种治疗有效药物如抗生素、甾体化合物、非甾体抗炎药等的液体气雾剂制剂。

气雾剂干粉制剂：也考虑将本发明的气雾剂制剂制备成包含细碎粉剂形式的粘附素抑制剂和分散剂的干粉制剂。

通过粉末吸入器装置分配的制剂将包括含有粘附素抑制剂(或其衍生物)的细碎干粉并且也可包含促进粉末从装置中分散量如占制剂重量 50-90% 的填充剂，如乳糖、山梨糖醇、蔗糖或甘露糖醇。最有利地，应该将粘附素抑制剂(或其衍生物)制备成平均粒径 10 $\mu$ m(或微米)，最优选 0.5-5 $\mu$ m 的微粒形式以便最有效地递送到远端肺。在另一实施方案中，所述干粉制剂可包括含有粘附素抑制剂、分散剂和填充剂的细碎干粉。与本发明制剂联合使用的填充剂可包含促进粉末从装置中分散量的乳糖、山梨糖醇、蔗糖或甘露糖醇。

本发明进一步考虑包含粘附素抑制剂和另一种治疗有效药物如抗生素、甾体化合物、非甾体抗炎药等的干粉制剂。

本文所涉及的口服固体剂量形式通常描述于 Remington's Pharmaceutical Science, 18th Ed. 1990(Mark Publishing Co. Easton PA 18042)第 18 章，将其引入本文供参考。固体剂量形式包括片剂、胶囊剂、丸剂、锭剂或糖锭剂、扁囊剂或小药丸。脂质体或类蛋白质包囊也可以用于制备本发明组合物的制剂(例如美国专利 4,925,673 中报道的类蛋白质微球)。脂质体包囊可以使用并且所述脂质体可用各种聚合

物衍生(例如美国专利 5,013,556). Marshall, K.在由 G.S. Banker 和 C. T. Rhodes 编辑的 *Modern Pharmaceutics* 第 10 章, 1979 中描述了治疗剂的可能的固体剂量形式, 将其引入本文供参考。通常, 所述制剂将包含所述组分(或其通过化学方法修饰的形式)和用于避免胃环境破坏和允许在肠中释放生物活性物质的惰性组分。

特别地, 本文也考虑上述衍生组分的口服剂量形式。所述组分可通过化学方法修饰使得经口递送所述衍生物是有效的。通常, 所述化学修饰是使至少一个部分连接到所述组分分子上, 其中所述部分可能(a)抑制蛋白水解; 和(b)从胃或肠摄入血流中。也期望增加所述组分的整体稳定性和增加体循环时间。所述部分的实例包括: 聚乙二醇、乙二醇和丙二醇的共聚物、羧甲基纤维素、葡聚糖、聚乙烯醇、聚乙烯吡咯烷酮和聚脯氨酸。Abuchowski 和 Davis, 1981, “可溶性聚合物-酶加合物” In: *Enzymes as Drugs*, Hocenberg 和 Roberts, 编者, Wiley-Interscience, New York, NY, pp. 367-383; Newmark 等人, 1982, *J. Appl. Biochem.* 4:185-189。可使用的其它聚合物是聚-1,3-二氧戊环和聚-1,3,6-tioxocane。如上所述, 药用优选的是聚乙二醇部分。

对于所述组分(或其衍生物)来说, 释放部位可以是胃、小肠(十二指肠、空肠或回肠)或大肠。本领域技术人员可获得在胃中不溶解而在十二指肠或肠道其它部位释放所述物质的制剂。优选地, 所述释放将通过保护蛋白质(或其衍生物)或通过将生物活性物质释放到胃环境以外如肠中, 避免胃环境的有害作用。

为了确保获得充分的胃耐受性, 不能渗透到至少 pH5.0 环境中的包衣是必需的。用作肠衣的更普通的惰性组分的实例是乙酸偏苯三酸纤维素(CAT)、邻苯二甲酸羟丙基甲基纤维素(HPMCP)、HPMCP 50、HPMCP55、聚乙酸邻苯二甲酸乙烯酯(PVAP)、Eudragit L30D、Aquateric、乙酸邻苯二甲酸纤维素(CAP)、Eudragit L、Eudragit S 和虫胶。这些包衣可以以混合膜的形式使用。

包衣或包衣混合物也可用于没有准备好预防被胃环境破坏的片剂。这可以包括糖衣或使片剂容易吞咽的包衣。胶囊剂可由用于递送干治疗剂, 即干粉的硬壳(如明胶)组成; 对于液体形式的治疗剂来说, 可以使用软明胶壳。扁囊剂的壳材料可以是稠淀粉或其它可食用纸。

对于丸剂、锭剂、模压片或片剂的磨碎物来说，可以使用湿法制粒技术。

肽治疗剂可以以颗粒或粒径大约为 1mm 的小药丸形式包含在由细小的多粒子组成的制剂中。通过胶囊给药的所述物质的制剂也可以是粉剂、轻轻压制的充填剂或者甚至可以是片剂。可通过压制法制备治疗剂。

也可以包含着色剂和调味剂。例如，可将蛋白质(或其衍生物)制备成制剂(如通过脂质体或微球包囊)，然后进一步包含在可食用产品，如含有着色剂和调味剂的冷藏饮料中。

可以用惰性材料稀释或增加治疗剂的体积。这些稀释剂可包括碳水化合物，特别是甘露糖醇、 $\alpha$ -乳糖、无水乳糖、纤维素、蔗糖、改性的葡聚糖和淀粉。某些无机盐也可以用作填充剂包括三磷酸钙、碳酸镁和氯化钠。一些可通过商业渠道获得的稀释剂为 Fast-Flo、Emdex、STA-Rx 1500、Emcompress 和 Avicell。

在固体剂量形式的治疗剂制剂中可包含崩解剂。用作崩解剂的材料包括但不限于淀粉，包括以淀粉为基础的商业用崩解剂 Explotab。也可以使用羟基乙酸淀粉钠、安伯来特、羧甲基纤维素钠、超支链淀粉、藻酸钠、明胶、桔子皮、酸性羧甲基纤维素、天然海绵状物和膨润土。另一种形式的崩解剂是不溶性阳离子交换树脂。粉状树胶可用作崩解剂和粘合剂并且这些可包括粉状树胶如琼脂、卡拉牙胶或西黄耆胶。藻酸及其钠盐也可以用作崩解剂。粘合剂可用于将治疗剂粘合在一起形成硬片剂并且包括来自天然产物如阿拉伯胶、西黄耆胶、淀粉和明胶的材料。其它的粘合剂包括甲基纤维素(MC)、乙基纤维素(EC)和羧甲基纤维素(CMC)。聚乙烯吡咯烷酮(PVP)和羟丙基甲基纤维素(HPMC)都可在乙醇溶液中用于将治疗剂制粒。

在治疗剂的制剂中可包含润滑剂用于预防在制备制剂过程中发生粘着。可在治疗剂和模具壁之间使用一层润滑剂，并且这些润滑剂包括但不限于硬脂酸包括其镁盐和钙盐、聚四氟乙烯(PTFE)、液体石蜡、植物油和蜡。也可以使用水溶液润滑剂如十二烷基硫酸钠、十二烷基硫酸镁、各种分子量的聚乙二醇, Carbowax 4000 和 6000。

可加入助流剂，它可以改善药物在制备制剂过程中的流动性并且有

助于压片过程中的重排。助流剂可包括淀粉、滑石粉、焦二氧化硅和水合硅铝酸盐。

为了有助于治疗剂溶解到水环境中，可加入表面活性剂作为润湿剂。表面活性剂可包括阴离子去污剂如十二烷基硫酸钠、琥珀酸二辛酯磺酸钠和二辛基磺酸钠。也可以使用阳离子去污剂并且可包括苯扎氯铵或苜索氯铵。可作为表面活性剂包含在制剂中的潜在的非离子去污剂包括 lauromacrogol 400、polyoxyl 40 stearate、聚氧乙烯氢化蓖麻油 10、50 和 60、单硬脂酸甘油酯，多乙氧基醚 40、60、65 和 80、蔗糖脂肪酸酯、甲基纤维素和羧甲基纤维素。这些表面活性剂可单独或以不同比例的混合物形式存在于蛋白质或其衍生物的制剂中。

例如潜在地增强多肽(或其衍生物)摄取的添加剂是脂肪酸如油酸、亚油酸和亚麻酸。

肺递送：本文也考虑通过肺递送本发明的多肽(或其衍生物)。当吸入时，所述多肽(或其衍生物)被递送到哺乳动物的肺中并且覆盖肺泡的粘膜表面。有关这方面的其它报道包括 Adjei 等人,1990,药物研究,7:565-569; Adjei 等人,1990, International Journal of Pharmaceutics, 63:135-144(醋酸亮丙瑞林); Braquet 等人,1989, Journal of Cardiovascular Pharmacology, 13(suppl. 5):143-146(内皮素-1); Hubbard 等人,1989, Annals of Internal Medicine, Vol. III, pp. 206-212(a1-抗胰蛋白酶); Smith 等人,1989, J.Clin. Invest. 84:1145-1146(a-1-蛋白酶); Oswein 等人,1990,“蛋白质的气雾化”, Proceedings of Symposium on Respiratory Drug Delivery II, Keystone, Colorado, March, (重组人生长激素); Debs 等人,1988, J. Immunol. 140:3482-3488(干扰素-g和肿瘤坏死因子- $\alpha$ )和 Platz 等人,美国专利 5,284,656(粒细胞集落刺激因子)。在 1995 年 9 月 19 日颁发给 Wong 等人的美国专利 5,451,569 中描述了用于通过肺递送药物的方法和组合物。

在本发明实施中考虑使用各种用于通过肺递送治疗剂产品的机械装置，包括但不限于雾化器、计量剂量的吸入器和粉剂吸入器，所有这些都是本领域技术人员所熟悉的。

典型地，适用于利用喷射式或超声波式喷雾器来应用的制剂将包含以每 ml 溶液大约 0.1-25mg 生物活性蛋白质的浓度溶解在水中的多肽

(或其衍生物)。所述制剂也可以包含缓冲剂和单糖(例如用于稳定蛋白质和调节渗透压)。所述喷雾器制剂也可以包含表面活性剂,用于减少或预防形成气雾剂时由溶液雾化引起的蛋白质在表面聚集。

利用计量剂量的吸入器装置来应用的制剂通常包含细碎的粉剂,其中所述粉剂包含在表面活性剂的帮助下悬浮在推进剂中的多肽(或其衍生物)。所述推进剂可以是用于该目的的任何常规材料如氯氟烃、氢氯氟烃、氢氟烃或烃,包括三氯氟甲烷、二氯二氟甲烷、二氯四氟乙醇和 1,1,1,2-四氟乙烷或其组合。适宜的表面活性剂包括脱水山梨糖醇三油酸酯和大豆卵磷脂。油酸也可以用作表面活性剂。

由粉末吸入器装置分配的制剂将包括含有多肽(或其衍生物)的细碎干粉并且也可包含促进粉末从装置中分散量,例如占制剂重量的 50-90% 的填充剂,如乳糖、山梨糖醇、蔗糖或甘露糖醇。最有利地,应该将蛋白质(或其衍生物)制备成平均粒径为 10 $\mu$ m(或微米),最优选 0.5-5 $\mu$ m 的微粒形式以便最有效地递送到远端肺中。

鼻递送:本文也考虑通过鼻或鼻咽递送多肽(或其衍生物)。通过鼻递送使得在将治疗剂产品给予到鼻后,多肽直接通过上呼吸道粘膜,不必使所述产品沉积在肺中。用于通过鼻递送的制剂包括与葡聚糖或环糊精形成的制剂。

下列实施例用于更充分地说明本发明的优选实施方案。然而,它们不以任何方式限制本发明范围。

### 实验细节部分

#### 实施例 1: 胆碱结合蛋白 A(CbpA)的肽截取物

产生包含 CbpA N-末端截取片段的多肽(血清型 4)。用 PCR 引物 SJ533 和 SJ537 扩增全长 CbpA,基于 CbpA 多肽衍生的 N-末端氨基酸序列来设计引物。5'正向引物 SJ533 = 5'GGC GGA TCC ATG GA(A, G) AA(C, T) GA(A, G) GG3'。这种简并引物是由氨基酸序列 XENEG 设计的,结合有 BamHI 和 NcoI 限制位点以及 ATG 起始密码子。3'反向引物 SJ537 = 5'GCC GTC GAC TTA GTT TAC CCA TTC ACC ATT GGC 3'。该引物结合有 SalI 限制位点以便克隆,并且结合有来自 CbpA 的天然终止密码子,该引物基于 4 型和 R6x 序列。

在 50℃ 退火温度下用高保真性酶(Boehringer Mannheim)以及用引物 SJ533 和 SJ537 扩增 30 次来由作为模板的基因组 DNA 产生 PCR 产物。所得 PCR 产物用 QIA 快速 PCR 纯化试剂盒(Qiagen, Inc.)纯化, 然后用 BamHI 和 Sall 限制酶消化并克隆到用 BamHI、XbaI 和 SmaI 限制酶消化的 pQE30 表达载体(Qiagen, Inc.)中。

#### 多肽 R2:

利用位于第二重复区(即, 图 1 所示的 C 区)末端的天然存在的 PvuII 位点(4 型序列的核酸 1228)来产生 CbpA 基因的截取形式, 仅仅含有该基因的 5' 部分。为了产生截取克隆, 用 PvuII 和 XbaI 消化全长克隆 PMI580(4 型)或 PMI581(R6x), 所得片段被连接到表达载体 pQE30 中并转化到合适的宿主中。表达和纯化该蛋白质。在该例中, 表达载体所利用的终止密码子位于插入物的下游, 所以被表达的蛋白由于在克隆位点的 5' 末端有附加的核酸而比插入物的预测尺寸要大。在 SEQ ID NO 1 中描述了多肽 R2 的氨基酸序列。

#### 多肽 R1:

使用类似的策略来表达 CbpA N-末端区域中的第一重复区, 即多肽 R1 的 A 区。这里利用了两个氨基重复区之间的天然存在的 XmnI 位点(4 型序列的核酸 856)。用 XmnI 和 AatII 消化 CbpA 全长克隆 PMI580。用 AatII 和 SmaI 消化载体 pQE30。再一次连接两个型号的片段, 转化到大肠杆菌中并筛选插入的克隆。选择一个阳性克隆并由该色斑纯化重组蛋白质。

使用大肠杆菌 pQE30 载体并用 Qia 表达系统(Qiagen)来表达和纯化所有多肽。使用抗组氨酸抗体和蛋白质特异性抗体, 通过宿主和 Western 分析检测 His 标记的蛋白质的氨基端。

#### R1 和 R2 的纯化:

为了诱导产生并纯化来自大肠杆菌的重组蛋白质, 单一集落选自含有重组质粒的平板培养的细菌并在 6.0ml 含有 50µg/ml 卡那霉素和 100µg/ml 氨苄青霉素的 LB 缓冲液中于 37℃ 温度下培养过夜。将该 6ml

培养物加到 1L 具有上述浓度下的抗生素的 LB 中。在 37℃ 温度下振摇培养物直到  $A_{600} = \sim 0.400$ 。在该 1L 培养物中加入 1M IPTG 到最终浓度为 1mM。然后在 37℃ 下振摇培养物 3-4 小时。在 J-6B 型离心机中以 4000rpm 转速旋转此 1L 培养物 15 分钟。弃去上清液并在 -20℃ 温度下贮存丸状沉淀。

将 1L 丸状沉淀重新悬浮在 25ml 的 50mM  $\text{NaH}_2\text{PO}_4$ 、10mM Tris、6M GuCl、300mM NaCl, pH8.0(缓冲液 A)中。将该混合物在室温下旋转 30 分钟并使用微型管尖以 50% Cuty Cycle 和 7 的输出功率在 VibraCell 超声处理器(Sonics and Materials, Inc., Danbury, CT)上超声两次共 30 秒。混合物在 JA20 旋转装置中以 10K 转速旋转 5 分钟并取出上清液。将上清液加载到与 GradiFrac 系统(Pharmacia Biotech, Upsala, 瑞典)连接的 10ml Talon(Clonetech, Palo Alto, CA)树脂柱上。将该柱用 100ml 缓冲液 A 平衡并用另外的 200ml 该缓冲液洗涤。以 100ml 的总体积使用 100% 50mM  $\text{NaH}_2\text{PO}_4$ 、8M 脲、20mM MES, pH6.0(缓冲液 B)作为最终靶缓冲液按照 pH 梯度来洗脱。以约 30% 缓冲液 B 洗脱蛋白质。收集洗脱峰并合并。

为了重折叠, 使用具有 1400 分子量截断值的透析管在室温下用 2L 体积的 PBS 透析约 3 小时。然后在 4℃ 温度下在 2L 的 PBS 中将样品透析过夜。使用 Centriprep-30 旋转柱通过将 PBS 加到旋转的余留物中并重新旋转来在蛋白质浓缩期间完成另外的缓冲液交换。使用 BCA 蛋白质测定方法确定蛋白质浓度并使用 Coomassie 染色的 4-20% SDS-PAGE 凝胶来显示纯度(图 3)。

## 实施例 2: 多肽 R1 和 R2 的凝集素活性

LNnt 是一种存在于真核细胞上的肺炎球菌受体的碳水化合物类似物。已经显示 CbpA 缺损的肺炎球菌突变体不能粘附到真核细胞或固定化糖上, 表明 CbpA 是粘着配体。CbpA 是一种模块蛋白, 它可被划分为两个区: N-末端功能区和 C-末端胆碱结合区(图 1)。分析多肽 R1 和 R2 的生物学活性以确定完整的 CbpA 的活性是否集中在唯一的 N 末端(由 R2 模拟的)或其片段(由 R1 模拟的)中。确定了在缺乏胆碱结合区(CBD)时仅有的 N-末端区 (R2) 是否包含凝集素结合生物活性。使



用全长 CbpA 和多肽 R2(缺失脯氨酸富含区中 PvuII 位点以外的 CBD 区的截取物)进行了试验。

测定中用已知的可被 CbpA 识别的糖缀合物: LNnT-白蛋白、3'唾液酸乳糖-白蛋白和阴性对照白蛋白包被组织培养孔。然后用白蛋白封闭培养板、洗涤并在 15 分钟内加入全长 CbpA 多肽 R2 或者多肽 R1(0.8 $\mu$ g/ml), 然后不用洗涤, 在 30 分钟内加入荧光素标记的 R6 肺炎球菌, 洗涤并目视记数粘附细菌。

在没有添加任何肽的情况下 R6 和碳水化合物的结合是阳性对照并且标定为 100% (表 1)。在三个独立的实验中, 全长 CbpA 或多肽 R2 竞争性地抑制肺炎球菌与 LNnT 包被的表面结合。与对照相比, 全长 CbpA 抑制率为 71%、64%和 63%; 多肽 R2 抑制率为 65%、53%和 74%。CbpA 和 R2 的等同活性表明胆碱结合区对于 CbpA 的 LNnT 凝集素活性来说不是必需的, 并且 R2 是候选的 LNnT 凝集素。

相对于与 LNnT 的结合来说, 与全长 CbpA(74 和 66%)相比, R2 不抑制肺炎球菌与 3'唾液酸乳糖的结合(79%和 101%)。这表明 CBD 缺失时丧失了唾液酸的识别活性。相比之下, R1 似乎具有识别唾液酸的活性, 享有 CbpA 的特性但在 R2 中明显掩盖了这种特性。这表明多肽折叠成功能区受该多肽组成和长度的影响。发现在其它菌株中序列有轻微的变化(见图 2)。假定 R1 和 R2 之间具有高度的序列同源性, 进一步的可能是对于凝集素活性同时需要 R1 和 R2 或者 R1 和 R2 是两种特异性稍微不同的凝集素(±唾液酸)。

表 1

CbpA 的可溶性形式对 R6 肺炎球菌与纯化糖缀合物的结合的抑制作用

Cbp 形式	LNnT		3'唾液酸乳糖	
	每单层的肺炎球菌数(SD)	% 对照	每单层的肺炎球菌数	% 对照(每孔)
无肽	3282 2421(489) 2210(350)	100%	2611 2115(125)	100%
全长 CbpA	2075 1740(167) 1415(50)	63, 71, 64	1933 1405(240)	74 66
多肽 R2	2461 1288(672) 1440(530)	74, 53, 65	2639 1670(420)	101 79
多肽 R1	3002 2245(182) 2500(310)	91, 92, 112	1052 1445(526)	40 68

N=LNnt 进行 3 次试验, 每次 3 个孔

N=唾液酸乳糖进行 2 次试验, 每次 3 个孔

## 凝集素活性与细胞结合活性的相关性

含有碳水化合物(糖蛋白和糖脂)和细菌的携带人细胞的表面分子通过碳水化合物与这些糖缀合物结合, 尽管蛋白质或脂的骨架非常不同。因此, 在体外具有凝集素活性的携带细菌的多肽可以粘附到人细胞表面。体外凝集素活性和细胞结合作用之间的这种直接的关联对于肺炎球菌来说是已知的。例如, LNnt 竞争性抑制肺炎球菌与 TNF 激活的 A549 人肺细胞的结合并阻止肺炎球菌在体内的发展。为了证实 CbpA 截取物的凝集素活性可反映细胞结合活性, 试验了 CbpA 和截取物对肺炎球菌与肺细胞结合的抑制作用(表 2)。全长 CbpA 和多肽 R2 竞争性抑制肺炎球菌与肺细胞的粘附, 与对照相比, 其抑制率分别为 58% 和 63%。多肽 R1 无效, 表明 R2 的 LNnt 结合活性对于肺炎球菌

与肺细胞的结合来说是必需的并且也解释了这种关系。

表 2

R6 肺炎球菌与 TNF 激活的人肺细胞的结合

Cbp 形式	A549 肺	
	每单层的肺炎球菌数(平均值)	% 对照
无肽	697, 704, 674 702, 722 (700)	100%
全长 CbpA	376, 431 (403)	58%
多肽 R2	517, 693 314, 342, 350 (443)	63%
多肽 R1	696, 642, 552 (630)	90%

N=每个试验进行 2 次, 每次 2 或 3 个孔

#### LNnT 凝集素活性取决于 R2

CbpA 的 N-末端区包含各自有大约 110 个氨基酸的两个重复区(见图 1, 多肽 R2 中的 A 区和 C 区)。为了研究两个区对生物活性 R1 的相对贡献, 将仅仅含有 A 区的与 R2 和全长 CbpA 进行比较。在粘附测定中试验时, 多肽 R1 完全未抑制与 LNnT 的粘附(野生型的 91、92 和 112%)。然而, 多肽 R1 显示出对与唾液酸乳糖的结合有某些抑制作用(对照的 68 和 40%)。这表明 LNnT 的凝集素活性需要多肽 R2 并且 R2 是候选的 LNnT 凝集素区。相比之下, R1 似乎具有识别唾液酸的活性。

#### CbpA N-末端区的抗体阻断细胞结合:

假定 CbpA 的 N-末端区结合细胞, 对 N-末端区活性的干扰会阻止或逆转细菌与细胞和纯化的糖缀合物的结合。一种这样的干扰机制是抗体。

表 3

## 抗 CbpA R2 抗体对 R6 肺炎球菌与 LNnt 包被表面结合的抑制作用

	每单层的肺炎球菌数(SD)	% 对照(平均值)
免疫前抗体	198(64); 88(4)	100%
R2 截取物的抗体	56(11); 9(2)	28%; 10%

5 $\mu$ l 未稀释的兔抗体 + 5 $\mu$ l  $2 \times 10^7$  R6x 预培养物, 室温下  $6 \times 30$  分钟, 然后加到 LNnT 包被孔中进行粘附测定。显示了两次独立的实验。

试验了产生于 CbpA 重组 N-末端区(R2) 的抗血清阻断肺炎球菌粘附 LNnT 的能力。在室温下培养兔多克隆抗 CbpA 抗血清(5 $\mu$ l)和 5 $\mu$ l  $2 \times 10^7$  标记的细菌 30 分钟。将混合物在固定化 LNnT 上涂敷 30 分钟, 然后用 PBS 洗涤 3 次以除去未结合的细菌。用显微镜数与培养板结合的细菌并以 6 个孔的平均值加标准偏差显示结果。表 3 所示结果表明针对 R2 多肽所产生的抗血清阻断了肺炎球菌与 LNnT 的结合。图 5 显示免疫前与抗 CbpA R2 抗体抑制肺炎球菌 R6x 与模型受体 LNnT 结合的滴定曲线。1:100 和 1:200 稀释度的抗 R2 对肺炎球菌粘附的阻断率超过 70%。进一步稀释到 1:400 消除了表明该作用特异性的活性。

用于制备表 3 和图 5 所示抗 CbpA 抗血清的 CbpA 是针对血清型 4 的 CbpA 产生的。在粘附抑制测定中使用的肺炎球菌 R6x 菌株是由血清型 2 衍生的。抗体阻断异种血清型细菌粘附的能力表明血清型之间的交叉保护活性。这种活性对于有效的疫苗免疫原来说是非常需要的。

## CbpA N-末端天然构象抗体的活性:

按 Rosenow 等人所述方法, 可由其天然宿主(肺炎球菌)在胆碱亲和柱上纯化 CbpA。另外, 多聚组氨酸标记可以连接到基因的末端以便通过几个组氨酸残基延伸转录的蛋白。这些残基便于在镍亲和基质上纯化。与较短的截取物相反, 全长多肽的纯化有利于保留天然三级结构。通过这些生化方法特别是由肺炎球菌也可由大肠杆菌或其它宿主细菌纯化的 CbpA 保留其天然三级结构。用作免疫原, 天然折叠的 CbpA 产生抗体, 它们潜在地不同于通过用可发生不同的折叠的截取物

免疫产生的抗体。类似地，用作治疗剂的 CbpA 可具有不同于截取物的三级结构，所述截取物会改善其阻断粘附的能力。基于这些考虑，作为允许其折叠成它的天然三级结构然后通过生化方法切掉 C 末端 (CBD) 生产 CbpA 可能是有利的全长蛋白质的。例如，用羟胺处理将在血清型 R6x 和血清型 4 胆碱结合蛋白 A 的氨基酸位点 475 切割 CbpA，分离 N 和 C 末端。N 末端片段适于用作治疗剂或免疫原。

另外，天然 CbpA 可用作免疫原和活性结构的抗血清。通过吸附除去 BD 抗体来在该混合物中富含生物活性的抗 N 末端抗体。通过在 R1 下与  $1 \times 10^8$  CbpA 缺陷性细菌一起培养 200 $\mu$ l 血清 1 小时来制备这种抗体。在该突变体上的其它胆碱结合蛋白吸附抗 CBD 抗体，然后通过离心和除去细菌来从抗血清中移出所述抗体。

为了证明吸附的抗 CbpA 抗体的生物活性，确定吸附的抗血清阻断肺炎球菌粘附到模型受体 LNnT 上的能力。与 1:600 稀释的抗血清一起培养 R6x 肺炎球菌然后加到用 LNnT 白蛋白包被的孔中。

表 4

吸附的抗 CbpA 抗血清阻断粘附

抗血清(1:600)	每孔肺炎球菌的数目 $\pm$ SD(对照%)
无抗体	563 $\pm$ 11(100%)
免疫前抗血清	479 $\pm$ 11(85%)
抗 CbpA 抗血清	294 $\pm$ 72(52%)
经吸附除去 CBD 抗体的 CbpA 抗血清	175 $\pm$ 38(31%)

这些结果表明其天然构象的 CbpA N 末端区域的抗体强烈阻断粘附。该活性大于图 5 中截取物的活性，截取物在 1:600 稀释度时没有活性。图 5 中的滴定研究进一步表明吸附的抗 CbpA 抗血清的这种活性。通过三角形显示肺炎球菌 4 型与 LNnT 包被孔基线粘附的。用各种稀释度的未吸附(正方形)或吸附(菱形)的抗血清预先培养肺炎球菌表明使粘附减少。两种抗血清在粘附上类似减少的事实表明大多数 N-末端 CbpA 残基的抗体(即，通过吸附除去胆碱结合区的抗体)的阻断活

性不降低生物活性。

### 实施例 3: 抗 R2 抗血清的被动保护

兔免疫血清的产生:

在 Covance(Denver, PA)产生了抗多肽 R2(CbpA 截取物)和 CbpA 的兔免疫血清。收集免疫前血清后,用 250 $\mu$ g 在完全弗氏佐剂中的含有氨基末端重复区(上文制备 483:58)的 R2 使新西兰白兔免疫。在第 21 天给该兔 125 $\mu$ g 在不完全弗氏佐剂中的 R2 进行加强并在第 31 天抽血。第二只兔子用纯化的 CbpA 进行类似的免疫。

### 小鼠的被动保护

腹膜内注射 100 $\mu$ l 1:2 稀释的在无菌 PBS 中的兔抗 R2 或免疫前血清(免疫前和第 31 天的免疫血清)使 C3H/HeJ 小鼠(5 只/组)被动免疫。给予血清 1 小时后,用 1600CFU 肺炎链球菌血清型 6B(菌株 SP317)攻击小鼠。监测小鼠 14 天的存活情况。用针对多肽 R2 产生的兔免疫血清免疫的小鼠在攻击后有 80% 存活(图 4)。所有用免疫前兔血清免疫的小鼠到第 7 天都死亡。

该数据表明特异于 CbpA 的抗体对系统肺炎球菌感染具有保护作用。该数据进一步表明胆碱结合区对于保护来说不是必需的,因为特异于截取的蛋白质多肽 R2 的抗体(缺失保守的胆碱结合重复区)对于保护来说是足够的。另外,针对 CbpA 血清型 4 的血清对血清型 6B 的攻击有保护作用。

### 实施例 4: 抗 R1 抗血清的自动保护作用

腹膜内注射 CbpA 截取蛋白质 R1(15 $\mu$ g 在 50 $\mu$ l PBS 中,加 50 $\mu$ l 完全弗氏佐剂)使 C3H/HeJ 小鼠(10 只/组)免疫。一组 10 只假免疫的小鼠接受 PBS 和佐剂。第二次免疫在 4 周后给予,腹膜内注射 15 $\mu$ g 有不完全弗氏佐剂的蛋白质(假免疫的接受 PBS 和 IFA)。在第 3、6 和 9 周抽取血液(后眼眶取血)进行免疫反应分析。在 9 周时从 10 只 CbpA 免疫的小鼠中收集的血清的 ELISA 终点抗 CbpA 截取物滴度为 4,096,000。在假免疫小鼠的血清中未检测到抗体。在第 10 周用 560CFU

肺炎链球菌血清型 6B(菌株 SPSJ2p, 由 P. Flynn, St. Jude Children's Research Hospital, Memphis, TN 提供)攻击小鼠。在 14 天内检测小鼠的存活情况。用 CbpA 截取蛋白质 R1 免疫的小鼠在攻击后有 80% 存活。所有假免疫的小鼠在第 8 天死亡(图 7)。

该数据表明用 CbpA 重组片段免疫诱导产生能够抵抗系统性肺炎链球菌感染和死亡的特异性抗体。该数据进一步表明胆碱结合区对于保护作用不是必需的, 因为免疫原是截取的蛋白质 R1。另外, 结果暗示一个氨基端重复区可能足以诱导保护反应。也表明了交叉保护作用, 因为基于血清型 4 DNA 序列产生了重组肺炎链球菌蛋白质, 用血清型 6B 分离物攻击后观察到了保护作用。

#### 实施例 5 预防幼年大鼠的鼻咽菌落集群

在体外, CbpA 的 N 末端区竞争性抑制肺炎链球菌附着。为了证明具有该活性的肽的治疗实用性, 给予幼年大鼠截取的肽, 然后用肺炎链球菌攻击并评价鼻咽的菌落集群情况。

用 10 $\mu$ l 含有 0.8 $\mu$ g 多肽 R2 或 R1 或不含蛋白质的 PBS 经鼻腔处理大鼠。15 分钟后, 经鼻腔引入 3 型肺炎链球菌(菌株 SIII)(10 $\mu$ l 含有  $1 \times 10^5$  cfer)。为了确定多肽竞争性抑制肺炎链球菌粘附和移生的能力, 在 72 小时时洗涤鼻腔并且确定每个动物中回收的肺炎链球菌的数目, 每组 4 只动物。仅仅接受 SIII 的大鼠每 10 $\mu$ l 有 2200、6500、6900 和 8700(平均 6075)个菌落。用截取物 R2 处理的动物显示最大的减少(3600、3500、2500、2100), 平均 10 $\mu$ l 有 2925 个细菌(对照的 48%)。用截取物 R1 处理的动物也显示菌落集群减少(5000、4800、3500、1600), 平均为 3725(对照的 61%)。

该实验表明将本发明的肽给予在治疗研究中指定的动物可使该动物抵抗随后的肺炎链球菌攻击。

#### 讨论:

通过实验表明, 多肽 R2: 1)作为疫苗抗原给药时诱导产生保护性抗体并且是一种用于疫苗制剂的优选组合物; 和 2)作为肽递送到呼吸道和/或鼻咽受侵袭体中时, 竞争性地预防肺炎链球菌附着并且是一种用于

对抗移生作用或侵袭性疾病防治剂的优选组合物。而且，CbpA 截取物起凝集素作用而没有 CBD。识别两种碳水化合物：LNnT，由含有图 1 中两个 N-末端重复区(A 和 C)的肽识别和唾液酸，由仅仅含有单个 N-最末端重复区(A)的肽识别。含有 N-末端重复区多肽 R1 和 R2 的截取物在细胞培养测定中也表明有凝集素活性。

多肽 R2 活性的重要特征包括：1)多肽 R2 和全长 CbpA 在识别纯化糖缀合物受体类似物、肺细胞和动物模型中有完全相关的生物活性。对于它们的抗体也证明了相关性；和 2)4 型衍生剂和在外测定中使用其它血清型(如，6B 和 2)的细菌之间具有交叉保护作用，这对于可以使用的疫苗、预防和治疗方式来说是重要的。

尽管本文通过提及各种具体的材料、方法和实施例描述和说明了本发明，但可以理解本发明不局限于材料的特定组合和为此目的所选择的方法。本领域技术人员会意识到可以暗含某些细节的各种变化。同样地，在本文中引用的任何参考文献从与本发明公开内容的相关程度来看均被考虑引入本文供参考。



## 序列表

<110> Tuomanen, Elaine I.

Wizemann, Theresa

Masure, H. R.

Johnson, Leslie S.

Koenig, Scott

<120> **包含胆碱结合蛋白 A N-末端截取物的氨基酸的多肽，  
由该多肽衍生的疫苗及其应用**

<130> 1340-1-017 msc

<140> 09/056,019

<141> 1998-04-07

<160> 39

<170> PatentIn Ver. 2.0

<210> 1

<211> 406

<212> PRT

<213> **肺炎链球菌**

<400> 1

Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn

1

5

10

15

Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu

20 25 30

Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val

35 40 45

Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val

50 55 60

Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys

65 70 75 80

Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu

85 90 95

Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser

100 105 110

Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp

115 120 125

Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu

130 135 140

Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys

145 150 155 160

Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu

	165	-170	175
Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu			
	180	185	190
Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys			
	195	200	205
Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu			
	210	215	220
Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg			
	225	230	235
Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg			
	245	250	255
Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala			
	260	265	270
Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser			
	275	280	285
Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu			
	290	295	300
Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr			
	305	310	315
			320

Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp  
 325 330 335

Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys  
 340 345 350

Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu  
 355 360 365

Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg  
 370 375 380

Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys  
 385 390 395 400

Val Lys Glu Lys Pro Ala  
 405

<210> 2

<211> 655

<212> PRT

<213> 肺炎链球菌

<400> 2

Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn  
 1 5 10 15

Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu

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

Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu  
 180 185 190

Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys  
 195 200 205

Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu  
 210 215 220

Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg  
 225 230 235 240

Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg  
 245 250 255

Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala  
 260 265 270

Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser  
 275 280 285

Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu  
 290 295 300

Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr  
 305 310 315 320

Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp  
 325 330 335

Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys  
 340 345 350

Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu  
 355 360 365

Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg  
 370 375 380

Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys  
 385 390 395 400

Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Lys  
 405 410 415

Ala Glu Lys Pro Ala Pro Ala Pro Lys Pro Glu Asn Pro Ala Glu Gln  
 420 425 430

Pro Lys Ala Glu Lys Pro Ala Asp Gln Gln Ala Glu Glu Asp Tyr Ala  
 435 440 445

Arg Arg Ser Glu Glu Glu Tyr Asn Arg Leu Thr Gln Gln Gln Pro Pro  
 450 455 460

Lys Thr Glu Lys Pro Ala Gln Pro Ser Thr Pro Lys Thr Gly Trp Lys  
 465 470 475 480

Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp Gly Ser Met Ala

	485	490	495
Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn			
	500	505	510
Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr			
	515	520	525
Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn Gly			
	530	535	540
Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu			
	545	550	555
Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala			
	565	570	575
Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn			
	580	585	590
Gly Asp Met Ala Thr Gly Trp Val Lys Asp Gly Asp Thr Trp Tyr Tyr			
	595	600	605
Leu Glu Ala Ser Gly Ala Met Lys Ala Ser Gln Trp Phe Lys Val Ser			
	610	615	620
Asp Lys Trp Tyr Tyr Val Asn Gly Ser Gly Ala Leu Ala Val Asn Thr			
	625	630	635
			640



Thr Val Asp Gly Tyr Gly Val Asn Ala Asn Gly Glu Trp Val Asn  
 645 650 655

<210> 3

<211> 284

<212> PRT

<213> 肺炎链球菌

<400> 3

Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn  
 1 5 10 15

Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu  
 20 25 30

Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val  
 35 40 45

Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val  
 50 55 60

Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys  
 65 70 75 80

Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu  
 85 90 95

Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser

	100		105		110
Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp					
	115		120		125
Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu					
	130		135		140
Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys					
	145		150		155
					160
Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu					
		165		170	175
Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu					
	180		185		190
Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys					
	195		200		205
Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu					
	210		215		220
Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg					
	225		230		235
					240
Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg					
		245		250	255

Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala  
 260 265 270

Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu  
 275 280

<210> 4

<211> 106

<212> PRT

<213> 肺炎链球菌

<400> 4

Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala  
 1 5 10 15

Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro  
 20 25 30

Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val  
 35 40 45

Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu  
 50 55 60

Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser  
 65 70 75 80

Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys

85 90 95  
 Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala  
 100 105  
 <210> 5  
 <211> 109  
 <212> PRT  
 <213> 肺炎链球菌  
 <400> 5  
 Thr Glu Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu  
 1 5 10 15  
 Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr  
 20 25 30  
 Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp  
 35 40 45  
 Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Val Lys Ala Asn  
 50 55 60  
 Glu Pro Arg Asp Glu Gln Lys Ile Lys Gln Ala Glu Ala Glu Val Glu  
 65 70 75 80  
 Ser Lys Gln Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp Arg  
 85 90 95

Glu Glu Ala Glu Glu Glu Ala Lys Arg Arg Ala Asp Ala  
 100 105

<210> 6

<211> 4

<212> PRT

<213> 肺炎链球菌

<400> 6

Lys Xaa Xaa Glu

1

<210> 7

<211> 376

<212> PRT

<213> 肺炎链球菌

<400> 7

Glu Asn Glu Gly Ser Thr Gln Ala Ala Thr Ser Ser Asn Met Ala Lys  
 1 5 10 15

Thr Glu His Arg Lys Ala Ala Lys Gln Val Val Asp Glu Tyr Ile Glu  
 20 25 30

Lys Met Leu Arg Glu Ile Gln Leu Asp Arg Arg Lys His Thr Gln Asn  
 35 40 45

Val Ala Leu Asn Ile Lys Leu Ser Ala Ile Lys Thr Lys Tyr Leu Arg  
 50 55 60

Glu Leu Asn Val Leu Glu Glu Lys Ser Lys Asp Glu Leu Pro Ser Glu  
 65 70 75 80

Ile Lys Ala Lys Leu Asp Ala Ala Phe Glu Lys Phe Lys Lys Asp Thr  
 85 90 95

Leu Lys Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu  
 100 105 110

Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr  
 115 120 125

Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Phe Asp  
 130 135 140

Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys  
 145 150 155 160

Glu Ser Arg Asn Glu Gly Thr Ile Lys Gln Ala Lys Glu Lys Val Glu  
 165 170 175

Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg  
 180 185 190

Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Asp Ala Lys Leu Lys

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

Asp Arg Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu  
 355 360 365

Asp Lys Val Lys Glu Lys Pro Ala  
 370 375

<210> 8

<211> 663

<212> PRT

<213> 肺炎链球菌

<400> 8

Glu Asn Glu Gly Ser Thr Gln Ala Ala Thr Ser Ser Asn Met Ala Lys  
 1 5 10 15

Thr Glu His Arg Lys Ala Ala Lys Gln Val Val Asp Glu Tyr Ile Glu  
 20 25 30

Lys Met Leu Arg Glu Ile Gln Leu Asp Arg Arg Lys His Thr Gln Asn  
 35 40 45

Val Ala Leu Asn Ile Lys Leu Ser Ala Ile Lys Thr Lys Tyr Leu Arg  
 50 55 60

Glu Leu Asn Val Leu Glu Glu Lys Ser Lys Asp Glu Leu Pro Ser Glu  
 65 70 75 80

Ile Lys Ala Lys Leu Asp Ala Ala Phe Glu Lys Phe Lys Lys Asp Thr



	85	90	95
Leu Lys Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu			
	100	105	110
Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr			
	115	120	125
Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Phe Asp			
	130	135	140
Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys			
	145	150	155
			160
Glu Ser Arg Asn Glu Gly Thr Ile Lys Gln Ala Lys Glu Lys Val Glu			
	165	170	175
Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg			
	180	185	190
Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Asp Ala Lys Leu Lys			
	195	200	205
Glu Ala Asn Val Ala Thr Ser Asp Gln Gly Lys Pro Lys Gly Arg Ala			
	210	215	220
Lys Arg Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn			
	225	230	235
			240

Asp Ala Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser  
 245 250 255

Ser Ser Leu Lys Ser Gly Lys Lys Val Ala Glu Ala Glu Lys Lys Val  
 260 265 270

Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg  
 275 280 285

Asn Tyr Pro Thr Asn Thr Tyr Lys Thr Leu Asp Leu Glu Ile Ala Glu  
 290 295 300

Ser Asp Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu  
 305 310 315 320

Ala Lys Glu Pro Arg Asp Glu Glu Lys Ile Lys Gln Ala Lys Ala Lys  
 325 330 335

Val Glu Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr  
 340 345 350

Asp Arg Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu  
 355 360 365

Asp Lys Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala  
 370 375 380

Thr Gln Pro Glu Lys Pro Ala Pro Lys Pro Glu Lys Pro Ala Glu Gln  
 385 390 395 400

Pro Lys Ala Glu Lys Thr Asp Asp Gln Gln Ala Glu Glu Asp Tyr Ala  
 405 410 415

Arg Arg Ser Glu Glu Glu Tyr Asn Arg Leu Thr Gln Gln Gln Pro Pro  
 420 425 430

Lys Thr Glu Lys Pro Ala Gln Pro Ser Thr Pro Lys Thr Gly Trp Lys  
 435 440 445

Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp Gly Ser Met Ala  
 450 455 460

Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn  
 465 470 475 480

Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr  
 485 490 495

Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn Gly  
 500 505 510

Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ala Met Ala Thr Gly Trp Leu  
 515 520 525

Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn Gly Ala Met Ala  
 530 535 540

Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn





---

145                            150                            155                            160

Glu Ser Arg Asn Glu Gly Thr Ile Lys Gln Ala Lys Glu Lys Val Glu  
                                  165                            170                            175

Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg  
                                  180                            185                            190

Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Asp Ala Lys Leu Lys  
                                  195                            200                            205

Glu Ala Asn Val Ala Thr Ser Asp Gln Gly Lys Pro Lys Gly Arg Ala  
                                  210                            215                            220

Lys Arg Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn  
225                            230                            235                            240

Asp Ala Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu  
                                  245                            250

<210> 10

<211> 106

<212> PRT

<213> 肺炎链球菌

<400> 10

Lys Ser Gly Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala  
1                            5                            10                            15

Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro  
 20 25 30

Thr Asn Thr Tyr Lys Thr Leu Asp Leu Glu Ile Ala Glu Ser Asp Val  
 35 40 45

Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu  
 50 55 60

Pro Arg Asp Glu Glu Lys Ile Lys Gln Ala Lys Ala Lys Val Glu Ser  
 65 70 75 80

Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg Lys  
 85 90 95

Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala  
 100 105

<210> 11

<211> 107

<212> PRT

<213> 肺炎链球菌

<400> 11

Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu Ala Lys  
 1 5 10 15

Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro Thr  
 20 25 30

Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Phe Asp Val Lys  
 35 40 45

Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu Ser  
 50 55 60

Arg Asn Glu Gly Thr Ile Lys Gln Ala Lys Glu Lys Val Glu Ser Lys  
 65 70 75 80

Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg Lys Lys  
 85 90 95

Ala Glu Glu Glu Ala Lys Arg Lys Ala Asp Ala  
 100 105

<210> 12

<211> 1219

<212> DNA

<213> 肺炎链球菌

<400> 12

gagaacgagg gagctacca agtaccact tcttctaata gggcaaatga aagtcaggca 60  
 gaacaaggag aacaacctaa aaaactcgat tcagaacgag ataaggcaag gaaagaggtc 120  
 gaggaatatg taataaaaaat agtgggtgag agctatgcaa aatcaactaa aaagcgacat 180  
 acaattactg tagctctagt taacgagttg aacaacatta agaacgagta tttgaataaa 240



```

atagttgaat caacctcaga aagccaacta cagatactga tgatggagag tcgatcaaaa 300
gtagatgaag ctgtgtctaa gtttgaaaag gactcatctt cttcgtcaag ttcagactct 360
tccactaac cggaagcttc agatacagcg aagccaaaca agccgacaga accaggagaa 420
aaggtagcag aagctaagaa gaaggttgaa gaagctgaga aaaaagccaa ggatcaaaaa 480
gaagaagatc gtcgtaacta cccaaccatt acttacaaaa cgcttgaact tgaaattgct 540
gagtccgatg tggaagttaa aaaagcggag cttgaaactag taaaagtgaa agctaacgaa 600
cctcgagacg agcaaaaaat taagcaagca gaagcgggaag ttgagagtaa acaagctgag 660
gctacaaggt taaaaaaaaat caagacagat cgtgaagaag cagaagaaga agctaaacga 720
agagcagatg ctaaagagca aggtaaacca aaggggcggg caaacgagg agttcctgga 780
gagctagcaa cacctgataa aaaagaaaat gatgcgaagt cttcagattc tagcgtaggt 840
gaagaaactc ttccaagccc atccctgaaa ccagaaaaaa aggtagcaga agctgagaag 900
aaggttgaag aagctaagaa aaaagccgag gatcaaaaag aagaagatcg ccgtaactac 960
ccaaccaata cttacaaaac gcttgaactt gaaattgctg agtccgatgt ggaagttaaa 1020
aaagcggagc ttgaaactagt aaaagaggaa gctaaggaaac ctcgaaacga ggaaaaagtt 1080
aagcaagcaa aagcgggaagt tgagagtaaa aaagctgagg ctacaagggt agaaaaaatc 1140
aagacagatc gtaaaaaagc agaagaagaa gctaaacgaa aagcagcaga agaagataaa 1200
gttaaagaaa aaccagctg                                     1219

```

<210> 13

<211> 1969

<212> DNA

<213> 肺炎链球菌

<400> 13

```

gagaacgagg gagctacca agtaccact tcttctaata gggcaaatga aagtcaggca 60
gaacaaggag aacaacctaa aaaactcgat tcagaacgag ataaggcaag gaaagaggtc 120
gaggaatatg taaaaaaaaat agtgggtgag agctatgcaa aatcaactaa aaagcgacat 180
acaattactg tagctctagt taacgagttg aacaacatta agaacgagta tttgaataaa 240
atagttgaat caacctcaga aagccaacta cagatactga tgatggagag tcgatcaaaa 300

```

gtagatgaag ctgtgtctaa gtttgaaaag gactcatctt cttcgtcaag ttcagactct 360  
 tccactaaac cggaagcttc agatacagcg aagccaaaca agccgacaga accaggagaa 420  
 aaggtagcag aagctaagaa gaaggttgaa gaagctgaga aaaaagccaa ggatcaaaaa 480  
 gaagaagatc gtcgtaacta cccaaccatt acttacaanaa cgcttgaact tgaaattgct 540  
 gagtccgatg tggaagttaa aaaagcggag cttgaactag taaaagttaa agctaacgaa 600  
 cctcgagacg agcaaaaaat taagcaagca gaagcgggaag ttgagagtaa acaagctgag 660  
 gctacaaggt taaaaaaat caagacagat cgtgaagaag cagaagaaga agctaaacga 720  
 agagcagatg ctaaagagca aggtaaacca aaggggaggg caaacgagg agttcctgga 780  
 gagctagcaa cacctgataa aaaagaaaat gatgcgaagt cttcagattc tagcgtaggt 840  
 gaagaaactc ttccaagccc atccctgaaa ccagaaaaaa aggtagcaga agctgagaag 900  
 aaggttgaag aagctaagaa aaaagccgag gatcaaaaag aagaagatcg ccgtaactac 960  
 ccaaccaata cttacaaaac gcttgaactt gaaattgctg agtccgatgt ggaagttaa 1020  
 aaagcggagg cttgaactag taaaagagga agctaaggaa cctcgaacg aggaaaaagt 1080  
 taagcaagca aaagcgggaag ttgagagtaa aaaagctgag gctacaaggt tagaaaaat 1140  
 caagacagat cgtaaaaaag cagaagaaga agctaaacga aaagcagcag aagaagataa 1200  
 agttaagaa aaaccagctg aacaaccaca accagcgccg gctccaaaag cagaaaaacc 1260  
 agtccagct ccaaaaccag agaatccagc tgaacaacca aaagcagaaa aaccagctga 1320  
 tcaacaagct gaagaagact atgctcgtag atcagaaga gaatataatc gcttgactca 1380  
 acagcaaccg ccaaaaactg aaaaaccagc acaaccatct actccaaaaa caggctggaa 1440  
 acaagaaaac ggatgtggt acttctacaa tactgatggt tcaatggcga caggatggct 1500  
 ccaaaacaat ggctcatggt actacctcaa cagcaatggc gctatggcga caggatggct 1560  
 ccaaaacaat ggctcatggt actatctaaa cgctaaggt tcaatggcaa caggatggct 1620  
 ccaaaacaat ggctcatggt actacctaaa cgctaaggt tcaatggcga caggatggct 1680  
 ccaatacaat ggctcatggt actacctaaa cgctaaggt tcaatggcga caggatggct 1740  
 ccaatacaat ggctcatggt actacctaaa cgctaaggt gatatggcga cagggtgggt 1800  
 gaaagatgga gatacctggt actatcttga agcatcaggt gctatgaaag caagccaatg 1860  
 gttcaaagta tcagataaat ggtactatgt caatggctca ggtgcccttg cagtcaacac 1920  
 aactgtagat ggctatggag tcaatgcaa tggatgaatgg gtaactaa 1969

<210> 14

<211> 853

<212> DNA

<213> 肺炎链球菌

<400> 14

```

gagaacgagg gagctacca agtaccact tcttctaata gggcaaatga aagtcaggca 60
gaacaaggag aacaacctaa aaaactcgat tcagaacgag ataaggcaag gaaagaggtc 120
gaggaatatg taaaaaaaaat agtgggtgag agctatgcaa aatcaactaa aaagcgacat 180
acaattactg tagctctagt taacgagttg aacaacatta agaacgagta tttgaataaa 240
atagttgaat caacctcaga aagccaacta cagatactga tgatggagag tcgatcaaaa 300
gtagatgaag ctgtgtctaa gtttgaaaag gactcatott cttegtcaag ttcagactct 360
tccactaac cggaagcttc agatacagcg aagccaaaca agccgacaga accaggagaa 420
aaggtagcag aagctaagaa gaaggttgaa gaagctgaga aaaaagccaa ggatcaaaaa 480
gaagaagatc gtcgtaacta cccaaccatt acttacaana cgcttgaact tgaatttgct 540
gagtcagatg tggaagttaa aaaagcggag cttgaactag taaaagttaa agctaacgaa 600
cctcgagacg agcaaaaaat taagcaagca gaagcgggag ttgagagtaa acaagctgag 660
gctacaaggt taaaaaaaaat caagacagat cgtgaagaag cagaagaaga agctaaacga 720
agagcagatg ctaaagagca aggtaaacca aaggggcggg caaaacgagg agttcctgga 780
gagctagcaa cacctgataa aaaagaaaat gatgcgaagt cttcagattc tagcgtaggt 840
gaagaaactc ttc

```

853

<210> 15

<211> 318

<212> DNA

<213> 肺炎链球菌

<400> 15

```

aaaccagaaa aaaaggtagc agaagctgag aagaaggttg aagaagctaa gaaaaagcc 60

```

gaggatcaaa aagaagaaga tcgccgtaac tacccaacca atacttacia aacgcttgaa 120  
 cttgaaattg ctgagtccga tgtggaagtt aaaaaagcgg agcttgaact agtaaaagag 180  
 gaagctaagg aacctcgaaa cgaggaaaaa gttaagcaag caaaagcgga agttgagagt 240  
 aaaaaagctg aggctacaag gttagaaaaa atcaagacag atcgtaaaaa agcagaagaa 300  
 gaagctaaac gaaaagca 318

<210> 16

<211> 327

<212> DNA

<213> 肺炎链球菌

<400> 16

acagaaccag gagaaaaggt agcagaagct aagaagaagg ttgaagaagc tgagaaaaaa 60  
 gccaaaggatc aaaaagaaga agatcgtcgt aactaccaa ccattactta caaaacgctt 120  
 gaacttgaaa ttgctgagtc cgatgtggaa gttaaaaaag cggagcttga actagtaaaa 180  
 gtgaaagcta acgaacctcg agacgagcaa aaaattaagc aagcagaagc ggaagttgag 240  
 agtaacaag ctgaggctac aaggttaaaa aaaatcaaga cagatcgtga agaagcagaa 300  
 gaagaagcta aacgaagagc agatgct 327

<210> 17

<211> 1129

<212> DNA

<213> 肺炎链球菌

<400> 17

gaaaacgaag gaagtaccca agcagccact tcttctaata tggcaaagac agaacatagg 60  
 aaagctgcta aacaagtcgt cgatgaatat atagaaaaaa tgttgagggg gattcaacta 120  
 gatagaagaa aacataccca aaatgctgcc ttaaacataa agttgagcgc aattaaaacg 180  
 aagtatttgc gtgaattaa tgttttagaa gagaagtcga aagatgagtt gccgtcagaa 240

ataaaagcaa agttagacgc agcttttgag aagtttaaaa aagatacatt gaaaccagga 300  
 gaaaaggtag cagaagctaa gaagaagggt gaagaagcta agaaaaaagc cgaggatcaa 360  
 aaagaagaag atcgtcgtaa ctaccaacc aatacttaca aaacgcttga acttgaaatt 420  
 gctgagttcg atgtgaaagt taaagaagcg gagcttgaac tagtaaaaga ggaagctaaa 480  
 gaatctcgaa acgagggcac aattaagcaa gcaaaagaga aagttgagag taaaaaagct 540  
 gaggctacaa ggtagaaaa catcaagaca gatcgtaaaa aagcagaaga agaagctaaa 600  
 cgaaaagcag atgctaagtt gaaggaagct aatgtagcga cttcagatca aggtaaacca 660  
 aaggggcggg caaacgagg agttcctgga gagctagcaa cacctgataa aaaagaaaat 720  
 gatgcgaagt cttcagattc tagcgtaggt gaagaaactc ttccaagctc atccctgaaa 780  
 tcaggaaaa aggtagcaga agctgagaag aaggttgaag aagctgagaa aaaagccaag 840  
 gatcaaaaag aagaagatcg ccgtaactac ccaaccaata cttacaaaac gcttgacctt 900  
 gaaattgctg agtccgatgt gaaagttaa gaagcggagc ttgaactagt aaaagaggaa 960  
 gctaaggaac ctcgagacga ggaaaaaatt aagcaagcaa aagcgaaggt tgagagtaaa 1020  
 aaagctgagg ctacaaggtt agaaaacatc aagacagatc gtaaaaaagc agaagaagaa 1080  
 gctaaacgaa aagcagcaga agaagataaa gttaaagaaa aaccagctg 1129

<210> 18

<211> 1992

<212> DNA

<213> 肺炎链球菌

<400> 18

gaaaacgaag gaagtaccca agcagccact tcttctaata tggcaaagac agaacatagg 60  
 aaagctgcta aacaagtcgt cgatgaatat atagaaaaaa tgttgaggga gattcaacta 120  
 gatagaagaa aacataccca aatgtcgcc ttaaacataa agttgagcgc aattaaaacg 180  
 aagtatttgc gtgaattaaa tgttttagaa gagaagtcga aagatgagtt gccgtcagaa 240  
 ataaaagcaa agttagacgc agcttttgag aagtttaaaa aagatacatt gaaaccagga 300  
 gaaaaggtag cagaagctaa gaagaagggt gaagaagcta agaaaaaagc cgaggatcaa 360  
 aaagaagaag atcgtcgtaa ctaccaacc aatacttaca aaacgcttga acttgaaatt 420

gctgagttcg atgtgaaagt taaagaagcg gagcttgaac tagtaaaaga ggaagctaaa 480  
gaatctcgaa acgagggcac aattaagcaa gcaaaagaga aagttgagag taaaaaagct 540  
gaggctacaa ggtagaaaa catcaagaca gatcgtaaaa aagcagaaga agaagctaaa 600  
cgaaaagcag atgctaagtt gaaggaagct aatgtagcga cttcagatca aggtaaacca 660  
aaggggicggg caaacgagg agttcctgga gagctagcaa cacctgataa aaaagaaaat 720  
gatgogaagt cttcagattc tagcgtaggt gaagaaactc ttccaagctc atccctgaaa 780  
tcaggaaaaa aggtagcaga agctgagaag aaggttgaag aagctgagaa aaaagccaag 840  
gatcaaaaag aagaagatcg ccgtaactac ccaaccaata cttacaaaac gcttgacctt 900  
gaaattgctg agtccgatgt gaaagttaa gaagcggagc ttgaactagt aaaagaggaa 960  
gctaaggaac ctcgagacga ggaaaaaatt aagcaagcaa aagcgaagt tgagagtaa 1020  
aaagctgagg ctacaaggtt agaaaacatc aagacagatc gtaaaaagc agaagaagaa 1080  
gctaaacgaa aagcagcaga agaagataaa gttaaagaaa aaccagctga acaaccacaa 1140  
ccagcgcggg ctactcaacc agaaaaacca gctccaaaac cagagaagcc agctgaacaa 1200  
ccaaaagcag aaaaacaga tgatcaacaa gctgaagaag actatgctcg tagatcagaa 1260  
gaagaatata atcgcttgac tcaacagcaa ccgcaaaaa ctgaaaaacc agcacaacca 1320  
tctactccaa aaacaggctg gaaacaagaa aacggtatgt ggtacttcta caatactgat 1380  
ggttcaatgg caacaggatg gctccaaaac aacggttcat ggtactatct aaacgctaat 1440  
ggtgctatgg cgacaggatg gctccaaaac aatggttcat ggtactatct aaacgctaat 1500  
ggttcaatgg caacaggatg gctccaaaac aatggttcat ggtactacct aaacgctaat 1560  
ggtgctatgg cgacaggatg gctccaatac aatggttcat ggtactacct aaacagcaat 1620  
ggcgctatgg cgacaggatg gctccaatac aatggttcat ggtactacct caacgctaat 1680  
ggtgatatgg cgacaggatg gctccaaaac aacggttcat ggtactacct caacgctaat 1740  
ggtgatatgg cgacaggatg gctccaatac aacggttcat ggtattacct caacgctaat 1800  
ggtgatatgg cgacaggttg ggtgaaagat ggagatacct ggtactatct tgaagcatca 1860  
ggtgctatga aagcaagcca atggttcaaa gtatcagata aatggtacta tgtcaatggc 1920  
tcaggtgccc ttgcagtcaa cacaactgta gatggctatg gagtcaatgc caatggtgaa 1980  
tgggtaaact aa 1992

&lt;210&gt; 19

<211> 763

<212> DNA

<213> 肺炎链球菌

<400> 19

```

gaaaacgaag gaagtaccca agcagccact tcttctaata tggcaaagac agaacatagg 60
aaagctgcta aacaagtcgt cgatgaatat atagaaaaaa tgttgagggga gattcaacta 120
gatagaagaa aacataccca aaatgtcgcg ttaaacataa agttgagcgc aattaaaacg 180
aagtatttgc gtgaattaaa tgttttagaa gagaagtcga aagatgagtt gccgtcagaa 240
ataaaagcaa agttagacgc agcttttgag aagtttaaaa aagatacatt gaaaccagga 300
gaaaaggtag cagaagctaa gaagaagggt gaagaagcta agaaaaagc cgaggatcaa 360
aaagaagaag atcgtcgtaa ctaccaacc aatacttaca aaacgcttga acttgaaatt 420
gctgagttcg atgtgaaagt taaagaagcg gagcttgaac tagtaaaaga ggaagctaaa 480
gaatctcgaa acgagggcac aattaagcaa gcaaaagaga aagttgagag taaaaaagct 540
gaggctacaa ggtagaaaa catcaagaca gatcgtaaaa aagcagaaga agaagctaaa 600
cgaaaagcag atgctaagtt gaaggaagct aatgtagcga cttcagatca aggtaaacca 660
aaggggcggg caaaacgagg agttcctgga gagctagcaa cacctgataa aaaagaaat 720
gatgcgaagt cttcagattc tagcgtaggt gaagaaactc ttc 763

```

<210> 20

<211> 318

<212> DNA

<213> 肺炎链球菌

<400> 20

```

aatcaggaa aaaaggtagc agaagctgag aagaaggttg aagaagctga gaaaaagcc 60
aaggatcaaa aagaagaaga tcgccgtaac taccacaacca atacttaca aacgcttgac 120
cttgaaattg ctgagtcgga tgtgaaagtt aaagaagcgg agcttgaact agtaaaagag 180
gaagctaagg aacctcgaga cgaggaaaa attaagcaag caaaagcga agttgagagt 240

```

aaaaagctg aggtacaag gttagaaaac atcaagacag atcgtaaaaa agcagaagaa 300  
gaagctaaac gaaaagca 318

<210> 21

<211> 321

<212> DNA

<213> 肺炎链球菌

<400> 21

ccaggagaaa aggtagcaga agctaagaag aaggttgaag aagctaagaa aaaagccgag 60  
gatcaaaaag aagaagatcg tcgtaactac ccaaccaata cttacaaaac gcttgaactt 120  
gaaattgctg agttcgatgt gaaagttaa gaagcggagc ttgaactagt aaaagaggaa 180  
gctaaagaat ctcgaaacga gggcacaatt aagcaagcaa aagagaaagt tgagagtaaa 240  
aaagctgagg ctacaaggtt agaaaacatc aagacagatc gtaaaaaagc agaagaagaa 300  
gctaaacgaa aagcagatgc t 321

<210> 22

<211> 121

<212> PRT

<213> 肺炎链球菌

<400> 22

Ser Pro Ser Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys

1 5 10 15

Val Glu Glu Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg

20 25 30

Arg Asn Tyr Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala





Arg Arg Asn Tyr Pro Thr Asn Thr Tyr Lys Thr Leu Asp Leu Glu Ile  
 35 40 45

Ala Glu Ser Asp Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys  
 50 55 60

Glu Glu Ala Lys Glu Pro Arg Asp Glu Glu Lys Ile Lys Gln Ala Lys  
 65 70 75 80

Ala Lys Val Glu Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile  
 85 90 95

Lys Thr Asp Arg Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala  
 100 105 110

Glu Glu Asp Lys Val Lys Glu Lys Arg Ala  
 115 120

<210> 24

<211> 428

<212> PRT

<213> 肺炎链球菌

<400> 24

Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn  
 1 5 10 15

---

Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu  
                   20                          25                          30

Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val  
           35                          40                          45

Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val  
       50                          55                          60

Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys  
       65                          70                          75                          80

Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu  
                           85                          90                          95

Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser  
           100                          105                          110

Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp  
           115                          120                          125

Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu  
       130                          135                          140

Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys  
       145                          150                          155                          160

Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu  
                           165                          170                          175

Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu  
 180 185 190

Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys  
 195 200 205

Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu  
 210 215 220

Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg  
 225 230 235 240

Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg  
 245 250 255

Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala  
 260 265 270

Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser  
 275 280 285

Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu  
 290 295 300

Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr  
 305 310 315 320

Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp

	325	330	335
Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys			
	340	345	350
Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu			
	355	360	365
Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg			
	370	375	380
Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys			
	385	390	395
			400
Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Lys			
	405	410	415
Ala Glu Lys Pro Ala Pro Ala Pro Lys Pro Glu Asn			
	420	425	

<210> 25

<211> 23

<212> DNA

<213> 肺炎链球菌

<400> 25

ggcggatcca tggaraayga rgg

23

<210> 26

<211> 33

<212> DNA

<213> 肺炎链球菌

<400> 26

gccgctcgact tagtttacc accaccatt ggc

33

<210> 27

<211> 5

<212> PRT

<213> 肺炎链球菌

<400> 27

Xaa Glu Asn Glu Gly

1 5

<210> 28

<211> 439

<212> PRT

<213> 肺炎链球菌

<400> 28

Ala Val Ala Ser Leu Phe Met Gly Ser Val Val His Ala Thr Glu Lys

1 5 10 15

Glu Val Thr Thr Gln Val Ala Thr Ser Ser Asn Lys Ala Asn Lys Ser

20 25 30

Gln Thr Glu His Met Lys Ala Ala Lys Gln Val Asp Glu Tyr Ile Lys  
 35 40 45

Lys Lys Leu Gln Leu Asp Arg Arg Lys His Thr Gln Asn Val Gly Leu  
 50 55 60

Leu Thr Lys Leu Gly Val Ile Lys Thr Glu Tyr Leu His Gly Leu Ser  
 65 70 75 80

Val Ser Lys Lys Lys Ser Glu Ala Glu Leu Pro Ser Glu Ile Lys Ala  
 85 90 95

Lys Leu Asp Ala Ala Phe Glu Gln Phe Lys Lys Asp Thr Leu Pro Thr  
 100 105 110

Glu Pro Gly Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala  
 115 120 125

Lys Lys Lys Ala Glu Asp Gln Lys Glu Lys Asp Leu Arg Asn Tyr Pro  
 130 135 140

Thr Asn Thr Tyr Lys Thr Leu Glu Leu Asp Ile Ala Glu Ser Asp Val  
 145 150 155 160

Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu  
 165 170 175

Ser Arg Asp Glu Lys Lys Ile Asn Gln Ala Lys Ala Lys Val Glu Asn

	180		185		190
Lys Lys Ala Glu Ala Thr Arg Leu Lys Asn Ile Lys Thr Asp Arg Glu					
	195		200		205
Lys Ala Glu Glu Ala Lys Arg Arg Ala Asp Ala Lys Leu Gln Glu Ala					
	210		215		220
Asn Val Ala Thr Ser Glu Gln Asp Lys Ser Lys Arg Arg Ala Lys Arg					
	225		230		235
					240
Glu Val Xaa Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala					
		245		250	
					255
Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Thr Ser Pro Ser					
	260		265		270
Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu					
	275		280		285
Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr					
	290		295		300
Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp					
	305		310		315
					320
Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys					
		325		330	
					335



Glu Ser Arg Asn Glu Glu Lys Ile Lys Gln Val Lys Ala Lys Val Glu  
 340 345 350

Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg  
 355 360 365

Lys Lys Ala Glu Glu Glu Glu Ala Lys Arg Arg Ala Ala Glu Glu Asp  
 370 375 380

Lys Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro  
 385 390 395 400

Gln Pro Glu Lys Pro Thr Glu Glu Pro Glu Asn Pro Ala Pro Ala Pro  
 405 410 415

Ala Pro Lys Pro Glu Asn Pro Ala Glu Lys Pro Lys Ala Glu Lys Pro  
 420 425 430

Ala Asp Gln Gln Ala Glu Glu  
 435

<210> 29

<211> 437

<212> PRT

<213> 肺炎链球菌

<400> 29

Ala Val Ala Ser Leu Phe Met Gly Ser Val Val His Ala Thr Glu Lys



Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu  
 165 170 175

Ser Arg Asp Glu Lys Lys Ile Asn Gln Ala Lys Ala Lys Val Glu Asn  
 180 185 190

Lys Lys Ala Glu Ala Thr Arg Leu Lys Asn Ile Lys Thr Asp Arg Glu  
 195 200 205

Lys Ala Glu Glu Ala Lys Arg Arg Ala Asp Ala Lys Leu Gln Glu Ala  
 210 215 220

Asn Val Ala Thr Ser Glu Gln Asp Lys Ser Lys Arg Arg Ala Lys Arg  
 225 230 235 240

Glu Val Leu Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala  
 245 250 255

Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Thr Ser Pro Ser  
 260 265 270

Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu  
 275 280 285

Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr  
 290 295 300

Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp  
 305 310 315 320

Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys  
 325 330 335

Glu Ser Arg Asn Glu Glu Lys Ile Lys Gln Val Lys Ala Lys Val Glu  
 340 345 350

Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg  
 355 360 365

Lys Lys Ala Glu Glu Glu Glu Ala Lys Arg Arg Ala Ala Glu Glu Asp  
 370 375 380

Lys Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro  
 385 390 395 400

Gln Pro Glu Lys Pro Thr Glu Glu Pro Glu Asn Pro Ala Pro Ala Pro  
 405 410 415

Ala Pro Lys Pro Glu Asn Pro Ala Glu Lys Pro Lys Ala Glu Lys Pro  
 420 425 430

Ala Asp Gln Gln Ala  
 435

<210> 30

<211> 439

<212> PRT





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

<210> 31

<211> 419

<212> PRT

<213> 肺炎链球菌

<400> 31

Ala Val Ala Ser Leu Val Met Gly Ser Val Val His Ala Thr Glu Asn

1 5 10 15

Glu Gly Thr Thr Gln Ala Pro Thr Ser Ser Asn Arg Gly Asn Glu Ser

20 25 30

Gln Ala Glu His Met Lys Ala Ala Lys Gln Val Asp Glu Tyr Ile Glu

35 40 45

Lys Met Leu Gln Leu Asp Arg Arg Lys His Thr Gln Asn Val Gly Leu

50 55 60

Leu Thr Lys Leu Gly Ala Ile Lys Thr Glu Tyr Leu Arg Gly Leu Ser

65 70 75 80

Val Ser Lys Glu Lys Ser Thr Ala Glu Leu Pro Ser Glu Ile Lys Glu

85 90 95

Lys Leu Thr Ala Ala Phe Lys Gln Phe Lys Lys Asp Thr Leu Lys Pro

100 105 110

Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Ala Glu Ala Lys Lys



115	120	125
Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile		
130	135	140
Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val Glu Val		
145	150	155
		160
Lys Lys Ala Glu Leu Glu Leu Val Lys Val Lys Ala Asn Glu Pro Arg		
	165	170
		175
Asp Glu Glu Lys Ile Lys Gln Ala Glu Ala Glu Val Glu Ser Lys Lys		
	180	185
		190
Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp Arg Glu Lys Ala		
	195	200
		205
Glu Glu Glu Ala Lys Arg Arg Val Asp Ala Lys Glu Gln Asp Glu Ser		
	210	215
		220
Ser Lys Arg Arg Lys Ser Arg Val Lys Arg Gly Asp Val Gly Glu Gln		
225	230	235
		240
Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala Lys Ser Ser Asp Ser Ser		
	245	250
		255
Val Gly Glu Glu Thr Leu Pro Ser Pro Ser Leu Lys Pro Gly Lys Lys		
	260	265
		270

Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala Asp Lys Lys Ala Lys  
 275 280 285

Ala Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro Thr Asn Thr Tyr Lys  
 290 295 300

Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala  
 305 310 315 320

Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu Pro Arg Asn Glu Glu  
 325 330 335

Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser Lys Lys Ala Glu Ala  
 340 345 350

Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys Lys Ala Glu Glu Glu  
 355 360 365

Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys Val Lys Glu Lys Pro Ala  
 370 375 380

Glu Gln Pro Lys Pro Ala Pro Ala Pro Gln Pro Glu Lys Pro Ala Pro  
 385 390 395 400

Lys Pro Glu Asn Pro Ala Glu Gln Pro Lys Ala Glu Lys Pro Ala Asp  
 405 410 415

Gln Gln Ala

<210> 32

<211> 437

<212> PRT

<213> 肺炎链球菌

<400> 32

Val Ala Ser Leu Phe Met Gly Ser Val Val His Ala Thr Glu Lys Glu  
1                    5                    10                    15

Val Thr Thr Gln Val Ala Thr Ser Ser Asn Lys Ala Asn Lys Ser Gln  
                  20                    25                    30

Thr Glu His Met Lys Ala Ala Lys Gln Val Asp Glu Tyr Ile Lys Lys  
                  35                    40                    45

Lys Leu Gln Leu Asp Arg Arg Lys His Thr Gln Asn Val Gly Leu Leu  
                  50                    55                    60

Thr Lys Leu Gly Val Ile Lys Thr Glu Tyr Leu His Gly Leu Ser Val  
65                    70                    75                    80

Ser Lys Lys Lys Ser Glu Ala Glu Leu Pro Ser Glu Ile Lys Ala Lys  
                  85                    90                    95

Leu Asp Ala Ala Phe Glu Gln Phe Lys Lys Asp Thr Leu Pro Thr Glu  
                  100                    105                    110

Pro Gly Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala Lys  
 115 120 125

Lys Lys Ala Glu Asp Gln Lys Glu Lys Asp Leu Arg Asn Tyr Pro Thr  
 130 135 140

Asn Thr Tyr Lys Thr Leu Glu Leu Asp Ile Ala Glu Ser Asp Val Glu  
 145 150 155 160

Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu Ser  
 165 170 175

Arg Asp Glu Lys Lys Ile Asn Gln Ala Lys Ala Lys Val Glu Asn Lys  
 180 185 190

Lys Ala Glu Ala Thr Arg Leu Lys Asn Ile Lys Thr Asp Arg Glu Lys  
 195 200 205

Ala Glu Glu Ala Lys Arg Arg Ala Asp Ala Lys Leu Gln Glu Ala Asn  
 210 215 220

Val Ala Thr Ser Glu Gln Asp Lys Ser Lys Arg Arg Ala Lys Arg Glu  
 225 230 235 240

Val Phe Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala Lys  
 245 250 255

Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Thr Ser Pro Ser Leu  
 260 265 270

Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala  
 275 280 285

Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro  
 290 295 300

Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val  
 305 310 315 320

Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu  
 325 330 335

Ser Arg Asn Glu Glu Lys Ile Lys Gln Val Lys Ala Lys Val Glu Ser  
 340 345 350

Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg Lys  
 355 360 365

Lys Ala Glu Glu Glu Glu Ala Lys Arg Arg Ala Ala Glu Glu Asp Lys  
 370 375 380

Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Gln  
 385 390 395 400

Pro Glu Lys Pro Thr Glu Glu Pro Glu Asn Pro Ala Pro Ala Pro Ala  
 405 410 415

Pro Lys Pro Glu Asn Pro Ala Glu Lys Pro Lys Ala Glu Lys Pro Ala

	420	425	430
Asp Gln Gln Ala Glu			
	435		
<210>	33		
<211>	433		
<212>	PRT		
<213>	肺炎链球菌		
<400>	33		
Cys Thr Val Ala Ser Leu Val Met Gly Ser Val Val His Ala Thr Glu			
1	5	10	15
Asn Glu Arg Thr Thr Gln Val Pro Thr Ser Ser Asn Arg Gly Lys Pro			
	20	25	30
Glu Arg Arg Lys Ala Ala Glu Gln Phe Asp Glu Tyr Ile Asn Lys Met			
	35	40	45
Ile Gln Leu Asp Lys Arg Lys His Thr Gln Asn Leu Ala Phe Asn Ile			
	50	55	60
Gln Leu Ser Arg Ile Lys Thr Glu Tyr Leu Asn Gly Leu Lys Glu Lys			
65	70	75	80
Ser Glu Ala Glu Leu Pro Ser Lys Ile Lys Ala Glu Leu Asp Ala Ala			
	85	90	95

Phe Lys Gln Phe Lys Lys Asp Thr Leu Pro Thr Glu Pro Glu Lys Lys  
 100 105 110

Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala Glu Lys Lys Val Ala  
 115 120 125

Glu Ala Lys Lys Lys Ala Lys Ala Gln Lys Glu Glu Asp His Arg Asn  
 130 135 140

Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Asp Leu Glu Ile Ala Glu Phe  
 145 150 155 160

Asp Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Lys Glu Ala  
 165 170 175

Asp Glu Ser Arg Asn Glu Gly Thr Ile Asn Gln Ala Lys Ala Lys Val  
 180 185 190

Glu Ser Glu Lys Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp  
 195 200 205

Arg Glu Lys Ala Glu Glu Glu Glu Ala Lys Arg Arg Ala Asp Ala Lys  
 210 215 220

Glu Gln Asp Glu Ser Lys Arg Arg Lys Ser Arg Gly Lys Arg Gly Ala  
 225 230 235 240

Leu Gly Glu Gln Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala Lys Ser

---

	245		250		255
Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser Leu Lys					
	260		265		270
Pro Gly Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala Asp					
	275		280		285
Lys Lys Ala Lys Ala Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro Thr					
	290		295		300
Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val Lys					
305		310		315	320
Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu Ser					
	325		330		335
Arg Asn Glu Glu Lys Ile Lys Gln Ala Lys Ala Lys Val Glu Ser Lys					
	340		345		350
Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys Lys					
	355		360		365
Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys Val Lys					
	370		375		380
Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Gln Pro Glu					
385		390		395	400



Lys Pro Ala Glu Glu Pro Glu Asn Pro Val Pro Ala Pro Lys Pro Glu  
 405 410 415

Asn Pro Ala Glu Gln Pro Lys Ala Glu Lys Pro Ala Asp Gln Gln Ala  
 420 425 430

Glu

<210> 34

<211> 427

<212> PRT

<213> 肺炎链球菌

<400> 34

Val Ala Val Ala Ser Leu Val Met Gly Ser Val Val His Ala Thr Glu  
 1 5 10 15

Lys Glu Val Thr Thr Gln Val Pro Thr Tyr Ser Asn Met Ala Lys Thr  
 20 25 30

Glu His Arg Lys Ala Ala Lys Gln Val Val Asp Glu Tyr Ile Glu Lys  
 35 40 45

Met Leu Arg Glu Ile Gln Leu Asp Arg Arg Lys His Thr Gln Asn Phe  
 50 55 60

Ala Phe Asn Met Lys Leu Ser Ala Ile Lys Thr Glu Tyr Leu Tyr Gly

65		70		75		80
Leu Lys Glu Lys Ser Glu Ala Glu Leu Pro Ser Glu Val Lys Ala Lys						
	85		90		95	
Leu Asp Ala Ala Phe Glu Gln Phe Lys Lys Asp Thr Leu Lys Leu Gly						
	100		105		110	
Glu Lys Val Ala Glu Ala Glu Lys Lys Val Ala Glu Ala Glu Lys Lys						
	115		120		125	
Ala Lys Ala Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro Thr Asn Thr						
	130		135		140	
Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys						
	145		150		155	160
Lys Ala Glu Leu Glu Leu Leu Lys Glu Glu Ala Lys Thr Arg Asn Glu						
	165		170		175	
Asp Thr Ile Asn Gln Ala Lys Ala Lys Val Glu Ser Lys Lys Ala Glu						
	180		185		190	
Ala Thr Lys Leu Glu Glu Ile Lys Thr Asp Arg Lys Lys Ala Glu Glu						
	195		200		205	
Glu Ala Lys Arg Lys Ala Glu Ala Glu Glu Asp Lys Val Lys Asp Lys						
	210		215		220	







Lys Arg Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn  
 225                                    230                                    235                                    240

Asp Ala Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser  
    245                                    250                                    255

Ser Ser Leu Lys Ser Gly Lys Lys Val Ala Glu Ala Glu Lys Lys Val  
    260                                    265                                    270

Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg  
    275                                    280                                    285

Asn Tyr Pro Thr Asn Thr Tyr Lys Thr Leu Asp Leu Glu Ile Ala Glu  
    290                                    295                                    300

Ser Asp Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu  
 305                                    310                                    315                                    320

Ala Lys Glu Pro Arg Asp Glu Glu Lys Ile Lys Gln Ala Lys Ala Lys  
    325                                    330                                    335

Val Glu Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr  
    340                                    345                                    350

Asp Arg Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu  
    355                                    360                                    365

Asp Lys Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala



Asn Val Leu Glu Asp Lys Ser Lys Ala Glu Leu Pro Ser Lys Ile Lys

85

90

95

Ala Glu Leu Asp Ala Ala Phe Glu Gln Phe Lys Lys Asp Thr Leu Pro

100

105

110

Thr Glu Pro Gly Lys Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu

115

120

125

Ala Glu Lys Lys Ala Lys Ala Gln Lys Glu Glu Asp Tyr Arg Asn Tyr

130

135

140

Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp

145

150

155

160

Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Lys Glu Ala Asp

165

170

175

Glu Ser Arg Asn Glu Gly Thr Ile Asn Gln Ala Lys Ala Lys Val Glu

180

185

190

Ser Glu Gln Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp Arg

195

200

205

Glu Lys Ala Glu Glu Glu Ala Lys Arg Arg Ala Asp Ala Lys Glu Gln

210

215

220

Asp Glu Ser Lys Arg Arg Lys Ser Arg Val Lys Arg Gly Asp Phe Gly



---

225	230	235	240
Glu Pro Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala Lys Ser Ser Asp			
	245	250	255
Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser Leu Lys Pro Gly			
	260	265	270
Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala Glu Lys Lys			
	275	280	285
Ala Lys Asp Gln Lys Glu Glu Asp His Arg Asn Tyr Pro Thr Ile Thr			
	290	295	300
Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys			
	305	310	315
Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Gly Ser Arg Asn			
	325	330	335
Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser Lys Lys Ala			
	340	345	350
Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys Lys Ala Glu			
	355	360	365
Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys Val Lys Glu Lys			
	370	375	380



65	70	75	80
Leu Glu Glu Lys Ser Glu Ala Glu Leu Thr Ser Lys Thr Lys Glu Thr			
	85	90	95
Lys Glu Glu Leu Thr Ala Ala Phe Glu Gln Phe Lys Lys Asp Thr Leu			
100	105	110	
Ser Thr Glu Pro Glu Lys Lys Val Ala Glu Ala Lys Lys Lys Val Glu			
115	120	125	
Glu Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Lys Asp Arg Arg Asn			
130	135	140	
Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser			
145	150	155	160
Asp Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Val Lys Ala			
	165	170	175
Asn Glu Pro Arg Asp Glu Glu Lys Ile Lys Gln Ala Glu Ala Lys Val			
	180	185	190
Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp			
195	200	205	
Arg Glu Gln Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr Asp Arg			
210	215	220	





Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val Gly  
 50 55 60

Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val Ala  
 65 70 75 80

Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys Ile  
 85 90 95

Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu Ser  
 100 105 110

Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser Ser  
 115 120 125

Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp Thr  
 130 135 140

Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu Ala  
 145 150 155 160

Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu  
 165 170 175

Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu  
 180 185 190

Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu Leu  
 195 200 205

Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys Gln  
 210 215 220

Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu Lys  
 225 230 235 240

Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg Arg  
 245 250 255

Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg Gly  
 260 265 270

Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala Lys  
 275 280 285

Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser Leu  
 290 295 300

Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala  
 305 310 315 320

Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro  
 325 330 335

Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val  
 340 345 350

Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu

	355		360			365									
Pro	Arg	Asn	Glu	Glu	Lys	Val	Lys	Gln	Ala	Lys	Ala	Glu	Val	Glu	Ser
	370					375								380	
Lys	Lys	Ala	Glu	Ala	Thr	Arg	Leu	Glu	Lys	Ile	Lys	Thr	Asp	Arg	Lys
385						390								395	400
Lys	Ala	Glu	Glu	Glu	Ala	Lys	Arg	Lys	Ala	Ala	Glu	Glu	Asp	Lys	Val
						405								410	415
Lys	Glu	Lys	Pro	Ala	Glu	Gln	Pro	Gln	Pro	Ala	Pro	Ala	Pro	Lys	Ala
						420								425	430
Glu	Lys	Pro	Ala	Pro	Ala	Pro	Lys	Pro	Glu	Asn	Pro	Ala	Glu	Gln	Pro
						435								440	445
Lys	Ala	Glu	Lys	Pro	Ala	Asp	Gln	Gln	Ala	Glu	Glu				
						450								455	460

<210> 39

<211> 459

<212> PRT

<213> 肺炎链球菌

<400> 39

Ile Val Ala Ser Leu Val Met Gly Ser Val Val His Ala Thr Glu Asn

1

5

10

15



Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn Glu Ser

20 25 30

Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu Arg Asp

35 40 45

Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val Gly Glu

50 55 60

Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val Ala Leu

65 70 75 80

Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys Ile Val

85 90 95

Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu Ser Arg

100 105 110

Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser Ser Ser

115 120 125

Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp Thr Ala

130 135 140

Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu Ala Lys

145 150 155 160

Lys Lys Val Glu Glu Val Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu

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

Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro Thr  
 325 330 335

Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val Glu  
 340 345 350

Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu Pro  
 355 360 365

Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser Lys  
 370 375 380

Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys Lys  
 385 390 395 400

Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys Val Lys  
 405 410 415

Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Lys Thr Glu  
 420 425 430

Lys Pro Ala Pro Ala Pro Lys Pro Glu Asn Pro Ala Glu Gln Pro Lys  
 435 440 445

Ala Glu Lys Pro Ala Asp Gln Gln Ala Glu Glu  
 450 455

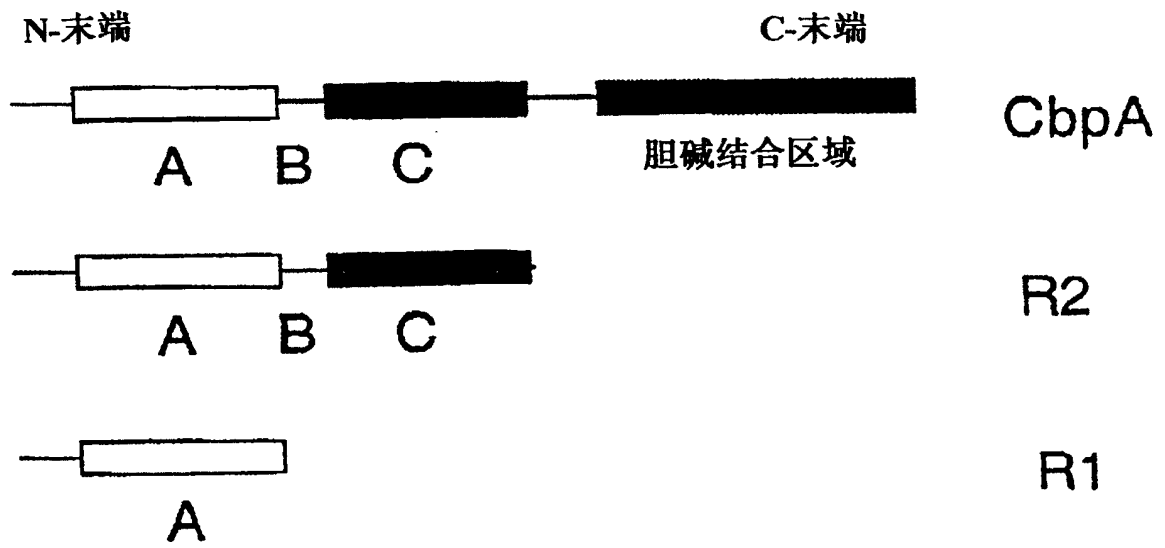


图 1

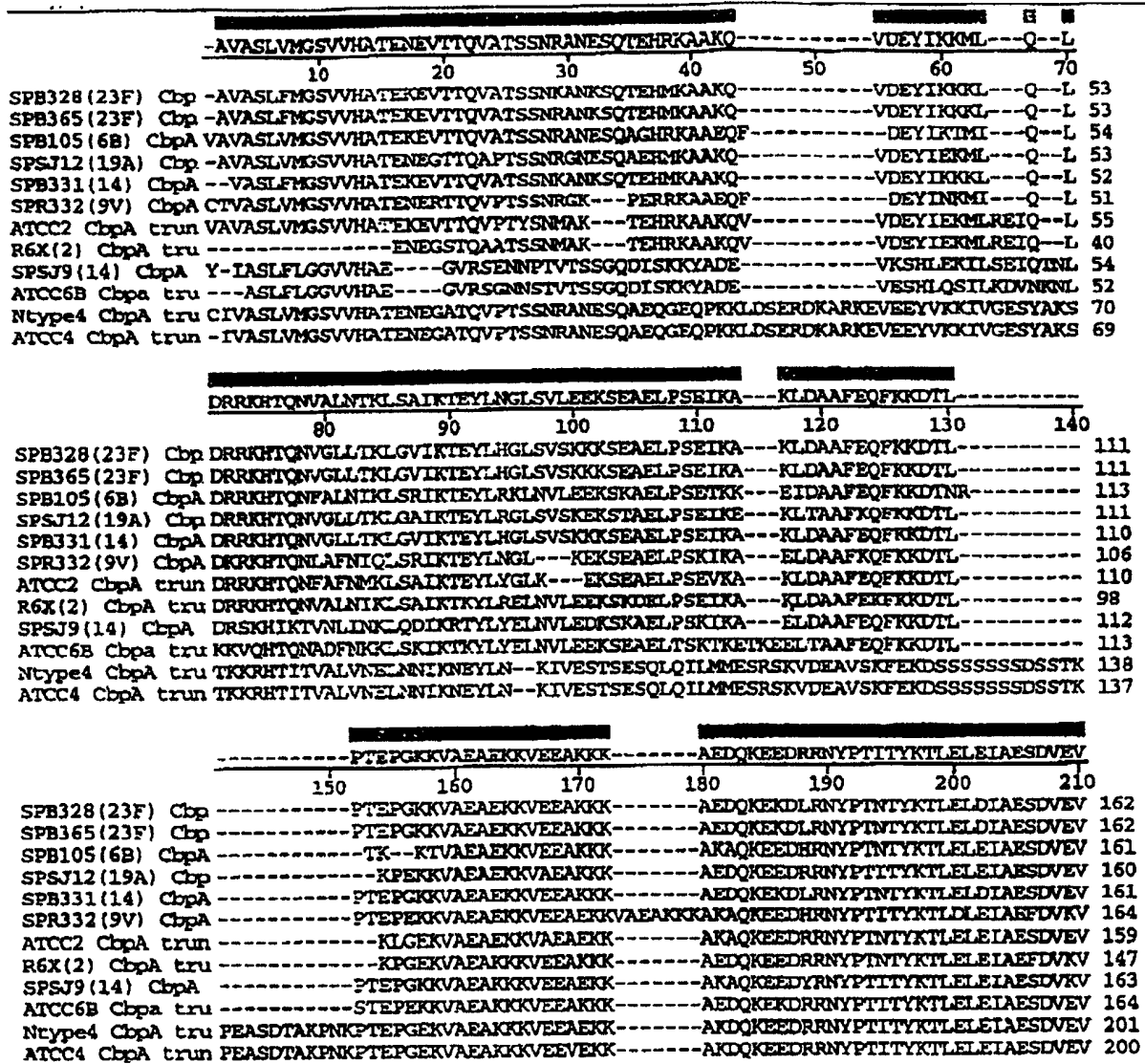


图 2A

KQAELELVKEEAKESRDEGKINQAKAKVESKQAEATRLKIKLREKAEFE-AKRRADAKLQEA---NVA  
 222 230 240 250 260 270 280

SPB328(23F) Cbp KQAELELVKEEAKESRDEGKINQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 227  
 SPB365(23F) Cbp KQAELELVKEEAKESRDEGKINQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 227  
 SPB105(6B) CbpA KQAELELVKEEAKESRDEGKINQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 230  
 SPSJ12(19A) Cbp KQAELELVKQWNEPRDEEKIKQAEAEVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 224  
 SPB331(14) CbpA KQAELELVKEEAKESRDEGKINQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 226  
 SPR332(9V) CbpA KQAELELVKCEADESRNEGTLNQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 229  
 ATCC2 CbpA trun KQAELELVKEEAK-TRNEDTLNQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 222  
 R6X(2) CbpA trun KEAELELVKEEAKESRNEGTLNQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 213  
 SPSJ9(14) CbpA KEAELELVKCEADESRNEGTLNQAKAKVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 227  
 ATCC6B CbpA trun KQAELELVKQWNEPRDEEKIKQAEAEVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 234  
 Ntype4 CbpA trun KQAELELVKQWNEPRDEEKIKQAEAEVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 261  
 ATCC4 CbpA trun KQAELELVKQWNEPRDEEKIKQAEAEVESKQAEATRLKIKLREKAEFE--AKRRADAKLQEA---NVA 260

-SEQKPKGRKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK  
 290 300 310 320 330 340 350

SPB328(23F) Cbp TSEQQKSKRAKREVLGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 297  
 SPB365(23F) Cbp TSEQQKSKRAKREVLGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 297  
 SPB105(6B) CbpA TSEQQKPKGRKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 300  
 SPSJ12(19A) Cbp -S-SKRRKSRGKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 291  
 SPB331(14) CbpA TSEQQKSKRAKREVLGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 296  
 SPR332(9V) CbpA ---SKRRKSRGKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 295  
 ATCC2 CbpA trun ---DKLKRTRKRVGPEATPDKQKNDKSSDSSVGEETLPSPSLKSGIKQVAEAEKQVEAEKQKAEQK 288  
 R6X(2) CbpA trun TSDQKPKGRKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKSGIKQVAEAEKQVEAEKQKAEQK 283  
 SPSJ9(14) CbpA -S--KRRKSRGKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 293  
 ATCC6B CbpA trun ---DEPKKCRKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 300  
 Ntype4 CbpA trun --EQGKPKGRKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 328  
 ATCC4 CbpA trun --EQGKPKGRKRGVPGELATPDKQKNDKSSDSSVGEETLPSPSLKPEKQVAEAEKQVEAEKQKAEQK 327

EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD  
 350 370 380 390 400 410 420

SPB328(23F) Cbp EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 367  
 SPB365(23F) Cbp EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 367  
 SPB105(6B) CbpA EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 370  
 SPSJ12(19A) Cbp EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 361  
 SPB331(14) CbpA EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 366  
 SPR332(9V) CbpA EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 365  
 ATCC2 CbpA trun EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 358  
 R6X(2) CbpA trun EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 353  
 SPSJ9(14) CbpA EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 363  
 ATCC6B CbpA trun EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 370  
 Ntype4 CbpA trun EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 398  
 ATCC4 CbpA trun EEDRRNYPTNYKILELEIAESDVEVKQAELELVKEEAKESRNEEKIKQAKAVESKQAEATRLKIKITD 397

RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA  
 430 440 450 460 470 480 490

SPB328(23F) Cbp RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 437  
 SPB365(23F) Cbp RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 437  
 SPB105(6B) CbpA RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 437  
 SPSJ12(19A) Cbp RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 419  
 SPB331(14) CbpA RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 436  
 SPR332(9V) CbpA RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 432  
 ATCC2 CbpA trun RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 425  
 R6X(2) CbpA trun RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 411  
 SPSJ9(14) CbpA RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 423  
 ATCC6B CbpA trun RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 437  
 Ntype4 CbpA trun RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 458  
 ATCC4 CbpA trun RKQAEFE-AKQKAAEEDKVKKEKPAEQQPAPAPQPEKPTTEEPENPAPAPKFPENPAEKPKAEKPADQQA 457

■  
 EE--

SPB328(23F) Cbp EE 439  
 SPB365(23F) Cbp EE 437  
 SPB105(6B) CbpA EE 439  
 SPSJ12(19A) Cbp EE 419  
 SPB331(14) CbpA E 437  
 SPR332(9V) CbpA E 433  
 ATCC2 CbpA trun EE 427  
 R6X(2) CbpA trun EE 413  
 SPSJ9(14) CbpA EE 425  
 ATCC6B CbpA trun EE 439  
 Ntype4 CbpA trun EE 460  
 ATCC4 CbpA trun EE 459

图 2B

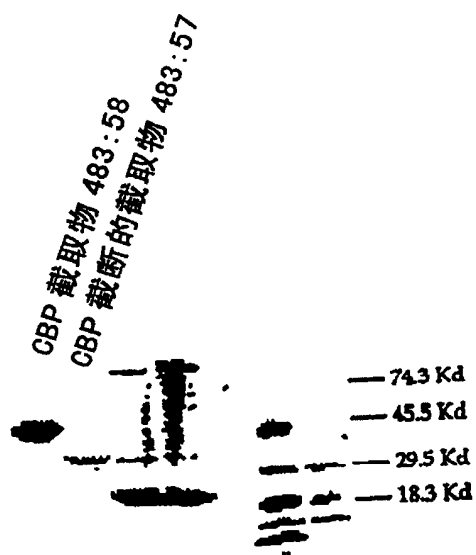


图 3

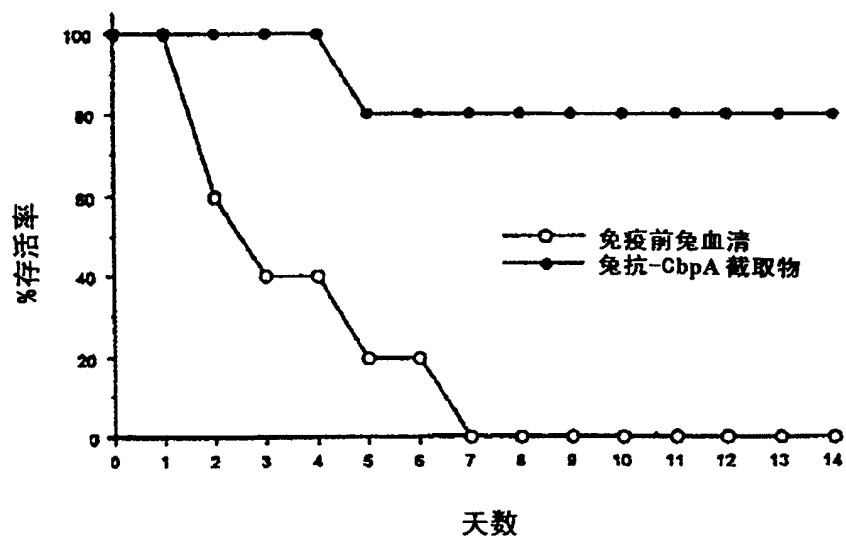


图 4



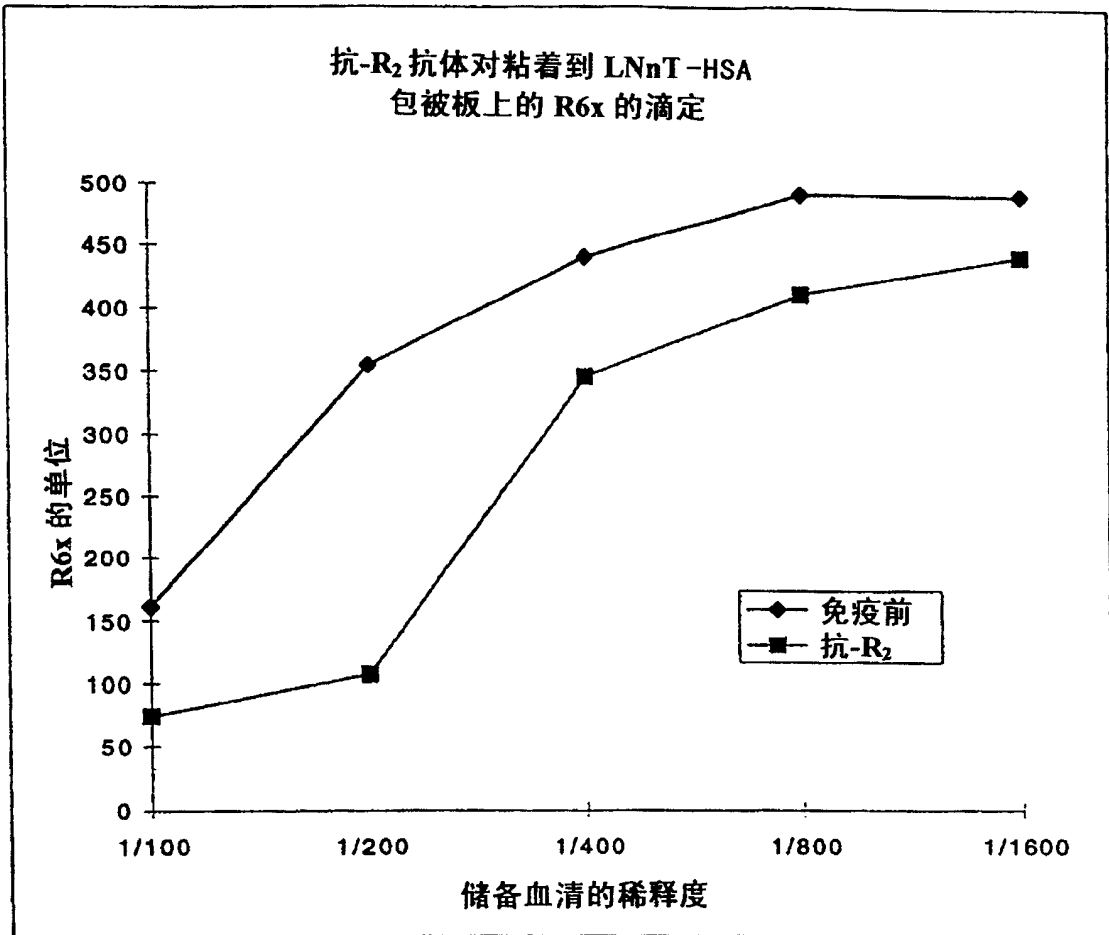


图 5

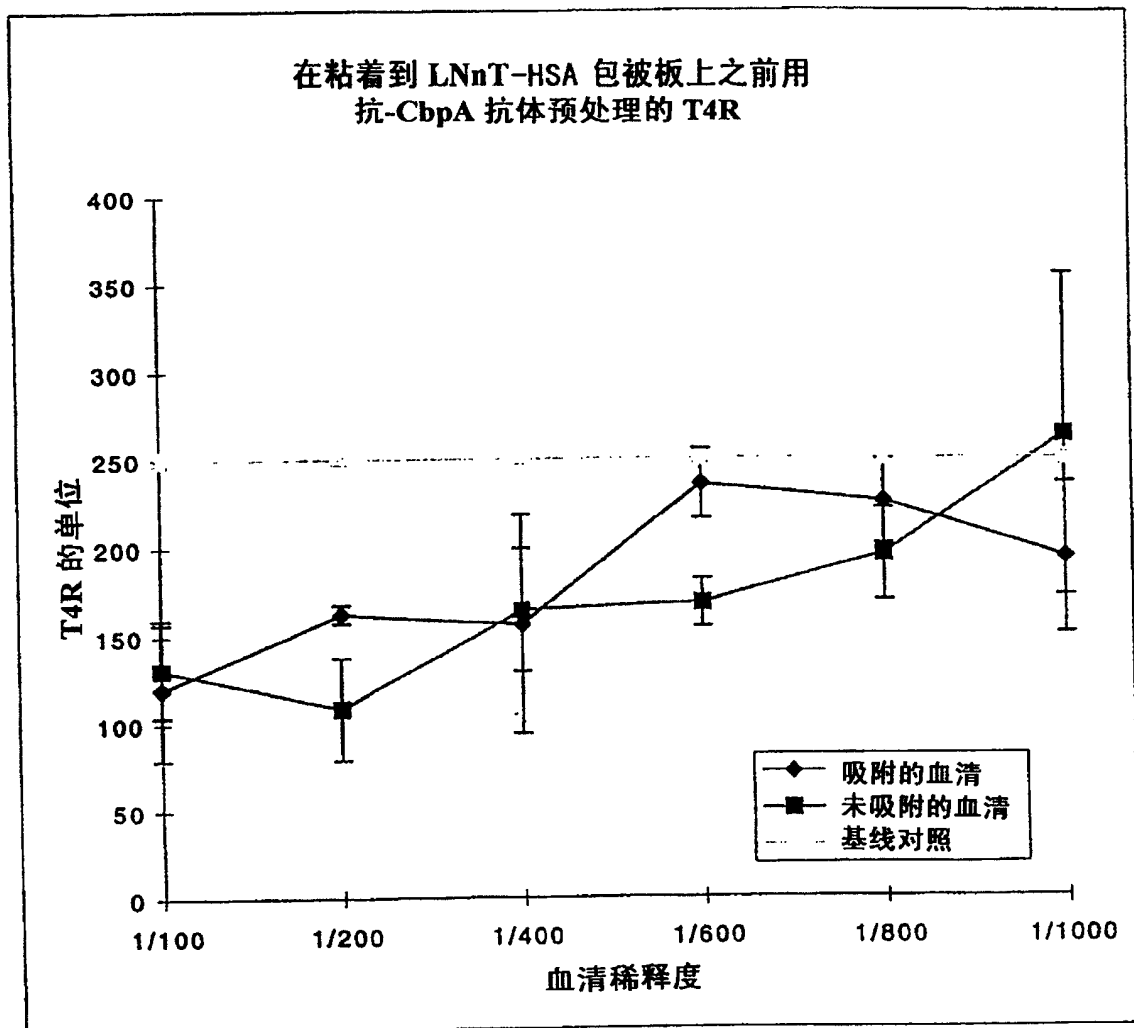


图 6

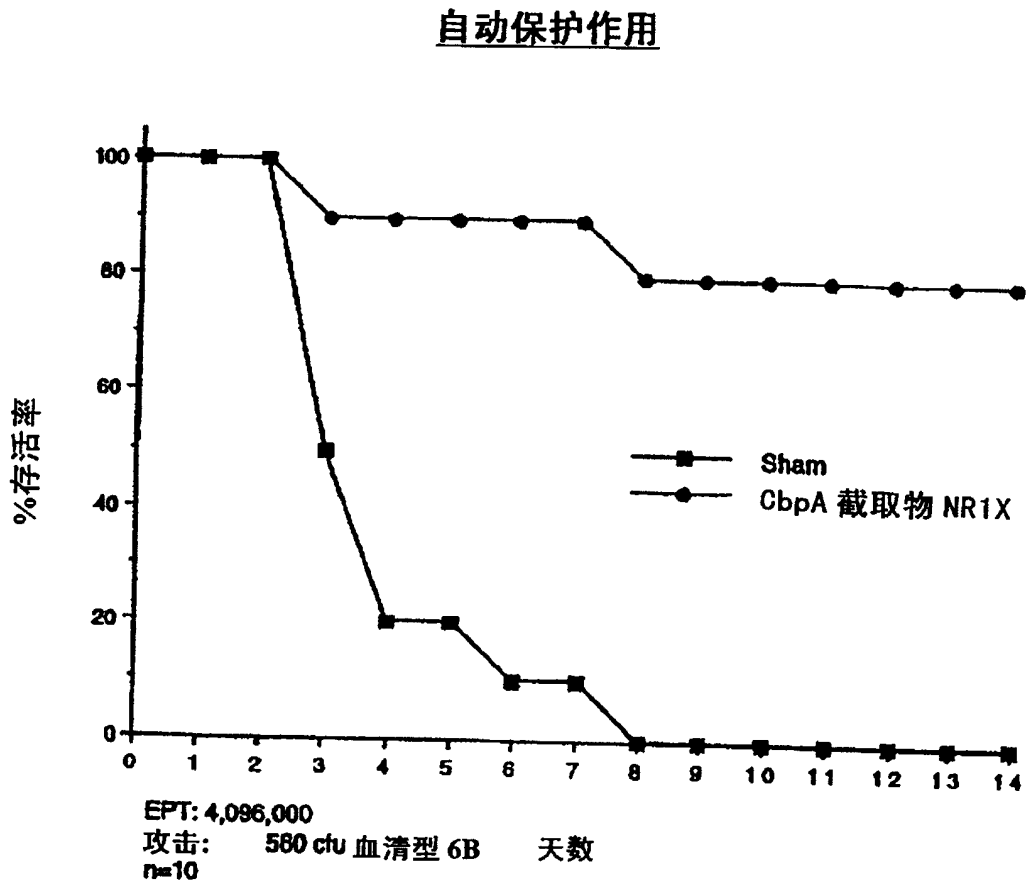


图 7