

[19] 中华人民共和国国家知识产权局

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

# [12] 发明专利申请公开说明书

C12N 15/57  
C12N 9/52 A61K 38/48  
C07K 16/40 C12N 1/21  
C12N 5/10

[21] 申请号 98809460.6

[43] 公开日 2001 年 4 月 11 日

[11] 公开号 CN 1291233A

[22] 申请日 1998.9.24 [21] 申请号 98809460.6

[30] 优先权

[32] 1997.9.24 [33] US [31] 60/059,907

[86] 国际申请 PCT/US98/20186 1998.9.24

[87] 国际公布 WO99/15675 英 1999.4.1

[85] 进入国家阶段日期 2000.3.24

[71] 申请人 美国明尼苏达州大学

地址 美国明尼苏达州

共同申请人 美国氰胺公司

[72] 发明人 马格列特 K·贺斯泰特 大卫 J·芬科尔

程 希 布鲁斯 A·葛林

爱美 W·马西

[74] 专利代理机构 上海专利商标事务所

代理人 徐 迅

权利要求书 4 页 说明书 20 页 附图页数 9 页

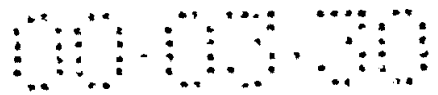
[54] 发明名称 来自肺炎链球菌之人类补体 C3 降解蛋白酶

[57] 摘要

本发明涉及鉴定及使用肺炎链球菌所表达之一群人类补体 C3 降解性蛋白酶。该蛋白质具有约 15kD 至约 25kD 之分子量。本发明较佳之蛋白酶包括 SE QIDNO:2 之氨基酸序列。

```
1 ATGTCAAGCC TTTACGTGA ATTGTATGCT AAACCCCTTAT CAGAAGCCCA
51 TGTAGAATCT GATGGTCTTA TTTTCGACCC AGCGCAAATC ACAAGTCCGAA
101 CCGCCAAATGG TGTGCTGTA CCGCAAGGAG ACCATTATCA CTTATTTCT
151 TATTCAACAAC TGTCACCTTT GGAAAGAAAA TTG GTCGTAT-TATTCCTT
201 CGTTATCGTT CAACCATTTG GGTACCAGATTCAAAGACCA GAACAACCCAG
251 TCACAATCG ACTCGGGAAA CCTAGTCCAA GTCCGAACCTCCACCAAAT
301 CCTCAACCAAG CTCGAAGCAA TCCAA TTSATGAAGAT TGGTCAAGAAAGC
351 TGTTCGAAAA GTAAGCGATG GTTATGCTTTGA GGAGAAT GGAGTTGCCT
401 CGTTATATCC CAAAGCAAGG ATCTTACAGCAGAAACA GCAAGCAGG CATTG
451 ATAGCAAACCT GCGCAAGCAG GAAA GTTTAT CTGATAAGCT AG
```

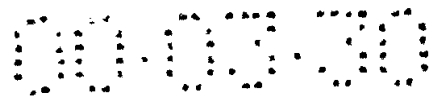
I S S N 1 0 0 8 - 4 2 7 4



## 权 利 要 求 书

---

1. 一种分离蛋白质，其特征在于，其与 SEQ ID NO:2 具至少 80%之序列一致性，其能够降解人类补体蛋白质 C3。
- 5       2. 如权利要求 1 所述的蛋白质，其特征在于，其系自肺炎链球菌分离。
3. 如权利要求 1 所述的蛋白质，其特征在于，其为重组蛋白质。
4. 如权利要求 1 所述的蛋白质，其特征在于，其分子量为约 15kDa 至约 25kDa。
5. 一种分离肽或多肽，其特征在于，其包括来自如权利要求 1 所述的蛋白质中至少 15 个连续氨基酸。
- 10       6. 一种分离肽或多肽，其特征在于，其包括 SEQ ID NO:2 中至少 15 个连续氨基酸。
7. 一种分离蛋白质，其特征在于，其包括 SEQ ID NO:2。
8. 如权利要求 7 所述的蛋白质，其特征在于，其分子量为约 15kDa 至约 25kDa。
9. 如权利要求 8 所述的蛋白质，其特征在于，其系自肺炎链球菌分离。
- 15       10. 如权利要求 8 所述的蛋白质，其特征在于，其降解人类补体蛋白质 C3。
11. 如权利要求 7 所述的蛋白质，其特征在于，其为 SEQ ID NO:2。
12. 一种分离蛋白质，其特征在于，其包括 SEQ ID NO:2 中自约第 1 个至约第 58 个氨基酸。
13. 如权利要求 12 所述的蛋白质，其特征在于，其进而包括 SEQ ID NO:2
- 20       中自约第 90 个至约第 132 个氨基酸。
14. 一种分离蛋白质，其特征在于，其包括 SEQ ID NO:5 中自约第 170 个至约第 227 个氨基酸。
15. 如权利要求 14 所述的蛋白质，其特征在于，其进而包括 SEQ ID NO:5 中自约第 258 个至约第 300 个氨基酸。
- 25       16. 一种降解人类补体蛋白质 C3 之分离蛋白质，其特征在于，编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格杂交条件下杂交。
17. 一种来自肺炎链球菌之约 15kDa 至约 25kDa 之分离蛋白质，其特征在于，其能够降解人类补体 C3。
18. 一种免疫系统刺激性组合物，其特征在于，其包括一有效量之免疫系统
- 30       刺激性肽或多肽，该肽或多肽具有衍生自一种蛋白质之至少 15 个连续氨基酸，该蛋白质与 SEQ ID NO:2 有至少 80% 序列一致性且能降解人类补体蛋白质 C3。



19. 如权利要求 18 所述的免疫系统刺激性组合物，其特征在于，该蛋白质系自肺炎链球菌分离。

20. 如权利要求 19 所述的免疫系统刺激性组合物，其特征在于，其进而包括至少一种来自肺炎链球菌分离之其它免疫系统刺激性蛋白质，肽或多肽。

5 21. 一种免疫系统刺激性组合物，其特征在于，其包括治疗有效量之一蛋白质的至少一部份，其中编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。

10 22. 一种免疫系统刺激性组合物，其特征在于，其包含有效份量之如权利要求 17 所述的蛋白质之至少一部份以及药学上可接受的载体，其中该蛋白质可有效地免疫或治疗哺乳动物个体以对抗肺炎链球菌的感染或侵入繁殖。

23. 如权利要求 22 所述的免疫系统刺激性组合物，其特征在于，该蛋白质系以能够对于该哺乳动物个体有效提供治疗效果的份量予以供给。

24. 一种抗体，其特征在于，其能够与 SEQ ID NO:2 有至少 80% 序列一致性且能降解人类补体蛋白质 C3 之蛋白质结合。

15 25. 如权利要求 24 所述的抗体，其特征在于，其为单株抗体。

26. 如权利要求 24 所述的抗体，其特征在于，其系得自小鼠，大鼠，山羊，鸡，人类，或兔。

27. 一种能够与一蛋白质之至少一部份结合的抗体，其特征在于，编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。

20 28. 一种分离核酸片段，其特征在于，其可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。

29. 如权利要求 28 所述的核酸片段，其特征在于，其系自肺炎链球菌分离。

30. 如权利要求 28 所述的核酸片段，其特征在于，该核酸片段编码一蛋白质之至少一部份。

25 31. 如权利要求 30 所述的核酸片段，其特征在于，该蛋白质降解人类补体 C3。

32. 如权利要求 28 所述的核酸片段，其特征在于，其在核酸载体中。

33. 如权利要求 32 所述的核酸片段，其特征在于，该载体为能够生产一蛋白质之至少一部份之表达载体。

34. 一种细胞，其特征在于，其包括如权利要求 28 所述的核酸。

30 35. 如权利要求 34 所述的细胞，其特征在于，该细胞为细菌或真核细胞。

36. 一种分离核酸片段，其特征在于，其包括 SEQ ID NO:1 或其互补链中



约第 1 个至约第 174 个核苷酸。

37. 如权利要求 34 所述的分离核酸片段,其特征在於,其进而包括 SEQ ID NO:1 或其互补链中约第 320 个至约第 492 个核苷酸。

5 38. 一种分离核酸片段,其特征在於,其包括 SEQ ID NO:1 或其互补链之核酸序列。

39. 一种 RNA 片段,其特征在於,其系由包括 SEQ ID NO:1 或其互补链之双链 DNA 序列所转录。

10 40. 一种在哺乳动物产生对抗肺炎链球菌的免疫反应的方法,其特征在於,其包括将一组合物施用至哺乳动物的步骤以产生对抗该蛋白质的免疫反应,该组合物包括治疗有效量之一蛋白质的至少一部份,其中编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。

41. 如权利要求 40 所述的方法,其特征在於,该免疫反应可为 B 细胞反应, T 细胞反应, 上皮细胞反应或内皮细胞反应。

15 42. 如权利要求 40 所述的方法,其特征在於,一蛋白质之至少一部份为至少 15 个氨基酸长度。

43. 如权利要求 40 所述的方法,其特征在於,该组合物进而包括至少一种来自肺炎链球菌之其它免疫系统刺激性肽,多肽或蛋白质。

44. 如权利要求 40 所述的方法,其特征在於,一蛋白质之至少一部份包括 SEQ ID NO:2 中至少 15 个氨基酸。

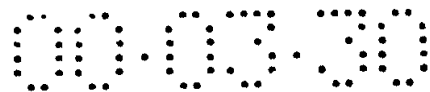
20 45. 一种抑制肺炎链球菌-媒介性 C3 降解作用的方法,其特征在於,其包括的步骤为:将肺炎链球菌细菌接触一抗体,该抗体能够与 SEQ ID NO:2 有至少 80% 序列一致性之蛋白质结合。

25 46. 一种抑制 C3-媒介性发炎反应及异种移植时排斥作用之方法,其特征在於,其包括在异种移植时所用的动物器官表面上表达具有 SEQ ID NO:2 氨基酸序列之蛋白质。

47. 一种分离核酸分子,其特征在於,其包含一至少 15 个核苷酸的区域而可与 SEQ ID NO:1 或其互补链中至少一区域在高度严格之杂交条件下杂交。

30 48. 一种分离核酸分子,其特征在於,其包含可与 SEQ ID NO:1 或其互补链中至少一区域在高度严格之杂交条件下杂交之序列,其中该区域系选自核苷酸 1-174 及 320-492。

49. 一种分离核酸分子,其特征在於,其包含一至少 15 个核苷酸的区域而可



与 SEQ ID NO:4 或其互补链中至少一区域在高度严格之杂交条件下杂交。

50. 一种分离核酸分子，其特征在于，其包含可与 SEQ ID NO:4 或其互补链中至少一区域在高度严格之杂交条件下杂交之序列，其中该区域系选自核苷酸 507-681 及 827-999。

5 51. 如权利要求 49 所述的核酸分子，其特征在于，其编码一蛋白质的至少一部份。

52. 如权利要求 51 所述的核酸分子，其特征在于，该蛋白质具有如 SEQ ID NO:5 所示之预测氨基酸序列。

10 53. 一种分离核酸片段，其特征在于，其具有选自包括 SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO:8 及 SEQ ID NO:9 之群组之核酸序列。

54. 一种包含治疗有效量之一蛋白质之至少一部份之免疫系统刺激性组合物，其中编码该蛋白质之核酸可与 SEQ ID NO:4 或其互补链在高度严格之杂交条件下杂交。

15 55. 一种在哺乳动物产生对抗肺炎链球菌的免疫反应的方法，其特征在于，其包括将一组合物施用至哺乳动物的步骤以产生对抗该蛋白质的免疫反应，该组合物包括治疗有效量之一蛋白质的至少一部份，其中编码该蛋白质之核酸可与 SEQ ID NO:4 或其互补链在高度严格之杂交条件下杂交。

20 56. 一种免疫系统刺激性组合物，其特征在于，其包含治疗有效量之如权利要求 51 所述的一蛋白质之至少一部份以及药学上可接受的载体，其中该蛋白质可有效地免疫接种或治疗哺乳动物以对抗肺炎链球菌的感染或侵入繁殖。

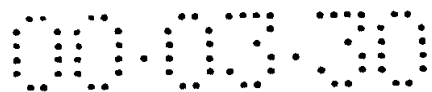
57. 如权利要求 56 所述的免疫系统刺激性组合物，其特征在于，该蛋白质系以能够对于该哺乳动物个体有效提供治疗效果的份量予以供给。

58. 如权利要求 56 所述的免疫系统刺激性组合物，其特征在于，该组合物为疫苗。

25 59. 一种多肽，其特征在于，其具 SEQ ID NO: 5。

60. 如权利要求 23 所述的免疫系统刺激性组合物，其特征在于，该核酸序列或其互补链所编码之蛋白质可抑制至少一个内源性肺炎链球菌核酸序列之转录或翻译。

30 61. 如权利要求 56 所述的免疫系统刺激性组合物，其特征在于，该核酸序列或其互补链所编码之蛋白质可抑制至少一个内源性肺炎链球菌核酸序列之转录或翻译。



# 说 明 书

来自肺炎链球菌之人类补体 C3 降解蛋白酶

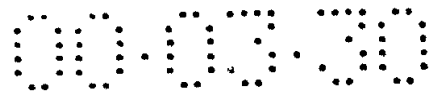
## 5 发明领域

本发明涉及肺炎链球菌 (*Streptococcus pneumoniae*) 且特别涉及一种肺炎链球菌蛋白质之鉴定, 该蛋白质能够降解人类补体蛋白质 C3。

## 发明背景

10 肺炎链球菌造成之呼吸道感染每年引起估计约 500, 000 例之肺炎及 47, 000 人死亡。对细菌性肺炎链球菌感染具有高危险性者为二岁以下之婴儿, 免疫系统功能不彰之个人及老年人。在这些族群中, 肺炎链球菌为细菌性肺炎及脑膜炎之主要原因。进而, 肺炎链球菌为所有年龄儿童之耳感染主因。儿童与老人两者在细菌侵入繁殖 (colonization), 局部或全身感染之后或以纯化多糖进行免疫接种  
15 后, 对于肺炎链球菌荚膜多糖之保护性抗体的合成有障碍。肺炎链球菌对于 HIV 感染之成人或儿童而言皆为侵入性细菌吸呼吸道疾病之主因, 并在这些病人身上引发出血性感染 (Connor 等人, *Current Topics in AIDS* 1987; 1:185-209 及 Janoff 等人, *Ann. Intern. Med.* 1992:117 (4) :314-324)。

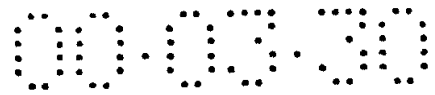
对于严重感染显示最大危险性之个体无法对现有荚膜多糖疫苗产生抗体。结  
20 果, 目前有四种结合 (conjugate) 疫苗进入临床试验之阶段。结合疫苗包括与蛋白质载体或佐剂结合之肺炎链球菌荚膜多糖, 希望能加强抗体反应。然而, 在临床试验中的结合疫苗另有其它潜在难题。例如, 在美国极为盛行之肺炎链球菌血清型与在例如以色列, 西欧, 南非或斯堪地那维亚等地最常见的血清型不同。因此, 在一个局部地域有利的疫苗在另一个地区可能为不利的。至于修正目前可用  
25 荚膜多糖疫苗或发展蛋白质结合物作为适应地域血清型差异之荚膜疫苗之潜在需求, 则产生令人却步之金钱及技术上复杂性。因此, 对于肺炎链球菌感染之预防及调制广泛保护性肺炎链球菌疫苗而言, 在不同有毒力之血清型中找寻具全世界保守性之免疫产生性表面暴露蛋白质是至为重要的。进而, 全世界性地出现青霉素及头芽孢素抗药性肺炎链球菌, 使得对于有效疫苗之需求至为迫切 (Baquero  
30 等人, *J. Antimicrob. Chemother.* 1991;28S;31-8)。



有数个肺炎链球菌蛋白质已被提出与肺炎链球菌荚膜多糖结合或作为单一免疫原来激发对抗肺炎链球菌之免疫反应。已有报导指出与肺炎链球菌黏附至呼吸道上皮细胞相关之表面蛋白质包括 PsaA, PspC/CBP112 及 IgA1 蛋白酶 (Sampson 等人, *Infect. Immun.* 1994; 62:319-324, Sheffield 等人, *Microb. Pathogen.* 5 1992; 13:261-9 及 Wani 等人 *Infect. Immun.* 1996; 64:3967-3974)。对抗这些黏附蛋白质之抗体可以抑制肺炎链球菌黏附至呼吸道上皮细胞, 且因此减抑侵入繁殖。其它如肺溶素 (pneumolysin), 自溶素 (autolysin), 神经胺苷酶或硫璃糖苷酶之细胞质肺炎链球菌蛋白质被提出作为疫苗抗原, 因为抗体可以潜在性阻止这些蛋白质在肺炎链球菌感染病人身上之毒性作用。然而, 这些蛋白质典型上并非位于肺炎链球菌的表面, 而是当细胞溶解及死亡时自细菌分泌或释出 (Lee 10 等人, *Vaccine* 1994;12:875-8 及 Berry 等人, *Infect. Immun.* 1994; 62:1101-1108)。虽然使用这些细胞质蛋白质作为免疫原可以缓和肺炎链球菌感染之晚期影响, 但对抗这些白质的抗体既不会促进肺炎链球菌死亡亦不会防止最初或后续之肺炎链球菌侵入繁殖。

正在进行肺炎链球菌疫苗试验之原型表面蛋白质为肺炎链球菌表面蛋白质 A (PspA)。PspA 为约 70-140kDa 之异源性蛋白质。PspA 的结构包括: 在胺基端之  $\alpha$  螺旋, 接着是富含脯氨酸之序列, 且在羧端以一系列之 11 个胆碱结合性重复序列作为结束。虽然与其结构相关的许多信息均为已知, PspA 在不同的肺炎链球菌血清型之间并无结构保守性, 且其功能完全未知 (Yother 等人, *J. Bacteriol.* 1992; 174:601-9 及 Yother, *J. Bacteriol.* 1994; 176:2976-2985)。研究已经证实 PspA 在动物之免疫产生性 (McDaniel 等人, *Microb. Pathogen.* 1994; 17:323-337)。虽然 PspA 具有免疫产生性, 其作为疫苗抗原之能力却因 PspA 系异源性、其以四个结构群存在以及其未予定性之功能而使情形变得复杂。

在无法对型-特异性多糖荚膜产生保护性抗体之病人身上, 补体之第三种成份 C3 及替代补体路径相关之蛋白质构成宿主对抗肺炎链球菌感染之第一道防线。25 因为补体蛋白质不能穿透肺炎链球菌之坚固细胞壁, 调理性 C3b 在肺炎链球菌表面之沉积成为清除肺炎链球菌之主要媒介作用。肺炎链球菌与血浆中 C3 之间的交互作用已知会在肺炎链球菌性菌血症时发生, 此时 C3b (C3 之调理活性片段) 之共价结合激活吞噬细胞之辨识及吞噬 (Johnston 等人, *J. Exp. Med.* 1969; 30 129:1275-1290, Hasin HE, *J. Immunol.* 1972; 109:26-31 及 Hostetter 等人, *J. Infect. Dis.* 1984; 150:653-61)。C3b 沉积在肺炎链球菌荚膜及细胞壁上。



然此控制肺炎链球菌感染之方法全然无效率。激发肺炎链球菌调理作用之方法能改善此微生物引发之疾病过程。目前对于限制肺炎链球菌感染之方法及治疗仍有强烈需求。

## 5 发明概述

本发明涉及鉴定及使用肺炎链球菌所表达之一群人类补体 C3 降解性蛋白质（蛋白酶）。该等蛋白质较佳具有约 15kD 至约 25kD 之分子量，该分子量例如系在 10% SDS 聚丙烯酰胺胶上测定。本发明包括自肺炎链球菌之不同 C3-降解性菌株所能分离之许多蛋白质。

10 就一方面而言，本发明涉及与 SEQ ID NO:2 具有至少 80% 序列一致性之分离蛋白质。在较佳之具体例中，该蛋白质系自肺炎链球菌分离，或者该蛋白质系一重组蛋白质。较佳的是，该分离蛋白质降解人类补体蛋白质 C3。本发明较佳之蛋白质为具有包含 SEQ ID NO:2 之氨基酸序列之分离蛋白质，更佳的是该氨基酸序列即 SEQ ID NO:2。术语“分离”在本文意指已由其自然环境移出之一天然发生物质或合成物质。术语“蛋白质”在本文中包括一或多个功能性单位，其包括一或多个肽或多肽。

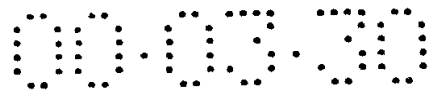
15 本发明亦相关自本发明 C3-降解性蛋白酶所分离之肽或多肽。较佳的是，本发明提供至少 15 个连续氨基酸之肽或多肽，其来自与 SEQ ID NO:2 具有 80% 序列一致性之分离蛋白质，且更佳的是提供 SEQ ID NO:2 中至少 15 个连续氨基酸之肽或多肽。在本发明另一方面，该肽或多肽能够降解人类补体蛋白质 C3。

20 本发明较佳之具体例包括具有 SEQ ID NO:2 中约第 1 个至约第 58 个氨基酸之分离蛋白质及具 SEQ ID NO:1 或其互补链中约第 1 至约第 174 个核苷酸。较佳的是，该分离核酸片段包括 SEQ ID NO:1 或其互补链中约第 150 至约第 174 个核苷酸。

25 在另一方面，本发明涉及降解人类补体蛋白质 C3 之分离蛋白质，其中编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。

30 本发明亦涉及免疫系统刺激性组合物（较佳为疫苗），其包括一有效量之免疫系统刺激性肽或多肽，该肽或多肽具有衍生自一种蛋白质之至少 15 个连续氨基酸，该蛋白质与 SEQ ID NO:2 有至少 80% 序列一致性且能降解人类补体蛋白质 C3。

较佳的是，该蛋白质系自肺炎链球菌分离。在一具体例中，该免疫系统刺激



性组合物或疫苗进而包括至少一种自肺炎链球菌所分离之其它免疫系统刺激性肽，多肽或蛋白质。

本发明进而涉及一种能够与 SEQ ID NO:2 有至少 80% 序列一致性且能降解人类补体蛋白质 C3 之蛋白质结合（典型上专一性结合）之抗体。在一具体例中，  
5 该抗体为单株抗体，且在另一具体例中，该抗体为多株抗体。在另一具体例中，该抗体为抗体片段。该抗体或抗体片段可得自小鼠，大鼠，山羊，鸡，人，或兔。

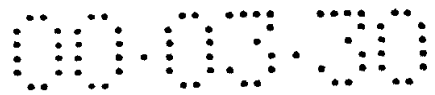
在另一具体例中，该抗体能够结合至一蛋白质之至少一部份，其中编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。

本发明亦涉及可与 SEQ ID NO:1 或其互补链以在高度严格之杂交条件下进行杂交之分离核酸片段（聚核苷酸）。在本文中所指的高度严格之杂交条件包括，  
10 例如，以 6X SSC, 5X Denhardt, 0.5% SDS, 及 100 微克/毫升片段化及变性之鲑鱼精子 DNA 在 65°C 进行杂交一整夜，并一次在 2X SSC, 0.1% SDS 中于室温下清洗约 10 分钟，接着一次于 65°C 清洗约 15 分钟，接着至少有一次在 0.2X SSC, 0.1% SDS 中于室温下清洗至少 3-5 分钟。在一具体例中，该核酸片段系自肺炎链球菌分离，且在另一具体例中，该核酸片段编码一蛋白质之至少一部份。在一  
15 具体例中，该蛋白质降解人类补体蛋白质 C3。在另一具体例中，该核酸片段编码一种不会降解人类补体 C3 之肽或多肽。

在另一具体例中，该核酸片段系在核酸载体中，且该载体为能够生产至少蛋白质的一部份的表达载体。包含该核酸片段的细胞亦于本发明中规划。在一具体  
20 例中，该细胞为细菌或真核细胞。

本发明进而涉及包括 SEQ ID NO:1 核酸序列或其互补链之分离核酸片段。本发明进而涉及由包括 SEQ ID NO:1 之双链 DNA 所转录之 RNA 片段。

在本发明另一方面，本发明涉及一种在哺乳动物（特别是人类）产生对抗肺炎链球菌的免疫反应的方法，其包括将一组合物施用至动物的步骤以产生对抗该  
25 蛋白质的免疫反应，该组合物包括治疗有效量之一蛋白质的至少一部份，其中编码该蛋白质之核酸可与 SEQ ID NO:1 或其互补链在高度严格之杂交条件下杂交。该免疫反应可为 B 细胞反应，T 细胞反应，上皮细胞反应或内皮细胞反应。在一较佳的具体例中，该组合物为疫苗组合物。较佳的是该蛋白质为至少 15 个氨基酸长度，且亦较佳的是该组合物进而包括至少一种来自肺炎链球菌之其它免疫系统  
30 刺激性肽，多肽或蛋白质。在一具体例中，该蛋白质包括 SEQ ID NO:2 中至少 15 个氨基酸。



本发明进而涉及来自肺炎链球菌之约 15kDa 至 约 25kDa 之能够降解人类补体 C3 之分离蛋白质，并涉及一种抑制肺炎链球菌-媒介性 C3 降解作用的方法，其包括的步骤为：将肺炎链球菌细菌接触一抗体，该抗体能够与 SEQ ID NO:2 有至少 80% 序列一致性之蛋白质结合。

5 本发明亦涉及一种抑制 C3-媒介性 发炎反应及异种移植时排斥作用之方法，其包括在异种移植时所用的动物器官表面上表达具有 SEQ ID NO:2 氨基酸序列之蛋白质。此方法特别有利于在例如猪肾引发本发明所述蛋白质之表达，且藉此抑制异种移植后之 C3 媒介性发炎。

10 本发明亦涉及一分离核酸分子，其包含一至少 15 个核苷酸的区域而可与 SEQ ID NO:1 或其互补链中至少一区域在高度严格之杂交条件下杂交。在一具体例中，分离核酸分子可与 SEQ ID NO:1 或其互补链中至少一区域在高度严格之杂交条件下杂交。较佳的是，至少一区域系包括 SEQ ID NO:1 之核苷酸 1-174 或 320-492。

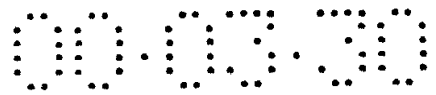
15 在另一具体例中，本发明亦涉及一分离核酸分子，其包含一至少 15 个核苷酸的区域而可与 SEQ ID NO:4 或其互补链中至少一区域在高度严格之杂交条件下杂交。在一具体例中，分离核酸分子可与 SEQ ID NO:4 或其互补链中至少一区域在高度严格之杂交条件下杂交。较佳的是，至少一区域系包括 SEQ ID NO:4 之核苷酸 507-681 或 827-999。

20 在另一具体例中，SEQ ID NO:4 中至少一部份的核酸分子编码一蛋白质的至少一部份。较佳的是，该部份具有如 SEQ ID NO:5 所示之预测氨基酸序列，且具有经例如 SDS-PAGE 测定中约 75kDa 至约 85kDa 之分子量。

本发明亦涉及具有如 SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO:8 及 SEQ ID NO:9 之核酸序列之分离 DNA 片段或引物。

25 在本发明另一方面，本发明涉及一免疫系统-刺激性组合物或疫苗，其包含治疗有效量之一蛋白质之至少一部份，其中编码该蛋白质之核酸可与 SEQ ID NO:4 或其互补链在高度严格之杂交条件下杂交。本发明涉及一种在哺乳动物（特别是人类）产生对抗肺炎链球菌的免疫反应的方法，其包括将一组合物施用至动物的步骤以产生对抗该蛋白质的免疫反应，该组合物包括治疗有效量之一蛋白质的至少一部份，其中编码该蛋白质之核酸可与 SEQ ID NO:4 或其互补链在高度严格之杂交条件下杂交。

30 在另一具体例中，本发明涉及一疫苗或免疫系统-刺激性组合物，其包含治



疗有效量之一蛋白质之至少一部份以及药学上可接受的载体，其中该蛋白质可有效地免疫接种或治疗哺乳动物以对抗肺炎链球菌的感染或侵入繁殖。该蛋白质系衍生自可与 SEQ ID NO:4 所示核酸序列或其互补链在高度严格之杂交条件下杂交之核酸分子。较佳的是，该蛋白质系以能够对于该哺乳动物个体（特别是人类个体）有效提供治疗效果的份量予以供给。

#### 图式简单说明

图 1 提供本发明之一 C3-降解性蛋白酶基因之翻译部份的核酸序列 (SEQ ID NO: 1)。

10 图 2 提供本发明之一 C3-降解性蛋白酶基因之氨基酸序列 (SEQ ID NO: 2)。

图 3 提供本发明之一 C3-降解性蛋白酶基因之氨基酸序列，其与编码本发明之一 C3-降解性蛋白酶基因之核酸序列并列 (SEQ ID NO: 2-3)。

图 4 提供预测 79kDa 氨基酸序列之核酸序列 (SEQ ID NO: 4)。

图 5 提供预测 79kDa 氨基酸序列 (SEQ ID NO: 5)。

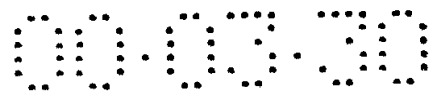
15 图 6 显示 SEQ ID NO: 1 与 SEQ ID NO: 4 之序列比对。

图 7 显示 SEQ ID NO: 2 与 SEQ ID NO: 5 中对应氨基酸 169-331 之序列比对。

#### 较佳具体例之详述

20 本发明涉及鉴定及分离分子量约为 20kDa ( $\pm 5$ kDa) (在 10% SDS-PAGE 胶上) 之人类补体 C3 降解性蛋白酶及编码 C3 降解性蛋白酶之核酸。已观察到来自数个血清型之肺炎链球菌在对数期生长阶段的培养物能首先降解 C3 之  $\alpha$ -链随后降解  $\beta$ -链，不会产生既有定义之 C3 切割片段 (Angel 等人, *J. Infect. Dis.* 170:600-608, 1994)。此类不发生切割之降解型态实质上不同于其它微生物之产物，如绿脓杆菌 (*Pseudomonas aeruginosa*) 之弹性蛋白酶及溶组织内阿米巴 (*Entamoeba histolytica*) 之半胱氨酸蛋白酶。

25 在本文中所使用的术语“降解”意指将氨基酸自蛋白质性分子移除，产生肽或多肽。如同在聚丙烯酰胺胶上的观察，本发明的蛋白质降解 C3 而不会产生特异性的切割片段。本发明之 C3-降解性蛋白酶对于 C3 有一些程度之较佳，例如，该  
30 C3-降解性蛋白酶并不会降解其它的蛋白质，如白蛋白。



约 20kDa 之 C3-降解性蛋白酶系自一插入性中断的肺炎链球菌基因库中鉴定出与野生型肺炎链球菌相比具有减弱之 C3 降解活性的克隆而予以分离。实施例 1 提供评估克隆之 C3-降解活性之例示性试验。鉴定具有减弱 C3 降解活性的克隆并基于该显示减弱 C3 降解活性的克隆选出一 546bp 之 SmaI 嵌入物。此 SmaI 片段用作由 CP1200 菌株制得的肺炎链球菌基因库的探针。来自肺炎链球菌基因库而能与该 SmaI 片段杂交之阳性克隆被分离，并鉴定与 C3-降解活性相关基因之开放阅读框。下列与 PspA 之一部份具有序列一致性之寡核苷酸 (SEQ ID NO:10) 在区别性杂交反应中被用于证实编码 C3-降解性蛋白酶之基因与编码 PspA 的基因截然不同。

10

SEQ ID NO:10

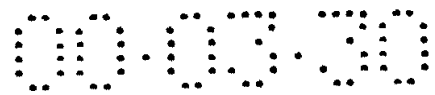
GAAAACAATAATGTAGAAGACTACTTTAAAGAAGCTTTAGA

20kDa 蛋白质之完整开放阅读框有 492 个碱基对 (SEQ ID NO:1)，预测分子量为 20kDa (+/-5kDa) 或约有 163 个氨基酸 (SEQ ID NO:2) 之蛋白质。编码 C3-降解性蛋白质之例示性基因序列示于图 1 (SEQ ID NO:1)，而该蛋白质之氨基酸序列示于图 2 (SEQ ID NO:2)。图 3 将较佳的基因序列与对应之较佳翻译蛋白质结合成为 SEQ ID NO: 3。

使用 SEQ ID NO: 2，确定该蛋白质之氨基酸序列与在 GenBank 或瑞士 Prot 数据库中的其它蛋白质无关。预测的蛋白质具有原核细胞膜蛋白质之富含脯氨酸 (特别是介于氨基酸第 80-108 个之间) 之序列特征，此建议该蛋白质系表达于细胞表面。该氨基酸序列并无显示明显的胆碱-结合性重复序列。将来自 CP1200 培养物之肺炎链球菌溶解物及上清液在以 C3 浸润之 SDS-PAGE 胶上进行电泳，在上清液及溶解物两者均鉴定出约 20kDa ( $\pm$ 5kDa) 之条带 (band)，证实具有如 SEQ ID NO: 2 所预测大小的蛋白质拥有 C3-降解活性 (见实施例 2)。如实施例 3 所述，编码该 20kDa 之 C3-降解性蛋白酶的基因在至少代表 5 个血清型 (血清型 1, 3, 4, 14, 及 19F) 之 24 个肺炎链球菌分离株中俱有保守性。

编码本发明 C3-降解性蛋白酶之全长基因被嵌入一基因表达载体以于大肠杆菌 (*E. coli*) 中表达。如实施例所述，将重组性 C3-降解性蛋白酶进行分离。本领域中一般技术人士将会了解，就一特定基因序列 (如 SEQ ID NO:1 所提供者) 而言，可以使用不同的表达载体来表达该基因。进而，可以使用本领域中不同的

30



熟知方法来生产及分离本发明之重组蛋白质，且本领域中一般技术人士亦会了解，本发明之 C3 降解性试验将测定除了在实施例中所提出的表达系统以外的特定表达系统是否具有功能，而毋须进一步的试验。在基础技术文献中可以发现许多不同的分子及免疫学技术，如 Sambrook 等人 (*Molecular Cloning-A Laboratory*  
5 *Manual*, 1989 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) 及 Harlow 等人 (*Antibodies;A Laboratory Manual*. Cold Spring Harbor, NY; Cold Spring Harbor Laboratory Press, 1988)。

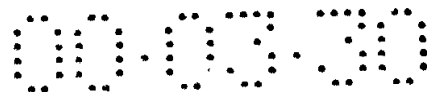
编码本发明 C3 降解性蛋白质之基因系使用来自 CP1200 菌株之肺炎链球菌基因组 DNA 片段所制得之质粒基因库进行鉴定。虽然已知有许多不同的方法可以得到质粒基因库，以较佳的策略而言，质粒基因库之构筑系将来自肺炎链球菌 CP1200  
10 菌株 (得自 D.A.Morrison, University of Illinois, Champagne-Urbana, Illinois, 且如 Havarstein LF 等人 *Proc. Natl. Acad. Sci. (USA)* 1995;92:11140-11144 所描述) 之经 *Sau3A* 切割之肺炎链球菌基因组 DNA 片段 (0.5-4.0kb)，插入一嵌入性穿梭 (shuttle) 载体 pVA891 (*erm<sup>r</sup>*, *cm<sup>r</sup>*; 具有大肠杆菌的复制源) 之 *Bam*HI 位置。将此基因库利用电击穿孔去转化大肠杆菌 DH5  
15  $\alpha$  MCR 菌株。自一些随机选择的大肠杆菌转化株中抽取质粒，显示所有转化株俱含有重组性质粒。

质粒基因库 DNA 可以自大肠杆菌转化株中抽取，并用于转化 CP1200 亲代肺炎链球菌株，该转化系使用藉由同构型重组之插入性突变技术完成。

肺炎链球菌菌株 CP1200 细胞可以在 CTM 培养基中使用以 HCl 移动 pH 之步骤  
20 转变为胜任细胞。将胜任细胞以少量分装方式贮于 -70°C 直到需要使用为止。

本发明之分离蛋白质与人类补体 C3 在 PBS 的存在下于 37°C 培育 4 小时以检测 C3 降解作用。不具有分离肺炎链球菌蛋白质之对照组样本可以作为具比较性目的之对照组。

本发明之蛋白质于 10% SDS-聚丙烯酰胺胶上具有明显分子量约 20kDa (±  
25 5kDa) 且较佳的是具有分子量约 15kDa 至约 25kDa。一示范性蛋白质序列示于 SEQ ID NO:2。本领域之一般技术人士将了解，在氨基酸序列中有一些变异是可以预期的，且这些变异不应自本发明之范围内排除。例如，保守性突变并不自本发明范围中排除，氨基酸序列一致性中少于约 80% 氨基酸序列一致性 (且较佳的是少于约 90%  
30 氨基酸序列一致性) 之变异亦不自本发明范围中排除，其中该蛋白质能够降解人类补体蛋白质 C3 且特别是该蛋白质系分离或原始得自一肺炎链球菌细菌。该蛋白



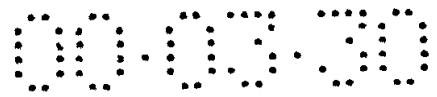
质之片段亦在本发明之范围中，特别是当这些片段能降解人类补体蛋白质 C3 时。

在肺炎链球菌菌株及血清型中可以预期有一些核酸序列变异，正如预期会有一些氨基酸变异一样。保守性氨基酸取代系本领域中已知的，且包括，例如，使用与该氨基酸属于同一类的其它成员氨基酸取代。例如，非极性氨基酸包括丙氨酸，白氨酸，异白氨酸，缬氨酸，脯氨酸，苯基丙氨酸，及色氨酸。极性电中性氨基酸包括甘氨酸，丝氨酸，苏氨酸，半胱氨酸，酪氨酸，天门冬酰胺及谷氨酰胺。阳性电荷（碱性）之氨基酸包括精氨酸，赖氨酸及组氨酸。阴性电荷（酸性）之氨基酸包括天门冬氨酸及谷氨酸。此类改变预期不会影响聚丙烯酰胺胶电泳所测定的分子量或等电点。特别较佳的保守性取代作用包括，但不限于，将 Lys 置换为 Arg（且反之亦然）以维持阳性电荷；将 Glu 置换为 Asp（且反之亦然）以维持阴性电荷；将 Ser 置换为 Thr 以维持自由态 OH；且将 Gln 置换为 Asn 以维持自由态 NH<sub>2</sub>。

本发明较佳的蛋白质包括具有氨基酸序列 SEQ ID NO:2 之蛋白质。在本发明中规划之其它蛋白质包括可降解人类补体蛋白质 C3 之蛋白质，且编码该蛋白质之核酸可与 SEQ ID NO:1 在高度严格之杂交条件下杂交，高度严格之杂交条件为例如以 6X SSC, 5X Denhardt, 0.5% SDS（十二烷基磺酸钠），及 100 微克/毫升片段化及变性之鲑鱼精子 DNA 在 65°C 进行杂交一整夜，并一次在 2X SSC, 0.1% SDS 中于室温下清洗约 10 分钟，接着一次于 65°C 清洗约 15 分钟，接着至少有一次在 0.2X SSC, 0.1% SDS 中于室温下清洗至少 3-5 分钟。典型上，SSC 溶液包含氯化钠，柠檬酸钠及水以制备贮存溶液。亦可使用该蛋白质之肽或多肽。本发明之较佳蛋白质包括 SEQ ID NO:2 中约第 1 个至约第 50 个氨基酸。

本发明蛋白质可以蛋白质的形式予以分离或制备。换言之，编码该蛋白质或该蛋白质一部份之核酸可以被嵌入表达载体或嵌入细胞的染色体中，以便在该细胞中表达该蛋白质。该蛋白质可以自细菌或另一细胞（较佳的是真核细胞，且更优选的是动物细胞）中纯化。或者，该蛋白质可以自表达该蛋白质之细胞中分离，如肺炎链球菌细胞。肽或多肽亦与本发明中虑及。该肽或多肽较佳为至少 15 个氨基酸的长度且较佳的肽或多肽具有至少来自 SEQ ID NO:2 之 15 个连续氨基酸。

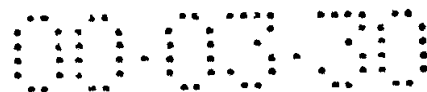
编码该 20kDa 蛋白质之核酸亦为本发明之一部份。SEQ ID NO:1 为编码 C3-降解性蛋白酶之优选核酸片段。本领域中一般技术人士将了解，一些取代并不会使 C3-降解性蛋白酶序列改变致使该 C3-降解性蛋白酶之特征及性质到达实质上改变的程度。例如，本发明亦规划与 SEQ ID NO: 1 具有至少 80%一致性的核酸。



决定一特别核酸序列是否落于本发明范围的方法应思考该特别核酸序列是否编码 C3-降解性蛋白酶及与 SEQ ID NO: 1 相较是否具有至少 80%之核酸一致性。编码 C3 蛋白酶的其它核酸序列包括那些编码 C3 蛋白酶的核酸，其中该 C3 蛋白酶与 SEQ ID NO:2 相较具有相同序列或至少 90%序列一致性，且包括涉及该核酸序列之简并性情形。简并性密码子意指有不同三字密码子可用于指定相同的氨基酸。例如，在本领域中熟知下列的 RNA 密码子（且因此，相关的 DNA 密码子，其中以 T 取代 U）可以互相交换用于编码每个特定氨基酸：

	苯丙氨酸 (Phe 或 F)	UUU, UUC, UUA 或 UUG
	白氨酸 (Leu 或 L)	CUU, CUC, CUA 或 CUG
10	异白氨酸 (Ile 或 I)	AUU, AUC 或 AUA
	甲硫氨酸 (Met 或 M)	AUG
	缬氨酸 (Val 或 V)	GUU, GUC, GUA, GUG
	丝氨酸 (Ser 或 S)	AGU 或 AGC
	脯氨酸 (Pro 或 P)	CCU, CCC, CCA, CCG
15	苏氨酸 (Thr 或 T)	ACU, ACC, ACA, ACG
	丙氨酸 (Ala 或 A)	GCU, GCG, GCA, GCC
	色氨酸 (Trp)	UGG
	酪氨酸 (Tyr 或 Y)	UAU 或 UAC
	组氨酸 (His 或 H)	CAU 或 CAC
20	谷胺酰胺 (Gln 或 Q)	CAA 或 CAG
	天门冬酰胺 (Asn 或 N)	AAU 或 AAC
	赖氨酸 (Lys 或 K)	AAA 或 AAG
	天门冬氨酸 (Asp 或 D)	GAU 或 GAC
	谷氨酸 (Glu 或 E)	GAA 或 GAG
25	半胱氨酸 (Cys 或 Q)	UGU 或 UGC
	精氨酸 (Arg 或 R)	AGA 或 AGG
	甘氨酸 (Gly 或 G)	GGU 或 GGC 或 GGA 或 GGG
	终止密码子	UAA, UAG 或 UGA

进而，可以将一特别的 DNA 序列进行修正而使用一特别细胞类别所偏好的密码子。例如，对于大肠杆菌之偏好密码子之使用，以及动物（包括人类）之偏好密码子均为已知。这些改变为本领域中一般技艺人士熟知，且因此这些基因序列

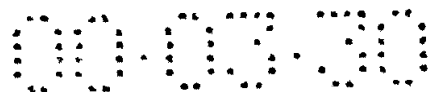


被认为是本发明之一部份。其它的核酸序列包括来自 SEQ ID NO:1 之长度为至少 15 个（且较佳为至少 30 个）核酸之核酸片段，或其它长度为至少 15 个（且较佳为至少 30 个）核酸之核酸片段，其中该片段可与 SEQ ID NO:1 在高度严格之杂交条件下杂交，高度严格之杂交条件例如以 6X SSC, 5X Denhardt, 0.5% SDS  
5 （十二烷基磺酸钠），及 100 微克/毫升片段化及变性之鲑鱼精子 DNA 在 65°C 进行杂交一整夜，并一次在 2X SSC, 0.1% SDS 于室温下清洗约 10 分钟，接着一次于 65°C 清洗约 15 分钟，接着至少一次在 0.2X SSC, 0.1% SDS 中于室温下清洗至少 3-5 分钟。

本发明之核酸片段可编码 SEQ ID NO:2 或 SEQ ID NO:5 之全部，可不含  
10 SEQ ID NO:2 或 SEQ ID NO:5（即无法转录的片段，包括基因调节性区域之片段或其它类似者）或可编码 SEQ ID NO:2 或 SEQ ID NO:5 之部份，较佳的是包含编码来自 SEQ ID NO:2 或 SEQ ID NO:5 之至少 9 个氨基酸之连续性核酸片段。因为在本发明中规划编码 C3 蛋白酶之一部份的核酸片段，将可了解的是并非所有的核酸片段都编码具 C3-降解活性之蛋白质或肽或多肽。进而，本发明之  
15 核酸可以进行突变以移除（或用其它方式不活化）该蛋白质之 C3 降解活性。因此，本发明亦规划符合上述杂交条件而具 C3-降解活性之片段。使核酸序列突变或以其它方式改变核酸序列的方法已于本领域中详述，且在免疫产生性及酵素特性方面不活化之蛋白质生产可以针对治疗用途进行测试。

本发明之核酸片段可以被并入核酸载体中或稳定地并入宿主基因组中以生产  
20 重组蛋白质，包括重组性嵌合（chimeric）蛋白质。在一具体例中，C3-降解性蛋白质系由在载体中的基因所编码，且该载体系在细胞中。较佳的是该细胞为原核性细胞，如细菌。基因及基因片段可以该基因的全部或部份与另一基因融合之方式存在，且该 C3-降解性蛋白质可以一个或多个蛋白质之融合蛋白质方式存在，该融合蛋白质系以单一蛋白质方式表达。本发明之许多相异核酸载体为本领域已  
25 知的且包括许多商业上可用的表达质粒或病毒载体。这些载体的使用为本领域中一般技术人士所已知的。在实施例中所使用者系一些例示性载体，但不应被认为对本发明范围有所限制。

本发明亦涉及能结合（典型上专一性结合）至来自肺炎链球菌之约 20kDa 之蛋白质（较佳的是约 15kDa 至约 25kDa 之蛋白质）之抗体，且较佳的是其中该蛋  
30 白质能够降解人类补体 C3。可以针对该蛋白质之全部或一部份制备多株抗体。类似地，可以针对本发明之约 20kDa 之 C3 降解性蛋白质之全部或其中一肽或多肽（片



段) 制备单株抗体。针对蛋白质制备抗体的方法详见于, 例如, Harlow 等人, (*Antibodies; A Laboratory Manual*. Cold Spring Harbor, NY; Cold Spring Harbor Laboratory Press, 1988)。在一较佳的实例中, 该抗体可自人类, 大鼠, 小鼠, 山羊, 鸡, 或兔子衍生。蛋白质结合性抗体片段及嵌合片段亦属熟知且在本发明之范围中。

5 本发明亦涉及免疫刺激性组合物的使用。术语“免疫刺激性”或“免疫系统刺激性”组合物意指本发明之蛋白质, 肽或多肽组合物, 它激活个体(如一哺乳动物)免疫系统中至少一种细胞型态, 较佳的是, 该免疫刺激性组合物在正常(未感染的)个体内可提供一种免疫接种反应或预防效果, 其典型上为一疫苗。然而, 在治疗应用或过程中任何可测得的免疫反应均有利于个体。免疫系统中优选的被活化细胞包括吞噬性细胞, 如嗜中性白血球或巨噬细胞, T 细胞, B 细胞, 上皮细胞及内皮细胞。包括本发明之肽, 多肽或蛋白质之免疫刺激性组合物可被用于在动物中产生抗体, 如大鼠, 小鼠, 山羊, 鸡, 兔子或人类, 或用于研究肺炎链球菌感染之动物模式。较佳的免疫刺激性组合物包括免疫刺激性份量(例如治疗有效量)之至少一种肽或多肽, 其包括来自该 C3 降解性蛋白酶之至少 15 个氨基酸。

10 术语“疫苗”意指用于免疫接种之组合物。免疫作用的过程可包括施用蛋白质, 肽, 多肽, 抗原, 核酸序列或互补性(例如反义)序列, 或抗体或其悬浮液, 其中施用该分子将可产生活性免疫力及提供保护以对抗肺炎链球菌感染或侵入繁殖。典型上, 此等疫苗被制成可注射之液体溶液或悬浮液。亦可制成适合在注射前溶解或悬浮于液体中的固体形式。该疫苗制备物可选择性地加以乳化, 或被包裹于微脂体中。

25 该免疫刺激性组合物(如一疫苗)可进而包括在药学上可接受的缓冲液或载剂中的其它蛋白质, 这些缓冲液或载剂如 PBS(磷酸盐缓冲盐水), 或在本领域中其它被认为适合及安全用于将蛋白质导入宿主以刺激免疫系统之缓冲液。该免疫刺激性组合物亦可包括其它免疫系统刺激性蛋白质, 如佐剂或来自肺炎链球菌或其它生物之免疫刺激性蛋白质, 肽或多肽。例如, 控制肺炎链球菌感染最有利的可能是肽或多肽之调配混合物(cocktail)。较佳的是在一疫苗制备物中使用本发明蛋白质之一个或多个片段以对抗或限制肺炎链球菌之侵入繁殖或肺炎链球菌侵入繁殖所带来之致病结果。

30 在本文中所使用“治疗有效量”意指有效产生一意欲结果之份量。该份量视个体免疫系统的健康或物理(即抗体合成)状况, 意欲保护的程

及其它相关因素而定。可预期的是该份量将落于一相当广泛的范围，且可经由例行性试验进行决定。

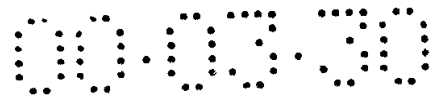
5 活性免疫刺激性成份经常与赋形剂或稀释剂混合，这些赋形剂或稀释剂为药学上可接受的载剂且与活性成份兼容。术语“药学上可接受的载剂”意指在与一组合中其它成份兼容且不危害接受者之意义上为“可接受的”载剂。合适的赋形剂为，例如，水，盐水，右旋糖，甘油，乙醇或类似物，以及其混合物。此外该免疫刺激性组合物（包括疫苗）可依喜好包含少量之辅助性物质，如湿润或乳化剂，pH 缓冲剂，和/或可以增进该免疫刺激性组合物效果之佐剂。

10 有效的佐剂或载剂包括但不限于氢氧化铝，N-乙酰基-壁酰基-L-苏胺酰基-D-异谷胺酰胺（thr-MDP）；N-乙酰基-新-壁酰基-L-丙胺酰基-D-异谷胺酰胺（CGP 11637，被称为 nor-UDP），N-乙酰基壁酰基-L-丙胺酰基-D-异谷胺酰基-L-丙氨酸-2-（1'，2'-二棕榈酰基-sn-甘油-3-羟基磷酸基氧）-乙基胺（CGP 19835A，被称为 MTP-PE），及 RIBI，其于 2%之角鲨烯/Tween 80 乳化液中包含三种自细菌抽取的成份，单磷酸基脂 A，二霉菌酸海藻糖及细胞壁架构（MPL+TDM+CWS）。

15 本发明亦涉及抑制肺炎链球菌-媒介性 C3 降解作用的方法，包括将肺炎链球菌与一蛋白质（如一抗体）接触，或另一能与来自肺炎链球菌之约 15kDa 至约 25kDa 之分离蛋白质结合之蛋白质。能与约 15kDa 至约 25kDa 之分离蛋白质结合之蛋白质可为一抗体或其片段，或该蛋白质可为嵌合蛋白质，该嵌合蛋白质包括来自抗体之抗体结合区域（如一可变性区域），该抗体能专一性辨识来自肺炎链球菌之约  
20 15kDa 至约 25kDa 且具有 C3 降解活性之分离蛋白质。

25 本发明之分离肺炎链球菌蛋白质可被分离，且可选择性地予以纯化，且可以使用该分离蛋白质或其免疫产生性片段来产生免疫学反应，在一实施例中，其包括在人类或实验动物身上之抗体反应。该蛋白质不具有 C3 降解活性之肽或多肽片段可针对其限制肺炎链球菌感染影响的能力进行试验。类似地，本发明之蛋白质之修正可例如经由突变来中断或不活化该蛋白质之 C3 降解活性。可以使用能够抑制本发明蛋白质之 C3-降解活性之抗体作为经由调理性路径来预防 C3 降解作用及促进肺炎链球菌清除之策略。该分离蛋白质可于试验中使用来检测对抗肺炎链球菌之抗体，或作为肺炎链球菌治疗用之疫苗之一部份，或作为包含一多价性或多个蛋白质，肽或多肽之疫苗。

30 因此，在本文所使用之术语“处理”意指对于正常哺乳动物个体或遭各种肺炎链球菌感染入侵，诊断为各种肺炎链球菌感染，或表达各种肺炎链球菌感染之



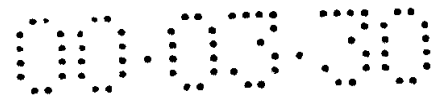
特征或症状之哺乳动物个体进行预防或治疗。术语“治疗”意指对于哺乳动物个体提供治疗性效果，使得该个体表达甚少或不表达肺炎链球菌感染或其它相关疾病之症状。此类处理可伴随本发明之核酸分子（具意义或反义者），蛋白质，肽或多肽或抗体之施用。

5 进一步的规划为本发明之蛋白质可以表达于脊椎动物细胞的表面且用于降解C3，例如，当补体沉积（或活化）成为难题时，例如在异种移植或补体-媒介性肾小球肾炎的情况。例如，完整的肺炎链球菌蛋白质，重组蛋白质，或以上两者任一之一部份，可以被并入异种移植细胞中，并以表面蛋白质或分泌蛋白质的方式表达以预防补体沉积（和/或补体-媒介性发炎反应）或使其减至最小。

10 本发明另一特别方面涉及使用表达分离蛋白质及由其而来之肽或多肽之疫苗载体。据此，在进而一方面，本发明提供一种在哺乳动物引发免疫反应的方法，其包括将表达本发明分离蛋白质和/或肽或多肽之至少一种或其混合物之疫苗载体提供至一哺乳动物。本发明之蛋白质及肽或多肽可使用活的疫苗载体传送至哺乳动物，特别是使用活的重组细菌，病毒或其它活的物质，其包含以外来多肽方式表达该蛋白质和/或肽或多肽时所需之遗传物质。特别是那些在胃肠道内繁殖的细菌已被发展成为疫苗载体，如沙门氏杆菌 (*Salmonella*)，志贺氏杆菌 (*Shigella*)，耶尔辛氏杆菌 (*Yersinia*)，弧菌 (*Vibrio*)，埃希氏菌 (*Escherichia*) 及 BCG，以上及其它实例在 J. Holmgren 等人, *Immunobiol.* 184, 157-179(1992)及 J. McGhee 等人, *Vaccine*, 10, 75-88 (1992) 中讨论。

20 本发明之额外具体例涉及在一个体（如哺乳动物）引发免疫反应的方法，包括将一份量之编码本发明之分离蛋白质和/或自其而来之肽或多肽之 DNA 分子（选择性联合一种感染-协助性因子）施用至该个体，其中该蛋白质和/或肽或多肽保留免疫产生性，且当该蛋白质和/或肽或多肽并入一免疫刺激性组合物（例如疫苗）并施用至人类时，可提供保护而不会在人类由肺炎链球菌病原引发后续感染时加重病情。感染-协助性因子为本领域中已知的。

25 进一步的规划为编码该 20kDa 蛋白质之基因之反义序列可作为抗肺炎链球菌感染之疫苗或治疗性处理。反义 DNA 被定义为非-编码性序列，其与 SEQ ID NO:1 之全部或部份互补（即一互补链）。例如，5'-ATGTCAAGC-3' 之反义序列为 3'-TACAGTTCG-5'。将反义序列或寡核苷酸传送至动物可造成动物产生抗体，或造成  
30 该序列并入活菌或其它细胞中致使该 20kDa 基因产物之全部或部份之转录和/或翻译受到抑制。



反义核酸序列的导入可例如经由将反义核酸装载于合适的载体（如微脂体）上而导入肺炎链球菌或受感染细胞内。典型上，具有 8 个或更多核苷酸之反义核酸序列能结合至细菌核酸或细菌讯息 RNA 上。反义核酸序列典型上包含至少约 15 个核苷酸（较佳为至少约 30 个核苷酸或更多核苷酸），以对于细菌核酸或细菌讯息 RNA 之杂交产物提供所需的稳定性。序列的导入较佳的是抑制至少一个内源性肺炎链球菌核酸序列之转录或翻译。装载反义核酸的方法在本领域中为熟知的，例如 U.S. 专利第 4, 242, 046 号。

本发明亦提供具有 2163 碱基之开放阅读框（SEQ ID NO: 4）之核酸，其包括开放阅读框（SEQ ID NO: 1）之核酸，其编码具有分子量约 20kDa（SEQ ID NO:2）之蛋白质。在本文所述之该 20kDa 蛋白质进而被定性为一 C3-降解性蛋白质。较大的开放阅读框，例如 2163bp（SEQ ID NO: 4），编码约 79kDa（SEQ ID NO:5）之假设性蛋白质。

在本文中所有引用的参考文献及发表文献俱并入本说明书作为参考资料。对于熟习该项技术者而言，有许多不同的可用替代技术及步骤可以类似地使实施者根据本说明书成功实施规划之发明。熟习该项技术者将会了解，虽然本发明已于上文中依据特定具体例及实例加以描述，本发明并未当然受到如此限制，且可以进行许多其它具体例，实例，用途，修正以及在这些具体例，实例及用途以外之情形，而不会偏离本申请案之发明范围。

## 20 实施例 1

### 具有减弱 C3-降解活性之嵌入性突变株之鉴定

嵌入性突变株系得自 Elaine Tuomanen 博士（Rockefeller Inst., New York, New York）。具有嵌入物之克隆于试验中进行测试以检测减弱 C3-降解活性。137 个克隆经由使该等细胞于 Todd Hewitt 培养液中在室温下于微效价盘上生长一整夜而进行测试。该等细胞以供肺炎链球菌之合成性培养基（见 Sicard A.M., *Genetics* 50:31-44, 1984）进行 1:10 之稀释，且剩余的细胞被冷冻于微效价盘中。将在 100 微升之培养基（包含 1 毫克/毫升在磷酸盐缓冲盐水（PBS）之 0.1% BSA）中的 63 纳克（ng）或 83 纳克之 C3 加入约 200 微升之经稀释细胞内。将细胞于 37 °C 培育 4 小时。将 100 微升之混合物加入 ELISA 盘中并于 4°C 培育一整夜。将该盘以清洗缓冲液进行清洗三次，且将在 PBS 中的 0.05% Tween 20 加入孔洞中，在每次清洗前进行 5 分钟培育。抗 C3 之 100 微升抗体（对人类 C3-IgG 部份具专



一性之多株西洋山俞菜过氧化氢酶-结合之山羊抗体, ICN Cappel, Costa Mesa, CA) 以在 PBS 中之 3% BSA 进行 1:1200 之稀释。该 ELISA 盘于 37°C 在黑暗中培育约 30 分钟至 1 小时, 并以上述清洗缓冲液进行清洗。将该试验使用 12 毫克 OPD (在 30 毫升 0.1M 柠檬酸钠缓冲液中) 以及 12 微升之 30%过氧化氢进行显影。试验结果于一 ELISA 盘读测仪上在 490 微毫米经由光学密度读数而决定。

每个克隆进行试验四次。与非突变对照组相较具有少于 40%之 C3 降解作用之 19 个克隆被选出。这 19 个克隆以上述试验进行筛选六次, 且由结果选出与非突变对照组相较具有少于 40%之 C3 降解作用之 6 个克隆。这 6 个克隆每个进行筛选 11 次, 且选出具有最低 C3-降解活性的 2 个克隆作进一步研究。

得到其中之一克隆之部份序列, 且将一 546bp 之 SmaI 片段以  $^{32}\text{P}$  进行随机引物标定 (得自 Stratagene 之套组, LaJolla, CA)。来自 SEQ ID NO:1 之 546bp 之 SmaI 片段与来自许多肺炎链球菌菌株经 EcoRI 及 KpnI 处理之切割物在 Southern 印迹上进行杂交。此相同的片段亦用于筛选得自肺炎链球菌菌株 CP1200 之基因组 DNA 之 Sau3A 片段基因库。

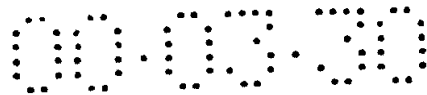
自 CP1200 基因库鉴定出 3.5kb 之嵌入物。将该嵌入物进行定序, 并鉴定出一 492 个碱基对之开放阅读框, 包括终止密码子。该开放阅读框编码 163 个氨基酸且预测分子量约为 18, 500 道尔顿之蛋白质。

构筑 PCR 引物以扩增开放阅读框; 该 5' PCR 引物包括一 BamHI 位置; 该 3' 引物包括一 PstI 位置。被扩增的嵌入物依符合阅读框方式连接至 His-标记之大肠杆菌表达载体 pQE30 (Qiagen, San Diego, CA)。将得到的质粒用于转化大肠杆菌菌株 BL21 (Novagen, Madison, WI), 其包含 lac 抑制物 (repressor) 质粒 pREP4 (Qiagen)。将大肠杆菌培养物进行诱发以表达 His-标记之蛋白质, 且该蛋白质以 Ni-NTA 树脂 (Qiagen) 进行管柱纯化。该纯化之蛋白质经由 SDS-PAGE 胶进行确认。

## 实施例 2

### 20kDa 之 C3-降解性蛋白酶之鉴定

为了测定 20kDa 蛋白质之 C3-降解能力, 将 0.5 毫克/毫升之 C3 (根据 Tack 等人, *Meth. Enzymol.* 80:64-101, 1984 进行制备) 于含有 15%丙烯酰胺之十二烷基磺酸钠 (SDS) 胶 (15% SDS-PAGE 胶) 上进行共聚。肺炎链球菌上清液系得自于 Todd Hewitt 培养液中生长至对数期之肺炎链球菌菌株 CP1200 培养物,



肺炎链球菌溶解物经由将  $5 \times 10^8$  个细胞与 5% SDS 在室温下培育 30 分钟而得到。经溶解物使用具 10,000MW 分离值之 Centricon 过滤装置 (Amicon, Beverly, MA) 进行浓缩 10 倍。样本在电泳前不进行加热。将上清液及溶解物的样本加入 15% 包含 C3 之 SDS-PAGE 胶, 且电泳系于  $4^\circ\text{C}$  在 150V 进行, 直到染料前缘跑出为止。

5 将胶连续以 50 毫升在水中的 2.5% Triton X-100 清洗 (2 次, 10 分钟), 以在 50mM Tris-HCl 中的 2.5% Triton X-100, pH7.4 (2 次, 10 分钟) 清洗, 及以 50mM Tris-HCl, pH7.4 清洗 (2 次, 10 分钟) 以移除 SDS。清洗后, 将 50 毫升之 50mM Tris-HCl, pH 7.4, 倒入含有这些胶的盘皿中, 将盘皿加盖并于  $37^\circ\text{C}$  培育 1.5 小时及一整夜 (约 16 小时)。将胶以考马士 (Coomassie) 蓝染色 10 分钟, 随后完全地脱色。

对照在溶解物及上清液中的暗蓝色背景, 可以观察到两个溶解性条带, 其中之一约为 20kDa 大小。在肺炎链球菌溶解物中的 C3 蛋白酶活性于  $37^\circ\text{C}$  培育 1.5 小时后观察到, 然而在 Pn 上清液中 C3 蛋白酶活性经一整夜培育后才观察到。因此, C3 蛋白酶活性显然主要与细胞相关。

15

### 实施例 3

编码该 20kD 蛋白质之基因在一些肺炎链球菌菌株中具保守性

自各种不同的肺炎链球菌菌株 (第 1 型, 第 3 型, L002 及 L003 (第 3 型), 第 4 型, 第 14 型之临床分离株及 CP1200, WU2, R6X, 6303, 109, 110, JY1119, 20 JY182, 及 JY53 之实验室分离株) 得到 DNA, 并将 SEQ ID NO:3 作为探针以检测编码该 20kD 蛋白质之核酸是否存在于来自这些菌株之 DNA 中。分离染色体 DNA 以 EcoRI 进行切割并经由电泳分开。将 DNA 转移至固体支持物上, 并与末端-标定的 SEQ ID NO:3 进行杂交, 杂交及清洗条件为以 6X SSC, 5X Denhardt's, 0.5% SDS 及 100 微克/毫升片段化且变性之鲑鱼精子 DNA 在  $65^\circ\text{C}$  进行杂交一整夜, 并一次在 2X SSC, 0.1% SDS 于室温下清洗约 10 分钟, 接着一次于  $65^\circ\text{C}$  清洗约 15 分钟, 接着两次在 0.2X SSC, 0.1% SDS 中于室温下清洗 3 分钟。

25

结果指出 SEQ ID NO:3 在每个受试 DNA 样本中同样有杂交反应, 指出该蛋白质显然在各菌株间具有保守性。在一些菌株中, 编码 20kDa 之 C3-降解性蛋白质之 DNA 显然为一较大之 2163bp 开放阅读框的一部份, 该开放阅读框假设性编码 30 79kDa 蛋白质。



#### 实施例 4

##### 肺炎链球菌 DNA/5F1 探针之 Southern 印迹分析

自 11 个肺炎链球菌菌株得到基因组 DNA 之 5 微克样本。每个样本以限制酶  
5 KpnI 切割。这些样本随后装载于琼脂糖胶上并经由电泳进行解析。包含在胶中的  
样本随后经由毛细移转而转移至得自 Amersham (Upsafla, Sweden) 之 Hybond-N+  
膜上。得自 5F1 分离株之一 540bp 之 SmaI 片段使用 <sup>32</sup>P QuickPrime 套组 (Pharmacia,  
Piscataway, NJ) 以 <sup>32</sup>P 进行随机引物标定, 并使用 NucTrap 管柱 (Stratagene,  
La Jolla, CA) 自非-并入的核苷酸中纯化, 并进行杂交。  
10 杂交条件为以 6X SSC, 5X Denhardt, 0.5% SDS, 及 100 微克/毫升片段  
化及变性鲑鱼精子 DNA 在 65°C 进行杂交一整夜, 并一次在 2X SSC, 0.1% SDS  
于室温下清洗约 10 分钟, 接着一次于 65°C 清洗约 15 分钟, 接着至少一次在 0.2X  
SSC, 0.1% SDS 中于室温下清洗至少 3-5 分钟。渍膜显示该 20kDa 基因系存在于  
所有肺炎链球菌的受试菌株。

15

#### 实施例 5

自 SEQ ID NO:1 制得两个 DNA 引物并用于扩增来自肺炎链球菌 (血清型 3)  
基因体 DNA 之 20kDa 基因序列。第一个引物为 5'-引物 (SEQ ID NO:6), 其自  
该 20kDa 基因之 ATG 起始密码子开展, 其中插入一 NcoI 位置, 且在 ATG 起始密码  
20 子之后具 Ala 残基以维持正确的阅读框。第二个引物为 3'-引物 (SEQ ID NO:7),  
其自该 20kDa 基因之终止密码子开展, 其中插入一 BamHI 位置。

5'-GGGGG CCA TGG CC TCA AGC CTT TTA CGT GAA TTG-3';

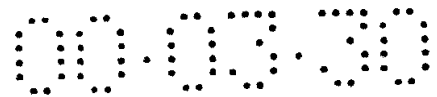
(SEQ ID NO:6)

25

5'-GGGGG GGA TCC CTA GCT ATA TGA GAT AAA CTT TCC TGC T-3';

(SEQ ID NO:7)

该两个引物系于 Applied Biosystems 380A DNA 合成仪 (Foster City,  
30 CA) 中使用购自 Glen Research (Sterling, VA) 之反应剂进行合成。扩增反应  
系使用 Perkin Elmer Thermocycler (ABI) 根据制造商的指示进行。经鉴定之



PCR 产物被连接至以 TA 为尾端之 PCR 选殖载体 pCR2.1 (其系得自 Invitrogen, Carlsbad, CA), 并用于转化 OneShot Top10F' 胜任细胞 (Invitrogen)。卡那霉素 (Kanamycin) 抗药性转化株经由将碱性溶解法制得之质粒 DNA 进行限制酶分析而加以筛选。鉴定出一约为 500bp 之嵌入物片段并随后以限制酶 NcoI 及 BamHI 进行切割。将该 500bp 片段自低融点琼脂糖胶中分离, 且随后连接至得自 Novagen (Madison, WI) 之 T7 激活性表达载体 pET28a 之 NcoI-BamHI 位置。

该连接混合物随后经转化送入 Top10F' 细胞 (Invitrogen), 且卡那霉素抗药性转化株经由将碱性溶解法制得之质粒 DNA 进行限制酶分析而加以筛选。重组性质粒 (pLP505) 随后经转化送入 BL21 (Novagen) 细胞并使其于添加 30 微克/毫升之卡那霉素之 SOB 培养基中生长。细胞生长至 O. D.<sub>600</sub> 值 0.6, 并随后以 0.4mM IPTG (Boehringer Mannheim, Indianapolis, Indiana) 进行诱发 2-4 小时。制备全细胞溶解物并于 14% SDS-PAGE 胶上进行电泳。将胶以考马士 (Coomassie) 蓝染色并检测到表达产物。该考马士蓝染色的胶显示在 28kDa 及 18kDa 分子量标记之间的条带, 且经测定约为 20kDa。

在该重组性 pLP505 质粒中嵌入物之 DNA 序列使用 ABI 370A DNA 测序仪得到。将该 DNA 序列与 SEQ ID NO: 1 之 DNA 序列使用 Pustell 之 MacVector DNA 矩阵点绘特征 (Oxford Molecular Group, Campbell, CA) 进行比对。将得自 pLP505 质粒, SEQ ID NO:1 及肺炎链球菌 (血清型 4) 基因组之 DNA 序列进行比对, 显示编码该 20kD 蛋白质之开放阅读框 (ORF) 可能为一较大 ORF 之部份, 即在血清型 4 之基因组中编码具有约 79kDa 预测分子量之蛋白质之 2163bp (SEQ ID NO:4)。DNA SEQ ID NO: 4 编码如 SEQ ID NO:5 所示之预测氨基酸序列。

肺炎链球菌 (血清型 4) 基因组序列系使用 MacVector 之 ClustalW 特征 (Oxford Molecular Group, Campbell, CA) 自网址为 [www.tigr.org](http://www.tigr.org) 之 Institute for Genomic Research 和/或经由网址为 [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) 之 NCBI 而得到。在该 20kDa 之氨基酸序列 (SEQ ID NO:2) 及预测为 79kDa 之氨基酸序列 (SEQ ID NO: 5) 之间进行序列比较。可以观察到 SEQ ID NO:2 中第 1-58 个氨基酸和第 90-132 个氨基酸分别与 SEQ ID NO:5 中第 170-227 个氨基酸和第 258-300 个氨基酸的序列基本上一致。包含这些特别区域的蛋白质及肽或多肽为本发明较佳的具体例。

基于可得的基因组 DNA (血清型 4) 序列, 设计在 2163bp ORF 两侧的引物并随后使用 ABI 380A DNA 合成仪进行合成 (SEQ ID NO:8 及 9)。SEQ ID NO:8 为一肺炎链球菌 5'-引物, 其具有插入的 NcoI 位置且在 ATG 起始密码子之后加入



Glu 残基以维持一正确的开放阅读框。SEQ ID NO: 9 为一肺炎链球菌 3'-引物，其具有插入的 HindIII 位置。

5' -AGA GCT CCT CCC ATG GAA GAT CCG AAT TAT CAG-3'; (SEQ ID NO:8)

5

5' -CCG GGC AAG CTT TTA CTT ACT CTC CT-3'; (SEQ ID NO:9)

10 一约 2100bpDNA 片段随后自 4 个不同的肺炎链球菌血清型 (血清型 3, 5, 6B 及 7) 扩增, 得到 4 个片段。4 个片段随后各自连接至 PCR 选殖性载体 PCR2.1 (Invitrogen), 并用于转化 OneShot Top10F' 细胞 (Invitrogen)。卡那霉素抗性转化株经由将碱性溶解法制得之质粒 DNA 进行限制酶分析而加以筛选。含有血清型 7 之 PCR 产物之重组性质粒被鉴定出来, 例如为 pLP512。使用 ABI 型号 370A DNA 测序仪得到血清型 7 之克隆之 DNA 序列。该 DNA 序列基本上同于 SEQ ID NO: 4 并编码一基本上同于 SEQ ID NO:5 之预测氨基酸序列。

15 熟习该项技术者将会了解, 虽然本发明已于上文中依据特定具体例及实施例加以描述, 本发明并未当然受到如此限制, 且可以进行许多其它具体例, 实例, 用途, 修正以及在上述具体例, 实例及用途以外之情形, 而不会偏离本申请案之发明范围。

## 说明书附图

1 ATGTCAAGCC TTTTACGTGA ATTGTATGCT AAACCCTTAT CAGAACGCCA  
 51 TGTAGAATCT GATGGTCTTA TTTTCGACCC AGCGCAAATC ACAAGTCGAA  
 101 CCGCCAATGG TGTTGCTGTA CCGCACGGAG ACCATTATCA CTTTATTCCT  
 151 TATTCACAAC TGTCACCTTTGGAAGAAAAA TTG GTCGTAT·TATTCCCCTT  
 201 CGTTATCGTT CAAACCATTG GGTACCAGAT TCAAAGACCAGAACAACCAG  
 251 TCCA CAATCG ACTCCGGGAA CCTAGTCCAA GTCCGAAACCTGCACCAAAT  
 301 CCTCAACCAG CTCCAAGCAA TCCAATTGATGAGAAATTGGTCAAAGAAAGC  
 351 TGTTCGAAAA GTAGGCGATG GTTATGTCTTTGAGGAGAAT GGAGTTGCCT  
 401 CGTTATATCC CAAGCCAAGG ATCTTACAGCAGAAACAGCAGCAGG CATTG  
 451 ATAGCAAAC TGGCCAAGCAG GAAA GTTTAT CTCATAAGCT AG

(SEQ ID NO: 1)

图 1

1 MSSLLRELYA KPLSERHVESDGLIFDPAQI TSRTANGVAV PHGDHYHFIP  
 51 YSQLSPLEEK LVLFFVIV QTIGYQIQRP EQPVHNRLRE PSPSPKPAPN  
 101 PQPAPSNPID EKLVEAVRK VGDGYVFEEN GVASLYPKPRILQOKQQQAL  
 151 LANWPSRKVY LIS\*

(SEQ ID NO: 2)

图 2



301 CCTCAACCAGCTCCAAGCAATCCAATTGATGAGAAATGGTCAAAGAA GCTGTTCGAAAA + 360

GGAGTTGGTCGAGGTTGGTTAGGTTAACTACTCTTTAACCAGTTCTTCGACAAGCTTTT

P Q P A P S N P I D E K L V K E A V R K -

361 GTAGGCGATGGTTATGTCTTTGAGGAGAATGGAGTTGCCTCGTTA TATCCAAGCCAAGG + 420

CATCCGCTACCAATACAGAACTCCTCTTACCTCAACGGAGCAATATAGGGTTCGGTTCC

V G D G Y V F E E N G V A S L Y P K P R -

421 ATCTTACAGCAAAACAGCAGCAGGCAATTGATAGCAAAGTGGCAAGCAAGAAAGTTTAT + 480

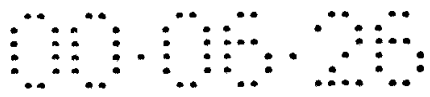
TAGAAATGTCGTCTTTGTCGTCGTCCGTAAGTATCGTTTGACCGGTTCCCTTTCAAATA

I L Q Q K Q Q Q A L I A N W P S R K V Y -

481 CTCATAAGCTAG + 492  
GAGTATTCGATC

L I S \*

图 3 (续)



ATGAAAGATCCGAATTATCAGTTGAAGGATTCAGACATTGTCAAT  
 GAAATCAAGGGTGGTTATGTTATCAAGGTAGATGGAAAATACTA  
 TGTTTACCTTAAGGATGCAGCTCATGCGGATAATATTCCGGACAAA  
 AGAAGAGATTAAACGTCAGAAGCAGGAACACAGTCATAATCACG  
 GGGGTGGTTCTAACGATCAAGCAGTAGTTGCAGCCAGAGCCCAA  
 GGACGCTATAACAACGGATGATGGTTATATCTTCAATGCATCTGAT  
 ATCATTGAGGACACGGGTGATGCTTATATCGTTCCTCACGGCGAC  
 CATTACCATTACATTCCTAAGAATGAGTTATCAGCTAGCGAGTTA  
 GCTGCTGCAGAAGCCTATTGGAATGGGAAGCAGGGATCTCGTCCT  
 TCTTCAAGTTCTAGTTATAATGCAAATCCAGCTCAACCAAGATTG  
 TCAGAGAACCACAATCTGACTGTCACTCCAACCTTATCATCAAAT  
 CAAGGGGAAAACATTTCAAGCCTTTTACGTGAATTGTATGCTAAA  
 CCCTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTCGAC  
 CCAGCGCAAATCACAAGTCGAACCGCCAGAGGTGTAGCTGTCCC  
 TCATGGTAACCATTACCACTTTATCCCTTATGAACAAATGTCTGA  
 ATTGGAAAACGAATTGCTCGTATTATCCCCTTCGTTATCGTTCA  
 AACCATTGGGTACCAGATTCAAGACCAGAACAACCAAGTCCACA  
 ATCGACTCCGGAACCTAGTCCAAGTCCGCAACCTGCACCAAATCC  
 TCAACCAGCTCCAAGCAATCCAATTGATGAGAAATTGGTCAAAG  
 AAGCTGTTTCGAAAAGTAGGCGATGGTTATGTCTTTGAGGAGAATG  
 GAGTTTCTCGTTATATCCCAGCCAAGGATCTTTCAGCAGAAACAG  
 CAGCAGGCATTGATAGCAAACCTGGCCAAGCAGGAAAGTTTATCT  
 CATAAGCTAGGAGCTAAGAAAACCTGACCTCCCATCTAGTGATCG  
 AGAATTTTACAATAAGGCTTATGACTTACTAGCAAGAATTCACCA  
 AGATTTACTTGATAATAAAGGTCGACAAGTTGATTTTGAGGCTTT  
 GGATAACCTGTTGGAACGACTCAAGGATGTCCCAAGTGATAAAG  
 TCAAGTTAGTGATGATATTCTTGCCTTCTTAGCTCCGATTGCTCA  
 TCCAGAACGTTTAGGAAAACCAAATGCGCAAATTACCTACACTG  
 ATGATGAGATTCAAGTAGCCAAGTTGGCAGGCAAGTACACAACA  
 GAAGACGGTTATATCTTTGATCCTCGTGATATAACCAGTGATGAG  
 GGGGATGCCTATGTAACCTCCACATATGACCCATAGCCACTGGATT  
 AAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGCAGCCCAGGC  
 TTATGCTAAAGAGAAAGGTTTGACCCCTCCTTCGACAGACCATCA  
 GGATTCAGGAAATACTGAGGCAAAGGAGCAGAAGCTATCTACA  
 ACCGCGTGAAAGCAGCTAAGAAGGTGCCACTTGATCGTATGCCTT  
 ACAATCTTCAATATACTGTAGAAGTCAAAAACGGTAGTTTAATCA

TACCTCATTATGACCATTACCATAACATCAAATTTGAGTGGTTTG  
ACGAAGGCCTTTATGAGGCACCTAAGGGGTATACTCTTGAGGATC  
TTTTGGCGACTGTCAAGTACTATGTTCGAACATCCAAACGAACGTC  
CGCATTTCAGATAATGGTTTTGGTAACGCTAGCGACCATGTTCAA  
GAAACAAAAATGGTCAAGCTGATACCAATCAAACGAAAAACCA  
AGCGAGGAGAAACCTCAGACAGAAAAACCTGAGGAAGAAACCC  
CTCGAGAAGAGAAACCGCAAAGCGAGAAACCAGAGTCTCCAAAA  
CCAACAGAGGAACCAGAAGAATCACCAGAGGAATCAGAAGAAC  
CTCAGGTCGAGACTGAAAAGGTTGAAGAAAAACTGAGAGAGGCT  
GAAGATTTACTTGGAAAAATCCAGGATCCAATTATCAAGTCCAAT  
GCCAAAGAGACTCTCACAGGATTAATAATAATTTACTATTTGGC  
ACCCAGGACAACAATACTATTATGGCAGAAGCTGAAAAACTATT  
GGCTTTATTAAAGGAGAGTAAG

图 4 (续)

MKDPNYQLKDS DIVNEIKGGYVIKVDGKYYVYLKDAAHADNIRTK  
 EEIKRQKQEHSNHGGSNDQAVVAARAQGRYTTDDGYIFNASDII  
 EDTGDAYIVPHGDHYHYIPKNELSAELAAAEAYWNGKQGSRPSS  
 SSYNANPAQPRLSENHNLTVTPTYHQNGENISSLLRELYAKPLSER  
 HVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARI  
 PLRYRSNHWPDSRPEQPSPQSTPEPSPSPQPAPNPQPAPSNPIDEKLV  
 KEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSKLAKQESLS  
 HKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALD  
 NLLERLKDVP SDKVKLVDDILAFLAPIRHPERLGKPN AQITYTDD EIQ  
 VAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTSHWIKKDSLSE  
 AERAAAQAYAKEKGLTPPSTDHQDSGNT EAKGAEAIYNRVKAAKK  
 VPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK  
 GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADT  
 NQTEKPSEEKQTEKPEEETPREEKQSEKPEPKPTEEP EESPEESEEP  
 QVETEKVEEKLREAE DLLGKIQDP IIKSNAKETLTGLKNLLFGTQD  
 NNTIMAEAEKLLALLKESK

图 5

## 序列对比

		10		20		30
20Kseq	ATG	TCAAGCCTTTT	ACGTGAATTGTATGCT			
79KfragDNA	ATT	TCAAGCCTTTT	ACGTGAATTGTATGCT			
		40		50		60
20Kseq	AAACCCTTATCAGAACGCCATGT	AGGAATCT				
79KfragDNA	AAACCCTTATCAGAACGCCATGT	GGGAATCT				
		70		80		90
20Kseq	GATGGT	CTTATTTTCGACCCAGCGCAAATC				
79KfragDNA	GATGGC	CTTATTTTCGACCCAGCGCAAATC				
		100		110		120
20Kseq	ACAAGTCGAACCGCCA	ATGGGTGT	TGCTGTA			
79KfragDNA	ACAAGTCGAACCGCCA	GAGGGTGT	AGCTGTC			
		130		140		150
20Kseq	CCGCACGGAG	ACCATTAT	CACTTTAT	TCCCT		
79KfragDNA	CCTCATGGTA	ACCATTAC	CACTTTAT	CCT		
		160		170		180
20Kseq	TATTCACAAC	TGTCACCT	TTGGAAGAA	AA	AA	
79KfragDNA	TATGAACAA	ATGTCGTGA	ATTGGA	AA	AA	CGA
		190		200		210
20Kseq	ATTGG	TCGTATTATTCCCTT	CGTTATCGT			
79KfragDNA	ATTGC	TCGTATTATTCCCTT	CGTTATCGT			

(SEQ ID NOS: 1 和 4)

图 6

	220	230	240
20Kseq	TCAAACCATTTGGGTACCAGATTCAAAGACC		
79KfragDNA	TCAAACCATTTGGGTACCAGATTCAAAGACC		
	250	260	270
20Kseq	AGAACCAACCAAGTCCACAATCGACTCCGGG		
79KfragDNA	AGAACCAACCAAGTCCACAATCGACTCCGGG		
	280	290	300
20Kseq	AACCTAGTCCAAGTCCGAACCTGCACCAAA		
79KfragDNA	AACCTAGTCCAAGTCCGAACCTGCACCAAA		
	310	320	330
20Kseq	ATCCTCAACCAGCTCCAAGCAATCCAATTG		
79KfragDNA	ATCCTCAACCAGCTCCAAGCAATCCAATTG		
	340	350	360
20Kseq	ATGAGAAATTGGTCAAAGAAGCTGTTCGAA		
79KfragDNA	ATGAGAAATTGGTCAAAGAAGCTGTTCGAA		
	370	380	390
20Kseq	AAGTAGGCGATGGTTATGTCCTTTGAGGAGA		
79KfragDNA	AAGTAGGCGATGGTTATGTCCTTTGAGGAGA		
	400	410	420
20Kseq	ATGGAGTTGCTCGTTATATCCCAAGCCAAG		
79KfragDNA	ATGGAGTTGCTCGTTATATCCCAAGCCAAG		
	430	440	450
20Kseq	GGATCTTACAGCAGAAACAGCAGCAGGCAT		
79KfragDNA	GGATCTTACAGCAGAAACAGCAGCAGGCAT		
	460	470	480
20Kseq	TGATAGCAAACCTGGCCAAGCAGGAAAGTTT		
79KfragDNA	TGATAGCAAACCTGGCCAAGCAGGAAAGTTT		
	490	500	510
20Kseq	ATCTCATAAGCTAG		
79KfragDNA	ATCTCATAAGCTAG		

(SEQ ID NOS: 1 和 4)

图 6 (续)

## 序列对比

		10	20	30																										
20Kseq	M	S	S	L	R	E	L	Y	A	K	P	L	S	E	R	H	V	E	S	D	G	L	I	F	D	P	A	Q	I	
79KfragDNA	I	S	S	L	R	E	L	Y	A	K	P	L	S	E	R	H	V	E	S	D	G	L	I	F	D	P	A	Q	I	
		40	50	60																										
20Kseq	T	S	R	T	A	N	G	V	A	V	P	H	G	D	H	Y	H	F	I	P	Y	S	Q	L	S	P	L	E	E	K
79KfragDNA	T	S	R	T	A	R	G	V	A	V	P	H	G	N	H	Y	H	F	I	P	Y	E	Q	M	S	E	L	E	K	R
		70	80	90																										
20Kseq	L	V	V	L	F	P	F	V	I	V	Q	T	I	G	Y	Q	I	Q	R	P	E	Q	P	V	H	N	R	L	R	E
79KfragDNA	I	A	R	I	I	P	L	R	Y	R	S	N	H	W	V	P	D	S	R	P	E	Q	P	S	P	Q	S	T	P	E
		100	110	120																										
20Kseq	P	S	P	S	P	K	P	A	P	N	P	Q	P	A	P	S	N	P	I	D	E	K	L	V	K	E	A	V	R	K
79KfragDNA	P	S	P	S	P	Q	P	A	P	N	P	Q	P	A	P	S	N	P	I	D	E	K	L	V	K	E	A	V	R	K
		130	140	150																										
20Kseq	V	G	D	G	Y	V	F	E	E	N	G	V	A	S	L	Y	P	K	P	R	I	L	Q	Q	K	Q	Q	-	-	Q
79KfragDNA	V	G	D	G	Y	V	F	E	E	N	G	V	S	R	Y	I	P	A	K	D	L	S	A	E	T	A	A	G	I	D
		160	170	180																										
20Kseq	A	L	L	A	N	W	P	S	R	K	V	Y	L	I	S															
79KfragDNA	S	K	L	A	K	Q	E	S	L	S	H	K	L																	

(SEQ ID NO:2 和一部分  
SEQ ID NO:5)

图 7