

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200780002688.5

[51] Int. Cl.

A61K 39/215 (2006.01)
A61K 39/42 (2006.01)
C07K 9/00 (2006.01)
C07K 14/005 (2006.01)
C12N 15/00 (2006.01)
C12P 21/00 (2006.01)

[43] 公开日 2009年2月18日

[11] 公开号 CN 101370516A

[22] 申请日 2007.1.19

[21] 申请号 200780002688.5

[30] 优先权

[32] 2006.1.19 [33] US [31] 11/335,197

[86] 国际申请 PCT/US2007/001270 2007.1.19

[87] 国际公布 WO2007/084583 英 2007.7.26

[85] 进入国家阶段日期 2008.7.18

[71] 申请人 美国政府 (由卫生和人类服务部国立卫生研究院的部长所代表)

地址 美国马里兰州

[72] 发明人 D·S·迪米特洛夫 X·肖
Z·仲羽

[74] 专利代理机构 北京北翔知识产权代理有限公司

代理人 张广育 姜建成

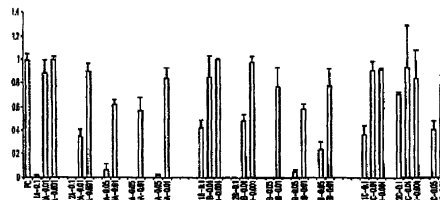
权利要求书 5 页 说明书 94 页 序列表 41 页
附图 16 页

[54] 发明名称

SARS - CoV 刺突糖蛋白的可溶片段

[57] 摘要

本发明涉及在病因学上与严重急性呼吸器官综合症 (SARS) 相关的病毒 (SARS - CoV) 的刺突蛋白; 刺突蛋白的多肽和肽片段; 编码刺突蛋白、刺突蛋白的多肽和肽片段、以及包括刺突蛋白或其部分的偶联蛋白的核酸片段和构建物; 模拟肽; 疫苗; 免疫和治疗严重急性呼吸器官综合症的方法; 抗体; 适配体; 和含有免疫组合物或与刺突蛋白结合的抗体 (或适配体) 的试剂盒。



1. 由 SEQ ID NO:66 或 69 构成的多肽，其中所述多肽当被接种给哺乳动物时能产生体液或细胞免疫应答。
2. 如权利要求 1 所述的多肽，其中所述多肽可溶于水溶液。
3. 如权利要求 1 所述的多肽，其中哺乳动物是人。
4. 如权利要求 1 所述的多肽，其中所述多肽是氨基端或羧基端加帽的。
5. 如权利要求 1 所述的多肽，其偶联到载体蛋白。
6. 如权利要求 5 所述的多肽，其中所述载体蛋白选自牛血清白蛋白、匙孔血蓝蛋白、卵白蛋白、小鼠血清白蛋白、兔血清白蛋白所组成的组。
7. 如权利要求 1 所述的多肽，其偶联到氨苯胂酸、磺胺酸、乙酰基或苦基。
8. 组合物，其包含载体和有效量的如权利要求 1 所述的多肽。
9. 如权利要求 8 所述的组合物，其中所述载体是选自氢氧化铝、脂质 A、灭活的细菌、多糖、矿油、弗氏不完全佐剂、弗氏完全佐剂、磷酸铝、铁、锌、钙的盐、酰基化酪氨酸、酰基化糖类、阳离子衍生的多糖、阴离子衍生的多糖、聚磷腈、生物可降解微球体、单磷酸脂质 A 和 quil A 所组成的组的佐剂。
10. 如权利要求 8 所述的组合物，其中所述有效量的所述多肽对于治疗 SARS-CoV 感染是有效的。
11. 如权利要求 8 所述的组合物，其中所述有效量的所述多肽对于抑制 SARS-CoV 与哺乳动物细胞的融合或进入哺乳动物细胞是有效的。
12. 核酸片段，其编码权利要求 1 所述的多肽。

13. 表达盒，其包括以可操作的方式连接到权利要求 12 所述核酸片段的启动子。

14. 根据权利要求 13 所述的表达盒，其中所述启动子是组成型启动子或调节型启动子。

15. 核酸构建物，其包括载体，其中所述载体包括对由 SEQ ID NOs:66 或 69 任何一个构成的多肽进行编码的核酸片段或根据权利要求 13 所述的表达盒。

16. 根据权利要求 15 所述的核酸构建物，其中所述载体选自质粒、粘粒、酵母人工染色体、细菌人工染色体、F 因子、病毒、表达载体和噬菌粒所组成的组。

17. 组合物，其包括药物载体和核酸片段或根据权利要求 13 所述的表达盒，所述核酸片段编码由 SEQ ID NOs:66 或 69 的任何一个构成的多肽。

18. 如权利要求 17 所述的组合物，其中所述有效量的所述核酸片段对于治疗 SARS-CoV 感染是有效的。

19. 如权利要求 17 所述的组合物，其中所述有效量的所述核酸片段对于抑制 SARS-CoV 与哺乳动物细胞融合或进入哺乳动物细胞是有效的。

20. 重组病毒，其包含病毒载体和核酸片段或根据权利要求 13 所述的表达盒，所述核酸片段编码由 SEQ ID NOs:66 或 69 的任何一个构成的多肽。

21. 如权利要求 20 所述的重组病毒，其中所述病毒载体选自痘苗病毒、金丝雀痘病毒、腺病毒和疱疹病毒所组成的组。

22. 组合物，其包含药物载体和有效量的如权利要求 21 所述的重组病毒。

23. 如权利要求 22 所述的组合物，其中所述有效量的所述重组病毒对于治疗或预防 SARS-CoV 感染是有效的。

24. 组合物，其包含药物载体和微生物，所述微生物含有编码由 SEQ ID NOs:66 或 69 中的任何一个构成的多肽的核酸片段；或根据权利要求 13 所述的表达盒。

25. 如权利要求 24 所述的组合物，其中所述微生物选自沙门氏菌和单核细胞增多性李斯特菌所组成的组。

26. 免疫组合物，包含药物载体和其内插入有如权利要求 13 所述表达盒的 DNA 载体。

27. 如权利要求 26 所述的组合物，其中所述载体选自质粒、粘粒、酵母人工染色体、细菌人工染色体、F 因子、病毒、表达载体和噬菌粒所组成的组。

28. 如权利要求 26 所述的组合物，其中所述组合物进一步包括肌肉坏死试剂。

29. 如权利要求 28 所述的组合物，其中所述肌肉坏死试剂是布比卡因或心脏毒素。

30. 抗体，其特异性地与含有由 SEQ ID NOs:66 或 69 中的任何一个构成的氨基酸序列的多肽相结合。

31. 如权利要求 30 所述的抗体，其中所述抗体特异性与由 SEQ ID NOs:66 或 69 的任何一个构成的表位相结合。

32. 如权利要求 30 所述的抗体，其中所述抗体是单克隆抗体、多克隆抗体、单链抗体、抗原结合抗体片段或人源抗体。

33. 如权利要求 30 所述的抗体，其中所述抗原结合抗体片段是 scFv、Fv、Fab'、双抗体、线性抗体或 F(ab')₂。

34. 如权利要求 30 所述的抗体，其中所述抗体偶联到可检测标记物。

35. 如权利要求 34 所述的抗体，其中所述可检测标记物是荧光蛋白、

荧光标记、放射标记、酶或亲和标记物。

36. 如权利要求 30 所述的抗体，其中所述抗体偶联到毒素。

37. 如权利要求 36 所述的抗体，其中所述毒素是 A 链毒素、核糖体失活蛋白、 α 帚曲霉素、多花白树毒蛋白、曲霉菌素、局限曲霉菌素、核糖核酸酶、表鬼白毒素、白喉毒素、假单细胞外毒素、蓖麻毒蛋白、阿霉素、柔红霉素、紫杉醇、溴化乙锭、丝裂霉素、表鬼白毒吡喃葡萄糖苷、鬼白噻吩甙、长春花新碱、长春花碱、秋水仙素、二羟炭疽菌素二酮、放线菌素 D、PE40、相思豆毒蛋白或糖皮质激素。

38. 如权利要求 30 所述的抗体，其中所述抗体具有由 SEQ ID NO:70 或 SEQ ID NO:71 构成的互补决定区(CBD)。

39. 药物组合物，其含有药物载体和有效量的如权利要求 30 所述抗体。

40. 在哺乳动物中产生针对严重急性呼吸器官综合症的免疫应答的方法，其包括将治疗有效量的如权利要求 1 所述多肽施用给所述哺乳动物。

41. 在哺乳动物中治疗严重急性呼吸器官综合症的的方法，其包括将治疗有效量的如权利要求 1 所述多肽施用给所述哺乳动物。

42. 在哺乳动物中治疗严重急性呼吸器官综合症的的方法，其包括将治疗有效量的与如权利要求 1 所述多肽相结合的抗体施用给所述哺乳动物。

43. 如权利要求 42 所述的方法，其中所述抗体与由 SEQ ID NO:66 或 69 构成的氨基酸序列特异性结合。

44. 在哺乳动物中治疗或抑制严重急性呼吸器官综合症的的方法，其包括将治疗有效量的核酸或根据权利要求 13 所述的表达盒施用给所述哺乳动物，其中所述核酸编码由 SEQ ID NOs:66 或 69 中任何一个构成的 S 多肽。

45. 如权利要求 40-44 中任一项所述的方法，其中所述哺乳动物是人。

46. 在哺乳动物中诊断严重急性呼吸器官综合症的方法，其包括：

(a) 使从所述哺乳动物获得的生物样品与结合到权利要求 1 所述多肽的抗体相接触；

(b) 确定是否所述抗体与所述生物样品结合。

47. 如权利要求 46 所述的方法，其中所述哺乳动物是人。

48. 制备抗体的方法，包括：获得用多肽免疫的动物，其中所述多肽具有由 SEQ ID NO:66 或 69 中的任何一个构成的氨基酸序列；和分离结合于所述多肽的抗体。

49. 试剂盒，含有包裹材料和抗体或适配体，所述抗体或适配体与具有由 SEQ ID NOs:66 或 69 的任何一个构成的氨基酸序列的多肽相结合。

50. 如权利要求 49 所述的试剂盒，进一步包含注射器。

51. 试剂盒，含有包裹材料和治疗有效量的如权利要求 1 所述多肽。

52. 如权利要求 51 所述的试剂盒，进一步包含注射器。

53. 肽，含有 NDLCFSNV(SEQ ID NO:72, S 蛋白位置 375-382)和 FELLNAPATVCG(SEQ ID NO:73, S 蛋白位置 501-512)。

54. 如权利要求 53 所述的肽，其中所述肽涉及建立 S 蛋白构象，其促进所述 S 蛋白受体结合结构域和其中和抗体之间的稳定性复合体的形成和保持。

SARS-CoV 刺突糖蛋白的可溶片段

[01] 本申请要求 2006 年 1 月 19 日提交的美国申请序列号 11/335,197 的优先权，因此其内容的全部被引入本文作为参考。本申请涉及 2004 年 7 月 20 日提交的 PCT 申请序列号 PCT/US2004/023345 和 2003 年 11 月 25 日提交的美国申请序列号 60/524,642，因此它们的内容的全部被引入本文作为参考。

政府资助

[02] 本文描述的发明在 Department of Health and Human Services 支持下开发。美国政府具有本发明的某些权利。

技术领域

[03] 一般而言，本发明涉及冠形病毒(本文是 SARS-CoV)编码的刺突多肽，其在病因学上与严重急性呼吸器官综合症(Severe Acute Respiratory Syndrome, SARS)相关。本发明进一步涉及具有与 SARS-CoV 的刺突蛋白片段相应的核酸序列的核酸和多肽，和其保守变体。本发明也涉及这些核酸、多肽、变体和片段的应用，以产生识别 SARS-CoV 的刺突蛋白的抗体，并用于产生对抗 SARS 的疫苗。本发明的另一方面涉及用于抑制 SARS-CoV 与动物细胞融合的刺突蛋白片段。

发明背景

[04] 严重急性呼吸器官综合症(SARS)是传染性的非典型肺炎，其近来以在 32 个国家和地区的患者中被发现。具有未知病因学的非典型肺炎最初在中国广东省被发现。该发现接下来出现在香港、越南、新加坡、加拿大和北京的严重呼吸系统疾病的报告中，这种疾病传播到家庭成员和卫生保健工作者。该疾病后

来被世界卫生组织(World Health Organization, WHO)诊断为“严重急性呼吸器官综合症(SARS)”。到2003年5月19日,29个国家向WHO报告了总共7,684个SARS病例。总共报告了643例死亡(病死率比例:8.2%)。

[05] 全世界的研究人员已经对从全球不同地区引起SARS的病毒基因组进行了测序。该病毒已被归类为冠形病毒。基于遗传数据支持的抗体的交叉反应性,冠形病毒被分为3类。两种以前已知的人类病毒和SARS-CoV属于不同的类群。引起SARS的冠形病毒不适合先前已知的任何群。而是,其自己形成新的类群。预测病毒蛋白的系统进化分析法表明不与冠形病毒3种先前已知的类群的任何一种十分相似。大多数冠形病毒引起呼吸器官或肠道疾病,其也通过粪-口(faecal-oral)途径传播。

[06] SARS的培养期通常为2到7天。感染特征是发热、干咳、呼吸短促和存在胸片上凝结的最小听诊发现。淋巴细胞减少、白细胞减少、血小板减少和肝酶以及肌酸酐激酶的增加也可在大多数病例中存在。与胃肠道相关的症状也在SARS患者中注意到。

[07] 广东、香港、北京和新加坡死于SARS的患者的病理研究表明在肺中弥漫性肺泡损伤(DAD)为最显著的特征。在具有严重疾病导致死亡的那些个体中,分散的II型肺细胞表明明显的细胞学改变,其包括多核化、巨细胞化(cytomegaly)、巨核化(nucleomegaly)、核染色质透明(clearing of nuclear chromatin)和核仁显著(prominent nucleoli)。尽管这些变化是严重的,但是它们是在其它弥漫性肺泡损伤的病例中可见的上皮细胞变化的光谱内。被识别的形态变化包括支气管上皮细胞脱落、纤毛损失(loss of cilia)和鳞状化生。其它发现包括局部肺泡内出血、吞噬血细胞作用、在小气道中坏死炎性碎片、组织肺炎(organizing pneumonia)或二级细菌性肺炎(secondary bacterial pneumonia)。

[08] 该种疾病的发病机理仍没有确定。然后,急性肺损伤的机制可包括通过靶向内皮细胞或上皮细胞,病毒对肺泡壁的直接损伤。或者,病毒可感染炎性细胞具有通过细胞因子、白细胞介素或肿瘤坏死因子- α 介导的损伤。也可能SARS的组织损伤与组织中病毒感染没有直接关系,而是靠近但不在肺组织的病毒感染诱导的细胞因子或其它因子的继发效应。

[09] 该致命病例的病理学评估表明肝细胞经历脂肪变性颗粒状变性、凋亡和

点坏死(dot necrosis), 并且 Kupffer 细胞增殖和淋巴细胞的门浸润(portal infiltrates)。在患者的胃肠壁有局部出血、血管充血(vascular congestion)和淋巴细胞性浸润。

[10] 由于 SARS-CoV 通过空气途径传播, SARS-CoV 对全世界大量人的健康存在特定的威胁, 因此, 极大需要在感染前使人免疫、诊断感染、在感染期间使人免疫和治疗感染有 SARS-CoV 的感染的人的方法。

发明概述

[11] 本文描述的发明满足了这些和其它的需要。本发明提供多肽; 肽片段; 病毒融合抑制剂; 偶联的蛋白; 免疫肽; 免疫组合物; 模拟肽(peptidomimetics); 核酸片段; 表达盒; 核酸构建物; 重组病毒(recombinant virus); 病毒疫苗; 肽疫苗; 微生物疫苗; DNA 疫苗; 抗体; 适配体(aptamer); 药物组合物; 使动物免疫的方法; 治疗严重急性呼吸器官综合症(SARS)的方法; 诊断 SARS 的方法; 和试剂盒(kits)。

[12] 本发明提供具有与 SARS 病因学相关的多肽相应的氨基酸序列的多肽。优选地, 该多肽是来自 SARS-CoV 的刺突蛋白, 其可抑制 SARS 与动物细胞融合和/或提高动物对抗 SARS-CoV 的免疫应答。在一些实施方式中, 多肽是可溶形式的来自 SARS-CoV 的刺突蛋白。在其它实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 17-757。在一些实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 762-1189。在其它实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 17-757。在一些实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 17-276。在其它实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 303-537。在一些实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 317-517。在其它实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 272-537。在一些实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 17-537。在其它实施方式中, 多肽包括来自 SARS-CoV 的刺突蛋白的氨基酸 17-1189(相对于 SEQ ID NO: 1)。本发明的多肽可抑制 SARS-CoV 与动物细胞融合。本发明的核酸和多肽当被用来接种动物时, 可诱发免疫应答。在一些实施方式中, 本发明的核酸和多肽当被用来接种动物

时，诱发细胞免疫应答。在其它实施方式中，本发明的核酸和多肽当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。有时，动物是人类。

[13] 本发明提供来自 SARS-CoV 的刺突蛋白的肽片段。优选的肽片段是可溶于水溶液的。本发明的肽片段可以缺少来自 SARS-CoV 的全长刺突蛋白的氨基酸序列的一个氨基酸残基。在一些实施方式中，肽片段长度为至少 3 个氨基酸。在其它实施方式中，肽片段长度为至少 10 个氨基酸。在一些实施方式中，肽片段长度为至少 20 个氨基酸。在其它实施方式中，肽片段长度为至少 30 个氨基酸。在一些实施方式中，肽片段长度为至少 20 个氨基酸。在其它实施方式中，肽片段长度为至少 50 个氨基酸。在一些实施方式中，肽片段长度为至少 60 个氨基酸。肽片段也可以是对给定长度片段的单氨基酸单位加成。例如，肽片段长度可以是 3, 4, 10, 11, 21, 22, 31 或 32 个氨基酸。本发明的肽片段可以抑制 SARS CoV 与动物细胞的融合或者当被用来接种动物时，可诱发免疫应答。接种动物后可诱发免疫应答的肽的实例包括例如具有序列 DVQAPNYTQHTSSMRGC (SEQ ID NO:58) 的 D24 肽、具有序列 PSSKRFQPQQFGRDC (SEQ ID NO:59) 的 P540 和肽 GFYTTTGIGYQ (SEQ ID NO:69)。在一些实施方式中，本发明的肽片段当被用来接种动物时，诱发细胞免疫应答。在其它实施方式中，本发明的肽片段当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[14] 本发明提供偶联的蛋白。偶联的蛋白包括被偶联到第二多肽的载体蛋白。优选地，载体蛋白是可溶的。在一些实施方式中，载体蛋白当被用来接种动物时，增加对偶联蛋白的第二多肽的免疫应答。在其它实施方式中，载体蛋白当被用来接种动物时，诱发对偶联蛋白的第二多肽的细胞免疫应答。在一些实施方式中，载体蛋白被用来接种动物时，诱发对偶联蛋白的第二多肽的体液免疫应答。第二多肽可以是本发明的多肽或肽片段，或其保守变体。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[15] 本发明提供包括本发明多肽或肽片段的免疫肽，或其保守变体，其中多肽或肽片段被偶联到乙酰基、苦基、氨基甲酸或磺胺酸。在一些实施方式中，

免疫肽被偶联到乙酰基或巯基。在其它实施方式中，免疫肽被偶联到氨基甲酸或磺胺酸。优选地，免疫肽是可溶的。优选地，免疫肽当被用来接种动物时，诱发免疫应答。在一些实施方式中，免疫肽当被用来接种动物时，诱发体液免疫应答。在其它实施方式中，免疫肽当被用来接种动物时，诱发细胞免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[16] 本发明提供为发明多肽或肽片段或其保守变体的模拟肽，其中，肽键被非肽键所取代。在一些实施方式中，模拟肽可抑制 SARS CoV 与动物细胞融合。在其它实施方式中，模拟肽当被用来接种动物时，诱发免疫应答。例如，模拟肽当被用来接种动物时，诱发细胞免疫应答。或者，模拟肽当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[17] 本发明提供含有佐剂和本发明核酸、多肽、肽片段或模拟肽的组合物。在一些实施方式中，组合物抑制 SARS CoV 与动物细胞融合。在其它实施方式中，组合物当被用来接种动物时，诱发免疫应答。在一些实施方式中，免疫组合物当被用来接种动物时，诱发细胞免疫应答。在其它实施方式中，免疫组合物当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[18] 本发明提供编码本发明多肽和肽片段的核酸片段，及其保守变体。

[19] 本发明提供具有可操作连接到本发明核酸片段的启动子的表达盒。在一些实施方式中，启动子是组成型的。在其它实施方式中，启动子是可诱导的。

[20] 本发明提供包括载体和本发明核酸片段的核酸构建物。该核酸构建物可包括本发明的表达盒。在一些实施方式中，载体可以是病毒。在其它实施方式中，载体是质粒。在进一步实施方式中，载体是表达载体。

[21] 本发明提供包括病毒载体和本发明的核酸片段的重组病毒。在一些实施方式中，病毒载体是疱疹病毒。在其它实施方式中，病毒载体是金丝雀痘病毒。在其它实施方式中，病毒载体是腺病毒。在进一步实施方式中，病毒载体是痘苗病毒。

[22] 本发明提供对抗 SARS 的病毒疫苗，其包括病毒载体、本发明的核酸片段和药物载体。在一些实施方式中，病毒载体是疱疹病毒。在其它实施方式中，病毒载体是金丝雀痘病毒。在其它实施方式中，病毒载体是腺病毒。在进一步实施方式中，病毒载体是痘苗病毒。优选地，配制药物载体用于注射。优选地，病毒疫苗当被用来接种动物时，诱发免疫应答。在一些实施方式中，病毒疫苗当被用来接种动物时，诱发细胞免疫应答。在其它实施方式中，病毒疫苗当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[23] 本发明提供对抗 SARS 的病毒疫苗，其包括本发明的模拟肽、多肽或肽片段，或其保守变体和药物载体。优选地，配制药物载体用于注射。优选地，肽疫苗被配制为单位剂型。优选地，肽疫苗当被用来接种动物时，诱发免疫应答。在一些实施方式中，肽疫苗当被用来接种动物时，诱发细胞免疫应答。在其它实施方式中，肽疫苗当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[24] 本发明提供对抗 SARS 的微生物疫苗，其包括表达本发明多肽或肽片段的微生物，或其保守变体和药物载体。优选地，微生物被减毒。在一些实施方式中，微生物是沙门氏菌。在其它实施方式中，微生物是李斯特菌属。在进一步实施方式中，微生物单核细胞增多性李斯特菌。在一些实施方式中，配制药物载体用于注射。在其它实施方式中，配制药物载体用于口服给药。优选地，微生物疫苗被配制为单位剂型。优选地，微生物疫苗当被用来接种动物时，诱发免疫应答。在一些实施方式中，微生物疫苗当被用来接种动物时，诱发细胞免疫应答。在其它实施方式中，微生物疫苗当被用来接种动物时，诱发体液免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[25] 本发明提供对抗 SARS 的 DNA 疫苗，其包括被插入本发明核酸片段的载体和药物载体。DNA 疫苗可包括佐剂。DNA 疫苗可包括肌肉坏死试剂 (myonecrotic agent)。例如，肌肉坏死试剂可以是布比卡因。在其它实施方式中，肌肉坏死试剂是心脏毒素。例如，载体可以是病毒。在其它实施方式中，载体

是噬菌体。在进一步实施方式中，载体是质粒。含有插入物的载体可在真核细胞中制备。然而，在一些实施方式中，含有插入物的载体在原核细胞中制备。例如，含有插入物的载体可以在细菌中制备。在一些实施方式中，配制药物载体用于粘膜输送。在其它实施方式中，配制药物载体用于注射。优选地，DNA疫苗被配制成单位剂型。优选地，DNA疫苗当被用来接种动物时，诱发免疫应答。在一些实施方式中，DNA疫苗当被用来接种动物时，诱发体液免疫应答。在其它实施方式中，DNA疫苗当被用来接种动物时，诱发细胞免疫应答。动物可以是爬行动物。在一些实施方式中，动物是鸟类。在其它实施方式中，动物是哺乳动物。在进一步实施方式，动物是人类。

[26] 本发明提供结合到本发明的多肽或肽片段的抗体或其保守变体。在一些实施方式中，抗体是抗原结合的抗体片段。在其它实施方式中，抗体是多克隆抗体。在进一步实施方式中，抗体是单链抗体。在其它实施方式中，抗体是单克隆抗体。在一些优选的实施方式中，抗体使人源抗体。抗体可以与可检测标记物偶联。例如，可检测标记物可以是放射标记。在一些实施方式中，可检测标记物是亲和标记。在其它实施方式中，可检测标记物是酶。在进一步实施方式中，可检测标记物是荧光蛋白。在一些优选的实施方式中，可检测标记物是荧光标记。抗体也可偶联到毒素。

[27] 本发明提供结合到本发明的多肽或肽片段的适配体或其保守变体。适配体可以与可检测标记物偶联。例如，可检测标记物可以是放射标记。在一些实施方式中，可检测标记物是亲和标记。在其它实施方式中，可检测标记物是酶。在进一步实施方式中，可检测标记物是荧光蛋白。在一些优选的实施方式中，可检测标记物是荧光标记。适配体也可偶联到毒素。

[28] 本发明提供含有本发明的抗体、S多肽或适配体和药物载体的药物组合物或试剂盒。优选地，配置药物组合物用于注射。

附图简述

[29] 该专利或申请文件可含有至少一幅彩色画出的附图。在要求和付清必要费用后，带有彩色附图(一幅或多幅)的该专利或专利申请公开的拷贝由办公室提供。

[30] 图 1A 阐明 DNA 构建物的琼脂糖凝胶电泳, 该构建物具有编码本发明的刺突蛋白的插入物。泳道从左到右: 第一泳道是 1kb DNA 序列梯(从底到顶的标记-0.5、1.0、1.6、2.0、3.0、4.0); 第 2 泳道表示用 BamHI/XbaI 消化的 DNA 构建物, 导致特征性的载体带(上边的带)和编码刺突蛋白的 DNA 片段(下面的带); 第 3 泳道表示用 HindIII 消化的 DNA 构建物, 由于在载体和编码刺突蛋白的 DNA 片段内存在 HindIII 位点其产生预期更小的带和更大的带。

[31] 图 1B 提供在除去信号序列(残基 1-16, SEQ ID NO:60)后表示多种可溶多肽片段的全长 SARS-CoV S 糖蛋白单体的示意图。可溶片段是称为“S”的刺突蛋白片段, 接下来是与刺突蛋白氨基酸——其构成片段的末端——相应的数量。因此, "S756"是开始于氨基酸 17(正好在信号序列后)、结束于氨基酸 756 的可溶性刺突蛋白片段。"TM"指跨膜片段, 箭头表示氨基酸位置 758-761(序列 RNTR)内可能的切割位点。"RBD"指在氨基酸位置 272-537(SEQ ID NO:57)内可能的受体-结合结构域, 可能是在位置 303 残基下游和位置 537 残基上游(SEQ ID NO:61)之间。

[32] 图 2 阐明在大肠杆菌内表达的 SARS-CoV 刺突蛋白的肽片段的变性聚丙烯酰胺凝胶电泳(SDS-PAGE)。肽片段相应于 SEQ ID NO: 1 的氨基酸 17-446。将编码氨基酸 17-446 的核酸片段克隆到 pRSET 载体, 以产生 pRSET-S(17-446), 其在 BL21DE3 细胞中被表达。在左边的数字和箭头表示千道尔顿的分子量标记。泳道含有下列多肽: M-分子量标记; 第 1 和第 2 泳道-含有 pRSET 载体而没有编码 SEQ ID NO: 1 的氨基酸残基 17-446 的核酸片段也没有异丙基硫代半乳糖苷(IPTG)诱导的对照大肠杆菌的多肽; 第 3 泳道-含有 pRSET 载体而没有编码 SEQ ID NO: 1 的氨基酸残基 17-446 的核酸片段, 但是具有 IPTG 诱导的对照大肠杆菌的多肽; 第 4 泳道-含有 pRSET 载体、编码 SEQ ID NO: 1 的氨基酸残基 17-446 的核酸片段和 IPTG 诱导的大肠杆菌的分析。右侧的箭头表示与在大肠杆菌中表达的 SEQ ID NO: 1 的氨基酸残基 17-446 相应的肽片段的位置。

[33] 图 3 阐明哺乳动物细胞中来自 SARS-CoV 的刺突蛋白的指定肽片段表达的槽印迹分析。将编码肽片段的核酸片段克隆到 pSecTag2B 载体以表达肽片段, 所述肽片段在 N-末端具有用于分泌的小鼠 k 链前导序列, 并且在 C 末端具有用于检测和亲和纯化的 c-Myc 表位和组氨酸标记。将核酸构建物转化入 HEK293 和 VeroE6 细胞。通过使用槽印迹分析和抗- c-Myc 抗体检测该指定肽片段的表

达。在左边和右边的数字表示包括在检测的肽片段内的氨基酸残基。左边的柱代表肽片段在 HEK293 细胞内的表达。右边的柱代表肽片段在 VeroE6 细胞内的表达。上边一半代表从细胞生长的介质中获得的样品(分泌蛋白),下边一半代表从细胞溶胞产物获得的样品(细胞内部分)。PC 是阳性对照,其由在 C 末端含有 PSA 和 c-Myc 标记的质粒的制造商提供。NC 是阴性对照,其含有缺少 c-Myc 表位或组氨酸标记的 SARS-CoV 的全长刺突蛋白。

[34] 图 4A 阐明人 293 或猴类 VeroE6 细胞中来自 SARS-CoV 的刺突蛋白的指定肽片段表达的槽印迹分析。在存在或不存在 T7 聚合酶的情况下用重组痘苗病毒(VTF7.3)表达的编码 S 片段(S276、S537 和 S756)的质粒转染的 293 和 Vero E6 细胞的上清液被转移到消化纤维素膜,并且用抗-c-Myc 表位抗体检测。在左边和右边的数字表示包括在检测的肽片段内的氨基酸残基。PSA PC 是在 C 末端含有 PSA 和 c-Myc 标记的阳性对照。pCDNA-S NC 是阴性对照,其含有缺少 c-Myc 表位或组氨酸标记的 SARS-CoV 的全长刺突蛋白。泳道如下:(1)人 293 细胞,其没有用 VTF7.3 痘苗病毒感染,(2)人 293 细胞,其用 VTF7.3 痘苗病毒感染,(3)猴类 VeroE6 细胞,其没有用 VTF7.3 痘苗病毒感染,和(4)猴类 VeroE6 细胞,其用 VTF7.3 痘苗病毒感染。

[35] 图 4B, 将按图 4A 描述的转染细胞的上清液与 Ni-NTA 琼脂糖小球温育,洗涤并且进行蛋白质印迹法,并且使用与图 4A 中同样的抗-c-Myc 表位抗体。

[36] 图 4C 阐明 S 片段的检测,这通过针对分别相应于在残基 24(D24, 中间面板)和 540(P540, 右侧面板)开始序列的肽培养的两个兔多克隆抗体。左边的面板表示抗-c-Myc 表位抗体检测的 S537 和 S756 的蛋白质印迹的比较。

[37] 图 5 阐明全长细胞膜相关联的 S 蛋白在细胞表面上表达,如使用兔多克隆抗体 P540 通过流式细胞计量术所示出的。使用编码全长 S 糖蛋白的核酸转染 293 细胞,然后用 VTF7.3 感染。收集细胞并用 P540 多克隆抗体和与 FITC 结合的抗兔第二抗体培养,洗涤并进行流式细胞计量术分析。使用相同的质粒表达 S,但是在命名为为阴性对照(NC)的对照试验没有使用 S 的核酸转染细胞;具有编码全长 S 糖蛋白的核酸的细胞被命名为 S。

[38] 图 6A 和 6B 阐明基本上没有天然发生的 S 糖蛋白切割。示出从表达 S756 的转染的 293 细胞、Se 和使用 P540 抗体表达 S 糖蛋白的 293 细胞的细胞溶胞产

物的得到的上清液的蛋白质印迹。观察到接近背景水平切割的 S 和 Se。图 6A 表示在分析前在 4℃ 保持 3 天的样品的蛋白质印迹,以监控非特异性蛋白酶活性对切割模式影响。与之相反,图 6B 表示在制备后立刻使用的样品的印迹。

[39] 图 7A-C 表示 S 糖蛋白介导的细胞融合。基于 pCDNA3 的没有 S 插入物的质粒被用作质粒对照,在表达 S 的细胞和表达 ACE2-ecto 的细胞之间的融合被用作阴性对照。pCDNA3-ACE2-ecto 构建物正好表达用 C9 肽标记的 ACE2 可溶的 ecto 结构域。图 7A 阐明在用 pSecTag2B-S 和 pCDNA3-ACE2-Ecto 转染的 293T 细胞之间没有合胞体形成。与之相反,图 7B 阐明在用 pSecTag2B-S 和 pCDNA3-ACE2 分别转染的 293T 细胞之间有合胞体形成。图 7C 图解说明通过基于报告基因的分析所测量的细胞融合。如所示,在 pCDNA3 和 pSecTag2B 载体中表达的 S 糖蛋白可在基于 β -gal 报告基因的细胞-细胞融合分析中被检测。

[40] 图 8A-C 表示 S 糖蛋白受体结合结构域(RBD)位于残基 272 和 537 之间。图 8A 阐明两种不同 S 可溶片段(S537 和 S756)结合到 293 和 Vero E6 细胞。图 8B 阐明将不同的 S 片段与 Vero E6 细胞相结合。阴性对照测量的背景 OD₄₀₅ 被从每一 S 片段的 OD₄₀₅ 值减去。然后,用 S537d 的 OD₄₀₅ 的百分比表示所得到的每一片段的 OD₄₀₅。图 8C 阐明 S 多肽片段与纯化的可溶 ACE2 的相互作用,如通过 ELISA 所测量。在所有实验中,阴性对照(NC)代表与其它精确相同的方式处理的样品,除了用于转染的质粒不编码任何蛋白。这里示出的数据代表至少 3 个独立的实验。对于所有样品,OD₄₀₅ 被表示为 S537s 的 OD₄₀₅ 的百分比。

[41] 图 9A-D 阐明在 SARS-CoV S 糖蛋白的 N 末端片段之间发生的二聚作用,如通过共免疫沉淀作用和交联所表示的。含有受体结合结构域的除了最小片段(S317-517)的所有 N-末端片段,与 P540 抗体通过 S756 共免疫沉淀。P540 抗体是兔多克隆抗体,其针对含有 S 糖蛋白的残基 540-555 的肽而被开发,并且其结合 S756 多肽,但是该结合不在 N-末端片段。

[42] 在图 9A 中,使用编码 N-末端片段的质粒(用编码氨基酸残基的数目或者开始和结束氨基酸的数字表示)单独转染 293T 细胞(左边六条泳道)或与编码 S756 的质粒联合转染 293T 细胞(右边四条泳道)。然后,用痘苗病毒 VTF7.3 感染这些细胞。培养后,收集培养基,使用识别所有片段的小鼠抗-c-Myc 表位抗体进行蛋白质印迹分析。

[43] 图 9B 表示含有受体结合结构域的除了最小片段(S317-517)的所有 N-末端 S 片段与 P540 抗体通过 S756 共免疫沉淀。图 9A 中使用的相同的培养基首先与仅识别 S576 的 P540 多克隆抗体共免疫沉淀。然后, 使用抗-c-Myc 表位抗体对这些共免疫沉淀物进行蛋白质印迹分析, 以证实 N-末端片段共免疫沉淀。

[44] 图 9C 示出在存在和不存在 DTT 的情况下, 分子量相应于二聚体形式的新的带。为了排除可导致共免疫沉淀的非特异性二硫键形成的可能性, 在共免疫沉淀实验的一个中, 包含 DTT。DTT 对于分泌的 S756(左边泳道)或 S756+S276(右边泳道)的免疫沉淀或共免疫沉淀没有影响。含有分泌的 S756(左边泳道)或 S756+S276(右边泳道)片段的培养基样品在存在 2mM DTT 下, 与 P450 进行共免疫沉淀。

[45] 图 9D 阐明 S 多肽寡聚体的大小。在 SDS-PAGE 分离后, S537 片段与 BS³(Pierce, Rockford, IL)交联, 如在实施例所描述的, 并且准备蛋白质印迹, 使用抗-c-Myc 抗体检测 S537 单体和它的寡聚体。如在图 9D 右边的泳道中所示, 当加入交联剂时, 出现新的带。该新的带具有与二聚体但不是更高的寡聚体相应的分子量。

[46] 图 10A 阐明 N 末端片段 S537 的二聚化, 如通过分子排阻色谱法所检测的。S537 和 S317-517 的洗提曲线被示出, 其中箭头和数字表示被洗提的标准校准蛋白的位置和分子量。

[47] 图 10B 提供通过使用抗-c-Myc 表位抗体收集的 S537 和 S317-517 片段的蛋白质印迹。

[48] 图 11A-B 阐明对于 S 糖蛋白介导的细胞-细胞融合需要 N 末端结构域。图 11A 提供 S 糖蛋白缺失突变体的图式和从细胞-细胞融合分析得到的数据, 其中 RBD 指受体结合结构域的临近位置。由于融合存在的信号用加号(+)表示, 缺乏高于背景可测量的信号用减号(-)表示。仅具有氨基酸 17-1255 的野生型多肽具有融合活性。具有氨基酸 103-1255 (Del1)的或具有 311- 1255 (Del2)的缺失突变体都不具有融合活性。图 11B 表示 S 糖蛋白全长和缺失突变体的表达水平, 如通过蛋白印迹分析所测量的。对于每一样品, 加载将等量的细胞溶胞产物, 并且使用兔多克隆抗体 P540 进行检测。图 11C 阐明全长 S 糖蛋白和 Del1 和 Del2 缺失突变体在细胞表面上表达, 如通过流失细胞计量术所测量的。尽管用空质粒

(empty plasmid)转染细胞的阴性对照明显不同于其它三条曲线的左侧，但是表面表达的水平是低的。

[49] 图 12A-B 阐明二聚的 S1 比含有受体结合结构域的单价片段更有效的与受体 ACE2 结合。图 12A 表示不同 S 片段的相对表达水平，如使用 200 μ l 用 S276、S319-518 和 S537 构建体转染的细胞的培养物上清液通过 ELISA 检测的。抗-His 和抗-c-Myc 表位抗体被用于夹心 ELISA，以检测分泌的标记 S 蛋白水平。图 12B 表示被 S 片段结合到 ACE2 的水平，如通过 ELISA 所测量的。将标记的 ACE2 通过先前已被包被在平板上的抗-9 抗体结合到平板。将用各种 S 蛋白转染细胞的细胞培养物的上清液混合，被在 ELISA 板中与抗-c-Myc 抗体一起培养(带阴影的条)或没有抗-c-Myc 抗体下培养(空白条)。表达或结合的最高水平假设为 100%。如所示，具有 N-末端二聚作用结构域和受体结合结构域的 S537 片段比仅具有受体结合结构域的 S319-518 片段更有效的与 ACE2 结合。

[50] 图 13A-B 阐明可溶的 S 外功能区(ectodomain)在分子排阻色谱的条件下是三聚体。在图 13A 中，纯化的 S 外功能区(Se, 残基 17-1189)在凝胶过滤柱上运行，所述凝胶过滤柱通过使用已知分子量的蛋白质而校准。等量的 BSA 被包括作为内部对照。在图 13B 中，不同部分被从凝胶过滤柱收集，并且通过蛋白质印迹分析。在含有指定分子量的 Se 片段的一些部分中检测两个带 S 多肽，其表示 Se 片段单独(下面的带)和其聚集物(上面的带)。

[51] 图 14A 阐明本发明的 DNA 疫苗在小鼠中可诱导非常高效价的抗-SARS-CoV 血清。小鼠 1A-5A 用编码含有刺突蛋白受体结合结构域(RBD)的 S319-518 片段的 DNA 免疫。小鼠 1B-5B 用融合到编码 Fc 片段核酸的编码 RBD 的 DNA(S319-518 片段)免疫。小鼠 1C-3C 仅接受质粒(没有 S 片段 DNA)。收集抗血清，并且通过 ELISA 检测，以确定不同分离群的效价。在图 14A 中，第一个数字指个体小鼠，字母表示各自的免疫组，最后的数字指使用的稀释。抗血清被稀释 50、250、1250 和 7250 倍，如在柱状图的 x-轴所示。这些数据表示用 S 蛋白的 DNA 编码和受体结合结构域免疫，诱导针对 SARS-CoV 的强的免疫应答。

[52] 图 14B 阐明从用编码 RBD 的 DNA 免疫的小鼠得道的抗血清可抑制 S 介导的细胞融合。将细胞(293T)与用编码刺突蛋白受体结合结构域多肽(S319-518)片段的 DNA 免疫的小鼠得道的抗血清一起培养，然后将细胞悬液与表达 S 蛋白

的细胞混合。按照在实施例 20 种所述的测量融合(也参见, Xiao et al. BBRC 2003)。对于每一融合反应活性的百分比(其中 1=100%)被划在 y-轴, 其中没有任何抑制的融合百分比被指定为 100%。PC 指没有加入血清的阳性对照。在每一组中, 对于小鼠血清#1 到#2, 使用 10(表示 0.1)、100(表示 0.01)和 1000(表示 0.001)的血清稀释因子。对于组 A 和组 B 中小鼠血清#3-#5 和对照组中#3, 使用 20(表示 0.05)和 100(表示 0.01)的稀释因子。这些数据表示用编码 S 蛋白受体结合结构域的 DNA 免疫, 可防止 SARS-CoV 感染。

[53] 图 15 阐明可溶性的 S 糖蛋白片段抑制 S-介导的细胞融合。首先, 将 10 ug/ml 的各种 S 片段与表达 ACE2 的细胞在室温下温育。然后, 将表达 ACE2 的细胞与表达 S 的细胞混合, 并且按照实施例中描述的融合分析。Y-轴是在减掉背景噪声后每一样品的 OD₅₉₅。每一构建物的数字表示各自多肽的起始和终止残基。

发明详述

[54] SARS 代表了一个重要的公共卫生忧虑。诊断和治疗被 SARS-CoV 感染的人的方法提供了防止或控制 SARS-CoV 感染进一步传播的机会。由于 SARS-CoV 通过空气途径感染人的能力, 这些方法是特别重要的。本发明提供编码 SARS-CoV 刺突蛋白氨基酸序列片段的核酸。本发明也提供与 SARS-CoV 刺突蛋白氨基酸序列片段相应的氨基酸序列的多肽。本发明也提供 SARS-CoV 刺突蛋白的肽片段和保守变体, 也提供偶联蛋白和模拟肽, 其具有与刺突蛋白氨基酸序列相应的部分。

[55] 刺突蛋白是重要的, 因为其存在于完整 SARS-CoV 的外面。因此, 在病毒有机会感染细胞之前, 其提供能供用来抑制或消除完整病毒的靶。

[56] 本发明的核酸和多肽提供了优于全长刺突蛋白的优点, 因为核酸容易制造并且本发明的多肽以大量、可溶的形式被制造。本发明的多肽提供额外的优于天然刺突蛋白的优点, 因为当其被施用给动物时, 对降解具有增加的抗性。也可配制本发明的多肽以增加其抗原性, 使它们当被施用给动物例如人时, 成为诱发免疫应答的更有效抗原。

[57] 因此, 本发明提供可用来配制疫苗和免疫组合物的核酸和多肽抗原, 所

述疫苗和免疫组合物可被用来免疫和治疗感染 SARS-CoV 的人。此外，本发明提供结合于 SARS-CoV 刺突蛋白的抗体，其可被用来诊断、免疫和治疗感染 SARS-CoV 的人。

定义：

[58] "佐剂"通常被定义为非特异性增强对抗原免疫应答的物质。多种佐剂可与本发明的免疫肽和 immunofragopeptides。大多数佐剂包含设计用来保护抗原免于快速分解代谢的物质例如氢氧化铝或矿油，和免疫应答刺激物例如脂质 A、百日咳杆菌(*Bordetella pertussis*)或结核分枝杆菌(*Mycobacterium tuberculosis*)源蛋白。适合的佐剂是商业上可获得的例如，弗氏不完全佐剂和完全佐剂(Difco Laboratories, Detroit, Mich.); Merck 佐剂 65(Merck and Company, Inc., Rahway, N.J.); 铝盐例如氢氧化铝凝胶(明矾)或磷酸铝; 钙、铁或锌的盐; 酰基化酪氨酸的不容悬液; 酰基化糖类; 阳离子或阴离子衍生的多糖; 聚磷腈; 生物可降解微球体; 单磷酸脂质 A 和 quil A。也可使用细胞因子例如 GM-CSF 或白细胞介素-2、-7 或-12 作为佐剂。

[59] "动物"指在抗原攻击后能增长免疫应答的生物。例如，爬行动物、鸟类和哺乳动物能在应答抗原攻击中产生抗体。在非人类生物中抗体增加被认为可用于减少或消除交叉反应性的诊断分析。

[60] "适配体"是和本发明的多肽或肽片段结合的肽、多肽或核酸(RNA 或 DNA)。

[61] "载体蛋白"指能与本发明的多肽或肽片段偶联形成偶联蛋白的多肽。载体蛋白可与多肽或肽片段，以便增加多肽或肽片段的溶解度或免疫原性。载体蛋白也可以与多肽或肽片段偶联以提供标记，该标记提供用于分离或检测偶联蛋白。例如，生物素可被用作载体蛋白，其与多肽或肽片段偶联，以产生偶联蛋白，然后该偶联蛋白通过与抗生物素蛋白相互作用而分离，或者通过使用荧光标记的抗生物素蛋白进行检测。在另一个实例中，抗体结合的载体蛋白可与多肽或肽片段偶联，以产生偶联蛋白，该偶联蛋白被与偶联蛋白的载体蛋白结合的抗体所结合。

[62] 本发明包括分离的或基本纯化的核酸、肽、多肽或蛋白质。在本发明的

上下文中，"分离的"核酸、DNA 或 RNA 分子或"分离的"多肽是核酸、DNA 分子、RNA 分子或多肽，其远离天然环境而存在，因此不是天然产品。分离的核酸、DNA 分子、RNA 分子或多肽可以以纯化的形式存在，或者可以在非天然环境例如转基因宿主细胞中存在。"纯化的"核酸分子、肽、多肽或蛋白质或其片段当通过重组技术被产生时，是基本上不含其它细胞物质或培养基的，或者当被化学合成时，是基本上不含化学前体或其它化学试剂的。在一个实施方式中，"分离的"核酸不含天然位于生物染色体 DNA 中核酸侧面的序列(即位于核酸 5'和 3'端的序列)，其中所述核酸来自所述生物。例如，在不同的实施方式中，分离的核酸分子可含有天然位于细胞染色体 DNA 中核酸分子侧面的大约 5 kb、4 kb、3 kb、2 kb、1 kb、0.5 kb 或 0.1 kb 以下的核苷酸序列，其中所述核酸来自所述细胞。基本不含细胞内物质的蛋白质、肽或多肽包括制备具有污染蛋白的大约 30%、20%、10%或 5%(干重)以下的蛋白质、肽或多肽。当本发明的蛋白质或其生物活性部分被重组制造时，优选地，培养基指化学前体或非目的蛋白的化学试剂在大约 30%、20%、10%或 5%(干重)以下。

[63] 术语多肽、肽和蛋白质在本文互换使用。

[64] 如本文使用的肽或多肽"片段"指小于全长肽、多肽或蛋白质。例如，肽或多肽片段的长度或其单一单位长度可具有至少大约 3 个、至少大约 4 个、至少大约 5 个、至少大约 10 个、至少大约 20 个、至少大约 30 个、至少大约 40 个氨基酸。例如，片段的长度可以是 6、7、8、9、10、11、12、13、14、15、16、17 或更多个氨基酸。在肽片段的大小上没有上限。然而，在一些实施方式中，肽片段的长度可以在大约 500 个氨基酸以下、大约 400 个氨基酸以下、大约 300 个氨基酸以下或大约 250 个氨基酸以下。优选地，肽片段当被用于接种动物时，可诱发免疫应答。通过用肽片段和佐剂、与佐剂偶联的肽片段、或与氨苯甲酸、磺胺酸、乙酰基或苦基偶联的肽片段联合接种动物，肽片段可诱发免疫应答。肽片段可包括非酰胺键，并且可以是模拟肽。

[65] 如本文使用的术语"可溶的"指多肽在水溶液中可被溶解的能力。例如，可溶肽可以与水介质混合，这样至少可检测部分的肽存在于水介质中。通过使用普通技术例如光吸收、荧光、结合染料的能力、减少银粒子的能力等可检测肽。

[66] 术语"特异性结合"指与单一表位结合，但不与一个以上表位结合的抗体。因此，特异性结合于多肽的抗体将与存在于多肽上的表位结合，但是不与存在

于其它多肽上的表位结合。

I. 本发明的多肽、肽片段、偶联蛋白、免疫肽和模拟肽

[67] 本发明提供具有如此氨基酸序列的多肽，所述氨基酸序列与病毒(SARS-CoV)的刺突蛋白的氨基酸序列相应，所述病毒与严重急性呼吸器官综合症(SARS)病因学上相关。代表性的氨基酸序列被 SEQ ID NO: 1 所提供，为了方便参考，下面提供了其序列。

```

1 MFIFLLFLTL TSGSDLDRCT TFDDVQAPNY TQHTSSMRGV
41 YYPDEIFRSD TLVLTQDLFL PFYSNVTFGH TINHTFGNPV
81 IPPKGGIYFA ATEKSNVVRG WVFGSTMNKK SQSVIIINNS
121 TNVVIRACNF ELCDNPFPAV SKPMGTQTHI MIFDNAFNCT
161 FEYISDAFSL DVSEKSGNFK HLREFVFKNK DGFLVYKGY
201 QPIDVVRDLP SGFNTLKPIF KLPLGINITN FRAILTAFSP
241 AQDIWGTSAA AYFVGYLKPT TFMLKYDENG TITDAVDCSQ
281 NPLAELKCSV KSFEIDKGIY QTSNFRVVPV GDVVRFPNIT
321 NLCPPGEVFN ATKFPSVYAW ERKKISNCVA DYSLVYNSTF
361 FSTFKCYGVS ATKLNOLCFS NVYADSFVVK GDDVRQIAPG
401 QTGVIADYNY KLPDDFMGCV LAWNTRNIDA TSTGNVNYKY
441 RYLRHGKLRP FERDISNVPF SPDGKPCPTP ALNICYWPLND
481 YGFYTTTGIG YQPYRVVVLV FELLNAPATV CGPKLSTDLI
521 KNQCVNFNFN GLTGGVLTTP SSKRFQPFQ FGRDVSDFD
561 SVRDPKTSEI LDISPCAFGG VSVITPGTNA SSEVAVLYQD
601 VNCTDVSTAI HADQLTFAWR IYSTGNVVFQ TQAGCLIGAE
641 HVDTSEYEDI PIGAGICASY HTVSLLRSTS QKSIVAYTMS
681 LGADSSIAYS NNTIAIPTNF SISITTEVMP VSMAKTSVDC
721 NMYICGDSTE CANLLQYGS FCTQLNRALS GIAEQDRNT
761 REVFAQVKQM YKTPTLKYFG GENFSQILEP PLKPTKRSFI
801 EDLLFNKVTI ADAGFMQYV ECLGDINARD LICAQKFNGL
841 TVLPPLLTDD MIAAYTAALV SGTATAGWTF GAGAALQIPF
881 AMQMAYRFNG IGVTQNVLYE NQKQIANQFN KAISQIQESL
921 TTTSTALGKL QDVVNQNAQA LNTLVKQLSS NFGAISSVLN
961 DILSRLDKVE AEVQIDRLIT GRLQSLQTYV TQQLIRAAEI
1001 RASANLAATK MSECVLGQSK RVDPCGKGYH LMSFPQAAPH
1041 GVVFLHVTYV PSQERNFTIA PAICHEGKAY FPREGVVFVN
1081 GTSWFITQRN FFSPQIITD NIFVSGNCDV VIGIINNTVY
1121 DPLQPELDSF KEELDKYFKN HTSPDVDLGD ISGINASVVN
1161 IQKEIDRLNE VAKNLNESLI DLQBLGKYEQ YIKWPWYVWL
1201 GFIAGLIAIV MVTILLCCMT SCCSCLKGAC SCGSCCKPDE
1241 DDSEPVKGV KLHYT

```

[68] 本发明也提供具有如此氨基酸序列的肽片段，所述氨基酸序列与病毒(SARS-CoV)的刺突蛋白的氨基酸序列相应，所述病毒与严重急性呼吸器官综合症(SARS)病因学上相关。这类氨基酸序列包括 SEQ ED NO: 13、14、15、20-59 和 61-63 表示的氨基酸序列。SEQ ID NO: 1 的肽片段的长度也可以是3个或更多个氨基酸，并且其当被用于免疫动物时，产生免疫应答。这些肽片段实例为长度具有3个氨基酸的肽，更长的单一氨基酸单位，例如长度为4、5、6、7、8、9、10个氨基酸，和从相应于 SEQ ID NO: 1 的氨基酸序列缺失一个氨基酸的氨基酸序列。

[69] 本发明也提供具有与本发明多肽或肽片段偶联的载体蛋白的偶联蛋白。载体蛋白可被用来增加偶联蛋白的溶解度。载体蛋白也可被用来增加偶联蛋白的免疫原性，以增加与本发明多肽或肽片段结合的抗体的产量。载体蛋白也可被用来提供分离或检测偶联蛋白。因此，可以通过将偶联蛋白与结合到偶联蛋白的载体蛋白部分的其它成分相互作用，检测或分离偶联蛋白。例如，具有抗生物素蛋白作为载体蛋白的偶联蛋白可通过使用已知的方法用生物素进行检测。多数载体蛋白可被用来产生本发明的偶联蛋白。这类载体蛋白的实例包括匙孔血蓝蛋白(keyhole limpet hemacyanin)、牛血清白蛋白、卵白蛋白、小鼠血清白蛋白、兔血清白蛋白等。通过使用重组方法产生融合蛋白，可将载体蛋白偶联到本发明的多肽或肽片段。也可通过化学结合方法或通过使用化学连接子，将载体蛋白偶联到本发明的多肽或肽片段。这类偶联方法是本领域已知的并已被描述。Harlow et al., *Antibodies: A Laboratory Manual*, page 319 (Cold Spring Harbor Pub. 1988); Taylor, *Protein Immobilization*, Marcel Dekker, Inc., New York, (1991)。

[70] 本发明提供具有本发明多肽或肽片段的免疫肽，所述多肽或肽片段被偶联到氨苯胂酸、磺胺酸、乙酰基和苦基。将这些基团偶联到多肽的方法是已知的并且已被报出。Weigle, *J. Exp. Med.* 116:913-928 (1962); Weigle, *J. EXP. Med.* 122:1049-1062 (1965); Weigle, *J. EXP. Med.* 121:289-308 (1965)。

[71] 本发明的多肽和肽片段也可以是糖基化形式或非糖基化形式。本发明的

多肽或肽片段可以溶于水溶液或不溶于水溶液。本发明的多肽和肽片段可以是保守变体。保守变体是多肽或肽片段，其源自全长多肽如以 SEQ ID NO: 1 为例，通过在全长多肽的 N-末端和/或 C-末端缺失(所谓的截短)、加成、减去一个或多个氨基酸；在全长多肽的一个或多个位点缺失、加成、减去一个或多个氨基酸。这样的变体可起因于例如遗传多态性或人为操作。这类操作的方法是本领域普遍知道的。例如，可通过诱变编码多肽的 DNA 制备 SEQ ID NO: 1 的氨基酸序列变体。诱变和核苷酸序列改变的方法是本领域公知的。参见，例如 Kunkel, Proc. Natl. Acad. Sci. USA, 82, 488 (1985); Kunkel et al., Methods in Enzymol., 154:367 (1987); 美国专利号 4,873,192; Walker and Gastra, eds., Techniques in Molecular Biology, MacMillan Publishing Company, New York (1983)和其中引用的参考文献。作为适当氨基酸取代的指导可在下述文献的模型中找到: Dayhoff et al., Atlas of Protein Sequence and Structure. Natl. Biomed. Res. Found., Washington, CD. (1978), 其被引入本文作为参考。优选保守取代，例如与具有相似特性的另一个氨基酸交换。例如，用疏水性氨基酸彼此取代或用亲水性氨基酸彼此取代。使用常规的筛选实验来确定来源于 SEQ ID NO: 1 的取代的多肽或肽片段当被施用给动物时，是否产生免疫应答。这类筛选实验的实例在本领域是公知的，并且包括酶联免疫吸附测定法、放射性免疫测定法、铬释放测定法等。这些测定法已被描述。Harlow et al., Antibodies: A Laboratory Manual, page 319 (Cold Spring Harbor Pub. 1988)。

[72] 本发明提供本发明的多肽和肽片段的模拟肽。模拟肽描述肽类似物，例如在制药工业作为非肽类药物的通常使用的肽类似物，其特性类似于模板肽的特性(Fauchere, J., Adv. Drug Res., 15: 29 (1986)和 Evans et al., J. Med. Chem., 30:1229 (1987))。模拟肽结构上与具有肽键的多肽或肽片段相似，但是其一个或多个肽键任选地被键例如 $-\text{CH}_2\text{NH}-$ 、 $-\text{CH}_2\text{S}-$ 、 $-\text{CH}_2\text{CH}_2-$ 、 $-\text{CH}=\text{CH}-$ (顺式和反式)、 $-\text{COCH}_2-$ 、 $-\text{CH}(\text{OH})\text{CH}_2-$ 和 $-\text{CH}_2\text{SO}-$ 所取代，其通过本领域已知的方法进行。模拟肽优于天然多肽实施方式的优点可包括更经济制造、更大的化学稳定性、改变的特异性和增强的药理性质例如半衰期、吸收、效价和功效(potency and efficacy)。

[73] 本发明的多肽、肽片段、偶联蛋白和模拟肽也可被修饰以用于体内应用，这通过在氨基酸或羧基端加成阻断剂以减少体内降解。这在那些取代中是有用

的，在那些取代中多肽末端倾向于在细胞摄取前被降解。这些阻断剂可包括但不限于另外相关的或不相关的肽序列，其可被连接到待被施用的多肽、肽片段、偶联蛋白和模拟肽的氨基和/或羧基端残基。这可在多肽、肽片段或偶联蛋白合成期间化学地进行，或者通过重组 DNA 技术使用本领域普通技术人员熟知的方法进行。或者，阻断剂例如吡咯谷氨酸或者本领域已知的其它分子可被连接到氨基和/或羧基端残基，或者氨基端的氨基或羧基端的羧基可用不同的部分取代。因此，本发明提供氨基末端加帽和羧基末端加帽的多肽和肽片段。

[74] 本发明的多肽或肽片段产生免疫应答的能力可以通过大量本领域认知的方法进行检测。例如，对于它们诱导抗体产生的能力，或刺激细胞毒性 T 淋巴细胞应答的能力。

[75] 本发明的多肽或肽片段可以在筛选实验中使用以识别或分离与本发明的多肽或肽片段结合的抗体，或 SARS-CoV 的刺突蛋白。例如，多肽或肽片段可被用于噬菌体展示实验以分离与多肽或肽片段结合的抗体。在另一个实例中，本发明的多肽或肽片段可被结合到抗体接触的固体载体，这样与多肽或肽片段结合的抗体变得固定于固体载体上。后来，这些抗体可从固体载体洗提出来。根据本领域其它许多已知方法，本发明的多肽和肽片段可被用来分离抗体。

[76] 可被用于小规模或大规模产生本发明的偶联蛋白、多肽或肽片段的表达系统包括但不限于用含有本发明核酸片段的重组核酸构建物转化的细胞或微生物。重组核酸构建物的实例可包括噬菌体 DNA、质粒 DNA、粘粒 DNA 或病毒表达载体。可被转化的细胞和微生物的实例包括细菌(例如，大肠杆菌或枯草芽孢杆菌)；酵母(例如，酵母菌和毕赤酵母菌)；昆虫细胞系统(例如，杆状病毒)；植物细胞系统；或哺乳动物细胞系统(例如，COS、CHO、BHK、293、VERO、HeLa、MDCK、W138 和 NIH 3T3 细胞)。也可用作宿主细胞的是直接从用质粒载体转染或用病毒载体转染的哺乳动物获得的初级细胞或刺激细胞。适当表达载体的实例包括但不限于质粒和病毒载体例如疱疹病毒、逆转录病毒、痘苗病毒、减毒痘苗病毒、金丝雀痘病毒、腺病毒、腺伴随病毒、慢病毒和疱疹病毒等。合成的方法也可被用来产生本发明的多肽和肽片段。这类方法是已知的并已被报告。Merrifield, Science. 85:2149 (1963)。

II. 本发明的核酸片段、表达盒和核酸构建物

[77] 本发明提供编码本发明的多肽、肽片段和偶联蛋白的分离的核酸片段。本发明的核酸片段也包括编码由于遗传密码简并的相同氨基酸的片段。例如，氨基酸苏氨酸被 ACU、ACC、ACA 和 ACG 所编码，因此其是简并的。本发明拟包括所有的编码相同氨基酸的多核苷酸的片段。这类突变是本领域已知的 (Watson et al, Molecular Biology of the Gene, Benjamin Cummings 1987)。突变也包括改变编码保守氨基酸变化的核酸片段，例如，用亮氨酸取代异亮氨酸等。这类突变也是本领域已知的。因此，本发明的基因和核苷酸序列包括天然发生的序列以及突变形式。

[78] 本发明的核酸片段可以被包含在载体内。载体可包括但不限于任何质粒、噬菌粒、F 因子、病毒、粘粒或噬菌体，它们可以是双链或单链线形或环状的，它们可以是或可以不是可自传染的或可移动的。载体也可以通过整合入细胞染色体组转化原核或真核宿主或者存在于染色体外(例如自主复制带有复制起始区的质粒)。

[79] 优选地，载体中的核酸片段在适当启动子或其它体外或宿主细胞内转录的调控元件的控制下，并且载体中的核酸片段可操作的连接到适当启动子或其它体外或宿主细胞内转录的调控元件，其中所述宿主例如真核细胞或微生物如细菌。载体可以是在多个宿主中行使功能的穿梭载体。载体也可以是克隆载体，其典型地含有一个或多个外源 DNA 序列可以以确定的方式插入的限制性内切酶位点。这类插入可以发生而没有损失克隆载体必要的生物功能。克隆载体也可以含有标记基因，所述标记适合用于识别和选择该克隆载体转化的细胞。标记基因的实例是四环素抗药性或对氨苄青霉素抗药性。许多克隆载体是商业可获得的 (Stratagene, New England Biolabs, Clontech)。

[80] 本发明的核酸片段也可被插入表达载体。典型地，表达载体含有编码细菌复制起点和抗生素抗性基因的原核 DNA 元件，以在宿主细胞提供表达载体的扩增和筛选；调控元件，其控制转录起始，例如启动子；和 DNA 元件，其控制转录物处理 (processing of transcripts)，例如内含子或转录中止/多腺苷酸化序列。

[81] 将核酸片段引导入载体的方法在本领域是可获得的 (Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd edition, Cold Spring Harbor Press,

Cold Spring Harbor, N. Y. (2001))。简言之，用一种或多种限制性内切酶(限制性内切核酸酶)处理待被插入核酸片段的载体，以产生具有平整末端、具有5'或3'悬垂部分的“粘性”末段或上述任意组合的线性载体。也可用限制性内切酶处理该载体，然后用另一种修饰酶例如聚合酶、核酸外切酶、磷酸酯酶或激酶处理，以产生具有可用于将核酸片段连接到该载体的特性的线性载体。用一种或多种限制性内切酶处理待被插入载体的核酸片段，以产生具有平整末端、具有5'或3'悬垂部分的“粘性”末段或上述任意组合的线性片段。也可用限制性内切酶处理该核酸片段，然后用另一种DNA修饰酶处理。这类DNA修饰酶包括但不限于聚合酶、核酸外切酶、磷酸酯酶或激酶，以产生具有可用于将核酸片段连接到载体的特性的核酸片段。

[82] 然后，将处理的载体和核酸片段连接在一起，形成含有根据本领域可获得方法(Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 3rd edition, Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (2001))的核酸片段的构建物。简言之，处理的核酸片段和处理的载体在存在适当缓冲液和连接酶的情况下进行组合。然后在允许连接酶将核酸片段连接到载体的适当条件下，温育该混合物。

[83] 本发明也提供含有在体外或宿主细胞内能指引本发明特定核酸片段如SEQ ID NO: 2表达的核酸序列的表达盒。同样，本发明的核酸片段也可被插入表达盒，以便产生反义信息。表达盒是分离的单位，以便表达盒可以是线性形式并且其行使体外转录和翻译分析。进行这些分析的材料和方法是从Promega Corp. (Madison, Wisconsin)商业上可获得的。例如，体外转录可如此产生：通过在T7启动子的控制下放置核酸序列，然后使用T7 RNA聚合酶来产生体外转录。然后可以在体外通过使用兔网织红细胞裂解物(reticulocyte lysate)对该转录进行翻译。或者，可将表达盒掺入允许该表达盒在宿主细胞中复制和扩增或者也在体外转录和翻译核酸片段的载体。

[84] 这样的表达盒可以包含一个或多个在调控调节序列中放置核酸片段的限制性位点。该表达盒也可以含有可操作连接到核酸片段的终止信号，以及恰当翻译该核酸片段所需的调节序列。含有核酸片段的表达盒可以是嵌合的，即就它的其它成分的至少一种而论，它的成分的至少一种是异源的。表达盒也可以是天然发生的表达盒，但是该表达盒以可用于异源表达的重组形式获得。核酸片段在表达盒中的表达可以在组成型启动子或诱导型启动子的控制下进行，所

述诱导型启动子仅当宿主细胞暴露于特定外界刺激时启动转录。

[85] 表达盒可以包括在体内和/或体外行使功能的 5'-3'方向的转录、转录和翻译的起始区、核酸片段和转录和翻译的中止区。该中止区可以天然具有转录起始区，可以天然具有核酸片段或者可以缘自另一种来源。

[86] 调节序列可以是位于编码序列的上游(5'非编码序列)、其中或下游(3'非编码区)的多核苷酸序列，并且其影响转录、转录后处理或稳定性、或者与编码序列相关的转录。调节序列可包括但不限于增强子、启动子、抑制物结合位点、转录前导序列、内含子和多腺苷酸信号序列。它们也可以包含天然和合成序列以及可以是合成和天然序列组合的序列。当调节序列不限于启动子、一些可用调节序列包括组成型启动子、诱导型启动子、调节型启动子 (regulated promoter)、组织特异性启动子、病毒启动子和合成启动子。

[87] 启动子是通过提供用于RNA聚合酶和蛋白转录需要的其它因子的识别来控制编码序列表达的核苷酸序列。启动子包括最小启动子，其仅仅由需要转录启动的所有基本元素例如 TATA 区和/或启动基因组成，所述启动基因是包括 TATA 区和用作指定转录启动位点的短 DNA 序列，向其中加入调节元素以控制表达。启动子也可以完全源自天然基因或者包含源自天然中发现的不同启动子，或者甚至包含合成 DNA 片段。启动子也可包含涉及蛋白因子结合的 DNA 序列，所述蛋白因子控制应答生理或发育条件的转录开始的功效。

[88] 本发明也提供含有载体和表达盒的构建物。载体可选自但不限于任何前面描述的载体。可以通过本领域已知的方法和先前描述的方法(Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 3rd edition, Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (2001))向载体中插入表达盒。在一个实施方式中，表达盒的调节序列可源自非插入了表达盒的载体的来源。在另一个实施方式中，含有载体和表达盒的构建物通过将本发明的核酸片段插入其自己含有调节序列的载体中形成。因此，表达盒通过将核酸片段插入载体形成。含有调节序列的载体是商业上可获得的，并且应用其的方法是本领域已知的(Clontech, Promega, Stratagene)。

III. 本发明的免疫组合物和疫苗

[89] 本发明提供免疫组合物和疫苗，其当被施用给动物时，可被用来产生针对病因学上与严重急性呼吸器官综合症相关的病毒的免疫应答。该免疫应答可以是体液免疫应答或细胞免疫应答。

[90] 本发明的免疫组合物可包括佐剂和本发明的核酸、多肽、肽片段、模拟肽、偶联蛋白、免疫肽或它们的任何组合。免疫组合物可包括与本发明的核酸、多肽、肽片段、模拟肽、偶联蛋白、免疫肽化学上不相关的佐剂。免疫组合物可包括与本发明的核酸、多肽、肽片段、模拟肽、偶联蛋白、免疫肽化学上相关的佐剂。本发明的免疫组合物也可包括药学上可接受的稀释剂或载体。

[91] 可以传统地配制免疫组合物。具体而言，被包含在组合物中的核酸、多肽、肽片段、模拟肽、偶联蛋白、免疫肽或它们的任何组合可以与药学上可接受的稀释剂或载体组合。药学上可接受的稀释剂或载体的实例包括水或盐水溶液，例如磷酸盐缓冲盐溶液(PBS)。一般而言，基于给药模式和途径以及标准药物实践(pharmaceutical practices)选择药学上可接受的稀释剂或载体。药学上可接受的稀释剂或载体以及所有必须应用于药物组合物中的被描述于 Remington's Pharmaceutical Sciences——本领域标准参考文件。

[92] 免疫组合物可包括本文公开的和本领域已知的佐剂。铝化合物可用作佐剂。这类铝化合物包括氢氧化铝、磷酸铝、羟基磷酸铝(aluminum hydroxyphosphate)等。根据标准的方法，可将核酸、多肽、肽片段、模拟肽、偶联蛋白、免疫肽或它们的任何组合吸附或沉淀在铝化合物上。其它佐剂包括聚磷腈(polyphosphazene) (WO 95/2415)、DC-chol(3-β-[N-(N', N'-二甲基氨基甲叉)氨基甲酰]胆固醇)(美国专利第 5,283,185 号和 WO 96/14831)、QS-21 (WO 88/9336)和得自 IrmunoChem (Hamilton, Montana)的 RIBI。含有未甲基化 CpG 二核苷酸("CpG")的免疫刺激寡核苷酸当通过全身和鼻途径施用时在本领域作为佐剂是已知的(WO 96/02555, EP 468520, Davis et al., J. Immunol., 160:870 (1998); McCluskie and Davis, J. Immunol., 161:4463 (1998)。CpG 当被配制成免疫组合物或疫苗时，通常在自由溶液(free solution)中与游离抗原一起施用(WO 96/02555; McCluskie and Davis, J. Immunol., 161 :4463 (1998))或者与抗原共价结合施用(PCT 公开号 WO 98/16247)，或者被与载体例如氢氧化铝一起配制(Brazolot-Millan et al., Proc.Natl.Acad.Sci.. 95:15553 (1998))。

[93] 本发明也提供疫苗，其包括本发明的核酸、多肽、肽片段、模拟肽、偶

联蛋白、免疫肽，核或它们的任何组合。按照本文描述的或者疫苗领域已知的配制这类疫苗。例如，根据本领域的已知方法可产生表达本发明多肽、肽片段或偶联蛋白的病毒疫苗。可被使用的病毒载体的实例包括腺病毒、疱疹病毒、痘苗病毒、金丝雀痘病毒等。疫苗也可被配制成酯质体。这类配制是本领域普通技术人员已知的。Liposomes: A Practical Approach. RRC New Ed, IRL press (1990)。

[94] 本发明也提供基于核酸的疫苗，其表达本发明的多肽、肽片段或偶联蛋白。例如，核酸疫苗可表达具有 SEQ ID NO: 1, 13, 14, 15, 20-59, 61-63 的多肽或 SEQ ID NO: 1 的片段。用编码本发明多肽、肽片段或偶联蛋白的核酸构建物接种动物可导致对编码抗原的体液或细胞介导免疫应答。认为一些骨髓源专职抗原递呈细胞被核酸构建物转染和编码的抗原被转录和翻译入诱发特意性应答的免疫源多肽。核酸疫苗的一个特征是它们提供诱发强的细胞毒性 T 淋巴细胞 (CTL) 应答。因为核酸-编码多肽在转染细胞的胞质溶胶中合成，所以这些应答发生。此外，在细菌中产生的核酸构建物富含未甲基化的 CpG 核苷酸，其被巨噬细胞识别为外源的。因此，它们诱发增强继承性免疫的内生免疫应答。因此，甚至当没有佐剂情况下施用，核酸疫苗是有效的。

[95] 将表达盒直接注射入活宿主细胞转化了大量细胞，并引起它们表达引入的核酸，从而表达基因产物。转化的细胞可在它们的细胞表面展示表达抗原的片段和主要组织相容性类 I (MHC I) 或类 II (MHC II) 复合物。

[96] 在将核酸构建物注入动物包括人之前，通过诱导肌肉降解可更有效地将核酸构建物引入细胞 (Vitadello et. al., Hum. Gene. Ther., 5:11 (1994); Danko and Wolff, Vaccine. 12:1499 (1994); Davis et. al., Hum. Gene. Ther., 4:733 (1993))。例如，认为这种治疗增加转化功效高达 40 倍。最通常使用的肌肉坏死试剂的两种是局部麻醉剂布比卡因 (bupivacaine) 和心脏毒素 (Danko and Wolff, Vaccine. 12:1499 (1994); Davis et. al., Hum. Gene. Ther., 4:733 (1993))。大量的其它技术已被使用来将核酸构建物转移到肌肉。这类其它技术包括逆转录病毒载体、腺病毒载体和脂质体。然而，在细胞中转移和表达外源核酸的这些输运机制中，直接注射裸核酸表现最有效。

[97] 核酸构建物可以在药学上可接受的载体中施用。药学上可接受的载体是适合施用给人或其它哺乳动物对象的生物相容性载体，例如生理盐水。治疗有

效量是能在治疗的动物中产生免疫应答(例如增强的 T 细胞应答或抗体产生)的核酸构建物的量。如在医学领域公知的,对于任何患者的剂量依赖于多种因素,包括患者体积、肌体表面积、年龄、施用的具体化合物、性别、给药时间和途径、一般健康情况和同时施用的其它药物。剂量将变化,但是核酸构建物给药的优选剂量是从大约核酸构建物的 10^6 到 10^{12} 个拷贝。当需要时,可重复给药。

[98] 许多给药途径可被用来施用核酸构建物。这些途径的实例包括肌肉内注射,静脉、腹膜内、皮内、鼻内和皮下注射核酸构建物都已经导致针对肌肉内流感病毒血凝素(HA)的免疫(在 Pardoll and Beckerleg, *Immunity* 3 (1995), 165-169 中的总述)。基于核酸的疫苗通过使用吞噬细胞例如巨噬细胞最佳吞噬大小的聚合体的、生物可降解微粒的或微胶囊的输运载体也可被施用。例如,可以使用直径大约 1-10 μ m 的 PLGA (聚-乳酸-共-乙交酯)微粒。核酸构建物被封装入这些微粒,其被巨噬细胞摄取并且逐步在细胞内生物降解,从而释放核酸构建物。释放后,在细胞内表达核酸。另一种达到核酸构建物摄取的方式是通过使用脂质体。这些脂质体可通过标准方法制备。核酸构建物可被单独掺入这些输运载体或者与组织特异性抗体共同掺入。或者,分子结合物可被制备,其包括通过静电或共价力(covalent force)连接到聚-L-赖氨酸的核酸构建物。聚-L-赖氨酸结合于配体,该配体可与靶细胞上受体相结合。Cristiano et al. (1995), *J. Mol. Med.* 73, 479。或者,淋巴组织特异性靶向可通过使用淋巴组织特异性转录调节元件(TRE)例如 B 淋巴细胞、T 淋巴细胞或树突细胞特异性 TRE 达到。淋巴组织特异性 TRE 是已知的(Thompson et al., *Mol. Cell. Biol.*, 12:1043 (1992); Todd et al., *J. Exp. Med.* 177:1663 (1993); Penix et al., *J. Exp. Med.*, 178:1483 (1993))。

[99] 本发明也提供基于微生物的疫苗。一般而言,这些疫苗与用核酸构建物转化的微生物相关,所述核酸构建物提供本发明多肽、肽片段或偶联蛋白的表达。例如,单核细胞增多性李斯特菌(*Listeria monocytogenes*)可被用作载体,以诱导 T-细胞免疫。这是因为其感染抗原递呈细胞,并且也因为感染开始于粘膜。Lieberman and Frankel, *Vaccine*, 20:2007-10 (2002)。因此,李斯特菌属可以用核酸构建物转化,所述核酸构建物提供诱导针对来自引起严重急性呼吸器官综合症的冠形病毒的刺突蛋白产生免疫应答的多肽、肽片段或偶联蛋白的表达。高度减毒形式的李斯特菌属可根据本领域报告的方法进行构建。Lieberman and Frankel. *Vaccine*. 20:2007 (2002)。沙门氏菌也可被用作载体以诱导针对引起严重

急性呼吸器官综合症的冠形病毒的细胞毒性 T 淋巴细胞(CTL)应答。Pasetti et al., *Infect Immun.* 70:4009 (2002)。

[100] 可通过在疫苗领域使用的任何传统途径施用免疫组合物或疫苗。例如，免疫组合物或疫苗可经口或静脉输注施用，或者皮下、肌肉内、腹膜内、直肠内、阴道内、鼻内、胃内气管内或肺内注射。给药途径的选择依赖于大量参数例如有效成分的性质；多肽、肽片段、模拟肽、偶联蛋白、免疫肽、DNA 疫苗的同源性；或与前述分子组合的佐剂。施用免疫组合物可以以单次剂量或以在一定时期内重复一次或多次的剂量进行。根据不同的参数变化适当的剂量。这些参数包括治疗的个体(成年或幼年)、免疫组合物或抗原本身、给药方式和频率、存在或不存在佐剂——如果存在，佐剂的类型——和期望的效果(例如预防或治疗)，如本领域普通技术人员所确定的。

IV. 本发明的抗体和适配体

[101] 本发明提供与如在 SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63, 66, 69 中提出的氨基酸序列结合的抗体或与 SEQ ID NO: 1 片段结合的抗体，或它们的保守变体。这类抗体可用于诊断、免疫和治疗严重急性呼吸器官综合症(SARS)。在一些实施方式中，抗体与具有 SEQ ID NO:58 或 59 的肽结合。与 P540 肽(SEQ ID NO:59)结合的抗体是高度有效的，甚至在高度稀释后，可检测刺突多肽。例如，以 1:10,000 稀释的 P540 抗体制备物仍可检测刺突多肽。

[102] 可使用作为免疫原的完整目的多肽或肽片段制备抗体。用来免疫动物的多肽或片段可源自翻译的 cDNA 或化学合成。如果需要，可将多肽或肽片段偶联到载体蛋白。这种通常使用的化学连接到肽的载体蛋白包括匙孔血蓝蛋白(KLH)、甲状腺球蛋白、牛血清白蛋白(BSA)和破伤风毒素。可使用偶联的蛋白使动物(例如，小鼠、大鼠或兔)免疫。

[103] 如果需要，可进一步纯化多克隆或单克隆抗体，例如通过结合到基质并从基质洗提，所述基质为增加抗体的多肽或肽片段所结合的。本领域普通技术人员将知道纯化或浓缩多克隆抗体以及单克隆抗体的免疫学领域的多种常规技术(Coligan, et al., Unit 9, *Current Protocols in Immunology*, Wiley Interscience, 1991, 其被引入作为参考)。

[104] 也可能使用抗个体基因型技术以产生模拟表位的单克隆抗体。例如，用于产生第一单克隆抗体的抗个体基因型的单克隆抗体，将在高可变区具有结合结构域，其是第一单克隆抗体结合的表位的“图像”。

[105] 适合与多肽或肽片段结合的抗体对于该多肽的至少一部分区域具有特异性。例如，本领域普通技术人员使用肽片段产生适当的本发明的抗体。本发明的抗体包括多克隆抗体、单克隆抗体以及多克隆和单克隆抗体的片段。

[106] 制备多克隆抗体是本领域普通技术人员已知的(Green et al., Production of Polyclonal Antisera, in Immunochemical Protocols (Manson, ed.), pages 1-5 (Humana Press 1992); Coligan et al., Production of Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters, in Current Protocols in Immunology, section 2.4.1 (1992), 它们被引入本文作为参考)。例如，优选根据预决定的进度表联合一种或多种促升免疫法，将多肽或肽片段注射入动物宿主，并且动物被周期放血。然后，从这些抗血清中纯化对于多肽或肽片段特异性的多克隆抗体，例如使用偶联到适当固体载体的多肽或肽片段的亲和色谱。

[107] 同样，制备单克隆抗体是常规的(Kohler & Milstein, Nature. 256:495 (1975); Coligan et al., sections 2.5.1-2.6.7; 和 Harlow et al., Antibodies: A Laboratory Manual, page 726 (Cold Spring Harbor Pub. 1988)), 它们被引入本为作为参考。简言之，通过将含有抗原的组合物注入小鼠、通过除去血浆样品证实存在抗体产生、除去脾脏以获得 B 淋巴细胞、将 B 淋巴细胞与骨髓瘤细胞融合以产生杂交瘤、克隆该杂交瘤、选择产生针对抗原的抗体的阳性克隆和从杂交瘤培养物分离抗体，可获得单克隆抗体。可通过大量已经建立的方法从杂交瘤培养物分离和纯化单克隆抗体。这类分离技术包括使用蛋白 A 琼脂糖凝胶的亲和色谱、分子排阻色谱法和离子交换柱层析(Coligan et al., sections 2.7.1-2.7.12 and sections 2.9.1-2.9.3; Barnes et al., Purification of Immunoglobulin G (IgG), in Methods in Molecular Biology, Vol. 10, pages 79-104 (Humana Press 1992))。单克隆抗体体外和体内增殖的方法是本领域普通技术人员公知的。在适合的培养基例如 Dulbecco's Modified Eagle Medium 或 RPMI 1640 培养基中可实行体外增殖，例如正常小鼠腹腔渗出细胞、脾细胞、骨髓巨噬细胞，所述培养基任选地补充哺乳动物血清例如胎牛血清或微量元素和助长补充物(growth-sustaining supplements)。体外产生提供相对纯的抗体制备物，并且允许增大到产生大量需

要的抗体。可通过在空气反应器(air reactor)、连续搅拌反应器、或固化活截留细胞培养物中，同族悬浮培养(homogenous suspension culture)实行大规模杂交瘤培养。可通过将细胞克隆注入与亲代细胞组织相容的哺乳动物例如 osyngeneic 小鼠以引起产生抗体的肿瘤的生长，实行体内增殖。任选地，在注入之前，动物用烃类特别是油例如原始状态的四甲基正十五烷预处理。在一到三周后，从动物的体液回收期望的单克隆抗体。

[108] 也可通过使用噬菌体呈现技术(phage display techniques)制备抗体。在一个实例中，用抗原例如本发明的多肽或肽片段免疫生物体。从免疫的生物体的脾脏分离淋巴细胞。从脾细胞分离总 RNA，并且包含在总 RNA 中的 mRNA 被反转录入互补的脱氧核糖核酸(cDNA)。编码免疫球蛋白轻链和重链的可变区的 cDNA 通过聚合酶链式反应(PCR)进行扩增。为了产生单链片段可变(scFV)抗体，轻链和重链扩增产物可被结合重叠延伸 PCR 连接以产生完整序列，并将其连接入适当载体。然后用编码 scFV 的载体转化大肠杆菌，并且用辅助噬菌体对大肠杆菌感染，以产生在其表面上展示抗体的噬菌体颗粒。或者，为了产生完整抗原结合片段(Fab)，重链扩增产物可与编码噬菌体外壳蛋白的核酸序列融合，并且将轻链扩增产物克隆入适当的载体。用编码轻链扩增产物的载体转化表达融合到噬菌体外壳蛋白的重链的大肠杆菌。在大肠杆菌的周质中产生轻链和重链之间的二硫键。该过程的结果是产生高达 10^9 个克隆的抗体文库。通过后来加入来自相同或不同宿主的另外的免疫生物体的免疫应答，文库的大小可被增加到 10^{18} 个噬菌体。识别特异性抗原的抗体可通过淘选进行选择。简言之，完整抗体文库可被暴露于抗体期望免疫的抗原。不表达结合于抗原的抗体的噬菌体被洗掉。表达期望抗体的噬菌体被固定于抗原上。然后，洗提这些噬菌体，并再一次在大肠杆菌中扩增。重复该过程以富集表达特异性结合于抗原的抗体的噬菌体的群落。在表达结合于抗原的抗体的噬菌体被分离后，含有抗体编码序列的载体可以从噬菌体颗粒分离，并且编码序列可被重新克隆入适合的载体，以产生可溶形式的抗体。在另一个实例中，可使用人噬菌体文库选择抗体，例如单克隆抗体，所述抗体与 SARS-CoV 的刺突蛋白结合。简言之，可以从 SARS-CoV 感染或没有感染的人中分离脾细胞，并且使用其根据上面描述和本领域已知的方法产生人噬菌体文库。可使用这些方法来获得与 SARS-CoV 的刺突蛋白结合的人单克隆抗体。分离抗原和抗体的噬菌体展示方法在本领域是已知的，并且

已被描述(Gram et al., Proc. Natl. Acad. Sci.. 89:3576 (1992); Kay et al., Phage display of peptides and proteins: A laboratory manual. San Diego: Academic Press (1996); Kermani et al., Hybrid.14:323 (1995); Schmitz et al., Placenta.21 Suppl. A:S106 (2000); Sanna et al., Proc. Natl. Acad. Sci.. 92:6439 (1995))。

[109] 本发明的抗体可以得自"人化的"单克隆抗体。通过将从小鼠免疫球蛋白的重或轻的可变链得到的小鼠互补决定区转移入人可变结构域,然后在鼠类对应物的框架区取代人的残基,产生人化的单克隆抗体。应用从人化单克隆抗体得到的抗体成分消除了与鼠类恒定区免疫原性相关的潜在问题。用来克隆鼠类免疫球蛋白可变结构域的一般技术被描述(Orlandi et al., Proc. Nat'l Acad. Sci. USA. 86:3833 (1989), 其整体被引入本文作为参考)。产生人化单克隆抗体的技术被描述(Jones et al., Nature, 321:522 (1986); Riechmann et al., Nature, 332:323 (1988); Verhoeyen et al, Science. 239:1534 (1988); Carter et al., Proc. Nat'l Acad. Sci. USA. 89:4285 (1992); Sandhu, Crit. Rev. Biotech.. 12:437 (1992); 和 Singer et al., J. Immunol., 150:2844 (1993), 它们被引入本文作为参考)。

[110] 另外,本发明的抗体也可源自人单克隆抗体。从转基因小鼠获得这类抗体,所述转基因小鼠已被"基因改造",以产生对抗原攻击应答的特异性人抗体。在该技术中,人重链和轻链位置要素被引入从胚胎干细胞系得到的小鼠的菌株,所述胚胎干细胞系含有内生的重链和轻链位置的靶向破裂。这些转基因小鼠可以合成对人抗原特异性的人抗体,并且该小鼠可被用来产生人抗体分泌型杂交瘤。从转基因小鼠获得人抗体的方法被描述(Green et al., Nature Genet., 7:13 (1994); Lonberg et al., Nature, 368:856 (1994); 和 Taylor et al., Int. Immunol.. 6:579 (1994), 它们被引入本文组为参考)。

[111] 本发明的抗体片段可通过抗体的蛋白水解作用或通过编码片段的DNA在大肠杆菌中的表达制备。抗体片段可通过常规方法用胃蛋白酶或木瓜蛋白酶消化整个抗体来获得。例如,抗体片段可以通过用胃蛋白酶酶切割抗体来产生,以提供F(ab')₂指示的5S片段。可使用硫醇还原剂进一步切割该片段,并且任选地使用二硫键切割产生的巯基的保护基,以产生3.5S Fab'单价片段。或者,使用胃蛋白酶酶切直接产生两个单价Fab'片段和Fc片段。这些方法被描述(美国专利号 4,036,945; 4,331,647; 和 6,342,221, 以及包含在其中的参考文献; Porter, Biochem. J., 73:119 (1959); Edelman et al., Methods in Enzymology, Vol. 1, page

422 (Academic Press 1967); 和 Coligan et al. at sections 2.8.1-2.8.10 and 2.10.1-2.10.4)。

[112] 也可使用其它切割抗体的方法, 例如分离重链以形成单价轻-重链片段, 进一步切割片段或其它酶、化学或遗传技术, 只要该片段与完整抗体识别的抗原结合。

[113] 例如, Fv 片段包括 V_H 和 V_L 链的连接。该连接可以是非共价的(Inbar et al., Proc. Nat'l Acad. Sci. USA, 69:2659 (1972))。或者, 可变链可以通过分子间二硫键连接, 或者通过化学物质例如戊二醛交联(Sandhu, Crit. Rev. Biotech., 12:437 (1992))。优选地, Fv 片段包括通过肽联结子(peptide linker)连接的 V_H 和 V_L 链。通过构建含有寡核苷酸连接的编码 V_H 和 V_L 结构域的 DNA 的结构基因, 制备这些单链抗原结合蛋白(sFv)。将该结构基因插入到表达载体, 其接下来被引入到宿主细胞例如大肠杆菌。该重组宿主细胞合成具有桥连两个 V 结构域的连接子肽(linker peptide)的单一多肽链。产生 sFvs 的方法被描述(Whitlow et al., Methods: A Companion to Methods in Enzymology, Vol. 2, page 97 (1991); Bird et al., Science. 242:423 (1988), Ladner et al., U.S. patent No. 4,946,778; Pack et al., Bio/Technology, [upsilon]:1271 (1993); 和 Sandhu, Crit. Rev. Biotech.. 12:437 (1992))。

[114] 另一形式的抗体片段是形成单一互补决定区(CDR)的肽。通过构建编码目的抗体 CDR 的基因, 可获得 CDR 肽(“最小识别单位”)。制备这类基因, 例如通过使用聚合酶链式反应从产生抗体细胞的 RNA 合成可变区(Larrick et al., Methods: A Companion to Methods in Enzymology, Vol. 2, page 106 (1991))。

[115] 本发明的抗体可与毒素偶联。可使用这类抗体治疗已被病因学上与严重急性呼吸器官综合症相关的病毒感染的动物包括人。例如, 和病因学上与严重急性呼吸器官综合症相关的冠形病毒的刺突蛋白相结合的抗体可与破伤风毒素相偶联, 并且将其施用给遭受前述病毒感染的动物。认为毒素偶联的抗体与呈现在感染细胞上的部分刺突蛋白相结合, 然后杀死感染的细胞。

[116] 可将本发明的抗体与可检测标记物偶联。这类抗体可被用于确定动物例如人是否感染 SARS-CoV 的诊断分析中。可检测标记物的实例包括荧光蛋白(即, 绿色荧光蛋白、红色荧光蛋白、黄色荧光蛋白)、荧光标记(即, 异硫氰酸荧光素、若丹明、德克萨斯红)、放射标记(即, 3H 、 ^{32}P 、 ^{125}I)、酶(即, β -半乳糖苷酶、辣

根过氧化物酶(horseradish peroxidase)、 β -葡糖苷酸酶、碱性磷酸酯酶)或者亲和标记物(即, 抗生物素蛋白、生物素、链霉抗生物素)。将抗体偶联到可检测标记物的方法是本领域已知的。Harlow et al., *Antibodies: A Laboratory Manual*, page 319 (Cold Spring Harbor Pub. 1988)。

[117] 本发明也提供本发明的多肽和肽片段的适配体。本发明的适配体可以是肽或核酸适配体。肽适配体是与本发明的多肽或肽片段结合的肽, 其亲和力通常比得上单克隆抗体-抗原复合体的亲和力。相似地, 核酸适配体是与本发明的多肽或肽片段结合的核酸, 其具有强的亲和力, 例如, 亲和力通常比得上单克隆抗体-抗原复合体的亲和力。

[118] 在一个实例中, 可以通过使用随机寡核苷酸序列的文库来分离核酸适配体。筛选该文库以确定哪一种寡核苷酸与本发明的 S 多肽和肽片段结合。结合的寡核苷酸被从固定的多肽或肽片段洗提, 然后通过 PCR 扩增。可以重复该过程以选择对本发明的多肽和肽片段具有高亲和力的适配体。然后, 可确定编码适配体的核酸序列, 并将其克隆到适当的载体以帮助产生和保持期望的适配体。

[119] 通过含有启动子、起始密码子、编码随机肽的核酸序列的文库的 mRNA 展示, 可以分离肽适配体。在一些实施方式中, DNA 文库也包括编码组氨酸标记物的核酸片段。在含有多聚腺苷酸连接子的嘌呤霉素被连接到新形成的 mRNA 的 3'端之后, 使用适合的聚合酶例如 T7 RNA 聚合酶转录该文库。当这些 mRNAs 在体外翻译时, 新生肽形成连接子的嘌呤霉素的共价键以形成 mRNA-肽融合分子。然后, 通过使用 Ni-NTA 琼脂糖和寡-dT-纤维素纯化该 mRNA-肽融合分子。然后, 反转录融合分子的 mRNA 部分。然后, 用本发明的多肽或肽片段温育双链 DNA/RNA-肽融合分子, 并且未结合的融合分子被洗掉。结合的融合分子从固定的多肽或肽片段洗提, 然后通过 PCR 扩增。可重复该过程以选择对本发明的多肽和肽片段具有高亲和性的适配体。然后可确定编码适配体的核酸序列, 并将其克隆入适当的载体。制备肽适配体的方法已被描述(Wilson et al., *Proc. Natl. Acad. Sci.*, 98:3750 (2001))。因此, 本发明提供识别本发明多肽和肽片段的适配体。

V. 本发明的药物组合物

[120] 本发明提供含有与如在 SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63, 66, 69 中提出的氨基酸序列结合的抗体或与 SEQ ID NO: 1 片段结合的抗体或它们的保守变体的药物组合物, 和药学上可接受的载体。在一些实施方式中, 抗体与具有 SEQ ID NO:58 或 59 的肽结合。与 P540 肽(SEQ ID NO:59)结合的抗体是高度有效的, 甚至在高度稀释后, 可检测刺突多肽。例如, 以 1:10,000 稀释的 P540 抗体制备物仍可检测刺突多肽。

[121] 本发明的药物组合物可以以多种形式制备, 其包括片剂、硬或软胶囊、水溶液、悬液和脂质体以及其它缓释剂型, 例如成型的聚合凝胶。可以配制口服剂型, 以便通过胃后将抗体释放到肠。这类剂型在美国专利号 6,306,434 和包含在其中的参考文献中描述。

[122] 口服液体药物组合物可以是下列形式, 例如, 水或油悬液、溶液、乳液、糖浆剂或酞剂, 或者可以表现为干燥产物在使用前与水或其它适合载体组合。这类药物组合物可以含有常规添加剂例如悬浮剂、乳化剂、无水载体(其可包含食用油)或防腐剂。

[123] 可以配制抗体用以肠胃外给药(例如, 通过注射诸如单次快速静脉注射或连续输注), 抗体可以以单位剂型存在于安瓿、预装注射器、小体积输注容器或添加防腐剂的多剂量容器中。药物组合物也可采用如此形式如油或水载体中的悬液、溶液或乳液, 其也可含有配制药剂例如悬浮剂、稳定剂和/或分散剂。适合直肠给药的药物组合物可被制备为单位剂量栓剂。适合的载体包括盐水溶液和其它在本领域共同使用的物质。

[124] 对于吸入给药, 抗体可便利地从吹入器、喷雾器或增压包(pressurized pack)或其它输送喷雾剂的便利方式输送。增压包可包括适当的推进剂例如二氯二氟甲烷、三氯氟甲烷、二氯四氟乙烷、二氧化碳或其它适当的气体。在增压气溶胶的情况下, 可以通过提供输送测定量的阈值来确定剂量单位。

[125] 或者, 对于吸入或喷入给药, 抗体可采用干粉组合物的形式, 例如调节剂和适当粉末基本成分例如乳糖或淀粉的粉末混合物。粉末组合物可以以单位剂型存在于例如胶囊或药筒例如明胶或发泡药包中, 粉末可以在吸入器或吹入器的帮助下从其中给药。对于鼻内给药, 可以通过液体喷雾器例如通过塑料瓶喷雾器施用抗体。

[126] 本发明的药物组合物也可包含其它成分例如调味剂、着色剂、抗微生物剂或防腐剂。应当理解用于治疗需要的抗体的量将不但随着选择的特定载体变化，而且随着给药途径、被治疗的病症的性质以及患者的年龄和病症而变化。最终，护理健康卫生的供应者可以确定实际的剂量。另外，也可配制药物组合物为单一单位剂型。

VI. 免疫、治疗和诊断严重急性呼吸器官综合症动物的方法

[127] 本发明提供免疫动物免于严重急性呼吸器官综合症的方法。该方法涉及将治疗有效量的与如在 SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63, 66, 69 中提出的氨基酸序列结合的抗体或与 SEQ ID NO: 1 片段结合的抗体或它们的保守变体施用给动物；将有效量的免疫组合物施用给动物；将有效量的病毒疫苗施用给动物；或将有效量的核酸疫苗施用给动物。该动物可以是哺乳动物例如人。施用疫苗和免疫组合物的方法已在本文描述并且在本领域是已知的。

[128] 通过根据本发明的被动免疫，也可以治疗 SARS-CoV 感染的动物。例如，可将与如在 SEQ ID NO: 1, 13, 14, 15, 20-55, 60, 61, 62, 63, 66, 69 中提出的氨基酸序列结合的抗体或与 SEQ ID NO: 1 片段结合的抗体或它们的保守变体施用给感染了 SARS-CoV 的动物例如人。这类给药在患者免疫妥协并且不能产生对抗 SARS-CoV 的或对疫苗或免疫组合物有效免疫应答的情况下是适合的。

[129] 本发明提供诊断动物内严重急性呼吸器官综合症的方法，其包括将从该动物获得的生物样品例如组织样品、血液、粘液或唾液与如此抗体接触，所述抗体与如在 SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63 中提出的氨基酸序列结合或与 SEQ ID NO: 1 片段结合；和确定抗体是否与生物样品结合。使用抗体检测生物样品中抗原存在的诊断分析在本领域是已知的。简言之，本发明的抗体可以固定在表面。然后，将生物样品与固定的抗体接触，以便包含在样品中的抗原与抗体结合以形成抗体-抗原复合物。然后，可任选洗涤生物样品以除去未结合的物质。然后，与可检测标记物例如酶或放射标记结合的本发明的第二抗体可被与抗体-抗原复合物接触，以便酶或放射标固定在该表面。然后，对可检测标记物进行检测以确定抗原是否存在于该生物样品。在另一个实例中，可将生物样品固定在表面。然后，与可检测标记物结合的本发明的抗体可被与固

定的生物样品接触，并且洗掉任何未结合的物质。然后，对可检测标记物的存在性进行检测以确定该生物样品是否含有抗原。这类分析的实例在本领域是已知的，并且包括酶联免疫吸附分析、放射性免疫测定等。

[130] 也可使用基于核酸的方法诊断严重急性呼吸器官综合症。在一个实例中，可使用聚合酶链式反应(PCR)诊断 SARS-CoV 感染。简言之，从动物获得生物样品例如组织样品、血液、粘液或唾液。然后，使用常规方法例如有机提取，提取该样品中的核酸。然后，将提取的核酸与正向和反向引物——其与编码 SARS 蛋白的核酸退火——、聚合酶、核苷酸和典型的缓冲液混合，其含有使用 SARS 核酸作为模板允许聚合酶延伸正向和反向引物的成分。然后检测在正向和反向引物之间扩增的 DNA 的存在，以确定该样品是否含有 SARS 源核酸。核酸杂交技术例如 RNA 印记法和 DNA 印记法可被用来检测在生物样品中存在 SARS 核酸。

VII. 试剂盒

[131] 本发明提供含有包裹材料和如此抗体的试剂盒，所述抗体与如在 SEQ ID NO: 1, 13, 14, 15, 45, 46, 或 47, 58, 59, 61, 62, 63, 66, 69 中提出的氨基酸序列结合或与 SEQ ID NO: 1 片段结合或与其保守变体结合。该试剂盒也可以包含注射器以将包含在试剂盒中的抗体注射到动物例如人中。在另一个实施方式中，本发明提供含有许多包裹材料和如此抗体的试剂盒，所述抗体与如在 SEQ ID NO: 1, 13, 14, 15, 20-59, 60, 61, 62, 63, 66, 69 中提出的氨基酸序列结合或与 SEQ ID NO: 1 片段结合或与其保守变体结合，配制该抗体用以施用给动物例如人。在一些实施方式中，该抗体与 SEQ ID NO: 59 列出的氨基酸序列结合。在其它实施方式中，该抗体与 SEQ ID NO: 58 列出的氨基酸序列结合。任选地，这样的试剂盒可包括注射器以将包含在试剂盒中的抗体注射到动物例如人中。

[132] 本发明也提供含有包裹材料和 DNA 疫苗的试剂盒，所述 DNA 疫苗具有编码如在 SEQ ID NO: 1, 13, 14, 15, 45, 46, 或 47, 58, 59, 61, 62, 63, 66, 69 中提出的氨基酸序列或 SEQ ID NO: 1 片段或其保守变体的 DNA 分子或表达载体。该试剂盒也可以含有用于施用 DNA 疫苗的仪器(例如注射器或基因枪)以将包含在试剂盒中的疫苗施用到动物例如人中。

[133] 本发明也提供含有包裹材料和疫苗组合物的试剂盒，所述疫苗组合物包含具有如在 SEQ ID NO: 1, 13, 14, 15, 45, 46, 或 47, 58, 59, 61, 62, 63, 66, 69 中提出的氨基酸序列或 SEQ ID NO: 1 片段的多肽，或其保守变体。该试剂盒也可以含有用于施用疫苗的仪器(例如注射器)以将包含在试剂盒中的疫苗施用到动物例如人中。

[134] 本发明也提供检测 SARS-CoV 感染的试剂盒，其含有包裹材料和具有如在 SEQ ID NO: 1, 13, 14, 15, 45, 46, 或 47, 58, 59, 61, 62, 63, 66, 69 中提出的氨基酸序列或 SEQ ID NO: 1 片段的多肽，或其保守变体。可将该多肽(一种或多种)固定在固体载体上。可使用这类试剂盒检测针对感染的动物或人的血浆中 SARS-CoV 的抗体。该试剂盒也可包含用于检测该抗体与 S 多肽(一种或多种)结合的手段。

VIII. 从 SARS-CoV 病毒的 Tor2 分离物得到的全长刺突(S)蛋白(氨基酸 1-1255)的氨基酸序列

```

MFIFLLFLTLTSGSDDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSD
TLYLTQDLFLPFYSNVTGFHTINHTFGNPFVIPFKDGIYFAATEKSNVVRG
WVFGSTMNKSQSVIINNSTNVVIRACNFELCDNPFFAVSKPMGTQHTHT
MIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKG
YQPIDVVRDLPSGFNTLKPFIKFLPLGINITNFRAILTAFSPAQDIWGTSAAA
YFVGYLKPTTFMLKYDENGITDAVDCSQNPLAELKCSVKSFEDKGIYQ
TSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVAD
YSVLYNSTFFSTFKCYGVSATKLNLDLCSNVYADSEFVVKGDDVRQIAPG
QTGVLIADYNYKLPDDFMGCVLAWNTRNIDATSTGNVNYKYRYLRHGK

```

LRPFERDISNVPFSPDGKPCPPALNCYWPLNDYGFYTTTGIGYQPYRVV
VLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTTPSSKRFQP
FQQFGRDVSDFDTSVRDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLY
QDVNCTDVSTAHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYE
CDPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTN
FSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRAL
SGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRFS
IEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLT
DDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQ
NVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTL
VKQLSSNFGAISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLI
RAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVV
FLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFS
PQITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSP
DVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKW
PWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSE
PVLKGVKLYHT (SEQ ID NO: 1)

IX. 全长刺突(S)蛋白(核苷酸 1-3768)的核酸序列

ATGTTTATTTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTG
ACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTCAAC
ATACTTCATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTATAGAT
CAGACACTCTTTATTTAACTCAGGATTTATTTCTTCCATTTTATTCTAA
TGTTACAGGGTTTCATACTATTAATCATAACGTTTGGCAACCCTGTCAT
ACCTTTTAAGGATGGTATTTATTTTGCTGCCACAGAGAAATCAAATGT
TGTCCGTGGTTGGGTTTTTGGTTCTACCATGAACAACAAGTCACAGTC
GGTGATTATTATTAACAATTCTACTAATGTTGTTATACGAGCATGTAA
CTTTGAATTGTGTGACAACCCTTTCTTTGCTGTTTCTAAACCCATGGG
TACACAGACACATACTATGATATTCGATAATGCATTTAATTGCACTTT
CGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAGG
TAATTTTAAACACTTACGAGAGITTTGTGTTTAAAAATAAAGATGGGTT
TCTCTATGTTTATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCT
ACCTTCTGGTTTTAACACTTTGAAACCTATTTTAAAGTTGCCTCTTGGT

ATTAACATTACAAATTTTAGAGCCATTCTTACAGCCTTTTCACCTGCT
CAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTGGCTATTTA
AAGCCAACTACATTTATGCTCAAGTATGATGAAAATGGTACAATCAC
AGATGCTGTTGATTGTTCTCAAATCCACTTGCTGAACTCAAATGCTC
TGTTAAGAGCTTTGAGATTGACAAAGGAATTTACCAGACCTCTAATTT
CAGGGTTGTTCCCTCAGGAGATGTTGTGAGATTCCCTAATATTACAAA
CTTGTGTCCTTTTGGAGAGGTTTTTAATGCTACTAAATTCCCTTCTGTC
TATGCATGGGAGAGAAAAAAATTTCTAATTGTGTTGCTGATTACTCT
GTGCTCTACAACCTCAACATTTTTTTCAACCTTTAAGTGCTATGGCGTT
TCTGCCACTAAGTTGAATGATCTTTGCTTCTCCAATGTCTATGCAGAT
TCTTTTGTAGTCAAGGGAGATGATGTAAGACAAATAGCGCCAGGACA
AACTGGTGTTATTGCTGATTATAATTATAAATTGCCAGATGATTTTCA
GGGTTGTGTCCTTGCTTGGAACTAGGAACATTGATGCTACTTCAAC
TGGTAATTATAATTATAAATATAGGTATCTTAGACATGGCAAGCTTA
GGCCCTTTGAGAGAGACATATCTAATGTGCCTTTCTCCCTGATGGCA
AACCTTGCACCCACCTGCTCTTAATTGTTATTGGOCATTAATGATT
ATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTG
TAGTACTTTCTTTTGAACCTTTTAAATGCACCGGCCACGGTTTGTGGAC
CAAATTATCCACTGACCTTATTAAGAACCAGTGTGTCAATTTTAATT
TTAATGGACTCACTGGTACTGGTGTGTTAACTCCTTCTTCAAAGAGAT
TTCAACCATTTCAACAATTTGGCCGTGATGTTTCTGATTTCACTGATT
CCGTTTCGAGATCCTAAAACATCTGAAATATTAGACATTTTCACCTTGCG
CTTTTGGGGGTGTAAGTGTAATTACACCTGGAACAAATGCTTCATCTG
AAGTTGCTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAG
CAATTCATGCAGATCAACTCACACCAGCTTGGCGCATATATTCTACTG
GAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTATAGGAGCTGAG
CATGTCGACACTTCTTATGAGTGCGACATTCCCTATTGGAGCTGGCATT
TGTGCTAGTTACCATAACAGTTTCTTTATTACGTAGTACTAGCCAAAAA
TCTATTGTGGCTTATACTATGTCTTTAGGTGCTGATAGTTCAATTGCTT
ACTCTAATAACACCATTGCTATACCTACTAACTTTTCAATTAGCATT
CTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCTCCGTAGATTGTA
ATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTTCTCC
AATATGGTAGCTTTTGCACACAACCTAAATCGTGCACTCTCAGGTATTG
CTGCTGAACAGGATCGCAACACACAGTGAAGTGTTCGCTCAAGTCAA

CAAATGTACAAAACCCCAACTTTGAAATATTTTGGTGGTTTTAATTTT
TCACAAATATTACCTGACCCTCTAAAGCCAACCTAAGAGGTCTTTTATT
GAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATG
AAGCAATATGGCGAATGCCTAGGTGATATTAATGCTAGAGATCTCAT
TTGTGCGCAGAAGTTCAATGGACTTACAGTGTTGCCACCTCTGCTCAC
TGATGATATGATTGCTGCCTACACTGCTGCTCTAGTTAGTGGTACTGC
CACTGCTGGATGGACATTTGGTGCTGGCGCTGCTCTTCAAATACCTTT
TGCTATGCAAATGGCATATAGGTTCATGGCATTGGAGTTACCCAAA
ATGTTCTCTATGAGAACCAAAAACAATCGCCAACCAATTTAACAAG
GCGATTAGTCAAATTCAAGAATCACTTACAACAACATCAACTGCATT
GGGCAAGCTGCAAGACGTTGTTAACCAGAATGCTCAAGCATTAAACA
CACTTGTTAAACAACCTAGCTCTAATTTTGGTGCAATTTCAAGTGTGC
TAAATGATATCCTTTCGCGACTTGATAAAGTCGAGGCGGAGGTACAA
ATTGACAGGTTAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGT
AACACAACAACCTAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATC
TTGCTGCTACTAAAATGTCTGAGTGTGTTCTTGGACAATCAAAAAGA
GTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCACAAGCA
GCCCCGCATGGTGTGTTCTTCTACATGTCACGTATGTGCCATCCCAG
GAGAGGAACCTCACACAGCGCCAGCAATTTGTCATGAAGGCAAAGC
ATACTTCCCTCGTGAAGGIGTTTTTGTGTTTAATGGCACTTCTTGGTTT
ATTACACAGAGGAACTTCTTTTCTCCACAAATAATTACTACAGACAAT
ACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAAACAAC
ACAGTTTATGATCCTCTGCAACCTGAGCTCGACTCATTCAAAGAAGA
GCTGGACAAGTACTTCAAAAATCATAACATCACCAGATGTTGATCTTG
GCGACATTTCAAGCATTAAACGCTTCTGTCGTCAACATTCAAAAAGAA
ATTGACCGCCTCAATGAGGTGCTAAAAATTTAAATGAATCACTCAT
TGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAATGGCCTT
GGTATGTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTCATGG
TTACAATCTTGCTTTGTTGCATGACTAGTTGTTGCAGTTGCCTCAAGG
GTGCATGCTCTTGTGGTTCTTGTGCTGCAAGTTTGATGAGGATGACTCTG
AGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAA (SEQ ID NO:
2)

实施例 1

刺突蛋白的克隆

[135] 通过使用重叠聚合酶链式反应(PCR), 获得编码全长刺突蛋白的核酸序列。从 British Columbia Cancer Agency (Vancouver, British Columbia)获得含有刺突蛋白片段的重叠克隆。在 PCR 反应中使用下列引物扩增编码 SARS-CoV 全长刺突蛋白的核酸序列: 克隆 1: 正向引物: 5'- A GTC GGA TCC GGT AGG CTT ATC ATT AGA G - 3' (SEQ ED NO: 3); 反向引物: 5'- CCA TCA GGG GAG AAA GGC AC-3 (SEQ ID NO: 4)。克隆 2: 正向引物: 5'- GTG CCT TTC TCC CCT GAT GG-3' (SEQ ID NO: 5); 反向引物: 5'- GAA GAG CAG CGC CAG CAC C-3' (SEQ ID NO: 6)。克隆 3: 正向引物: 5'- GGT GCT GGC GCT GCT CTT C-3' (SEQ ID NO: 7); 反向引物: 5'- A CTG TCT AGA GTT CGT TTA TGT GTA ATG-3 (SEQ ID NO: 8)。

[136] 上述引物对产生的核酸片段之间重叠 PCR 产生的核酸片段含有与严重急性性呼吸器官综合症病因学上相关的病毒(SARS-CoV)的刺突蛋白的第 1 到第 1255 的氨基酸残基。下划线引物序列表示用来将扩增片段克隆入 pCDNA3(+)(Invitrogen, Carlsbad, California)的 BamHI 和 XbaI 的限制性内切酶切割位点。

[137] 全长刺突蛋白基因如图 1 所示克隆。图 1 示出用于编码插入 pCDNA3.1(+)载体的全长刺突蛋白的核酸片段的凝胶(第 2 泳道: BamHI 和 XbaI; 第 3 泳道: HindIII), 所述载体已用限制性内切酶消化。

实施例 2

全长刺突蛋白的氨基末端(S1)和羧基末端(S2)片段的产生

[138] 计算机分析鉴别刺突蛋白的氨基末端(S1)和羧基末端(S2)之间的潜在的功能性分离位点(potential functional separation site)。S1 和 S2 之间的分离位点是在相关于 SEQ ID NO: 1 的 758 和 761(⁷⁵⁸RNTR⁷⁶¹)之间。使用 PCR 产生编码刺突蛋白的氨基末端片段(S1)和羧基末端片段(S2)的核酸。

[139] 使用下列引物, S1 正向引物: 5'-AGTC GGA TCC GAC CGG TGC ACC

ACT TTT G-3' (SEQ ID NO: 9)和反向引物, S1 反向引物: 5'-AGTC GGG CCC CTG TTC AGC AGC AAT ACC-3'(SEQ ID NO: 10)制备编码该刺突蛋白 17-757 氨基酸残基的核酸片段。使用两个限制性位点 BamHI 和 ApaI——两个引物的下划线部分, 将编码刺突蛋白的氨基末端片段(S1)的核酸片段基因克隆到表达用的 pSecTag2B 质粒。

[140] 在 PCR 反应中, 使用下列引物, S2 正向引物: 5'-ACTG GGATCC GAA GTG TTC GCT CAA GTC-3' (SEQ ID NO: 11)和 S2 反向引物: 5'-ACTG TCTAGA TTG CTC ATA TTT TCC C-3' (SEQ ID NO: 12) 制备编码该刺突蛋白 762-1189 氨基酸残基的核酸片段。使用两个限制性位点 BamHI 和 XbaI——两个引物的下划线部分, 将编码刺突蛋白的羧基末端片段(S2)的核酸片段基因克隆到表达用的 pCDNA3.1(+)质粒。

[141] 为了克隆含有残基 272-537 的片段, 使用下列引物对用于 PCR 扩增: 引物 5' GATCGGATCCGGTACAATCACAG 3' (SEQ ID NO: 64)和引物 5' GATCGGGCCCGACACACTGGTTC 3' (SEQ ID NO: 65)。用 BamHI 和 ApaI 消化扩增的片段, 并将其连接到用相同的限制性内切酶消化的 pSecTag2B。全长刺突蛋白中许多可溶性刺突蛋白片段的位置的示意图在图 1B 中提供。

[142] 在一些情况下, 编码 S 片段和全长 S 多肽的核酸具有用小鼠 k 链的前导序列(METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16)取代内生的前导序列(刺突蛋白氨基酸 1-16, MFIFLLFLTLTSGSDL (SEQ ID NO:60))以允许分泌, 如下所述。

实施例 3

缺少胞质尾区和跨膜结构域的完整可溶刺突蛋白(sS)的产生

[143] 使用下列引物对, 产生编码缺少胞质尾区的刺突蛋白(sS)片段的核酸片段, 所述胞质尾区具有 SEQ ID NO: 1 的氨基酸 17-1189: S1 正向引物: 5'-AGTC GGATCC GAC CGG TGC ACC ACT TTT G-3'(SEQ ID NO: 9)和反向引物: 5' ACTG TCTAGA TTG CTC ATA TTT TCC C-3' (SEQ ID NO: 12)。

实施例 4

刺突蛋白的氨基末端和羧基末端片段的表达

[144] 通过将含有 pSecTag2B 或 pCDNA3.1(+)质粒和核酸插入物的表达构建物转染到 293 或 Vero E6 细胞, 进行表达, 所述核酸插入物编码氨基末端(S1)、羧基末端(S2)片段或缺少胞质尾区和跨膜结构域的 SARS-CoV 的刺突蛋白的片段。认为跨膜结构域的消除允许多肽和肽片段可溶于水溶液。然后, 检测编码片段的表达功效。在得到阳性信号之后——如使用凝胶分析确定的, 可产生稳定的转染细胞系。根据常规用于其它高度糖基化蛋白的方法, 纯化全长刺突蛋白及其片段。例如使用用于大量生产的扁豆凝集素(lentil lectin)柱。所形成的蛋白质: 可溶 S1(sS1)、可溶 S2(sS2)和完整可溶 S(sS), 将具有下列氨基酸序列。黑体字表示可被分离的信号肽, 所以分泌的蛋白将不包括它。

刺突蛋白的可溶氨基末端片段的氨基酸序列(氨基酸 17-757)

DRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSN
 VTGFHTINHTFGNPVIFPKDGIYFAATEKSNVVRGWVFGSTMNKSQSVI
 IINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHMIFDNAFNCTFEYISD
 AFSLDVSEKSGNFKHLREFVFKNKDGFYVYKGYQPIDVVRDLPSGFNT
 LKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAFYVGYLKPTTFMLKY
 DENGTTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPVSGDVVRF
 PNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKC
 YGVSATKLNLDLCSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPD
 DFMGCVLAWNTRNIDATSTGNVNYKYRYLRHGKLRPFERDISNVPFSPD
 GKPCPTPALNCYWPLNDYGFYTTTIGIGYQPYRVVVLSELLNAPATVCG
 PKLSTDLIKNQCVNFNENGLTGTGVLTPSSKRFQPFQFGRDVSDFDTSV
 RDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHAD
 QLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTV
 SLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMMAK
 TSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQ (SEQ ID
 NO: 13)

刺突蛋白的可溶羧基末端片段的氨基酸序列(氨基酸 762-1189)

EVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLA
DAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALV
SGTATAGWTFGAGAALQIPFAMQMA YRFNGIGVTQNVLYENQKQIANQ
FNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSV
LNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAAT
KMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNF
TTAPAICHEGKAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTFVSGNC
DVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVV
NIQKEIDRLNEVAKNLNESLIDLQELGKYEQ (SEQ ID NO: 14)

具有 SEO ID NO: 1 的氨基酸 17-757 和 762-1189 的可溶刺突蛋白的氨基酸序列
(缺少信号肽和潜在切割位点)

DRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTDLFLPFYSN
VTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNKSQSVI
IINNSTNVVIRACNFELCDNPFFAVSKPMGTQHTMIFDNAFNCTFEYISD
AFSLDVSEKSGNFKHLREFVFKNKDGFYVYKGYQPIDVVRDLPSGFNT
LKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKY
DENGTTTDAVDCSQNPLAELKCSVKSFEDKGIYQTSNFRVVPSGDVVRF

PNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKC
 YGVSATKLNLDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLFD
 DFMGCVLAWNTRNIDATSTGNVNYKYRYLRHGKLRPFERDISNVPFSPD
 GKPCPPALNCYWPLNDYGFYTTTGIGYQPYRVVLSFELLNAPATVCG
 PKLSTDLIKNQCVNFNFNGLTGTGVLTTPSSKRFQPFQFGRDVSDFDTSV
 RDPKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHAD
 QLTPAWRIYSTGNVVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTV
 SLLRSTSQKSIVAYTMSLGADSSLAYSNNTIAIPTNFSISITTEVMPVSMK
 TSVDCNMYICGDSTECANLLQYGSFCTQLNRALSGIAAEQDEVFAQVK
 QMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQ
 YGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAG
 WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQI
 QESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRL
 DKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVL
 GQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICH
 EGKAYFPREGVVFVNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINN
 TVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRL
 NEVAKNLNESLIDLQELGKYEQ (SEQ ID NO: 15)

实施例 5

刺突蛋白另外的可溶片段的产生

[145] 编码含有 SEQ ID NO: 1 的氨基酸 17-757 的多肽的核酸序列通过使用聚合酶链式反应(PCR)获得。在 PCR 反应期间使用下列引物扩增核酸序列: 正向引物: 5' AGCT GGA TCC GAC CGG TGC ACC ACT TTT G 3' (SEQ ID NO: 9); 和反向引物: 5' AGCT GGG CCC CTG TTC AGC AGC AAT ACC 3' (SEQ ID NO: 10)。用 BamHI 和 ApaI 消化所形成的 PCR 产物, 并且编码具有与 SEQ ID NO: 43 相应的氨基酸序列的多肽。然后, 将消化的 PCR 产物连接到用相同酶消化的 pSecTag2B (Invitrogen, Carlsbad, California)。含有 PCR 产物插入物的 pSecTag2B 构建物编码具有 SEQ ID NO: 46、在 N 末端用于分泌的小鼠 k 链前导序列 (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16)和在 C 末端含有用于亲和纯化的组氨酸标记(HHHHHH)(SEQ ID NO: 18)的 myc 表位(EQKLISEEDL)(SEQ ID

NO: 17)的多肽。

[146] 编码含有 SEQ ID NO: 1 的氨基酸 17-276 的多肽的核酸序列通过使用聚合酶链式反应(PCR)获得。在 PCR 反应期间使用下列引物扩增核酸序列: 正向引物: 5' AGCT GGA TCC GAC CGG TGC ACC ACT TTT G 3' (SEQ ID NO: 9); 和反向引物: 5' CTAG CTC GAG CAA CAG CAT CTG TG 3' (SEQ ID NO: 19)。用 BamHI 和 XhoI 消化所形成的 PCR 产物, 并且编码具有 SEQ ID NO: 44 的氨基酸。然后, 将消化的 PCR 产物连接到用相同酶消化的 pSecTag2B (Invitrogen, Carlsbad, California)。含有 PCR 产物插入物的 pSecTag2B 构建物编码具有 SEQ ID NO: 47、在 N 末端用于分泌的小鼠 k 链前导序列 (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16)和在 C 末端含有用于亲和纯化的组氨酸标记(HHHHHH)(SEQ ID NO: 18)的 myc 表位(EQKLISEEDL)(SEQ ID NO: 17)的多肽。

[147] 通过用 BamHI 和 HincII 消化编码 SEQ ID NO: 43(如前所述)的核酸序列, 得到编码含有 SEQ ID NO: 1 的氨基酸 17-537 的多肽的核酸序列。该产生的核酸片段编码具有 SEQ ID NO: 45 的多肽。将该核酸片段连接到用 BamHI 和 EcoRV 消化的 pSecTag2B 载体。含有 PCR 产物插入物的 pSecTag2B 构建物编码具有 SEQ ID NO: 48、在 N 末端用于分泌的小鼠 k 链前导序列 (METDTLLLWVLLLWVPGSTGD) (SEQ ID NO: 16)和在 C 末端含有用于亲和纯化的组氨酸标记(HHHHHH)(SEQ ID NO: 18)的 myc 表位(EQKLISEEDL)(SEQ ID NO: 17)的多肽。

[148] 在哺乳动物细胞中这些肽片段的表达在图 3 中阐明。该图示出肽片段可以被分泌到表达该肽片段的细胞生长的培养基中。图 3 也表明肽片段可溶于水性介质。

表 1

本发明另外的肽片段的实例

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸位 置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
20	1-100	MFIFLLEFLTLTSGSDLDRCTTFDDVQAP NYTQHTSSMRGVVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRG
21	101-200	WVFGSTMNKSQS VIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFYVYKGY
22	201-300	QPIDVVRDLPSGFNTLKPIFKLPLGINITN FRAILTAFSPAQDIWGTSAAYFVGYLK PTTFMLKYDENGTTTDAVDCSQNPLAEL KCSVKSFEIDKGIY
23	301-400	QTSNFRVVP SGDVVRFPNITNLCPFGEVF NATKFPSVYA WERKKISNCVADYSVLY NSTFFSTFKCYGVSATKLNLDLCSNVYA DSFVVKGDDVRQIAPG
24	401-500	QTGV IADYNYKLPDDFMGCVLAWNTRN IDATSTGNYNKYRYLRHGKLRPFERDI SNVPFSPDGKPCPPALNCYWPLNDYGF YTTTGIGYQPYRVVLS
25	501-600	FELLNAPATVCGPKLSTD LIKNQCVNFN ENGLTGTGVLTPSSKRFQPFQQFGRDVS DFTDSVRDPKTSEILDISPCAFGGVSVITP GTNASSEVAVLYQD
26	601-700	VNCTDVSTAIHADQLTPAWRIYSTGNNV FQTQAGCLIGAEHVDTSYECDIPIGAGIC ASYHTVSLLRSTSQKSIVAYTMSLGADS SIAYSNNTIAIPTNF
27	701-800	SISITTEVMPVSMAKTSVDCNMYICGDST ECANLLQYGSFCTQLNRALSGIAAEQD RNTREVFAQVKQMYKTPTLKYFGGFNF SQLPDPLKPTKRSFI
28	801-900	EDLLFNKVT LADAGFMKQYGECLGDIN ARDLICAQKENGLTVLPPLLTDDMLAAY TAALVSGTATAGWTFGAGAALQIPFAM QMAYRFNGIGVTQNVLYE

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
29	901-1000	NQKQIANQFNKAISQIQESLTTTSTALGK LQDVVNQNAQALNTLVKQLSSNFGAISS VLNDILSRLDKVEAEVQIDRLITGRQLQSL QTYVTQQLIRAAEI
30	1001-1100	RASANLAATKMSECVLGQSKRVDFCGK GYHLMSEFPQAAPHGVVFLHVTVVPSQE RNFTTAPAICHEGKAYFPREGVVFVNGT SWFITQRNFFSPQIITTD
31	1101-1189	NTFVSGNCDVVIGIINNTVYDPLQPELDS FKEELDKYFKNHTSPDVDLGDISGINASV VNIQKEIDRLNEVAKNLNESLIDLQELGK YEQ
32	1-200	MFIFLLFLTLTSGSDLDRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLQ DLFLPFYSNVTGFHTINHTFGNPVIFPKD GIYFAATEKSNVVRGWVFGSTMNKSQ SVIINNSTNVVIRACNFELCDNPFPAVSK PMGTQHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFVYK GY
33	201-400	QPIDVVRDLPSGFNTLKPIFKLPLGINITN FRAILTAFSPAQDIWGTSAAYFVGYLK PTTFMLKYDENGTTTDAVDCSQNPLAEL KCSVKSEIDKGIYQTSNFRVPSGDVVR FPNITNLCPFGEVFNATKFPSVYAWERK KISNCVADYSVLYNSTFFSTFKCYGUSA TKLNDLCFSNVYADSFVVKGDDVVRQIAP G
34	401-600	QTGVIADYNYKLPDDFMGCVLAWNTRN IDATSTGNVNYKYRYLRHGKLRPFERDI SNVPFSPDGKPCPPALNCYWPLNDYGF YTTTGIGYQPYRVVVLSELLNAPATVC GPKLSTDLIKNCVNFNFENGLTGTGVLT PSSKRFPFQQFGRDVSDFTDSVRDPKTS EILDISPCAFGGVSVITPGTNASSEVAVLY QD

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
35	601-800	VNCTDVSTAIHADQLTPAWRIYSTGNNV FQTQAGCLIGAEHVDTSYECDPIGAGIC ASYHTVSLLRSTSQKSIVAYTMSLGADS SLAYSNNTIAIPTNFSISITTEVMPVSMK TSVDCNMYICGDSTECANLLLQYGSFCT QLNRALSGIAAEQDRNTREVFAQVKQM YKTPTLKYFGGFNFSQILPDPLKPTKRSFI
36	801-1000	EDLLFNKVTLADAGFMKQYGECLGDIN ARDLICAQKFNGLTVLPPLLTDDMIAAY TAALVSGTATAGWTFGAGAALQIPFAM QMA YRFNGIGVTQNVLYENQKQIANQF NKAISQIQESLTTTSTALGKLQDVVNQN AQALNTLVKQLSSNFGAISSVLNDILSRL DKVEAEVQIDRLITGRLQSLQTYVTQQLI RAAEI
37	1001-1189	RASANLAATKMSECVLGQSKRVDFCGK GYHLMSFPQAAPHGVVFLHVTYVPSQE RNFTTAPAICHEGKAYFPREGVVFVNGT SWFITQRNFFSPQIITTDNTFVSGNCDVVI GIINNTVYDPLQPELDSFKEELDKYFKNH TSPVDLGDISGINASVVNIQKEIDRLNE VAKNLNESLIDLQELGKYEQ
38	1-400	MFIFLLFLTLTSGSDLDRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTD DLFLPFYSNVTGFHTINHTFGNPVIPPFD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIINNSTNVVIRACNFELCDNPPFAVSK PMGTQTHMFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAYFVGY LKPTTFMLKYDENGTTTDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVPDGDV VRFNITNLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNLCFSNVYADSFVVKGDDVRQI APG

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
39	1-600	MFIFLLEFLTLTSGSDDLDRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTEGPNVIPFKD GIYFAATEKSNVVRGWVFGSTMNKSQ SVIINNSTNVVIRACNFELCDNPFFAVSK PMGTQHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAAYFVGY LKPTTFMLKYDENGITTDAVDCSQNPLA ELKCSVKSFEDKGIYQTSNFRVVPBGDV VRFPNITNLCPFGEVFNATKFPSVYAW RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNVNYKYRYLRHGKLRPFE RDISNVFSPDGKPCPPALNCYWPLND YGFYTTTGIGYQPYRVVLSFELLNAPA TVCGPKLSTDLIKNCVNFNFNGLTGTG VLTPSSKRFQPFQFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQD

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
40	1-800	MFIFLLFLTLTSGSDLDRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYLTQ DLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNKSQ SVIINNSTNVVIRACNFELCDNPPFAVSK PMGTQHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGLYVYK GYQPIDVVRDLPSGFNTLKPFLPLGINI TNFRAILTAFSPAQDIWGTSAAYFVGY LKPTTFMLKYDENGITDAVDCSQNPLA ELKCSVKSFEIDKGIYQTSNFRVVP SGDVVRFPNITNLCPFGEVFNATKFPSVY AWERKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNLCFSNVYADSFVVKGDDVVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNVNYKYRYLRHGKLRPFE RDISNVFPSPDGKPCPPALNCYWPLND YGFYTTTGIGYQPYRVVLSFELLNAPA TVCGPKLSTDLIKNOQVNFNFNGLTGTG VLTPSSKRFQPFQFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVIIPGTNASSEVA VLYQDVNCTDVSTAHADQLTPAWRIYS TGNNVFQTQAGCLIGAEHVDTSYECDIPI GAGICASYHTVSLRSTSQKSIVAYTMSL GADSSIAYSNTIAIPTNFSISITTEVMPVS MAKTSVDCNMYICGDSTECANLLLQYG SFCTQLNRALSGIAAEQDRNTREVFAQV KQMYKTPTLKYFGGFNFSQILPDPLKPT KRSFI

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
41	1-1000	MFIFLLFLTLTSGSDLRCTTFDDVQAP NYTQHTSSMRGVVYPDEIFRSDTLYL DLFLPFYSNVTGFHTINHTFGNPVIP FDGIYFAATEKSNVVRGWVFGSTMNKSQ SVIIIINNSTNVIRACNFELCDNPFFAVSK PMGTQHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFNKNDGFLYVYK GYQPIDVVRDLPSGFNTLKPIFKLPLGINI TNFRAILTAFSPAQDIWGTSAAYFVGY LKPTTFMLKYDENGTTIDA VDCSQNPLA ELKCSVKSEIDKGIYQTSNFRVPSGDV VRFPNITNLCPFGEVFNATKFPSVYAW RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNLCFSNVYADSFVVKGDDVVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNVNYKYRYLRHGKLRPFE RDISNVPFSPDGKPCPPALNCYWPLND YGFYTTTIGIGYQPYR VVLSFELLNAPA TVCGPKLSTD LIKNQCVNFNFNGLTGTG VLTPSSKRFQPFQPFGRDVSDFTDSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQDVNCTDVSTAIHADQLTPAWRIYS TGNNVFQTQAGCLIGAEHVDTSYECDIPI GAGICASYHTVSLLRSTSQKSIVAYTMSL GADSSIAYSNNTIAIPTNFSISITTEVMPVS MAKTSVDCNMYICGDSTECANLLLQYG SFCTQLNRALSGIAAEQDRNTREVFAQV KQMYKTPTLKYFGGFNFSQILPDPLKPT KRSFIEDLLFNKVTLADAGFMKQYGECL GDINARDLICAQKFNGLTVLPPLLTDDMI AAYTAALVSGTATAGWTFGAGAALQIP FAMQMAYRFNGIGVTONVLYENQKQIA NQFNKAISQIQESLTTTSTALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILS RLDKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEI

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
42	1-1189	MFIFLLFLTLTSGSDLDRCTTFDDVQAP NYTQHTSSMRGVYYPDEIFRSDTLYL TLQDLFLPFYSNVTGFHTINHTFGNPVIPFKD GIYFAATEKSNVVRGWVFGSTMNNKSQ SVIIIINNSTNVVIRACNFELCDNPFFAVSK PMGTQHTMIFDNAFNCTFEYISDAFSLD VSEKSGNFKHLREFVFKNKDGFLYVYK GYQPIDVVRDLPSGFNTLKPFIKPLGINI TNFRAILTAFSPAQDIWGTSAAYFVG YLKPTTFMLKYDENGTTTDAVDCSQNPLA ELKCSVKSFEDKGIYQTSNFRVVP SGDVVRFPNIINLCPFGEVFNATKFPSVYAWE RKKISNCVADYSVLYNSTFFSTFKCYGV SATKLNDLCFSNVYADSFVVKGDDVRQI APGQTGVIADYNYKLPDDFMGCVLAWN TRNIDATSTGNYNKYRYLRHGKLRPFE RDISNVPFSPDGKPCPPALNCYWPLND YGFYTTTGIGYQPYRVVLSFELLNAPA TVCGPKLSTDLIKNCVNFENGLTGTG VLTPSSKRFQPFQFGRDVSDFDTSVRDP KTSEILDISPCAFGGVSVITPGTNASSEVA VLYQDVNCTDVSTAIHADQLTPAWRIYS TGNNVFQTQAGCLIGAEHVDTSYECDIPI GAGICASYHTVSLRSTSQKSIVAYTMSL GADSSIAYSNNITIAIPTNFSISITTEVMPVS MAKTSVDCNMYICGDSTECANLLLOYG SFCTQLNRALSGIAAEQDRNTREVFAQV KQMYKTPTLK YFGGFNFSQILPDPLKPT KRSFIEDLLFNKVTLADAGFMKQYGECL GDINARDLICAQKFNGLTVLPLLTDDMI AAAYTAALVSGTATAGWTFGAGAALQIP FAMQMA YRFNGIGVTQNVLYENQKQIA NQFNKAISQIQESLTTTSTALGKLQDVVN QNAQALNTLVKQLSSNFGAISSVLNDILS RLDKVEAEVQIDRLITGRLQSLQTYVTQ QLIRAAEIRASANLAATKMSECVLGQSK RVDFCGKGYHLMSFPQAAPHGVVFLHV TYVPSQERNFTTAPAICHEGKAYFPREG VFVFNGTSWFITQRNFFSPQIITDNTFVS GNCDVVIGHNNTVYDPLQFELDSFKEEL DKYFKNHTSPDVDLGDISGINASVVNIQ KEIDRLNEVAKNLNESLIDLQELGKYEQ
43	17-100	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTLQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
44	17-200	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHITFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIINNSTNVVIRACN FELCDNPFPAVSKPMGTQHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGY
45	17-400	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHITFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIINNSTNVVIRACN FELCDNPFPAVSKPMGTQHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGITIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNLCFSNVYADS FVVKGDDVRQIAPG
46	17-600	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHITFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIINNSTNVVIRACN FELCDNPFPAVSKPMGTQHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGITIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVPSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVFSPDGKPCPTP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNQCV NFNFNGLTGTGVLTPSSKRFQPFQFGR DVSDFTDSDVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQD

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
47	17-800	<p>DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNKSQSVIIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFLYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTTT DAVDCSQNPLAELKCSVKSFEDKGIYQ TSNFRVVPDGDVVRFPNITNLCPFGEVEN ATKFPSVYAWERKKISNCVADYSVLVNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVFSPDGKPCPTP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNCV NFNFNGLTGTGVLTPSSKRFQPFQFGR DVSDFDTSVRDPKTSEILDSPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNVVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAIP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLQYGSFCTQLNRALSGIAA EQDRNTREVFAQVKQMYKTPTLKYFGG FNFSQILPDPLKPTKRSFI</p>

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
48	17-1000	<p> DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNKSQSVIINNSTNVVIRACN FELCDNPPFAVSKPMGTQHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTEMLKYDENGITIT DAVDCSQNPLAELKCSVKSFEDKGIYQ TSNFRVVPBGDVVRFNITNLCPFGEVEN ATKFPVYAWERKKISNCVADYSVLYNS TFSTFKCYGVSATKLNLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVPFSPDGKCTPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNOCV NFNENGLTGTGVLTPSSKRFQPFQQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNVVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLLQYGSFCTQLNRALSGIAA EQDRNTREVFAQVKQMYKTPTLKYFGG FNFSQILPDPLKPTKRSFIEDLLFNKVTLA DAGFMKQYGECLGDINARDLICAQKFN GLTVLPPLLTDDMIAAYTAALVSGTATA GWTFGAGAALQIPFAMQMAYRFNGIGV TQNVLYENQKQIANQFNKAISQIQESLTT TSTALGKLQDVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRLDKVEAEVQIDRLI TGRLQSLQTYVTQQLIRAAEI </p>

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
49	17-1189	<p> DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTQDLFLPFYSNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIIIINNSTNVVIRACN FELCDNPFFAVSKPMGTQHTMIEDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTTT DAVDCSQNPLAELKCSVKSFIDKGIYQ TSNFRVVPVSGDVVRFNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVFSPDGKPCPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNOCV NFNENGLTGTGVLTPSSKRFPFQFQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNVVFQTQAGCLIG AEHVDTSYECDIPGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNTIAIP TNFSISITTEVMPVSMKTSVDCNMYICG DSTECANLLQYGSFCTQLNRALSGIAA EQDRNTREVFAQVKQMYKTPTLKYFGG FNFSQILPDPLKPTKRSFIEDLLENKVTLA DAGFMKQYGECLGDINARDLICAQKFN GLTVLPPLTDDMIAAYTAALVSGTATA GWTFGAGAALQIPFAMQMAYRFNGIGV TQNVLYENQKQIANQFNKAISQIQESLTT TSTALGKLQDVVNQNAQALNTLVKQLS SNFGAISSVLNDILSRDKVBAEVQIDRLI TGRLQSLQTYVTQQLIRAABIRASANLA ATKMSECVLGQSKRVDFCGKGYHLMSF PQAAPHGVVFLHVTVVPSQERNFTTAPA ICHEGKAYFPREGVVFVNGTISWFITQRNF FSPQIITDNTFVSGNCDVVIGIINNTVYD PLQPELDSFKEELDKYFKNHTSPDVLG DISGINASVVNIQKEIDRLNEVAKNLNES LIDLQELGKYEQ </p>

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
50	17-276	DRCTTFDDVQAPNYTQHTSSMRGVVYP DEIFRSDTLYL TQDLFLPFYS SNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMN KSQSVI INNSTNVVIRACN FELCDNPFFAVSKPMGTQ HTMIFD NAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGF LYVYKGY QPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTA FSPAQ DIW GTSAAAYFVG YLKPTTF MLKYDENGTTT DAV
51	17-446	DRCTTFDDVQAPNYTQHTSSMRGVVYP DEIFRSDTLYL TQDLFLPFYS SNVTGFHTI NHTFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMN KSQSVI INNSTNVVIRACN FELCDNPFFAVSKPMGTQ HTMIFD NAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGF LYVYKGY QPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTA FSPAQ DIW GTSAAAYFVG YLKPTTF MLKYDENGTTT DAVDCSQNPLAELKCSVK SFEIDKGIY Q TSNFRVVPSGDVVRFPNITNLC PFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFPSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGN YNYKYR YLRHG

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
52	17-537	<p>DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTDLFLPFYSNVTGFHTI NEITFGNPVIPFKDGIYFAATEKSNVVRG WVFGSTMNNKSQSVIINNSTNVIRACN FELCDNPFPAVSKPMGTQTHTMIFDNAF NCTFEYISDAFSLDVSEKSGNFKHLREFV FKNKDGFYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRALTAFSQAQDIW GTSAAAYFVGYLKPTTFMLKYDENGTTT DAVDCSQNPLAELKCSVKSFIDKGIYQ TSNFRVVPVSGDVVRFPNITNLCPFGEVFN ATKFPSVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVIADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVPFSPDGKPCPP ALNCYWPLNDYGFYTTTGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNCQCV NFNENGLTGTGV</p>

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
53	17-757 和 N-末端小鼠 K 链前导序列以及 C-末端 myc 表位和聚多组氨酸标记	METD TLLLWVLLLWVPGSTGDDRCTTF DDVQ APNYTQHTSSMRGVYYPDEIFRSD TLYL TQDLFLPFYSNVTGFHTINHTFGNP VIPFK DGIYFAATEKSNVVRGWVFGSTM NNKS QSVIINNSTNVVIRACNFELCDNPF FAVSK PMGTQHTMIFDNAFNCTFEYIS DAFSL DVSEKSGNFKHLREFVFKNKDGF LYVYK GYQPIDVVRDLPSGFNTLKPIFKL PLGIN ITNFRAILTAFSPAQDIWGTSAAA YFVG YLKPTTFMLKYDENGTTTDAVDCS QNPLA EKCSVKSFEDKGIYQTSNFRVV PSGD VVRFNITNLCFFGEVFNATKFPVS YAWER KKISNCVADYSVLYNSTFFSTFK CYGV SATKLNLCFSNVYADSFVVKGD DVRQI APGQTGVIADYNYKLPDDFMGC VLAWN TRNIDATSTGNVNYKYRYLRHG KLRPF ERDISNVPPSPDGKPCCTPPALNCY WPLND YGFYTTTGIGYQPYRVVVLSEFEL LNAP ATVCGPKLSTDLIKNCVNFNFNG LTGTG VLTSSKRFQPFQQFGRDVSDFT DSVR DPKTSELDISPCAFGGVSVITPGTN ASSE AVLYQDVNCTDVSTAIHADQLTP AWRI YSTGNNVPQTQAGCLIGAEHVDTS YECD IPGAGICASYHTVSLLRSTSQKSIV AYTMS LGADSSIAYSNNTIAIPTNFSISITT EVMP PVSMAKTSVDCNMYICGDSTECAN LLQY GSFCTQLNRALSGIAAEQEQLIS EEDL HHHHHH
54	17-276 和 N-末端小鼠 K 链前导序列以及 C-末端 myc 表位和聚多组氨酸标记	METD TLLLWVLLLWVPGSTGDDRCTTF DDVQ APNYTQHTSSMRGVYYPDEIFRSD TLYL TQDLFLPFYSNVTGFHTINHTFGNP VIPFK DGIYFAATEKSNVVRGWVFGSTM NNKS QSVIINNSTNVVIRACNFELCDNPF FAVSK PMGTQHTMIFDNAFNCTFEYIS DAFSL DVSEKSGNFKHLREFVFKNKDGF LYVYK GYQPIDVVRDLPSGFNTLKPIFKL PLGIN ITNFRAILTAFSPAQDIWGTSAAA YFVG YLKPTTFMLKYDENGTTTDAVEQK LISE DLHHHHHH

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
55	17-537 和 N-末 端小鼠 K 链前 导序列以及 C- 末端 myc 表位 和聚多组氨酸 标记	METD TL LLWV LL LWV PG STG DD RCT TF DDVQ AP NYTQ HT SSMR GV YYP DE IFR SD TLYL TQ DLFL PF YSNV TG FHTI N HTF GNP VIPF KD GIYFA A TEK SN VVRG W VFG STM NNKS Q SVIII NN STNV V IRAC N FELC DN PF FAV SK PMGT Q THIM IF DNA F NCT FE YIS DAF SLD VSEK SG NFKH L REF V FKN KD GF LYV Y KGYQ PI DVVR DL PSG F NIL K PIPK L PLG IN ITNF RA ILTA F SPAQ DI WGT S AAA YFV GY LKPT TF MLK Y DENG TI TTDA V DCS QNE LA ELK CS VKS FE IDK GI YQTS N FRV V PSG D VVR FP NITN LC PFGE V FNAT K FPSV YAW ER KKIS NC VADY SV LYN ST FFST E K CYG VS ATKL ND LCFS NY ADSF V VKGD DVR Q LAPG QT GVI AD YNY K L P DDF M G C VLA WN TRNI D ATST G NYNY K YR Y L R HG KLR PF ERDIS N VFP SP DGK P CTPP AL NCY WPL ND YGFY TT TGIG Y QPYR V VVLS F EL LN AP ATVC GP KLST DL IKN Q CVNF N FNG LTGT GV BQKL ISE EDL HH HHHH

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸 位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
56	17-756 N-末端 没有信号肽	DRCTTFDDVQAPNYTQHTSSMRGVYYP DEIFRSDTLYLTLQDLFLPFYSNVTGFHTI NHTEGPNVIFPKDGIYFAATEKSNVVVRG WVFGSTMNNKSQSVIINNSTNVVIRACN FELCDNPFFAVSKPMGTQIHTMIFDNAF NCTFEYISDAFSLDVSEKSGNEFKHLREFV FKNKDGFYVYKGYQPIDVVRDLPSGFN TLKPIFKLPLGINITNFRAILTAFSPAQDIW GTSAAAYFVGYLKPTTFMLKYDENGITIT DAVDCSQNPLAELKCSVKSFEIDKGIYQ TSNFRVVPDGDVVRFPNITNLCPFGEVFN ATKFPVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVFSPDGKPCPTP ALNCYWPLNDYGFYTTTIGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNOVCV NFNENGLTGTGVLTPSSKRFQPFQFGR DVSDFTDSVRDPKTSEILDISPCAFGGVS VITPGTNASSEVAVLYQDVNCTDVSTAI HADQLTPAWRIYSTGNNVFQTQAGCLIG AEHVDTSYECDIPIGAGICASYHTVSLLR STSQKSIVAYTMSLGADSSIAYSNNTIAP TNFSISITTEVMPVSMAKTSVDCNMYICG DSTECANLLQYGSFCTQLNRALSGIAA E
57	272-537	ITDAVDCSQNPLAELKCSVKSFEIDKGIY QTSNFRVVPDGDVVRFPNITNLCPFGEVFN ATKFPVYAWERKKISNCVADYSVLYNS TFFSTFKCYGVSATKLNDLCFSNVYADS FVVKGDDVRQIAPGQTGVADYNYKLPD DFMGCVLAWNTRNIDATSTGNVNYKYR YLRHGKLRPFERDISNVFSPDGKPCPTP ALNCYWPLNDYGFYTTTIGIGYQPYRVV VLSFELLNAPATVCGPKLSTDLIKNOVCV NFNENGLTGTGV
58	24-39 D24 肽	DVQAPNYTQH TSSMRGC
59	540-555 P540 肽	PSSKRFQPFQFGRDC

SEQ ID 编号	相对于 SEQ ID NO:1 的氨基酸位置	氨基酸序列 黑体表示的氨基酸残基(氨基酸残基 1-16) 识别信号序列
60	1-16 刺突信号序列	MFIFLLFLTLTSGSDL
61	303-537 含有受体结合结构域	SNFRVVP SGDVVRFPNITNLC PFGEVFNATK FPSVYA WERKKISNC CVADYSVLYNST FFSTFKCYGVSATKLN DLCFSNVYADSFVVK GDDVRQIAPGQT GVIADYNYK LPDDFMGCVLAWNTRNID ATSTGNINYKYR YLRHGKLRPFERDISN VPFSPDGK PCTPPALNCYWPLNDYGFYTTGIGYQP YRVVLS FELLNAPATVCGPKLST DLIK NQCVN FN FENGLTGTGV
62	319-517 含有受体结合结构域	ITNLC PFGEVFNAT KF P S VYA WERKKISNC CVADYSVLYNST FFSTFKCYGVSATKLN DLCFSNVYADSFVVK GDDVRQIAPGQT GVIADYNYK LPDDFMGCVLAWNTRNID ATSTGNINYKYR YLRHGKLRPFERDISN VPFSPDGK PCTPPALNCYWPLNDYGFYTTGIGYQP YRVVLS FELLNAPATVCGPKLST
63	319-518 含有受体结合结构域	ITNLC PFGEVFNAT KF P S VYA WERKKISNC CVADYSVLYNST FFSTFKCYGVSATKLN DLCFSNVYADSFVVK GDDVRQIAPGQT GVIADYNYK LPDDFMGCVLAWNTRNID ATSTGNINYKYR YLRHGKLRPFERDISN VPFSPDGK PCTPPALNCYWPLNDYGFYTTGIGYQP YRVVLS FELLNAPATVCGPKLST D
66	317-517 含有受体结合结构域	NLC PFGEVFNAT KF P S VYA WERKKISNC VADYSVLYNST FFSTFKCYGVSATKLN DLCFSNVYADSFVVK GDDVRQIAPGQT GVIADYNYK LPDDFMGCVLAWNTRNID ATSTGNINYKYR YLRHGKLRPFERDISN VPFSPDGK PCTPPALNCYWPLNDYGFYTTGIGYQP YRVVLS FELLNAPATVCGPKLST

实施例 6

刺突蛋白的结构

[149] 为了表征 SARS-CoV S 蛋白质的特性和功能, 编码全长 Tor2 分离物的核

酸被克隆到表达载体,如上所述。Tor2 分离物被进一步描述于 Marra 等 *The genome sequence of the SARS-associated coronavirus*, *Science* 300:1399- 1404 (2003)。产生的克隆包括全长 S 蛋白质(1255 个残基)、外功能区 Se (残基 17-1189)——其仅具有 S 蛋白的细胞外结构域并且刺突蛋白的推定的跨膜结构域和胞质尾区缺失、含有 N 末端 276(SEQ ID NO:50)、537(SEQ ID NO:52)和 756(SEQ ID NO: 56)氨基酸残基(分别为 S276、S537 和 S756)的片段——其包括推定的 16-残基信号序列或小鼠 k 链前导序列、和含有残基 272-537(SEQ ID NO:57)的内部片段(参见图 1B)。

[150] 氨基酸残基 758-761(RNTR)形成部分下列普通基序(general motif),用于前体转化酶切割:



其中 X 是任何氨基酸残基, $n=0、2、4$ 或 6。

[151] S1 亚单位大概被包括在 S756 片段里。该发现与鼠类冠状病毒的 S1 亚单元——例如菌株其 S1 为 769 个残基——和人类冠状病毒 OC43 (778 个残基)的 S1 亚单元的大小相一致。参见 Gallagher & Buchmeier, *Coronavirus spike proteins in viral entry and pathogenesis*, *Virology* 279: 371-374 (2001); Kunkel & Herrler, *Structural and functional analysis of the surface protein of human coronavirus OC43*, *Virology* 195 417: 195-202 (1993)。然而,对于人类冠状病毒 229E,认为 S1 由更短的 547 个残基片段构成,其与 S537 相应。Bonavia et al., *Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E*, *J. Virol.* 77: 2530-2538 (2003)。

[152] 所有的 S 糖蛋白片段和全长 S 糖蛋白在 SDS-PAGE 凝胶上跑胶,为比估计分子量明显高的位置,这表明这些多肽可能翻译后修饰。S276 多肽具有明显为大约 75 kDa 的分子量, S537 具有明显为大约 100-110 kDa 的分子量, S756 具有明显为大约 130-140 kDa 的分子量,以及 Se 和 S 具有明显为大约 200 kDa 或更高的分子量(图 4 和 6)。甚至当低曝光下观察,相应于这些多肽的条带是宽的(图 6; 一些数据没有示出)。这些数据表明 S 糖蛋白和其片段明显观察到糖基化。基于分子量的大概估计,表明 S2 亚单位不像 S756(组成 S1 亚单位)一样严重糖基化。值得注意地,如果假设仅糖基化促成分子质量增加,那么 S276 被严

重糖基化。

[153] 尽管在 SDS-PAGE 凝胶上观察到由于小的蛋白得到的弱的条带，但是从细胞培养物上清液获得的大多数 SARS-CoV S 糖蛋白没有被切割。这些弱的条带的一条在与 S756 的相同位置跑开，这表明不充分切割的可能性(图 4 和 6)。蛋白酶随机消化可以发生，需要进一步的研究来确定 S 糖蛋白的切割对其功能是否是必须的。

实施例 7

大肠杆菌中肽片的表达

[154] 将编码 SEQ ID NO:51 肽片段的核酸片段——其包含 SEQ ID NO: 1 的氨基酸残基——克隆到 pRSET 载体 (Invitrogen, San Diego, CA) 以产生质粒 pRSET-S(17-446)。用 pRSET-S(17-446) 转化大肠杆菌 BL21DE3 细胞，然后用 IPTG 诱导。诱导的结果在图 2 中示出。

实施例 8

使用 T7 启动子驱使本发明的克隆的肽片段的表达

[155] 在转染前一天，人类 293 细胞或猴类 Vero E6 细胞在 5 ml 的 DMEM+10% FBS 培养基中生长到 1.2×10^6 细胞/T25 瓶(60 mm 培养皿)的密度。然后使用 Polyfect (Qiagen) 转染试剂盒根据制造商的方案，用含有编码刺突蛋白多种肽片段插入物的 pSecTag2B 构建物(6 μ g 每种)转染细胞。如上所述制备这些构建物。

[156] 转染 4 个小时后，携带 T7 聚合酶的 VTF7.3 痘苗病毒被用来以 20 MOI(感染复数)感染转染的细胞(Fuerst et al., Proc. Natl. Acad. Sci., 93:11371 (1986))。该方法提供在 pSecTag2B 载体中用 T7 启动子代替 CMV 启动子，其弱的更多(Nussbaum et al., J. Virol., 68:5411 (1994))。感染 3 个小时后，将 1.5 ml 的新鲜培养基加入到细胞，然后将细胞转移到 31°C 培养箱中。细胞另外培养 24 小时，然后，收集培养基。

[157] 在用任何 S 核酸构建物转染的细胞中，没有观察到可测量的致细胞病变性，这表明全长 S 糖蛋白和 S 糖蛋白的可溶片段可能不具有明显的细胞毒性作

用。然后，在更高水平的表达下，这些作用是可能的，并且如下描述的合胞体的形成可导致细胞死亡。

实施例 9

刺突-特异性抗体

[158] 用 0.1 mg 的通过计算机程序选择其免疫原性的多种肽免疫新西兰兔。用 ELISA 和蛋白质印迹检测从免疫兔得到的血清的反应性。从用两种肽免疫的兔得到的血清表现对刺突糖蛋白的最高和特异性的活性，对其进行选择以进一步研究。分别通过肽 DVQAPNYTQH TSSMRGC (SEQ ID NO:58) 和 PSSKRFQPFQQFGRDC (SEQ ID NO:59) 诱发名称为 D24 和 P540 的抗体。另一种抗-SARS-CoV S 糖蛋白多克隆抗体 IMG-542——其识别 S 糖蛋白的氨基酸 288-303，购自 IMGENEX (San Diego, CA)。

实施例 10

刺突多肽的免疫沉淀和纯化

[159] 可溶性刺突多肽片段从 Vero E6 或 293 细胞培养基获得。然而，仅在细胞溶胞产物中检测到全长刺突糖蛋白。

[160] 用编码多种可溶 S 片段的核酸转染的细胞的培养基被收集并在 1000g 下经历离心 10 分钟，以除去细胞碎片。清除的培养基与 Ni-NTA 琼脂糖小球(Qiagen, Valencia, CA)或免疫沉淀抗体加糖蛋白 G-Sepharose 小球(Sigma, St. Louis, MO)一起在 4°C 下温育 2 h。然后混合小球和等体积的 SDS 凝胶样品缓冲液，煮沸 3 分钟，并且进行凝胶分析。对于全长 S 糖蛋白，首先在补充有 1% NP-40 和 0.5mM PMSF 的 PBS 中，在 4°C 下裂解细胞 1h，在桌面 Eppendorf 离心机中 14,000 rpm 下，离心 20 min。清除的溶胞产物首先被免疫沉淀或者直接用于蛋白质印迹法。

实施例 11

蛋白质印迹和槽印迹

[161] 首先用基于 PBS 的 NP40S 裂解缓冲液裂解糖蛋白，如上所述，并且通过离心清除碎片。对于可溶性 S 片段，收集培养基并清除，如上所述。对于槽印迹，根据制造商 (Bio-Rad, Hercules, CA) 建议的方案，上清液的清除的溶胞产物或培养基被直接用来印迹消化纤维膜，并且该膜经历抗体检测，如常规蛋白质印迹。对于蛋白质印迹，单克隆抗-c-Myc 表位抗体 (Invitrogen, Carlsbad, CA) 或抗-刺突蛋白兔多克隆抗体——通过刺突肽免疫兔得到的——被稀释于 TBST 缓冲液。抗体和该膜温育 2 h，洗涤，然后该膜与和 HRP 结合的第二抗体一起温育 1 h，洗涤四次(每次 15 min)，然后使用 ECL 试剂 (Pierce, Rockford, IL) 研究。

实施例 12

细胞结合分析和 ELISA

[162] 通过离心收集并清除含有可溶性 S 片段的培养基。将 Vero E6 或其它细胞 (5×10^6) 与 0.5 ml 清除的含有可溶性 S 片段和 $2 \mu\text{g}$ 的与 HRP 结合的抗-c-Myc 表位抗体的培养基在 4°C 下一起温育 2 小时。然后用冰冷 PBS 洗涤细胞三次，并通过离心收集。将细胞沉淀物与从 Roche (Indianapolis, IN) 得到的 ABTS 底物一起在室温下温育 10 min，通过离心清除底物，并且测量 405nm 下的光密度。槽印迹分析结果在图 4 中表示，并借在下面进一步研究。

[163] 对于 ELISA，将纯化的 ACE2 (R&D, Minneapolis, MN) 吸附到 Maxisorp ELISA 平板上 pH 9.6 的缓冲液中，浓度为 100 ng 每孔。含有多种可溶性 S 片段和 $0.6 \mu\text{g}$ 的与 HRP 结合的抗-c-155Myc 表位抗体的培养基 154 ($150 \mu\text{l}$) 在 37°C 的每个孔中温育 2 h。洗涤孔，并且向每个孔加入 $60 \mu\text{l}$ 的 ABTS 底物。20 分钟后，测量光密度 (OD405)。

实施例 13

荧光染料重新分配细胞融合分析

[164] 用编码 S 糖蛋白的质粒转染的 HeLa 或 293T 细胞，被负载有 Calcein AM (分子探针)，其在细胞内被转化为钙荧光素绿 (calcein green)。将细胞与含有 $1 \mu\text{g/ml}$ 的 Calcein AM 的培养基一起在 37°C 、5% CO_2 下温育 1 小时，然后洗涤

并重悬于新鲜培养基中。铺板的靶细胞 Vero E6 被用 CMAC(分子探针)染色, 这通过将 1 μ g/ml 的 CMAC 温育在 37 $^{\circ}$ C、5% CO₂ 的培养基中 30 分钟进行。然后用培养基洗涤细胞两次, 在新鲜培养基中温育 20 分钟, 再次洗涤, 并且每孔覆盖有 0.5 ml 培养基。负载有钙荧光素的 S-表达细胞被加入到靶细胞中, 并且在 37 $^{\circ}$ C、5% CO₂ 下温育 1、2 或 4 小时。测量融合, 其为具有双链的细胞和与 S 糖蛋白表达细胞接触的靶细胞总数之间的比例。使用 Universal Imaging 的 MethaMorph 4.0 软件拍摄显微照片。

实施例 14

基于 β -半乳糖苷酶报告基因的细胞-细胞融合分析

[165] 将 293T 细胞(1.5×10^6)接种于 T25 瓶。第二天, 用 pCDNA3-S、pSectag2B-S、pCDNA3-ACE2 和 pCDNA3-ACE2-Ecto 使用 Polyfect 转染试剂盒(Qiagen, Valencia, CA)根据制造商建议的方案分别地转染这些细胞。转染 4 小时后, 使用表达 T7 聚合酶的痘苗病毒 VTF7.3 感染用 S 构建物转染的细胞, 使用编码 β -gal 的痘苗病毒(VCB21R)感染用 ACE-2 构建物转染的细胞。感染 2 小时后, 细胞和新鲜培养基一起温育, 并且转移到 31 $^{\circ}$ C, 温育整晚。第二天, 表达 S 糖蛋白的细胞和表达 ACE-2 的细胞以 1:1 的比例混合, 并在 37 $^{\circ}$ C 下温育。三个小时后, 通过加入 NP-40 至 0.5% 的终浓度, 裂解细胞。将细胞溶胞产物(50 μ l)和等体积 CPRG 底物混合, 1 小时后, 测量 OD₅₉₅。

实施例 15

刺突多肽在哺乳动物细胞中的表达

[166] 对于某些实验, 除全长 S 糖蛋白外的所有蛋白被标记有 c-Myc 表位和组氨酸标记。在用相应质粒转染然后用表达 T7 聚合酶的痘苗病毒感染后, 在 293 和 Vero E6 细胞中表达这些蛋白。

[167] 使用抗-c-Myc 单克隆抗体检测标记的蛋白(图 4)。如在图 4 中所示, T7 启动子对于表达 S 糖蛋白是高度有效的启动子。在这些实验中, T7 启动子产生比 CMV 启动子更高的表达水平, 在大多数情况下, 其为强启动子(图 4A)。如在

图 4A 中所示, S 片段是可溶性, 并且其培养物上清液中的浓度反比于它们的大小。

实施例 16

抗刺突抗体

[168] 为了能检测未标记蛋白质、确认从抗-c-Myc 抗体获得的数据和定位可能的抗原位点, 研究兔多克隆抗体。针对肽, 培养这些抗体中的两种 D24 和 P540, 其开始残基分别为残基 24 和残基 540。D24 和 P540 抗体制备物特异性识别某些可溶片段(图 4C)。如期望的, D24 识别所有片段; P540 识别 S756、Se 和 S, 但是不识别较小的片段(图 4C; 一些数据没有示出)。D24 抗体制备物是相对较弱。然而, P540 制备物甚至在 1:10,000 稀释时仍是非常灵敏的, 其被大量用于本文描述的实验中。

[169] P540 抗体制备物被用来检测 S 糖蛋白是否细胞内、细胞外或在细胞表面上表达。如在图 5 中所示, 尽管非常低的水平, 但是全长 S 糖蛋白在细胞表面表达, 如流式细胞仪所测量的。

实施例 17

刺突蛋白介导细胞融合

[170] 全长 S 糖蛋白在中性 pH 下介导与表达受体分子的细胞的融合。实施细胞-细胞融合分析以证实全长重组 S 糖蛋白是有功能的, 并且确定 S 蛋白是否需要其它病毒蛋白和/或低 pH 以发挥其融合活性。

[171] 用 pCDNA3-S 和 pSectag2B-S 表达全长 S 糖蛋白, 有效支持与表达 ACE2 的细胞融合, 如通过多种尺寸的合胞体形成和基于 β -gal 报告基因分析所证明的(图 7)。令人感兴趣地, pSectag2B-S 构建物——其中 S 糖蛋白前导肽被小鼠 k 链前导序列所替代——诱导更快的合胞体形成。而且, 形成的合胞体比用 pCDNA3-S 诱导的合胞体更大和更多, 所述 pCDNA3-S 编码内生 S 糖蛋白(数据没有示出)。S 表达型 pSectag2B-S 介导的融合程度也比 pCDNA3-S 的更高, 如通过基于报告基因分析所测量的(图 7B)。这些数据表明天然 S 糖蛋白不能被有

效地输送到细胞表面。这些研究也表示本文描述的 β -gal 分析可用作鉴别 SAR-CoV 进入细胞的抑制剂的快速和定量方法，以及作为研究 SARS-CoV 进入机制的工具。

[172] 值得注意地，当 Vero E6 细胞没有用编码 ACE2 的质粒转染时，使用 β -gal 分析和合胞体形成分析，没有检测到 Vero E6 细胞融合，细胞仅表达天然浓度的受体。为了研究由于这两种分析低敏感度的可能性，使用另一种分析。该新的分析基于能检测单细胞融合的荧光染料重新分配。甚至使用这种基于荧光的分析，没有检测到用编码全长 S 糖蛋白的质粒转化的细胞和多种阴性对照之间的统计学显著差异。一些阴性对照包括不同 pH 下使用编码可溶性 S 片段的质粒的转染(数据没有示出)。今当用编码 ACE2 的质粒转染细胞时，检测到明显的细胞-细胞融合，这表示通过重组 ACE2 表达达到的受体的高水平表达对于细胞-细胞融合可能是重要的。总之，这些结果表明重组 S 糖蛋白可介导细胞融合，该融合可在中性 pH 下发生，并且其功效取决于受体分子的浓度。

[173] 而且，S 糖蛋白的可溶片段抑制 S-介导的细胞融合。如在图 15 中所示，加入 S 片段 S272-537 和 S17-537——其具有如下描述的受体结合结构域，抑制 S-介导的细胞融合。在该分析中，S272-537 (SEQ ID NO:57) 片段表现最大的抑制性。不具有受体结合结构域的 S17-276 片段对 S-介导的细胞融合具有很少或没有抑制。这些数据表明具有受体结合结构域的 S 多肽片段可抑制 SARS-CoV 与动物细胞的融合，从而抑制或防止 SARS-CoV 感染。

[174] 因此，用抗 RBD 抗体、S 多肽、S 肽或适配体阻断、调节或抑制刺突蛋白受体结合结构域的活性可有效预防或治疗 SARS-CoV 感染。

实施例 18

刺突蛋白受体结合结构域的鉴别

[175] 该实施例阐明刺突蛋白受体结合结构域位于残基 272 到 537 (SEQ ID NO:57) 之内，并且可能位于残基 303-537 (SEQ ID NO:61) 之内。后面的实验已经表明含有残基 319-517 (SEQ ID NO:62) 的片段也具有受体结合活性。

[176] 基于多种可溶片段与表达受体的 Vero E6 细胞结合的分析被开发以对 S 糖

蛋白的受体结合结构域(RBD)定位。该分析涉及测量与针对 S 多肽的抗体与 Vero E6 细胞结合相关的荧光,并且在 SARS-CoV 受体识别前,进行该分析。易于 SARS-CoV 感染的 Vero E6 细胞与全长 S 多肽和多种可溶性 S 片段一起温育。相似地,易于 SARS-CoV 感染的几种细胞系与全长 S 多肽和其可溶片段一起温育。

[177] 如在图 8A 和 8B 中所示,除了最小的一个 S 片段(S276),所有的片段、S 片段与 Vero E6 细胞结合。当易于 SARS-CoV 感染的几种细胞系与全长 S 多肽和其可溶片段一起温育时,没有检测到这种结合。与 Vero E6 细胞结合正比于片段的表达水平,并近似反比于片段的尺寸。这些发现表明 RBD 位于残基 272 和 537 之间。

[178] 为了进一步确定 RBD 的位置,使用抗体(IMG 542)——其为使用含有残基 288-303 的肽产生的。尽管该抗体不与 S537 片段结合,但是该抗体不抑制 S537 片段与 Vero E6 细胞结合(图 8B;一些数据没有示出),这表示 RBD 位于残基 303 和 537 之间。因为相对大的抗体尺寸和空间位阻可能性,可能 RBD 在残基 303 的下游。近来, HCoV-229E 的 RBD 被定位于含有氨基酸残基 407-547 的片段。Ksiazek 等, *A novel coronavirus associated with severe acute respiratory syndrome*, N. Engl. J. Med. 348: 1953-1966 (2003); Rota 等, *Characterization of a novel coronavirus associated with severe acute respiratory syndrome*, Science 300: 1394-1399 (2003)。与之相反,小鼠肝炎病毒的 RBD 位于 N 末端 330 氨基酸。

[179] 仍需要证实,在 SARS-CoV S1 糖蛋白(例如 S272-537)的含有 RBD 的片段和 HCoV-229E 或小鼠肝炎病毒 RBD 之间是否存在结构相似性,并且这种相似性是否与用于复制的相同宿主的使用相关。这两种病毒使用不同的受体。这里描述的直接细胞结合方法可有助于识别其它病毒受体。

[180] 近来,工作人员已经报道:鉴别 ACE2 为 SARS-CoV 的功能性受体。Li 等, *Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus*. Nature 426: 450-54 (2003)。该鉴别 ACE2 为受体允许进一步证实本文提供的结果是正确的。如在图 8C 中所示,当纯化的 ACE2 被用于 ELISA 以检测结合时,观察到相同的结合方式,如细胞结合分析的。对于所用检测的 S 片段,这是真的(图 8C)。

[181] 本文提供的结果不仅提供研究 SARS 病毒进入细胞——证实 ACE2 是 SARS-CoV S1 糖蛋白受体和确定 RBD 的位置——的新工具,而且促进开发新型疫苗免疫原和治疗剂,以预防和治疗 SARS。

实施例 19

S 糖蛋白的 N-末端和 C-末端寡居作用

[182] 该实施例阐明 S 糖蛋白的 N-末端片段——RBD 的上游,可能在融合中起重要作用,并且 S 外功能区形成三聚物,其能通过 6 个螺旋束中间体介导融合。

材料和方法

[183] *抗体和质粒*。兔抗-S 血清用于蛋白质印迹和 FACS 分析,本发明人开发 P450,如上所述。也参见, Xiao 等, Biochem. Biophys. Res. Comm. 312: 1159-65 (2003)。抗-Myc 表位抗体购自 Invitrogen (Carlsbad, CA)。抗-ACE2 山羊多克隆抗体购自 R&D 系统(Minneapolis, MN),并用蛋白质印迹确定。

[184] 位点定向诱变被用来产生共有切割位点,其相应于 pCDNA3 中全长 SARS-CoV S 糖蛋白基因内的 HTV-1 被膜糖蛋白(Env)和一些冠形病毒的切割位点。使用制造商提供的方法,应用 Stratagene(La Jolla, CA)的 QuickChange 试剂盒。对于各种 N 末端 S 片段的表达,通过 PCR 扩增相应的基因片段,并将其克隆到 pSecTag2 表达载体(Invitrogen, Carlsbad)。表达标记有 C9 肽的 ACE2 可溶性外功能区的质粒 pCDNA3-ACE2-ecto 由 Michael Farzan (Harvard University, Boston MA)友好提供。

[185] *蛋白表达和纯化*。S 糖蛋白的多种 N 末端片段被亚克隆入 pSecTag2 表达载体并转染入 293T 细胞,然后用 VTF7.3 进行感染,如在 Xiao 等, Biochem. Biophys. Res. Comm. 312: 1159-65 (2003)中描述的。表达并被分泌入培养基的蛋白通过使用 HiTrap Ni⁺⁺螯合柱(Pharmacia)在天然条件下纯化。通过 PBS 缓冲液透析纯化的蛋白质,并将其储存用于进一步分析。

[186] *通过共免疫沉淀检查 S 糖蛋白二聚化和其与 ACE2 的相互作用*。对于 S 片段二聚化,不同 S 糖蛋白构建物单独或组合来转染 293T 细胞,如在 Xiao 等,

Biochem. Biophys. Res. Comm. 312: 1159-65 (2003)中描述的。含有 S 片段的培养基经过用兔抗-S 多克隆抗血清 P540 免疫沉淀。对于一些共免疫沉淀实验,加入 DTT 以产生还原条件,来消除分子间通过二硫键的相互作用。通过蛋白质印迹使用抗-Myc 表位单克隆抗体检测免疫沉淀的 S 片段。相似地,表达可溶性 ACE2-C9。分泌入培养基的 ACE2-C9 被直接与多种 S 片段在 4℃ 下一起温育 2 小时。然后,将 ACE2 与 1D4 抗-C9 单克隆抗体和蛋白质 G-Sepharose 小珠在 4℃ 下一起温育 1 小时,对 ACE2 免疫沉淀。用 PBS 洗涤小珠四次,将其悬浮于 SDS-PAGE 样品缓冲液中,煮沸 3 分钟,并进行凝胶分离。在样品中 ACE2 或 S 的存在性通过蛋白质印迹检查,如在 Xiao 等, Biochem. Biophys. Res. Comm. 312: 1159-65 (2003)中所描述的。

[187] **流失细胞计量术。**用全长 S 糖蛋白或具有不同 N 末端缺失的 S 糖蛋白转染并用 VTF7.3 感染的细胞与 P540 兔抗-S 多克隆抗体和与 FITC 结合的山羊抗兔抗体一起在含有 1%BSA 的 PBS 中在 4℃ 下温育 2 小时。然后,在冰冷 PBS 中洗涤 4 次细胞,并且用 FacsCalibur (Becton Dickinson, San Jose, California)分析。

[188] **S 片段的凝胶过滤分析。**在被纯化到 Ni-螯合柱并且与 PBS 缓冲液交换后,将 S 片段样品加样到 Superose 12 10/300 GL 柱(Pharmacia, Uppsala, Sweden)上,该柱已经用 PBS 预平衡。用 PBS 以 0.5 ml/min 洗脱蛋白质,并且收集 0.5 ml 分馏物。Superose 12 柱用 669、440、232、158、67、44 和 25 kD 的蛋白质分子质量标准校准。10 μ l 等分物从每一分馏物中获得用于蛋白质印迹分析。

[189] **交联。**纯化的 S537 片段在 PBS 中被稀释到 0.2 μ g/ml 的浓度。将 BS³ (Pierce, Rockford, IL)加入到 S537 溶液中,至 1 mg/ml 的终浓度,并在冰上温育 1 分钟。然后,将样品与等体积 4X SDS-PAGE 负载缓冲液混合,并通过蛋白质印迹分析。

[190] **细胞融合 β -gal 报告基因分析。**用 pSecTag2B-S 或 pCDNA3-ACE2 转染和分别 VTF7.3 和 VCB21R 感染的细胞通过胰蛋白酶消化收集,并用 PBS 洗涤一次。然后,将细胞悬浮于 pH 7.4 的常规 DMEM 培养基中,并混合。在 4 小时温育后,裂解细胞,并且使用 CPRG 作为底物(Roche)测量 β -gal 活性,如在 Xiao 等, Biochem. Biophys. Res. Comm. 312: 1159-65 (2003)中描述的。

[191] **ELISA。**使用两种 ELISA 分析。在夹心 ELISA 中,用抗组氨酸标记抗体包被平板,然后加入 S 片段,并且用抗-c-Myc 表位抗体检测。使用该分析检测

S 片段。在第二种 ELISA 分析中, C9 标记受体 ACE2 通过抗-C9 抗体(ID4)被包被在平板上, 加入 S 片段, 并且在洗涤后, 用抗-c-Myc 表位抗体检测。在所有的实验中, 和 c-Myc 表位抗体的温育在室温下进行 2 小时。测量光密度(OD), 并根据最大值标准化。

结果

[192] S 糖蛋白的 RBD 上游的 N-末端片段形成二聚体。对于另一种冠形病毒(MHV)先前已经表明可溶性 S1(与 SU 相似)片段形成二聚体, 含有受体结合结构域的 N-末端 330 个氨基酸残基区参加二聚化作用, 并且仅二聚体结合于受体 CEACAM。参见 Lewicki & Gallagher, J. Biol. Chem. 277:19727-34 (2002)。然而, 本发明人和其他人已经确定 SARS-CoV 受体结合结构域的位置为从 N-末端的下游。Xiao et al. Biochem. Biophys. Res. Comm. 312: 1159-65 (2003); Wong et al. J. Biol. Chem. 279: 3197-3201 (2004); Babcock et al. J. Virol. 78: 4552-4560 (2004)。

[193] 为了致力于含有受体结合结构域寡聚作用的可能性和评估其在介导膜融合中的功能, 检测几种 S 片段的寡聚作用。这些 S 片段包括不与受体 ACE2 结合的 N 末端片段(残基 17 到 276, 被称为 S276, SEQ ID NO:50)、与 ACE2 结合的几种 S 片段(S756、S537、S272-537)以及保持受体结合活性的含有残基 319 到 517 的片段(被称为 S319-517, SEQ ID NO:62)。选择这些片段部分是因为其独立折叠并且被分泌到细胞培养物上清液中, 尽管它们的表达效力明显变化(图 9 A, 左边)并且当与 S756 共表达时, 它们的浓度降低(图 9 A, 右边)。

[194] 为了找出这些片段是否与最大的片段(S756)——其包括被膜糖蛋白的受体结合亚单位的同等物(一般为 SU, 对于冠形病毒为 S1)——寡聚合, 共表达多肽片段, 然后用抗体 P540 免疫沉淀细胞培养物上清液中的混合物。如在前面的实施例所述, 针对含有 S 糖蛋白残基 540-555(SEQ ID NO:59)的肽, 开发兔多克隆抗体制备物。该 P450 抗体与 S756 多肽结合, 而不与其它片段结合(图 9B, 左边)。除了最小的片段(S319-517)外, 所有含有受体结合结构域的 N-末端片段通过 P540 与 S756 共免疫沉淀(图 9B, 右边)。为了消除可导致共免疫沉淀的非特异性二硫键的可能性, 在一个共免疫沉淀实验中包含 DTT。DTT 对分泌的 S756 (左边泳道)或 S756+S276(右边泳道)的免疫沉淀或共免疫沉淀没有影响(图 9C,

左边面板)。

[195] 为了找出寡聚体的大小，一种片段(S537)与 BS³ 交联。图 9C 右边的面板表明出现了新的带，起分子量相应于二聚体，而不是更高阶的寡聚物。为了排除由于交联造成的假象的可能性并进一步证实行程二聚体，也通过凝胶过滤分析 S537 片段。观察到两个凝胶过滤洗脱峰：一个由于分子量大约 230 kDa 的种类，另一个大约 110 kDa(图 10A，上边面板)，其分别对应于二聚体大小寡聚物合单体。与之相反，含有受体结合结构域的最小的片段(S319-517)仅作为大约 35 kDa 分子量的单体被洗脱出(图 2 A，下面的面板)。总之，这些结果表明可溶性 SU 是二聚体，并且二聚化结构域在从残基 317 和受体结合结构域的 N-末端的区域的上游。

[196] 对于 S 介导细胞-细胞融合，需要二聚的 N 末端区域。因为推定二聚化结构域在 S1 内的受体结合结构域的上游，并且融合结构在 S2 中，所以可以假设对于 S 介导的融合，可能不需要二聚化。为了检测该假设，产生全长 S 糖蛋白的两种缺失突变体。一个片段缺失 N-末端 103 残基，另一个片段缺失 N-末端 311 残基(图 9A)，从而消除了假定的二聚化结构域。与野生型全长 S 糖蛋白表现活性相比，两种突变体没有表现融合活性(图 9A)。为了检测特定表达水平是否可解释缺乏可观察到的融合基因活性，通过流式细胞计量术和蛋白质印迹测量表面和综合表达水平。从这两个分析得到的数据表明两种确实突变体的表达水平与野生型的表达水平很难区分(图 11B 和 C)。这些结果表明对于融合需要 N-末端，这可能涉及或可能不涉及二聚化的机制。

[197] 含有受体结合结构域的二聚化的 S1 比单体片段与 ACE2 结合更有效的多。另一种冠形病毒(MHV)的先前的研究表明仅二聚化的 S1 与受体 CEACAM 结合。Lewicki & Gallagher, J. Biol. Chem. 277:19727-734 (2002)。对 SARS-CoV 片段进行实验以了解二聚化状态的 S1 可如何影响融合。具体而言，通过使用用以将单价 S1 片段转变为二价 S1 片段的抗-c-Myc 表位抗体，观察单价和二价形式的 S1 片段与 ACE2 的结合。这些 S1 片段的一种(S319-517, SEQ ID NO:62)与表面固定的 ACE2 结合没有达到任何可测量的程度，除非通过抗-c-Myc 表位抗体，其在与受体一起温育之前或期间，在溶液中将片段转变为二价分子(图 12)。与之相反，S537 再没有抗体情况下与 ACE2 结合，尽管抗体存在增加了其结合(图 12)。这些结果表明 S1 的二聚化状态促进可增强融合效力的总的亲和力的增加。

[198] 可溶性S外功能区是三聚体。病毒I类融合蛋白被膜糖蛋白例如流行性感
冒的红细胞凝集素(HA)是通过跨末结构域三聚化的。因为 SARS-CoV S糖蛋白
近来被发现是I类融合蛋白,所以S2亚单位可能促进整个S糖蛋白的三聚化。
然而,二聚化S1与三聚化S2可导致更高阶寡聚物,其形成取决于天然S糖蛋
白中二聚化结合位点的可用性。为了检测这种可能性,可溶性S外功能区(Se)
的大小通过凝胶过滤估计,其中跨膜结构域和胞质尾区被缺失。如在图13中所
示,具有大约三聚体大小(MW为512 kDa)的复合体被检测到。没有检测到更高
阶得寡聚物。这些结果不但表明Se片段及或许全长的膜相关的S在天然未结合
的状态是三聚化的,而且表明S1中的二聚化位点对于三聚体间相互作用不是简
单可用得。

[199] 这些结果表明下列:1) SARS-CoV S糖蛋白的SU亚单元(S1)形成二聚体,
2)二聚化结构域不重叠并且在受体结合结构域的上游,3)二聚化结构域的缺失废
止融合,4)含有受体结合结构域的二聚化的S1比单体片段与受体分子结合更有
效的多,和5)可溶性S外功能区在凝胶过滤条件下形成三聚体。

[200] 先前已经报了一些I类融合蛋白(其结合受体分子)可形成二聚体,包括例
如逆转录病毒HIV-I的gpl20和冠形病毒MHV的S1。Center等, J. Virol. 74:
4448-55 (2000); Lewicki等, J. Biol. Chem. 277: 19727-34 (2002)。直到本工作,
S1二聚化对膜融合的调节的作用仍然不清楚。现在一般接受可溶性外功能区例
如HIV-I的gpl40蛋白和SIV被膜糖蛋白(Env)形成三聚体,尽管可观察到二聚
体和四聚体。Center等, Proc. Nat'l Acad. Sci. U.S.A. 98: 14877-82 (2001)。相似
地,表明包括S2的SARS-CoV的冠形病毒至少一种可能的融合中间体的四级结
构是三聚化的。Liu等, Lancet 363: 938-947 (2004); Bosch等, Proc. Nat'l Acad. Sci.
U.S.A. 101: 8455-60 (2004)。与之相反,一些数据表明MHV S2蛋白在从S1解
离后是单体的。Lewicki等, J. Biol. Chem. 277: 19727-34 (2002)。二聚体到三聚
体的转变在II类融合蛋白介导的融合的机制中起关键作用。因此,已经提出一
些冠形病毒的四级结构中的变化在融合机制中起重要作用。同样地(Id.),应该注
意HIV-I Env和MHV S糖蛋白被切割,并且SU可从跨膜亚单位解离,然而,
这种解离对于融合可能不是重要的。与之相反,SARS-CoV S当以膜关联或可溶
性形式表达时没有被切割,切割对于融合可能是不需要的。因此,尽管SARS-CoV
S糖蛋白是I类融合蛋白,但是缺乏切割是如下规则的例外:I类融合蛋白的Envs

推测被切割以赋予高能的亚稳态，其可驱使融合反应。

[201] SARS-CoV S 糖蛋白的 SU(S1)结构域与跨膜结构域(S2)的外功能区可形成二聚体也可形成三聚体的发现拥有令人感兴趣的拓扑情形(topological situation)。因此，如果三聚体中的两个单体也形成二聚体，那么三个单体将保持自由与另一个三聚体的“游离”单体相互作用，并且形成两个三聚体的二聚体。在另一种情况中，三聚体中每一个单体的方位不允许在三聚体中形成二聚体，但是留下“游离”结合位点以与另一个三聚体的单体二聚化。在这种情况下，可能期望形成三聚体的网络。最后，三聚体的三维结构不可能允许单体二聚化位点与同一或不同三聚体中的其它单体相互作用。本文提供的初步数据支持了后一种可能性，其中，使用描述得凝胶过滤条件没有检测到更高阶的寡聚物。在那些条件下，或者三聚体间二聚化发生，但是三个单体构象不允许单体与其它三聚体相互作用或这种相互作用太弱而不能检测到，或者三聚体三维结构识这样：其不允许二聚化相互作用。

[202] 本文提供的数据表明在二聚化结构域部分缺失后，缺乏融合，并且标明二聚化区域在融合中器重要作用，尽管其机制可能不是通过二聚化相互作用。此外，在天然条件下，其中 S 糖蛋白的表面浓度可非常高，如电子显微照片所示，可能二聚化相互作用在稳定相互作用分子的“网络”——其可能与介导 II 类融合蛋白进入的网络有些相似——中起作用。这类网络如果有可增加与受体分子相互作用的亲和力，并且可能通过提供预先装配的 Env 分子网络促进融合孔结构的形成，或者甚至提供能量以在不存在 S 切割——其产生高能亚稳态——的情况下驱使融合反应。

实施例 20

编码 RBD 多肽的 DNA 免疫的小鼠的血清抑制 S-介导的细胞融合

[203] 该实施例阐明用编码受体结合结构域多肽的 DNA 免疫动物可预防 SARS 感染。

材料和方法

[204] 将小鼠分为三组：组 A 小鼠 1 到 5 用编码含有刺突蛋白受体结合结构域 (RBD) 的 S319-518 片段的质粒 pSecTag-SRBD 免疫；组 B 小鼠 1 到 5 用编码融合到 Fc 的 RBD (S 319-518) 片段的融合蛋白的质粒 pEAK-10-RBD-Fc 免疫以及组 C 小鼠 1 到 3 用对照质粒免疫。每组 5 只 BALB/C 小鼠在第 0 天、第 14 天和第 28 天免疫。使用基因枪，小鼠每次免疫接受 2 μ g 以下的 DNA。在第 56 天收集血清。在图 14A-B 中，第一个数字指单独小鼠，字母指各自的免疫组，最后的数字指使用的稀释。

[205] 将细胞(293T)与从免疫的小鼠得到的抗血清一起温育，然后与表达 S 蛋白的细胞混合。如在前面实施例所述的测量融合(也参见，Xiao 等，BBRC 2003)。PC 指阳性对照，其中没有加入血清。对于每组中小鼠 1 到 2，使用 10、100 和 1000 的血清稀释因子。对于组 A 和组 B 中的小鼠 3 到 5 以及对照组中的小鼠 3，使用 20 和 100 的稀释因子。

结果

[206] 从小鼠获得的抗血清的抗体效价在图 14A 中示出。如所示，用编码刺突蛋白受体结合结构域(S319-518，组 A 和 B)的 DNA 免疫的小鼠具有非常高的抗血清效价——稀释高达 1:7250 在 ELISA 分析中仍旧强烈与抗原反应。

[207] 如在图 14B 中所示，从用编码刺突蛋白受体结合结构域的 DNA 免疫的小鼠得到的抗-血清以剂量依赖的方式抑制表达 S 蛋白的细胞的融合。因此，从小鼠 1A 和 2A 得到的抗-血清——该小鼠用编码 S 受体结合结构域的 DNA 免疫——当以 1: 10 稀释使用时，基本上消除 S 蛋白介导的细胞融合。该抗-血清更高的稀释(1:100 和 1:1000)功效更低。在从小鼠 3A(1:20 稀释)、从小鼠 4A(1:20 稀释)和从小鼠 5A(1:20 稀释)得到的抗-血清抑制细胞融合上，观察到相似的结果。

[208] 这些数据表明用编码 S 蛋白受体结合结构域多肽的 DNA 免疫的动物可增加对刺突蛋白的强免疫应答，并且可预防 SARS 感染。如上所述，具有受体结合结构域的 S 糖蛋白的可溶片段抑制 S-介导的细胞融合(参见图 15)。

实施例 21

与中和抗体复合的 SARS CoV 受体结合结构域的结构

[209] 该实施例阐明允许 SARS CoV 受体结合结构域(RBD)与中和抗体结合的结构特征。

材料和方法

[210] *RBD 的表达和纯化*。使用 BamHI 和 EcoRI 限制性位点, 将含有 S 糖蛋白的残基 317-518 的片段克隆入 pSecTag2B (Invitrogen), 如上所述。也参见 Xiao 等, Biochem. Biophys. Res. Commun. 312, 1159-1164 (2003); Chakraborti 等, Virol. J 2, 73 (2005)。使用正向引物 5' ACT GTC TAG ATG GTA CCG AGC TCG GAT CC 3' (XbaI, SEQ ID NO:67)和反向引物 5' CAG TAG ATC TCG AGG CTG ATC AGC G 3' (BglIII, SEQ ID NO:68)进一步讲该插入物克隆入 pAcGP67-A。pAcGP67-S 与 BaculoGold 线性杆状病毒共转染入 SF9 细胞。通过多重扩增, 制备高效价的重组杆状病毒原种。在 SF9 细胞中表达蛋白, 并且将其培养在不含血清的 HyQ-SFX-插入物培养基(Hyclone)中, 用 HiTrap Ni 螯合柱从条件培养基中纯化蛋白。洗脱的单体蛋白被浓缩, 进一步用 Superdex 75 10/300GL 柱纯化——所述柱用 PBS+0.2 M NaCl 平衡, 并且在 PBS+0.2 M NaCl 中浓缩至 5-10 mg/ml 的浓度。

[211] *高亲和性 RBD-特异性 Fab m396 和的选择、表达和纯化以及其转变为 IgG1*。从 10 个健康捐献者的外周血 B 细胞构建天然人类 Fab 噬菌体展示文库(总共大约 10^{10} 个成员), 并且使用它针对纯化的可溶性单体 RBD 选择 Fabs, 结合于磁珠(Dynabeads M-270 Epoxy, DYNAL Inc., New Hyde Park, New York)。将扩增的 10^{12} 噬菌体展示的 Fabs 文库与 5、3 和 1 μ g 的 RBD 分别在第 1 轮、第二轮和第三轮生物淘选(biopanning)期间在 500 μ l 体积中在室温下一起温育 2 小时。在第三轮生物淘选后, 从感染的 TG1 细胞中随机选取 95 个克隆, 并且使用噬菌体 ELISA 识别具有高结合亲和性的噬菌体展示 Fabs 的克隆。结合于 $A_{450} > 1.0$ 的 RBD 的 8 个克隆被选择进一步鉴定。这些克隆的 VH 和 VL 被测序。它们是相同的, 并且选择的 Fab 被称为 m396。

[212] m396 抗体重链 CDR3 的序列是 DTVMGGMDV (SEQ ID NO:70), m396

轻链 CDR3 为 QVWDSSSDYV (SEQ ID NO:71)。

[213] 用于结晶的 Fab 首先用 HiTrap Ni 螯合柱纯化, 然后使用 PBS+0.2 M NaCl 用 Superdex 75 10/300GL 柱进一步纯化, 并浓缩至 10-20 mg/ml。将 Fab 重链和轻链扩增, 并重新克隆到 pDR12 载体(D. Burton, the Scripps Research Institute, La Jolla, CA 提供), 具有 cDNA 序列的 Fc 基因片段取代基因组 DNA。

[214] 结晶和结构确定。通过混合 1:1 摩尔比的各种成分并在 4°C 下温育整晚, 形成 SCV RBD-Fab m396 复合体。仅对于 1:2 比例的复合体和储存溶液(reservoir solution)通过蒸汽渗透技术(vapor diffusion technique)使用 pH6.5 的 15 v/v 甘油、20% PEG 6000、100 mM MES 钠, 在 2-3 周内获得晶体。在 Advanced Photon Source (APS)的 Southeast Regional Collaborative Access Team(SER-CAT)光束线设备(beamline facility) 22-ID, Argonne National Laboratory, 收集高达 2 埃分辨率的数据组。使用 HKL2000 程序组实行数据处理(Otwinowski,Z. & Minor ,W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276, 307-326 (1997))。

SCV RBD/Fab m396

晶体数据和处理统计在表 2 中总结

数据收集、处理和加细(refinement)统计

数据收集	
空间群	P2 ₁
细胞量纲	
a、b、c(埃)	64.7、68.6、80.1
α 、 β 、 γ (°)	90.0、98.1、90.0
分辨率范围(埃)	30.0-2.30
R _{sym} 或 R _{合并} (%)	0.060(0.263)
I/ σ I	16.1(2.2)
完整性	91.0(72.8)
丰余性	3.0(2.0)
加细	
分辨率(埃)	30.0-2.3
反射次数	27719
R _{工作} /R _{游离}	19.8/26.1
R.m.s 偏差	
键长(埃)	0.000
键角(°)	0.0

[215] 通过使用从受体复合物(PDB 密码 2AJF)得到的 SCV RBD 和四个独立结构域 V_H、V_L、C_H和 C_L分别作为研究模型(在 Fab m396 的 C_H和 C_L结构域之间观察到 7.9° 的旋转角)的 PHASER 分子取代解决该结构问题(参见 Storoni 等, *Acta Crystallogr. D. Biol. Crystallogr.* 60, 432-438 (2004))。SCV RBD 的 RBM(大约 60 残基, 430-490)和 Fab 模型的大多数 CDRs 没有包括在研究模型中, 并且从电子密度建立。使用 CNS 加细(refine)该复合物(参见 Brunger 等, *Structure.* 5, 325-336 (1997)), 并且最终模型被精炼至 2.3 埃分辨率。在 Asn330, 总共 299 个水分子、磷酸盐例子和一个 N-连接葡萄糖胺部分在加细的最后阶段被加入。最终的 R 和 R_{游离}分别为 19.8 和 26.1(表 2)。

结果

[216] 与本文描述的潜在中和抗体 Fab m396 复合的 SCV RBD 结构识别主要的中和决定子，与受体识别的关系，消除中和结构机制以及提供 SCV 进入机制的理解。

[217] 用于产生抗-SARS CoV 抗体的主要表位在 S 蛋白的位置 482-492 包括肽基序列 GFYTTTGIGYQ (SEQ ID NO:69)。

[218] 与 Fab m306 复合的 RBD 结构和与 ACE2 复合的 RBD 的结构相似(Li 等, Science 309, 1864-1868 (2005)), 尽管更高的分辨率允许识别一些先前混乱或没有定位的残基。RBD 由核——其包括 5 个(β 1- β 4 和 β 7)链的反平行层, 和具有两个链的反平行(β 5- β 6)层的长的延伸的环——在其中部连接到核——构成。抗体复合的 RBD 结构含有 8 个半胱氨酸, 其在核形成三个二硫键, 在延伸环形成一个二硫键。

[219] 在与抗体复合的 RBD 结构的 C- α 原子位置和与受体复合的结构之间的方均根差的值为 1.3 埃。因为高的分辨率, 抗体-结合的 RBD 结构被相对好地定义, 其包括先前未分辨的残基 376 到 381 和 503 到 511, 其明白地位于 378 和 511 之间的另外的二硫桥。因此, 肽基序列 NDLCFSNV (SEQ ID NO:72, S 蛋白位置 375-382)和 FELLNAPATVCG (SEQ ID NO:73, S 蛋白位置 501-512)可涉及确立 S 蛋白构象 (protein conformation), 其促进在 S 蛋白 RBD 和其中和抗体之间的稳定复合体的形成和保持。

[220] 形状关联统计参数(S_c , Lawrence & Colman, J. Mol. Biol. 234: 946-50 (1993))——相邻表面之间几何拟合的量度, 其最大值为 1——, 对于 RBD-抗体界面测量为 0.66, 其表示高度的形状互补性。在复合体界面, 1760\AA^2 总表面积被隐藏接近等于 RBD 的 870\AA^2 和抗体的 890\AA^2 的贡献, 如 1.4\AA 探针所确定。抗体-结合 β 6- β 7 环单独占 RBD-抗体界面的 63%, 这表明环残基的主要作用涉及结合。重链 CDRs 促成抗体结合位点的总面积的 66%。结合界面的大小接近于其它抗原-抗体复合体的平均(Davies & Cohen, Proc. Natl. Acad. Sci U. S. A. 93, 7-12 (1996))。

[221] Fab m396 抗体主要识别沿 β 6- β 7 环位置 482 到 491 的 10 个残基, 该环从

RBD 表面显著突出。该环接触 Fab m396 的 CDRs(互补决定区)的四个: H1、H2、H3 和 L3。该 4 个 CDRs 在抗体的可变区的表面上形成浅的缺口, 所述可变区提供进入 $\beta 6$ - $\beta 7$ 环紧密装配的深结合袋(binding pocket)。 $\beta 6$ - $\beta 7$ 环的大多数残基在结合袋与 Fab m396 相互作用。具体而言, 残基 Ile489 和 Tyr491 插入抗体结合位点的表面上的深袋中。RBD 的 15 到 17 个残基和 Fab m396 参与相互作用, 并形成两个分子之间 3.5 Å 接触距离限制内限定的 RBD-抗体界面。这些残基在表 3 中确定, 并且包括下面的 S 蛋白 RBD 残基: Thr-363, Lys-365, Lys-390, Gly-391, Asp-392, Arg-395, Tyr-436, Arg-426, Gly-482, Tyr-484, Thr-485, Thr-486, Thr-487, Gly-488, Ile489, Tyr491, Gln-492 和 Tyr-494。

[222] 跨约结合表面存在的分子间相互作用具有从范德瓦尔斯接触和直接以及水介导的氢键的贡献(表 3 和 4)。

表 3: 在 SCV RBD/Fab m396 界面的残基之间的接触

范德瓦尔斯接触 ^a		氢键 ^b		距离(Å)
抗体	SCV RBD	抗体	SCV RBD	
H1		H1		
Ser ^{H31}	Tyr ^{S484} (3), Thr ^{S486} (7), Thr ^{S487} (3)	Ser ^{H31} (O)	Thr ^{S486} (Oγ1)	2.61
Tyr ^{H32} Thr ^{H33}	Thr ^{S486} (2) Thr ^{S487} (3), Gly ^{S488} (4), Tyr ^{S491} (1)	Thr ^{H33} (Oγ1)	Gly ^{S488} (N)	2.67
H2		H2		3.19
Gly ^{H50} Ile ^{H51} Thr ^{H52}	Tyr ^{S491} (2) Tyr ^{S491} (3) Tyr ^{S487} (1), Tyr ^{S491} (4)	Thr ^{H52} (Oγ1)	Tyr ^{S491} (Oγ)	3.26
Ile ^{H53} Leu ^{H54}	Tyr ^{S484} (4) Tyr ^{S436} (1), Gly ^{S482} (3), Tyr ^{S484} (1)	Thr ^{H52} (Oγ1)	Tyr ^{S491} (Cδ1)	3.16
Ile ^{H56}	Lys ^{S390} (1), Tyr ^{S491} (4)	Thr ^{H52} (Oγ1)	Tyr ^{S491} (Cε1)	3.33
Asn ^{H58}	Asp ^{S392} (6), Tyr ^{S491} (7)	Thr ^{H52} (Oγ1)	Tyr ^{S491} (Cε2)	3.24
H3		Thr ^{H52} (Oγ1)	Tyr ^{S491} (Cζ)	3.32
Val ^{H97}	Arg ^{S426} (6), Thr ^{S485} (1), Gln ^{S492} (4)	Ile ^{H56} (Cδ1)	Lys ^{S390} (Nζ)	3.01
L3		Asn ^{H58} (Nδ2)	Tyr ^{S491} (OH)	2.81
Trp ^{L91}	Gly ^{S391} (1), Asp ^{S392} (1), Ile ^{S489} (7), Tyr ^{S494} (1)	Asn ^{H58} (Nδ2)	Tyr ^{S491} (Cζ)	3.11
Ser ^{L93}	Thr ^{S363} (2), Arg ^{S395} (4)	H3		2.76
		Val ^{H97} (O)	Gln ^{S492} (Nε2)	
		L3		3.06
			Lys ^{S365} (Nζ)	
Ser ^{L94}	Lys ^{S365} (1), Arg ^{S395} (1)	Ser ^{L94} (N)		
		Asp ^{L95} (Oδ1)	Arg ^{S395} (NH2)	3.00

^a 范德瓦尔斯接触具有原子间距离 ≤ 4.0 Å

^b 氢键标准基于供体-受体距离($\leq 3.5 \text{ \AA}$)

表 4: 水介导的氢键^b

抗体	SCV RBD	10 距离(\AA)
H2		
Asn ^{H58} (Oδ1)	Lys ^{S390} (Nε)	2.55, 2.765
Asn ^{H58} (Nδ2)	Asp ^{S392} (Oδ2)	2.53, 2.55
L1		
Ser ^{L30} (Ca)	Thr ^{S359} (O)	3.14, 2.620
L3		
Trp ^{L91} (Ne1)	Asp ^{S392} (Oδ1)	2.88, 2.67
Ser ^{L93} (O)	Gly ^{S391} (O)	2.68, 2.585
Tyr ^{L96} (OH)	Gly ^{S490} (N)	2.83, 2.73

^b 氢键标准基于供体-受体距离($\leq 3.5 \text{ \AA}$)

[223] 复合体中抗体和 SCV RBD 之间界面上隐藏表面积的具体情况在表 5 中给出。

表 5: 在 SCV RBD/Fab m396 界面的残基被隐藏的表面积(\AA^2)

抗体	隐藏面积(\AA^2)	SCV RBD	隐藏面积(\AA^2)
H1		β2-链	
Ser ^{H31}	58.14	Thr ^{S363}	13.78
Tyr ^{H32}	19.42	Lys ^{S365}	36.94
Thr ^{H33}	36.39		10
H2		3₁₀-螺旋	
Gly ^{H50}	2.0	Lys ^{S390}	31.33
Ile ^{H51}	2.31	Gly ^{S391}	15.13
Thr ^{H52}	20.72	Asp ^{S392}	38.29
Ile ^{H53}	34.36	Arg ^{S395}	86.08
Leu ^{H54}	88.19		15
Ile ^{H56}	40.69	其它残基	
Asn ^{H58}	52.03	Arg ^{S426}	33.18
H3		Thr ^{S436}	13.86
Val ^{H97}	108.12		20
L3		β6-β7	
Trp ^{L91}	99.32	环	
Ser ^{L93}	93.11	Gly ^{S482}	20.29
Ser ^{L94}	45.4	Tyr ^{S484}	55.3
Asp ^{L95}	28.26	Thr ^{S485}	4.01
		Thr ^{S486}	55.41
		Thr ^{S487}	52.85
		Gly ^{S488}	46.39
		Ile ^{S489}	96.93
		Tyr ^{S491}	122.2
		Gln ^{S492}	15.42
		Tyr ^{S494}	12.74

[224] 在 SCV RBD-抗体复合体中主要相互作用大部分在 RBD 的 β 6- β 7 环和抗体 Fab m396 的 4 个 CDRs: H1、H2、H3 和 L3 之间。在电子密度图(没有示出)上, 清楚阐述这些相互作用。因此, H1 与疏水残基 Tyr484、Thr486 和 Thr487 接触。具体而言, H1 的 Thr33 直接与 RBD 主链的 Gly488 残基的酰胺氮原子接触。该类型的相互作用也在 RBD-ACE2 复合体中发现, 其中 S 蛋白 RBD Gly488 的酰胺参与主链和 ACE2 的 Lys35 的羧基的氢键。与其它 CDRs 相比, H2 的主要作用是与 RBD 结合并且使用大量残基进行接触; 最显著的特征是在 H2 导致的浅的缺口中 RBD(122 \AA^2)的 Tyr491 的隐藏, 其中 H2 中的 Asn58 的氨基接触

RBD Tyr491 的酚氧原子。另一个主要相互作用是在 H2 Thr52 和 Tyr491 的苯环之间。Val97 是涉及 RBD 相互作用的 H3 区域的唯一残基；然而，其隐藏所用 CDR 残基的最大表面积(108 \AA^2)。因此，Val97 的羧基氧原子产生与距离 2.7 \AA 内的 RBD Gln492 的氨基单原子的强的氢键。涉及定位氢键的主链和侧链残基之间的接触，如在 H1、H2 和 H35 中，在确定复合体中 RBD 和抗体相对位置方面起重要作用，并且有助于相互作用的特异性。

[225] 计算重链 RBD 相互作用的 S_c 参数具有 0.74 的高值，其表明对于重链 RBD 识别的高度相关的界面几何形状。L3-RBD 相互作用包括水介导和其它次要的结合位点，包括 RBD 的 Arg395(表 3 和 4)。L3 的残基 Trp91 与芳香族残基 Ile489 堆积，这是 RBD 中主要热点；在界面，Trp91 和 Ile489 残基的每一个隐藏大约 100 \AA^2 的表面积。

[226] RBD 上次要结合位点包括在 β_2 (Thr363 和 Lys365)、 3_{10} 螺旋接着 β_3 (Lys390、Gly391、Asp392 和 Arg395)中的残基和 Arg426 与 Tyr436 的两个残基。这些残基除了对抗体结合的次要贡献，它们大部分在稳定 β_6 - β_7 环的形成中起重要作用。特别是氢键——包括在 Gly391 氮原子和 Gly490 与 Gln492 的羧基氧原子之间的氢键、Arg426 的氨基和 Thr485 的主链羧基之间的氢键以及 Tyr436 和 Tyr484 的分氧原子之间的氢键——稳定 RBD 中 β_6 - β_7 环的形成。

[227] RBD-抗体结合界面具有两个主要的特有特征：第一，界面之间的高水平的互补性，第二，主要热点 RBD 残基 Tyr491 锚定于抗体结合位点。通过大量的疏水残基和包括亲水和极性残基的网络，产生 RBD-抗体相互作用。RBD 的 β_6 - β_7 环的两个热点 Ile489 和 Tyr491 形成良好的突出脊(protruding ridge)，而抗体结合袋(antibody binding pocket)包括通过重链形成的浅的缺口中的空穴。因此，互补位和表位结构是高度互补的，这可是它们相互作用高亲和力的主要因素。RBD-抗体相互作用的另一个特有特征是在抗体结合位点，RBD 的 Tyr491 插入结合袋的底部，其中 H2 的 Thr52 和 Asn58 以特定的方式与 Tyr491 相互作用。Tyr491 残基以如此方式——Tyr491 的酚羟基与 Asn52 侧链的氮原子形成氢键——被强力保持在两个残基之间，同时 Tyr491 的苯环作为 Thr52 侧链氧原子的极好的氢键受体。被 Tyr491 优选识别的这两个 H2 残基的结构基础——其在抗体中与结合位点袋排成一行——包括侧链特异性，其是 RBD-抗体 Fab m396 识别的独特结构特征。

[228] RBD-ACE2 和 RBD-m396 结构的比较为理解抗体介导中和的分子基础和 SCV 进入的机制提供了重要线索。抗体和受体占由 RBD 上 $\beta 6$ - $\beta 7$ 环(Thr484、Thr486、Thr487、G488 和 Y491)和 Arg426 构成的共同区域。发现这些共同残基对 RBD 与抗体和 ACE2 的结合是关键。抗体和受体结合之间的主要差异与限定受体结合决定子而不是共同结合区域($\beta 6$ - $\beta 7$ 环)相关。中和决定子连续位于 $\beta 6$ - $\beta 7$ 环的主要片段中, 而受体 ACE2 具有在出现在 RBD 顶部的延伸环大部分之上的决定子。位于 $\beta 6$ - $\beta 7$ 环中心的重链与 ACE2 的高水平重叠示出中性和受体识别的共同决定子。这些观察表明通过对 S 蛋白 RBD 的 $\beta 6$ - $\beta 7$ 环中相同残基的竞争和空间位阻——其阻断 RBD 上的受体结合位点, 抗体中和 SARX-CoV。

[229] 近来, 报道大鼠作为 SARS 样冠形病毒的储存宿主。Li, W. 等, Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310, 676-679 (2005); Lau, S.K. 等, Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U. S. A* 102, 14040-14045 (2005)。人和麝猫(civet)分离物的序列大鼠分离物的序列明显不同, 尽管它们种族遗传相关的。值得注意地, SCV RBD 的残基 Arg426、Ile489 和 Tyr491——其涉及抗体结合到 RBD——在已知的大鼠分离物中是保守的, 这表明 m396 潜在的中和活性, 尽管这种可能性应该被实验评估。

[230] 与抗体的复合体中的 RBD 结构意外地与 ACE2-复合的 RBD 结构相似。尽管未结合(unliganded)的 RBD 的结构目前不知晓, 但是其可能与结合到抗体 m396 或 ACE2 的 RBD 结构十分相似。然而, ACE2 和 m396 的结合不可能在 RBD 中诱导极度相似的构象变化, 以产生基本相同的 RBD 结构, 特别是根据它们交叠而不是不同结合位点和它们结合的其它分子特异性。在结合结构中意外的相似性挑战了目前的范例(paradigm)——SARS CoV 进入通过 ACE2-活化机制, 尽管可能膜-结合的 ACE2 可通过多价结合诱导三聚体 S 糖蛋白构象变化。另一种可能性是通过特异性结合到 S 糖蛋白, 然后结合到可诱导构象变化的复合受体(一个或多个, coreceptor(s)), ACE2 行使功能, 所述构象变化活化 S 糖蛋白的融合结构。

[231] 这些结果暗示开发对抗 SARS CoV 的疫苗和治疗剂。而且, 这些结果增强对抗体介导病毒中和以及病毒进入机制的理解。新识别的抗体 m396 本身具有治疗潜力; 其目前正被评估对感染病毒的效力, 并将被在动物模型中检测。基

于其结构(未结合的 m396 的结构也被确定, 其与结合的没有明显不同(数据没有示出)), 可以设计其它治疗方式。

[232] 抗体表位的结构可被用于设计可诱导 m396 或 m396-类似抗体的疫苗免疫原(反向疫苗学(retrovaccinology)方法, Burton, D.R. Antibodies, viruses and vaccines. Nat. Rev. Immunol. 2, 706-13 (2002))。特别吸引人的是主要中和决定子 β 6- β 7 环和基于其序列的约束肽(constraint peptides)的潜在应用为疫苗免疫原。其突出性质、暴露和容易被抗体接近表明其在中和机制中的关键作用, 可能除了受体 ACE2 和 m396, 其也与其它抗体结合。

参考文献

- D.S. Dimitrov, *Cell biology of virus entry*, *Cell* 101 (2000) 697–702.
- D.S. Dimitrov, *Virus entry: molecular mechanisms and biomedical applications*, *Nat. Rev. Microbiol.* 2 (2004) 109-122.
- K.V. Holmes, SARS-associated coronavirus, *N. Engl. J. Med.* 348 (2003) 1948–1951.
- M.M. Lai, D. Cavanagh, *The molecular biology of coronaviruses*, *Adv. Virus Res.* 48 (1997) 1–100.
- T.G. Ksiazek, D. Erdman, C.S. Goldsmith, S.R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J.A. Comer, W. Lim, P.E. Rollin, S.F. Dowell, A.E. Ling, C.D. Humphrey, W.J. Shieh, J. Guarner, C.D. Paddock, P. Rota, B. Fields, J. DeRisi, J.Y. Yang, N. Cox, J.M. Hughes, J.W. LeDuc, W.J. Bellini, L.J. Anderson, A novel coronavirus associated with severe acute respiratory syndrome, *N. Engl. J. Med.* 348 (2003) 1953–1966.
- P.A. Rota, M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M.H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D.Wang, D.D. Erdman, T.C. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Li.ck, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A.D. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson, W.J. Bellini, Characterization of a novel coronavirus associated with severe acute respiratory syndrome, *Science* 300 (2003) 1394–1399.
- M.A. Marra, S.J. Jones, C.R. Astell, R.A. Holt, A. Brooks-Wilson, Y.S. Butterfield, J. Khattra, J.K. Asano, S.A. Barber, S.Y. Chan, A. Cloutier,

- S.M. Coughlin, D. Freeman, N. Ginn, O.L. Grith, S.R. Leach, M. Mayo, H. McDonald, S.B. Montgomery, P.K. Pandoh, A.S. Petrescu, A.G. Robertson, J.E. Schein, A. Siddiqui, D.E. Smailus, J.M. Stott, G.S. Yang, F. Plummer, A. Andonov, H. Artsob, N. Bastien, K. Bernard, T.F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Strober, G.A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R.C. Brunham, M. Kraiden, M. Petric, D.M. Skowronski, C. Upton, R.L. Roper, The genome sequence of the SARS-associated coronavirus, *Science* 300 (2003) 1399–1404.
- T.M. Gallagher, M.J. Buchmeier, Coronavirus spike proteins in viral entry and pathogenesis, *Virology* 279 (2001) 371–374.
- F. Kunkel, G. Herrier, Structural and functional analysis of the surface protein of human coronavirus OC43, *Virology* 195 (1993) 195–202.
- A. Bonavia, B.D. Zelus, D.E. Wentworth, P.J. Talbot, K.V. Holmes, Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E, *J. Virol.* 77 (2003) 2530–2538.
- W. Li, M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, H. Choe, M. Farzan, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus, *Nature* 426: 450–54 (2003).
- P.L. Earl, B. Moss, Mutational analysis of the assembly domain of the HIV-1 envelope glycoprotein, *AIDS Res.Hum.Retroviruses.* 9 (1993) 589–94.
- R.J. Center, P. Schuck, R.D. Leapman, L.O. Arthur, P.L. Earl, B. Moss, J. Lebowitz, Oligomeric structure of virion-associated and soluble forms of the simian immunodeficiency virus envelope protein in the prefusion activated conformation, *Proc.Natl.Acad.Sci.U.S.A* 98 (2001) 14877–14882.
- R.J. Center, P.L. Earl, J. Lebowitz, P. Schuck, B. Moss, The human immunodeficiency virus type 1 gp120 V2 domain mediates gp41-independent intersubunit contacts, *J.Virol.* 74 (2000) 4448–4455.
- D.N. Lewicki, T.M. Gallagher, Quaternary structure of coronavirus spikes in complex with carcinoembryonic antigen-related cell adhesion molecule cellular receptors, *J. Biol. Chem.* 277 (2002) 19727–19734.

- B. Tripet, M.W. Howard, M. Jobling, R.K. Holmes, K.V. Holmes, R.S. Hodges, Structural characterization of the SARS-coronavirus spike S fusion protein core, *J Biol. Chem.* 279 (2004) 20836-20849.
- J. Zhu, G. Xiao, Y. Xu, F. Yuan, C. Zheng, Y. Liu, H. Yan, D.K. Cole, J.L. Bell, Z. Rao, P. Tien, G.F. Gao, Following the rule: formation of the 6-helix bundle of the fusion core from severe acute respiratory syndrome coronavirus spike protein and identification of potent peptide inhibitors, *Biochem.Biophys.Res.Comm.* 319 (2004) 283-288.
- B.J. Bosch, B.E. Martina, Z.R. van der, J. Lepault, B.J. Haijema, C. Versluis, A.J. Heck, R. de Groot, A.D. Osterhaus, P.J. Rottier, Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides, *Proc. Natl. Acad. Sci. U.S.A* 101 (2004) 8455-8460.
- S. Liu, G. Xiao, Y. Chen, Y. He, J. Niu, C.R. Escalante, H. Xiong, J. Farmer, A.K. Debnath, P. Tien, S. Jiang, Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanism and identification of fusion inhibitors, *Lancet* 363 (2004) 938-947.
- B.J. Bosch, Z.R. van der, C.A. de Haan, P.J. Rottier, The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex, *J.Virol.* 77 (2003) 8801-8811.
- H. Bisht, A. Roberts, L. Vogel, A. Bukreyev, P.L. Collins, B.R. Murphy, K. Subbarao, B. Moss, Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice, *Proc. Natl. Acad. Sci. U.S.A* 101 (2004) 6641-6646.
- G.J. Babcock, D.J. Eshaki, W.D. Thomas, Jr., D.M. Ambrosino, Amino acids 270 to 510 of the severe acute respiratory syndrome coronavirus spike protein are required for interaction with receptor, *J.Virol.* 78 (2004) 4552-4560.
- G. Simmons, J.D. Reeves, A.J. Rennekamp, S.M. Amberg, A.J. Piefer, P. Bates, Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry, *Proc.Natl.Acad.Sci.U.S.A* 101 (2004) 4240-4245.

- S.K. Wong, W. Li, M.J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2, *J.Biol.Chem.* 279 (2004) 3197-3201.
- Smith,A.E. & Helenius,A. How viruses enter animal cells. *Science* 304, 237-242 (2004).
- Peiris,J.S. *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361, 1319-1325 (2003).
- Drosten,C. *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1967-1976 (2003).
- Xiao,X., Chakraborti,S., Dimitrov,A.S., Gramatikoff,K. & Dimitrov,D.S. The SARS-CoV S glycoprotein: expression and functional characterization. *Biochem. Biophys. Res. Commun.* 312, 1159-1164 (2003).
- Xiao,X. & Dimitrov,D.S. The SARS-CoV S glycoprotein. *Cell Mol. Life Sci.* 61, 2428-2430 (2004).
- He,Y. *et al.* Receptor-binding domain of SARS-CoV spike protein induces highly potent neutralizing antibodies: implication for developing subunit vaccine. *Biochem. Biophys. Res. Commun.* 324, 773-781 (2004).
- Zhang,M.Y., Choudhry,V., Xiao,X. & Dimitrov,D.S. Human monoclonal antibodies to the S glycoprotein and related proteins as potential therapeutics for SARS. *Curr. Opin. Mol. Ther.* 7, 151-156 (2005).
- Jiang,S., He,Y. & Liu,S. SARS vaccine development. *Emerg. Infect. Dis.* 11, 1016-1020 (2005).
- He,Y., Lu,H., Siddiqui,P., Zhou,Y. & Jiang,S. Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. *J Immunol.* 174, 4908-4915 (2005).
- He,Y. *et al.* Identification of a critical neutralization determinant of severe acute respiratory syndrome (SARS)-associated coronavirus: importance for designing SARS vaccines. *Virology* 334, 74-82 (2005).
- Chen,Z. *et al.* Recombinant modified vaccinia virus Ankara expressing the spike glycoprotein of severe acute respiratory syndrome coronavirus induces protective neutralizing antibodies primarily targeting the receptor binding region. *J Virol.* 79, 2678-2688 (2005).

- Yi,C.E., Ba,L., Zhang,L., Ho,D.D. & Chen,Z. Single amino acid substitutions in the severe acute respiratory syndrome coronavirus spike glycoprotein determine viral entry and immunogenicity of a major neutralizing domain. *J Virol.* 79, 11638-11646 (2005).
- Li,F., Li,W., Farzan,M. & Harrison,S.C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309, 1864-1868 (2005).
- Burton,D.R. Antibodies, viruses and vaccines. *Nat. Rev. Immunol.* 2, 706-713 (2002).
- Chakraborti,S., Prabakaran,P., Xiao,X. & Dimitrov,D.S. The SARS coronavirus S glycoprotein receptor binding domain: fine mapping and functional characterization. *Virology* 332, 73 (2005).
- Nybakken,G.E. *et al.* Structural basis of West Nile virus neutralization by a therapeutic antibody. *Nature* 437, 764-769 (2005).
- Lawrence,M.C. & Colman,P.M. Shape complementarity at protein/protein interfaces. *J Mol Biol.* 234, 946-950 (1993).
- Davies,D.R. & Cohen,G.H. Interactions of protein antigens with antibodies. *Proc. Natl. Acad. Sci. U. S. A.* 93, 7-12 (1996).
- Li,W. *et al.* Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310, 676-679 (2005).
- Lau,S.K. *et al.* Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U. S. A.* 102, 14040-14045 (2005).
- Otwinowski,Z. & Minor,W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276, 307-326 (1997).
- Storoni,L.C., McCoy,A.J. & Read,R.J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D. Biol. Crystallogr.* 60, 432-438 (2004).
- Brunger,A.T., Adams,P.D. & Rice,L.M. New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure.* 5, 325-336 (1997).

[233] 本文所有参考或提到的专利和出版物表示本发明涉及领域的本领域普通技术人员的技术水平，并且因此每一参考的专利或出版物以相同程度并入本文

作为参考，好象其以整体单独并入作为参考或者以本文提出的整体并入作为参考。申请人保留将从这类引用专利或出版物的任何和所有材料和信息物理上(physically)并入的权利。

[234] 本文描述的具体方法和组合物是代表性的优选实施方式，其被示例但不拟作为对本发明范围的限制。在考虑本说明书后，本领域普通技术人员将联想到其它目的、方面和实施方式，这些其它目的、方面和实施方式被包含在权利要求所限定的范围的本发明的精神内。对本领域技术人员十分明显可对本文公开的发明进行不同的代替和修改，而没有背离本发明的精神。本文说明性描述的本发明可适当在缺少任何一种或多种要素、或一种或多种限制的情况下实行，其没有在本文作为必需的具体公开。本文说明性描述的方法和过程可适当以不同步骤次序实行，并且不必限于本文和在权利要求中制定的步骤次序。如在本文和所附权利要求中使用的，单数形式“一(a, an 和 the)”包括复数形式参考，除非上下文明确使出不同。因此，例如，“宿主细胞(一个宿主细胞)”包括多个(例如，培养物或积落)这类宿主细胞等。本专利绝不可解释为限于本文特别公开的具体实例或实施方式或者方法。专利绝不可解释为被专利和商标办公室的任何检查人员或任何其他官员或职工所作的任何声明所限定，除非该声明是申请人的回应著述中特别地和没有限制地采用或保留表达采用的。

[235] 已被使用的术语和表达被用作描述性术语而不是限制性的，在这些术语和表达的使用中，不拟排除任何示出和描述或其一部分的特征的相等物，但是要知多种修改可能在本发明所要求的范围内。因此，应该理解尽管本发明已通过优选的实施方式和任选特征具体公开，但是本领域普通技术人员可以进行本文公开概念的修改和改变，考虑这种修改和改变在本发明的范围内，如所附权利要求所限定。

[236] 本文发明已被宽广地和一般地描述。包括在一般公开内容中的更狭窄的种类和亚种基团的每一个也构成本发明的一部分。这包括本发明的一般描述，其限制性条款或反面限制除去了该类中任何主题，不管切离的物质是否在本文特定引用。

[237] 其它实施方式在下面的权利要求中。此外，根据 Markush 群(Markush group)描述了本发明的特征和方面，本领域普通技术人员将意识到本发明因此也根据 Markush 群的任何独立成员或亚群进行描述。

<110> 国立卫生研究院
 D. S. 迪米特洛夫
 X. 肖
 Z. 仲羽

<120> SARS-CoV 刺突糖蛋白的可溶片段

<130> 1662.053W01

<140> US 11/335,197
 <141> 2006-01-19

<150> PCT/US2004/023345
 <151> 2004-07-20

<150> US 60/489,166
 <151> 2003-07-21

<150> US 60/524,642
 <151> 2003-11-25

<160> 73

<170> FastSEQ for Windows Version 4.0

<210> 1
 <211> 1255
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 1
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala

370 375 380
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445

 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 450 455 460
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800

 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 805 810 815
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 820 825 830
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 835 840 845
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 850 855 860
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 865 870 875 880
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 885 890 895
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 900 905 910
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Ser Thr Ala Leu Gly
 915 920 925
 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 930 935 940
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 945 950 955 960
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 965 970 975
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 980 985 990
 Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala

995 1000 1005
 Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
 1010 1015 1020
 Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His
 1025 1030 1035 1040
 Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
 1045 1050 1055
 Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
 1060 1065 1070
 Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
 1075 1080 1085
 Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
 1090 1095 1100
 Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
 1105 1110 1115 1120
 Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
 1125 1130 1135
 Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
 1140 1145 1150

 Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
 1155 1160 1165
 Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
 1170 1175 1180
 Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu
 1185 1190 1195 1200
 Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu
 1205 1210 1215
 Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys
 1220 1225 1230
 Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
 1235 1240 1245
 Gly Val Lys Leu His Tyr Thr
 1250 1255

- <210> 2
- <211> 3768
- <212> DNA
- <213> SARS 冠状病毒 (SARS coronavirus)

<400> 2
 atgtttatt tcttattatt tcttactctc actagtggta gtgaccttga ccggtgcacc 60
 actttttagt atgttcaagc tcctaattac actcaacata cttcatctat gagggggggt 120
 tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttatttctt 180
 ccattttatt ctaatgttac agggtttcat actattaate atacgtttgg caaccctgtc 240
 atacctttta aggatgggat ttattttgct gccacagaga aatcaaatgt tgtccctggt 300
 tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
 actaatgitt ttatacagagc atgtaacttt gaattgtgtg acaacccttt ctttgcgtgt 420
 tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
 ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaa 540
 cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggtctat 600
 caacctatag atgtagtctg tgatctacct tctggtttta acactttgaa acctattttt 660
 aagttgcctc ttggatttaa cattacaaat ttagagcca ttcttacagc cttttcactt 720
 gctcaagaca atttcagggt gtcagctgca gcctattttt tggctattt aaagccaact 780
 acatttatgc tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa 840
 aatccacttg ctgaactcaa atgctctggt aagagctttg agattgacaa aggaatttac 900
 cagacctcta atttcagggt tglttccctca ggagatgttg tgagattccc taatattaca 960
 aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
 gagagaaaaa aaatttctaa ttgtgttgcg gattactctg tgctctacaa ctcaacattt 1080
 ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140
 aatgtctatg cagattcttt ttagtcaag ggagatgatg taagacaaat agcgcagga 1200
 caaacgtggt ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc 1260
 cttgcttggg aactagggaa cattgatgct acttcaactg gtaattataa ttataaatat 1320
 aggtatctta gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc 1380
 tccccgatg gcaaaccttg caccaccctt gctcttaatt gttattggcc attaaatgat 1440
 tatggttttt acaccactac tggcattggc taccaacctt acagagtgt agtactttct 1500
 ttgaaacttt taaatgcacc ggccacgggt tgggaccaa aattatccac tgaccttatt 1560
 aagaaccagt gtgtcaatt taattttaat ggactcactg gtactgtgtg gttactcctt 1620
 tcttcaaga gatttcaacc atttcaacaa tttagccgtg atgtttctga tttcactgat 1680
 tccgttcgag atctcaaaa atctgaata ttagacatt caccttgcgc ttttgggggt 1740
 gtaagtgtaa ttacactgg aacaaatgct tcacttgaag ttgctgttct atatcaagat 1800
 gttactgca ctgatgttc tacagcaatt catgcagatc aactcacacc agcttggcgc 1860
 atatatctta ctggaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
 catgtcgaca ctcttatga gtgcagcatt cctattggag ctggcatttg tgcctgttac 1980
 caticagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
 ttaggtgctg atagtccaat tgcttactct aataacacca ttgctatacc tactaacctt 2100
 tcaattagca ttactcagga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160
 aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgccttccca atatggtagc 2220
 ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tcgcaacaca 2280

cgatgaagt	tcgctcaagt	caaacaaatg	tacaaaacc	caactttgaa	atattttggt	2340
ggttttaatt	tttcacaaat	attacctgac	cctctaaagc	caactaagag	gtcttttatt	2400
gaggacttgc	lctttaataa	ggtgacactc	gctgatgctg	gcttcatgaa	gcaatatggc	2460
gaatgcctag	gtgatattaa	tgctagagat	ctcatttgtg	cgcagaagt	caatggactt	2520
acagtgttc	cacctctgct	caactgatgat	atgattgctg	cctacactgc	tgctctagtt	2580
agtggtaactg	ccactgctgg	atggacatt	ggtgctggcg	ctgctcttca	aatacctttt	2640
gctatgcaaa	tggcatatag	gttcaatgac	attggagtt	cccaaatgt	tctctatgag	2700
aaccaaaaac	aaatcgccaa	ccaatttaac	aaggcgatta	gtcaaatca	agaatcactt	2760
acaacaacat	caactgcatt	gggcaagctg	caagacgttg	ttaaccagaa	tgctcaagca	2820
ftaaacacac	ttgttaaaca	acttagctct	aattttggtg	caatttcaag	tgtgctaaat	2880
gataicctti	cgcgacttga	taaagtcgag	gcggagggtac	aaattgacag	gttaattaca	2940
ggcagacttc	aaagccttca	aacctatgta	acacaacaac	taatcagggc	tgctgaaatc	3000
agggtctctg	ctaactcttc	tgctactaaa	atgtctgagt	gtgttcttgg	acaatcaaaa	3060
agagtgtact	tttgggaaa	gggtaccac	cttatgtcct	tcccacaagc	agccccgcat	3120
ggtgtgtct	tcctacatgt	cacgtatgtg	ccatcccagg	agaggaact	caccacagcg	3180
ccagcaattt	gtcatgaagg	caaagcatac	ttccctcgtg	aagggttttt	tgtgtttaat	3240
ggcacttctt	ggtttattac	acagagggaac	ttcttttctc	cacaaataat	tactacagac	3300
aatacatttg	tctcaggaaa	ttgtgatgac	gttattggca	tcattaacaa	cacagtttat	3360
gatcctctgc	aacctgagct	cgactcattc	aaagaagagc	tggaacaagta	cttcaaaaat	3420
catacatcac	cagatgttga	tcttggcgac	atttcaggca	ttaacgcttc	tgctgctcaac	3480
attcaaaaag	aaattgaccg	cctcaatgag	gtcgcataaa	atttaaatga	atcactcatt	3540
gaccttcaag	aattgggaaa	atatgagcaa	tatattaaat	ggccttggt	tggttggctc	3600
ggcttcattg	ctggactaat	tgccatcgct	atggttacia	tcttgctttg	ttgcatgact	3660
agttgttga	gttgctcaaa	gggtgcatgc	tcttgtggt	cttgcgcaa	gtttgatgag	3720
gatgactctg	agccagttct	caagggtgct	aaattacatt	acacataa		3768

<210> 3
 <211> 29
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物

<400> 3
 agtcggatcc ggtaggctta tcattagag 29

<210> 4
 <211> 20
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物

<400> 4
 ccatcagggg agaaaggcac 20

<210> 5
 <211> 20
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物

<400> 5
 gtgcctttct cccctgatgg 20

<210> 6
 <211> 19
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物

<400> 6
 gaagagcagc gccagcacc 19

<210> 7
 <211> 19
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物

<400> 7

ggtgctggcg ctgctcttc	19
<210> 8	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物	
<400> 8	
actgtctaga gttcgtttat gtgtaatg	28
<210> 9	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物	
<400> 9	
agtcggatcc gaccggtgca ccacttttg	29
<210> 10	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物	
<400> 10	
agtcgggccc ctgttcagca gcaatacc	28
<210> 11	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物	
<400> 11	
actgggatcc gaagtgttcg ctcaagtc	28
<210> 12	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成引物	
<400> 12	
actgtctaga ttgctcatat ttcc	26
<210> 13	
<211> 740	
<212> PRT	
<213> SARS 冠状病毒 (SARS coronavirus)	
<400> 13	
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln	
1 5 10 15	
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg	
20 25 30	
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser	
35 40 45	
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val	
50 55 60	
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn	
65 70 75 80	
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln	
85 90 95	
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys	
100 105 110	
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met	

115 120 125
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 165 170 175
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 180 185 190
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 195 200 205
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 210 215 220
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 225 230 235 240
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 245 250 255
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 260 265 270
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 275 280 285
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 290 295 300
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 305 310 315 320
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 325 330 335
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 340 345 350
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 355 360 365
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 370 375 380
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 385 390 395 400
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 405 410 415
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 420 425 430
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 435 440 445
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 450 455 460
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 465 470 475 480
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 485 490 495
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 500 505 510
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 515 520 525
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 530 535 540
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 545 550 555 560
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 565 570 575
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 580 585 590
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 595 600 605
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 610 615 620
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 625 630 635 640
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 645 650 655
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 660 665 670
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 675 680 685
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 690 695 700
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 705 710 715 720
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 725 730 735
 Ala Ala Glu Gln

740

<210> 14
 <211> 429
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

 <400> 14
 Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys
 1 5 10 15
 Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys
 20 25 30
 Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr
 35 40 45

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

 Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu
 405 410 415
 Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln
 420 425

<210> 15
 <211> 1170
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50

```

50          55          60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65          70          75          80
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
          85          90          95
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
          100          105          110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
          115          120          125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
          130          135          140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145          150          155          160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
          165          170          175
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
          180          185          190
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
          195          200          205
Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
          210          215          220
Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225          230          235          240
Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
          245          250          255

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
          260          265          270
Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
          275          280          285
Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
          290          295          300
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305          310          315          320
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
          325          330          335
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
          340          345          350
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
          355          360          365
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
          370          375          380
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385          390          395          400
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
          405          410          415
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
          420          425          430
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
          435          440          445
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
          450          455          460
Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465          470          475          480
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
          485          490          495
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
          500          505          510
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
          515          520          525
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
          530          535          540
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
545          550          555          560
Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
          565          570          575
Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
          580          585          590
Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
          595          600          605

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
          610          615          620
His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
625          630          635          640
Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
          645          650          655
Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
          660          665          670
Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile

```



```

        675                680                685
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 690                695                700
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 705                710                715                720
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
        725                730                735
Ala Ala Glu Gln Asp Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys
        740                745                750
Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu
        755                760                765
Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu
        770                775                780
Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly
 785                790                795                800
Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys
        805                810                815
Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile
        820                825                830
Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp
        835                840                845
Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met
        850                855                860
Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu
 865                870                875                880
Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile
        885                890                895
Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp
        900                905                910
Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu
        915                920                925
Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser
        930                935                940
Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr
 945                950                955                960
Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
        965                970                975
Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser
        980                985                990
Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly
        995                1000                1005
Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe
 1010                1015                1020
Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala
 1025                1030                1035                1040
Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val
        1045                1050                1055
Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe
        1060                1065                1070
Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys
        1075                1080                1085
Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln
 1090                1095                1100
Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
 1105                1110                1115                1120
His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala
        1125                1130                1135
Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala
        1140                1145                1150
Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
 1155                1160                1165
Glu Gln
 1170

```

<210> 16

<211> 21

<212> PRT

<213> 人工序列

<220>

<223> 合成K链前导序列

<400> 16

```

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1                5                10                15

```

```

Gly Ser Thr Gly Asp
 20

```

<210> 17
 <211> 10
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成myc表位

<400> 17
 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
 1 5 10

<210> 18
 <211> 6
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成组氨酸标签

<400> 18
 His His His His His His
 1 5

<210> 19
 <211> 24
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成引物

<400> 19
 ctagctcgag caacagcatc tgtg

24

<210> 20
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 20
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly
 100

<210> 21
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 21
 Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile
 1 5 10 15
 Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu
 20 25 30
 Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr
 35 40 45
 His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile
 50 55 60
 Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys
 65 70 75 80
 His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val
 85 90 95
 Tyr Lys Gly Tyr
 100

<210> 22
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 22
 Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu
 1 5 10 15
 Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg
 20 25 30
 Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser
 35 40 45
 Ala Ala Ala Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu
 50 55 60
 Lys Tyr Asp Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln
 65 70 75 80
 Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp
 85 90 95
 Lys Gly Ile Tyr
 100

<210> 23
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 23
 Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe
 1 5 10 15
 Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr
 20 25 30
 Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys
 35 40 45
 Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe
 50 55 60
 Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser
 65 70 75 80
 Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln
 85 90 95
 Ile Ala Pro Gly
 100

<210> 24
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 24
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 1 5 10 15
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 20 25 30
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 35 40 45
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 50 55 60
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 65 70 75 80
 Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 85 90 95
 Val Val Leu Ser
 100

<210> 25
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 25
 Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser
 1 5 10 15
 Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu
 20 25 30
 Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg Phe Gln Pro Phe

35 40 45
 Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp Ser Val Arg Asp
 50 55 60
 Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ala Phe Gly Gly
 65 70 75 80
 Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val
 85 90 95
 Leu Tyr Gln Asp
 100

<210> 26
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 26
 Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr
 1 5 10 15
 Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln
 20 25 30
 Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys
 35 40 45
 Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser
 50 55 60
 Leu Leu Arg Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr Thr Met Ser
 65 70 75 80
 Leu Gly Ala Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile
 85 90 95
 Pro Thr Asn Phe
 100

<210> 27
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 27
 Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr
 1 5 10 15
 Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala
 20 25 30
 Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala
 35 40 45
 Leu Ser Gly Ile Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe
 50 55 60
 Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly
 65 70 75 80
 Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys
 85 90 95
 Arg Ser Phe Ile
 100

<210> 28
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 28
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 1 5 10 15
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 20 25 30
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 35 40 45
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 50 55 60
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 65 70 75 80
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 85 90 95
 Val Leu Tyr Glu
 100

<210> 29

<211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

 <400> 29
 Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile
 1 5 10 15
 Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp
 20 25 30
 Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu
 35 40 45
 Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser
 50 55 60
 Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr
 65 70 75 80
 Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg
 85 90 95
 Ala Ala Glu Ile
 100

<210> 30
 <211> 100
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

 <400> 30
 Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu
 1 5 10 15
 Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met
 20 25 30
 Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe Leu His Val Thr
 35 40 45
 Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys
 50 55 60
 His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn
 65 70 75 80
 Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile
 85 90 95
 Ile Thr Thr Asp
 100

<210> 31
 <211> 90
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

 <400> 31
 Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn
 1 5 10 15
 Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu
 20 25 30
 Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu
 35 40 45
 Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu
 50 55 60
 Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile
 65 70 75 80
 Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln
 85 90

<210> 32
 <211> 200
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

 <400> 32
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val

```

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

```

```

<210> 33
<211> 200
<212> PRT
<213> SARS 冠状病毒 (SARS coronavirus)

```

```

<400> 33
Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu
1          5          10          15
Lys Pro Ile Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg
20         25         30
Ala Ile Leu Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser
35         40         45

Ala Ala Ala Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu
50         55         60
Lys Tyr Asp Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln
65         70         75         80
Asn Pro Leu Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp
85         90         95
Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp
100        105        110
Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val
115        120        125
Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys
130        135        140
Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe
145        150        155        160
Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp
165        170        175
Leu Cys Phe Ser Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp
180        185        190
Asp Val Arg Gln Ile Ala Pro Gly
195        200

```

```

<210> 34
<211> 200
<212> PRT
<213> SARS 冠状病毒 (SARS coronavirus)

```

```

<400> 34
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
1          5          10          15
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
20         25         30
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
35         40         45
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
50         55         60
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
65         70         75         80
Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
85         90         95
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
100        105        110
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
115        120        125
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
130        135        140
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
145        150        155        160

```

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 165 170 175
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 180 185 190

Glu Val Ala Val Leu Tyr Gln Asp
 195 200

<210> 35
 <211> 200
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 35
 Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr
 1 5 10 15
 Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln
 20 25 30
 Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys
 35 40 45
 Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser
 50 55 60
 Leu Leu Arg Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr Thr Met Ser
 65 70 75 80
 Leu Gly Ala Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile
 85 90 95
 Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser
 100 105 110
 Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser
 115 120 125
 Thr Glu Cys Ala Asn Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln
 130 135 140
 Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp Arg Asn Thr
 145 150 155 160
 Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu
 165 170 175
 Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu
 180 185 190
 Lys Pro Thr Lys Arg Ser Phe Ile
 195 200

<210> 36
 <211> 200
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 36
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 1 5 10 15
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 20 25 30
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 35 40 45
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 50 55 60
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 65 70 75 80
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 85 90 95
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 100 105 110
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
 115 120 125
 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 130 135 140
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 145 150 155 160
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 165 170 175
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 180 185 190
 Gln Leu Ile Arg Ala Ala Glu Ile
 195 200

<210> 37

<211> 190
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 37
 Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu
 1 5 10 15
 Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met
 20 25 30
 Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe Leu His Val Thr
 35 40 45
 Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys
 50 55 60
 His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn
 65 70 75 80
 Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile
 85 90 95
 Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile
 100 105 110
 Gly Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
 115 120 125
 Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro
 130 135 140
 Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn
 145 150 155 160
 Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
 165 170 175
 Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln
 180 185 190

<210> 38
 <211> 400
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 38
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400

<210> 39
 <211> 600
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 39
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 450 455 460
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480
 Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590
 Glu Val Ala Val Leu Tyr Gln Asp
 595 600

<210> 40

<211> 800

<212> PRT

<213> SARS 冠状病毒 (SARS coronavirus)

<400> 40

Met Phe Ile Phe Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Thr Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95

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

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 450 455 460
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480
 Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800

<210> 41

<211> 1000

<212> PRT

<213> SARS 冠状病毒 (SARS coronavirus)

<400> 41

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300

 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 450 455 460
 Lys Pro Cys Thr Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 805 810 815
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 820 825 830
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 850 855 860
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 865 870 875 880
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 885 890 895
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 900 905 910
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
 915 920 925
 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 930 935 940
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 945 950 955 960
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 965 970 975
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 980 985 990
 Gln Leu Ile Arg Ala Ala Glu Ile
 995 1000

<210> 42

<211> 1190

<212> PRT

<213> SARS 冠状病毒 (SARS coronavirus)

<400> 42

Met Phe Ile Phe Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300

 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe

Met Gly Cys Val 405 410 415
 Thr Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 450 455 460
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 805 810 815
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 820 825 830
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 835 840 845
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 850 855 860
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 865 870 875 880
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 885 890 895
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 900 905 910
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Ser Thr Ala Leu Gly
 915 920 925
 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 930 935 940
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 945 950 955 960
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 965 970 975
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 980 985 990
 Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
 995 1000 1005

 Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
 1010 1015 1020
 Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His

1025 1030 1035 1040
 Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
 1045 1050 1055
 Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
 1060 1065 1070
 Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
 1075 1080 1085
 Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
 1090 1095 1100
 Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
 1105 1110 1115 1120
 Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
 1125 1130 1135
 Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
 1140 1145 1150
 Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
 1155 1160 1165
 Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
 1170 1175 1180
 Leu Gly Lys Tyr Glu Gln
 1185 1190

<210> 43
 <211> 84
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 43
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly

<210> 44
 <211> 184
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 44
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 115 120 125
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 165 170 175
 Phe Leu Tyr Val Tyr Lys Gly Tyr
 180

<210> 45
 <211> 384
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 45
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 115 120 125
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 165 170 175
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 180 185 190
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 195 200 205
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 210 215 220
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 225 230 235 240
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 245 250 255
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 260 265 270
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 275 280 285
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 290 295 300
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 305 310 315 320
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 325 330 335
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 340 345 350
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 355 360 365
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 370 375 380

<210> 46
 <211> 584
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 46
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 115 120 125
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly

Phe Leu Tyr Val 165 Tyr Lys Gly Tyr Gln 170 Pro Ile Asp Val Val Arg Asp
 180 185 190
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 195 200 205
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 210 215 220
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 225 230 235 240
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 245 250 255
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 260 265 270
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 275 280 285
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 290 295 300

 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 305 310 315 320
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 325 330 335
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 340 345 350
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 355 360 365
 Asp Ser Phe Val Val Lys Gly Asp Val Arg Gln Ile Ala Pro Gly
 370 375 380
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 385 390 395 400
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 405 410 415
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 420 425 430
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 435 440 445
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 450 455 460
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 465 470 475 480
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 485 490 495
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 500 505 510
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 515 520 525
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 530 535 540
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 545 550 555 560
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 565 570 575
 Glu Val Ala Val Leu Tyr Gln Asp
 580

<210> 47
 <211> 784
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 47
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 115 120 125

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 165 170 175
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 180 185 190
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 195 200 205
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 210 215 220
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 225 230 235 240
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 245 250 255
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 260 265 270
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 275 280 285
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 290 295 300
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 305 310 315 320
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 325 330 335
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 340 345 350
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 355 360 365
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 370 375 380
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 385 390 395 400
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 405 410 415
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 420 425 430
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 435 440 445
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 450 455 460
 Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 465 470 475 480
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 485 490 495
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 500 505 510
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 515 520 525
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 530 535 540
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 545 550 555 560
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 565 570 575
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 580 585 590
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 595 600 605
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 610 615 620
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 625 630 635 640
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 645 650 655
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 660 665 670
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 675 680 685
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 690 695 700
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 705 710 715 720
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 725 730 735
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 740 745 750
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe

Ser Gln Ile 755 Leu Pro Asp Pro 760 Leu Lys Pro Thr Lys 765 Arg Ser Phe Ile
770 775 780

<210> 48
<211> 984
<212> PRT
<213> SARS 冠状病毒 (SARS coronavirus)

<400> 48
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
1 5 10 15
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
20 25 30
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
35 40 45
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
50 55 60
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
65 70 75 80
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
85 90 95
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
100 105 110
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
115 120 125
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
130 135 140
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
145 150 155 160
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
165 170 175
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
180 185 190
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
195 200 205

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
210 215 220
Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
225 230 235 240
Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
245 250 255
Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
260 265 270
Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
275 280 285
Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
290 295 300
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
305 310 315 320
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
325 330 335
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
340 345 350
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
355 360 365
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
370 375 380
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
385 390 395 400
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
405 410 415
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
420 425 430

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
435 440 445
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
450 455 460
Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
465 470 475 480
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
485 490 495
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
500 505 510
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
515 520 525
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp

530 535 540
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 545 550 555 560
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 565 570 575
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 580 585 590
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 595 600 605
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 610 615 620
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 625 630 635 640
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 645 650 655
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 660 665 670
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 675 680 685
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 690 695 700
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 705 710 715 720
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 725 730 735
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 740 745 750
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 755 760 765
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 770 775 780
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 785 790 795 800
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 805 810 815
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 820 825 830
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 835 840 845
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 850 855 860
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 865 870 875 880
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 885 890 895
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
 900 905 910

 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 915 920 925
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 930 935 940
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 945 950 955 960
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 965 970 975
 Gln Leu Ile Arg Ala Ala Glu Ile
 980

<210> 49

<211> 1174

<212> PRT

<213> SARS 冠状病毒 (SARS coronavirus)

<400> 49

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 115 120 125
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 165 170 175
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 180 185 190
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 195 200 205

 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 210 215 220
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 225 230 235 240
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 245 250 255
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 260 265 270
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 275 280 285
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 290 295 300
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 305 310 315 320
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 325 330 335
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 340 345 350
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 355 360 365
 Asp Ser Phe Val Val Lys Gly Asp Val Arg Gln Ile Ala Pro Gly
 370 375 380
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 385 390 395 400
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 405 410 415
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 420 425 430
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 435 440 445
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 450 455 460
 Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 465 470 475 480
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 485 490 495
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 500 505 510
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 515 520 525
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 530 535 540
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 545 550 555 560

 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 565 570 575
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 580 585 590
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 595 600 605
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 610 615 620
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 625 630 635 640
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 645 650 655
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 660 665 670
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 675 680 685
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 690 695 700
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 705 710 715 720
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 725 730 735

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 740 745 750
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 755 760 765
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 770 775 780
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 785 790 795 800
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 805 810 815
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 820 825 830
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 835 840 845
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 850 855 860
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 865 870 875 880
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 885 890 895
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
 900 905 910

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 915 920 925
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 930 935 940
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 945 950 955 960
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 965 970 975
 Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
 980 985 990
 Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
 995 1000 1005
 Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His
 1010 1015 1020
 Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
 1025 1030 1035 1040
 Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
 1045 1050 1055
 Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
 1060 1065 1070
 Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
 1075 1080 1085
 Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
 1090 1095 1100
 Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
 1105 1110 1115 1120
 Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
 1125 1130 1135
 Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
 1140 1145 1150
 Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
 1155 1160 1165
 Leu Gly Lys Tyr Glu Gln
 1170

<210> 50
 <211> 260
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 50
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110

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

<210> 51
 <211> 430
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 51
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30

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

 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 385 390 395 400

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 405 410 415
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
 420 425 430

<210> 52
 <211> 521
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 52
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 100 105 110
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 115 120 125
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 130 135 140
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 145 150 155 160
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 165 170 175
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 180 185 190
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 195 200 205
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 210 215 220
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 225 230 235 240
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 245 250 255
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 260 265 270
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 275 280 285
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 290 295 300
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 305 310 315 320
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 325 330 335
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 340 345 350
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 355 360 365
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 370 375 380
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 385 390 395 400
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 405 410 415
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 420 425 430
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 435 440 445
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 450 455 460
 Tyr Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 465 470 475 480
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 485 490 495
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 500 505 510
 Phe Asn Gly Leu Thr Gly Thr Gly Val
 515 520

<210> 53
 <211> 777
 <212> PRT
 <213> 人工序列

<220>

<223> SEQ ID NO:1的氨基酸17-757加上N端小鼠K链前导序列和C端myc表位以及多组氨酸标签的合成序列

<400> 53

```

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1          5          10          15
Gly Ser Thr Gly Asp Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala
 20          25          30
Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro
 35          40          45
Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe
 50          55          60
Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile Asn His Thr
 65          70          75          80
Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala
 85          90          95
Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly Ser Thr Met
100         105         110
Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val
115         120         125

Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala
130         135         140
Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn
145         150         155         160
Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp
165         170         175
Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe
180         185         190
Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile
195         200         205
Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile
210         215         220
Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu
225         230         235         240
Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala
245         250         255
Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp
260         265         270
Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu
275         280         285
Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile
290         295         300

Tyr Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg
305         310         315         320
Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala
325         330         335
Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn
340         345         350
Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr
355         360         365
Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe
370         375         380
Ser Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg
385         390         395         400
Gln Ile Ala Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys
405         410         415
Leu Pro Asp Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn
420         425         430
Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu
435         440         445
Arg His Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro
450         455         460
Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro Ala Leu Asn Cys Tyr
465         470         475         480
Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr
485         490         495
Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro
500         505         510
Ala Thr Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln
515         520         525
Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr
530         535         540

```

Pro Ser Ser Lys Arg Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val
 545 550 555 560
 Ser Asp Phe Thr Asp Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu
 565 570 575
 Asp Ile Ser Pro Cys Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly
 580 585 590
 Thr Asn Ala Ser Ser Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys
 595 600 605
 Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp
 610 615 620
 Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys
 625 630 635 640
 Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro
 645 650 655

 Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg
 660 665 670
 Ser Thr Ser Gln Lys Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala
 675 680 685
 Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn
 690 695 700
 Phe Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser Met Ala Lys
 705 710 715 720
 Thr Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys
 725 730 735
 Ala Asn Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg
 740 745 750
 Ala Leu Ser Gly Ile Ala Ala Glu Gln Glu Gln Lys Leu Ile Ser Glu
 755 760 765
 Glu Asp Leu His His His His His His
 770 775

<210> 54
 <211> 297
 <212> PRT
 <213> 人工序列

<220>
 <223> SEQ ID NO:1的氨基酸17-276加上N端小鼠K链前导序列和C端myc表位以及多组氨酸标记的合成序列
 <400> 54

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
 1 5 10 15
 Gly Ser Thr Gly Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala
 20 25 30
 Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro
 35 40 45
 Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe
 50 55 60
 Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile Asn His Thr
 65 70 75 80
 Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala
 85 90 95
 Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly Ser Thr Met
 100 105 110

 Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val
 115 120 125
 Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala
 130 135 140
 Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn
 145 150 155 160
 Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp
 165 170 175
 Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe
 180 185 190
 Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile
 195 200 205
 Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile
 210 215 220

 Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu
 225 230 235 240
 Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala
 245 250 255
 Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp
 260 265 270
 Glu Asn Gly Thr Ile Thr Asp Ala Val Glu Gln Lys Leu Ile Ser Glu
 275 280 285
 Glu Asp Leu His His His His His His

290

<210> 55
<211> 558
<212> PRT
<213> 人工序列

<220>
<223> SEQ ID NO:1的氨基酸17-537加上N端小鼠K链前导序列和C端myc表位以及多组氨酸标记的合成序列

<400> 55
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly Asp Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala
20 25 30
Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro
35 40 45
Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe
50 55 60
Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile Asn His Thr
65 70 75 80
Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala
85 90 95
Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly Ser Thr Met
100 105 110
Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val
115 120 125
Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala
130 135 140
Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn
145 150 155 160
Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp
165 170 175
Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe
180 185 190
Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile
195 200 205
Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile
210 215 220
Phe Lys Leu Pro Leu Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu
225 230 235 240
Thr Ala Phe Ser Pro Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala
245 250 255
Tyr Phe Val Gly Tyr Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp
260 265 270
Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu
275 280 285
Ala Glu Leu Lys Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile
290 295 300

Tyr Gln Thr Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg
305 310 315 320
Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala
325 330 335
Thr Lys Phe Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn
340 345 350
Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr
355 360 365
Phe Lys Cys Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe
370 375 380
Ser Asn Val Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg
385 390 395 400
Gln Ile Ala Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys
405 410 415
Leu Pro Asp Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn
420 425 430
Ile Asp Ala Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu
435 440 445
Arg His Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro
450 455 460
Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr
465 470 475 480
Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr
485 490 495
Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro
500 505 510
Ala Thr Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln
515 520 525
Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val Glu Gln

530
 Lys Leu Ile Ser Glu Glu Asp Leu His His His His His His
 545 550 555

<210> 56
 <211> 739
 <212> PRT
 <213> 人工序列

<220>
 <223> N端没有信号肽的SEQ ID NO:1的氨基酸17-756的合成序列

<400> 56
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 1 5 10 15
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 20 25 30
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 35 40 45
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 50 55 60
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 65 70 75 80
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 85 90 95

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

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 450 455 460
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 465 470 475 480
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 485 490 495
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 500 505 510

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 515 520
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 530 535 540
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 545 550 555 560
 Ala Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 565 570 575
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 580 585 590
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 595 600 605
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 610 615 620
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 625 630 635 640
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 645 650 655
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 660 665 670
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 675 680 685
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 690 695 700
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 705 710 715 720
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 725 730 735
 Ala Ala Glu

<210> 57
 <211> 265
 <212> PRT
 <213> 人工序列

<220>
 <223> SEQ ID NO:1的氨基酸272-537的合成序列

<400> 57
 Ile Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys
 1 5 10 15
 Cys Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser
 20 25 30
 Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Asn Ile Thr
 35 40 45
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 50 55 60
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 65 70 75 80
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 85 90 95
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 100 105 110
 Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 115 120 125
 Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 130 135 140
 Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 145 150 155 160
 Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 165 170 175
 Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 180 185 190
 Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 195 200 205
 Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 210 215 220
 Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 225 230 235 240
 Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 245 250 255
 Phe Asn Gly Leu Thr Gly Thr Gly Val
 260 265

<210> 58
 <211> 17

<212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)
 <400> 58
 Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly
 1 5 10 15
 Cys

<210> 59
 <211> 15
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 59
 Pro Ser Ser Lys Arg Phe Gln Pro Gln Gln Phe Gly Arg Asp Cys
 1 5 10 15

<210> 60
 <211> 16
 <212> PRT
 <213> SARS 冠状病毒 (SARS coronavirus)

<400> 60
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15

<210> 61
 <211> 235
 <212> PRT
 <213> 人工序列

<220>
 <223> 含有受体结合结构域的SEQ ID NO: 1的氨基酸303-537的合成序列

<400> 61
 Ser Asn Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn
 1 5 10 15
 Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe
 20 25 30
 Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala
 35 40 45
 Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys
 50 55 60
 Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val
 65 70 75 80
 Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala
 85 90 95
 Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
 100 105 110
 Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala
 115 120 125
 Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
 130 135 140
 Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro
 145 150 155 160
 Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu
 165 170 175
 Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr
 180 185 190
 Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val
 195 200 205
 Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn
 210 215 220
 Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly Val
 225 230 235

<210> 62
 <211> 199
 <212> PRT
 <213> 人工序列

<220>
 <223> 含有受体结合结构域的SEQ ID NO: 1的氨基酸319-517的合成序列

<400> 62

```

Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe
1      5      10      15
Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala
20      25      30
Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys
35      40      45
Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val
50      55      60
Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala
65      70      75      80

Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
85      90      95
Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala
100     105     110
Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
115     120     125
Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro
130     135     140
Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu
145     150     155     160
Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr
165     170     175
Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val
180     185     190
Cys Gly Pro Lys Leu Ser Thr
195

```

<210> 63
<211> 200
<212> PRT
<213> 人工序列

<220>
<223> 含有受体结合结构域的SEQ ID NO: 1的氨基酸319-518的合成序列

```

<400> 63
Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe
1      5      10      15
Pro Ser Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala
20      25      30
Asp Tyr Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys
35      40      45
Tyr Gly Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val
50      55      60
Tyr Ala Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala
65      70      75      80
Pro Gly Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp
85      90      95
Asp Phe Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala
100     105     110
Thr Ser Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly
115     120     125

Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro
130     135     140
Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu
145     150     155     160
Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr
165     170     175
Arg Val Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val
180     185     190
Cys Gly Pro Lys Leu Ser Thr Asp
195     200

```

<210> 64
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 64
gatcggatcc ggtacaatca cag

23

<210> 65
<211> 23

<212> DNA
<213> 人工序列

<220>
<223> 合成引物

<400> 65
gatcgggccc gacacactgg ttc

23

<210> 66
<211> 197
<212> PRT
<213> 人工序列

<220>
<223> 含有受体结构域的SEQ ID NO: 1的氨基酸317-517的合成序列。

<400> 66
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
1 5 10 15
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
20 25 30
Ser Val Leu Tyr Asn Ser Thr Phe Ser Thr Phe Lys Cys Tyr Gly
35 40 45
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
50 55 60
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
65 70 75 80
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
85 90 95
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
100 105 110
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
115 120 125
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
130 135 140
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
145 150 155 160
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
165 170 175
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
180 185 190
Pro Lys Leu Ser Thr
195

<210> 67
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 合成引物.

<400> 67
actgtctaga tggtagcgag ctggatcc

29

<210> 68
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 合成引物.

<400> 68
cagtagatct cgagctgat cagcg

25

<210> 69
<211> 11
<212> PRT
<213> SARS 冠状病毒 (SARS coronavirus)

<400> 69
Gly Phe Tyr Thr Thr Gly Ile Gly Tyr Gln
1 5 10

<210> 70
<211> 9
<212> PRT
<213> 人 (Homo sapiens)

<400> 70
Asp Thr Val Met Gly Gly Met Asp Val
1 5

<210> 71
<211> 10
<212> PRT
<213> 人 (Homo sapiens)

<400> 71
Gln Val Trp Asp Ser Ser Ser Asp Tyr Val
1 5 10

<210> 72
<211> 8
<212> PRT
<213> SARS 冠状病毒 (SARS coronavirus)

<400> 72
Asn Asp Leu Cys Phe Ser Asn Val
1 5

<210> 73
<211> 12
<212> PKT
<213> SARS 冠状病毒 (SARS coronavirus)

<400> 73
Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
1 5 10

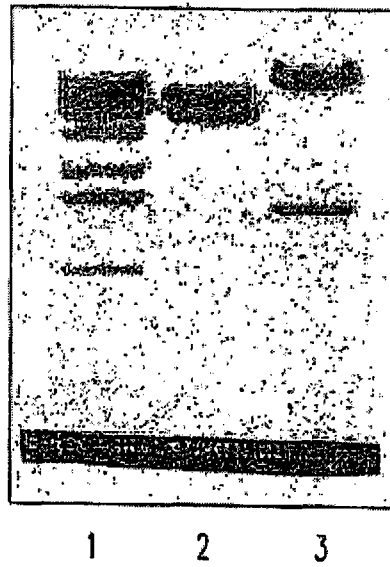


图 1A

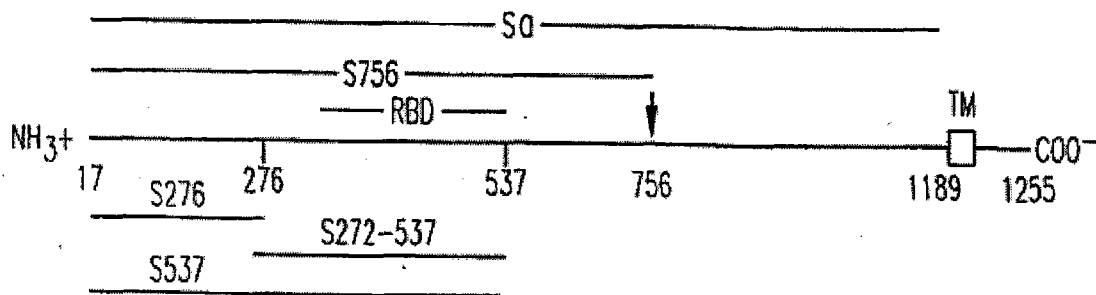


图 1B

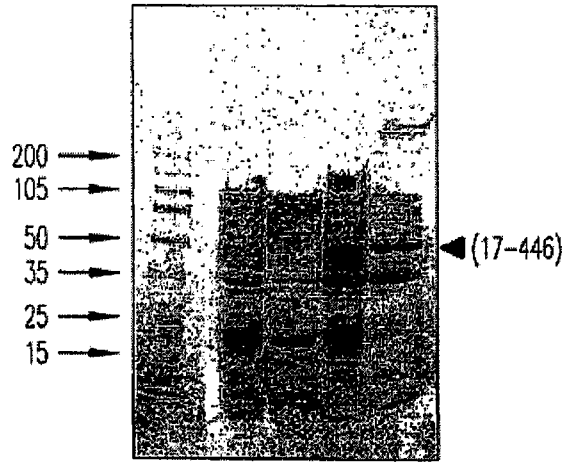


图2

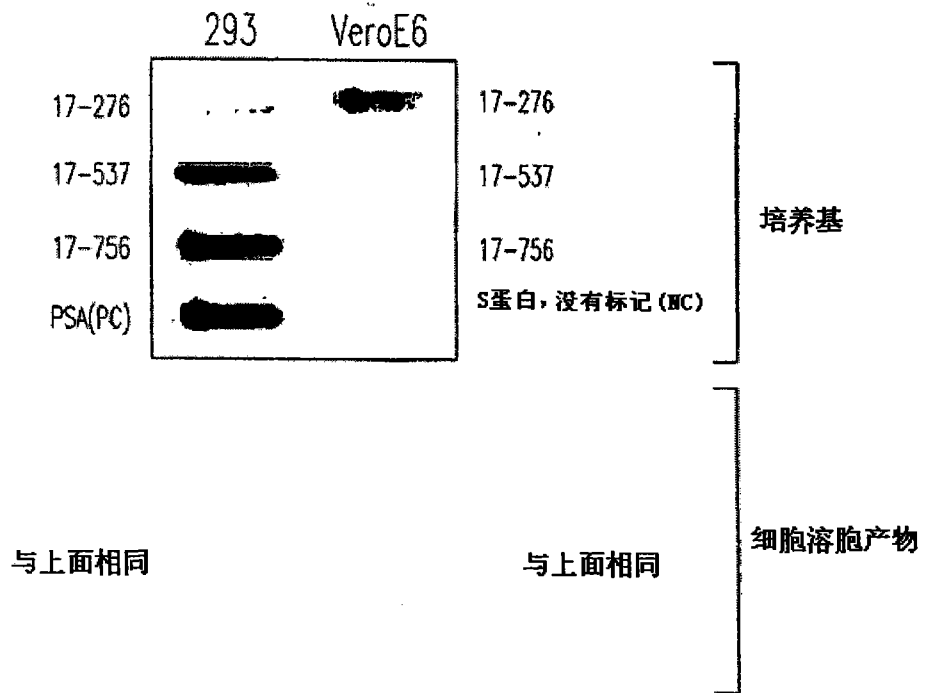


图3



图4A

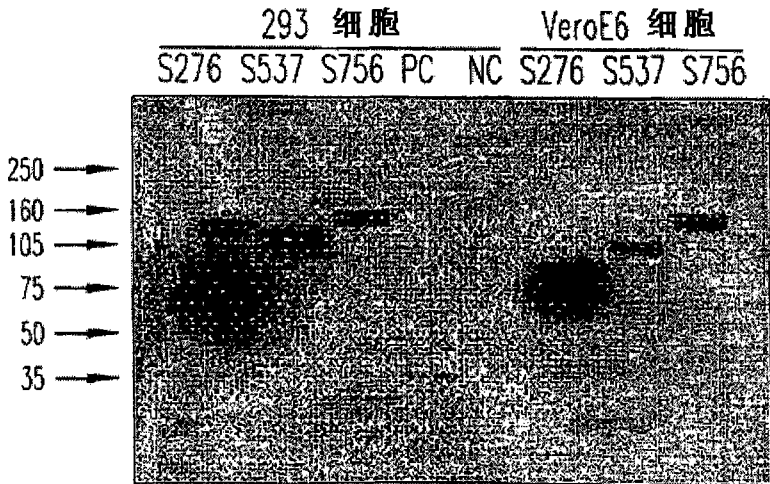


图4B

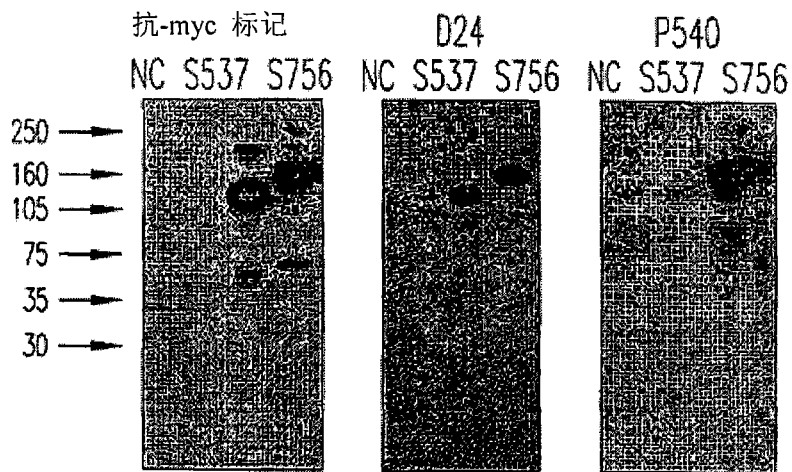


图4C

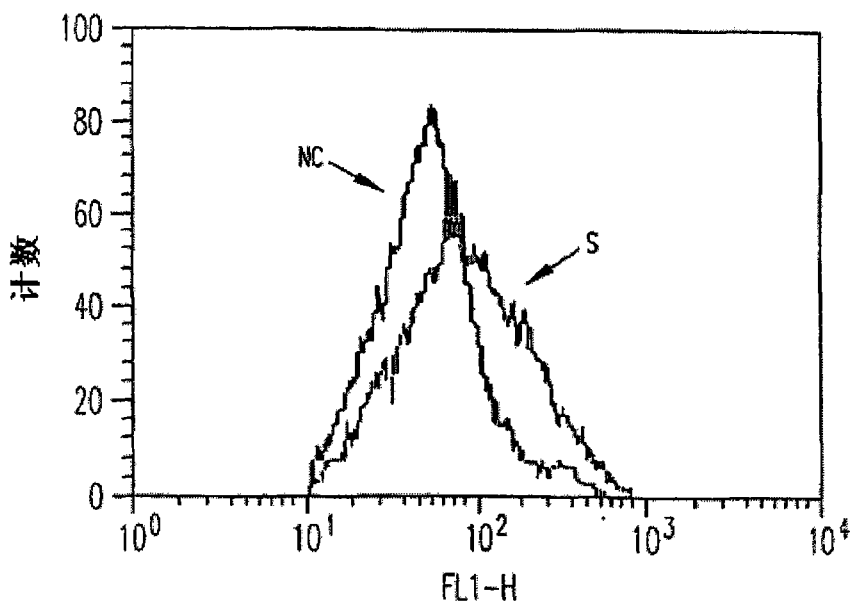


图5

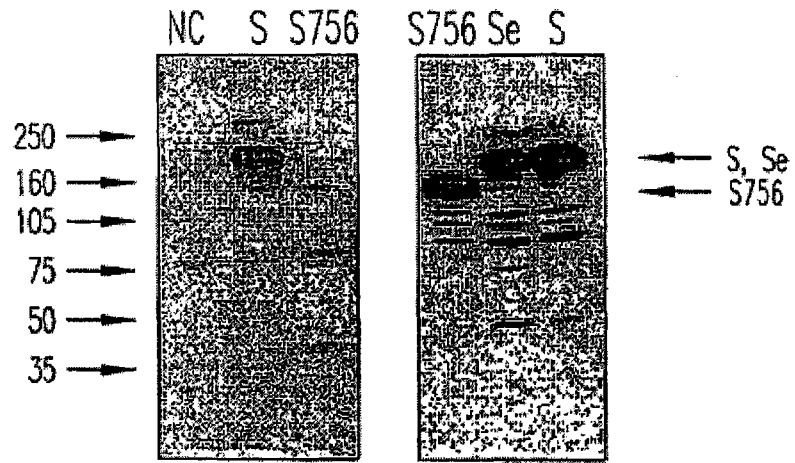


图6A

图6B

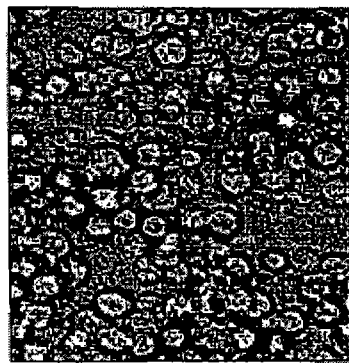


图7A

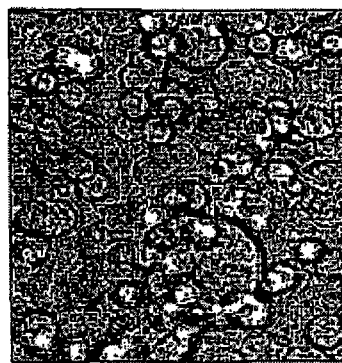


图7B

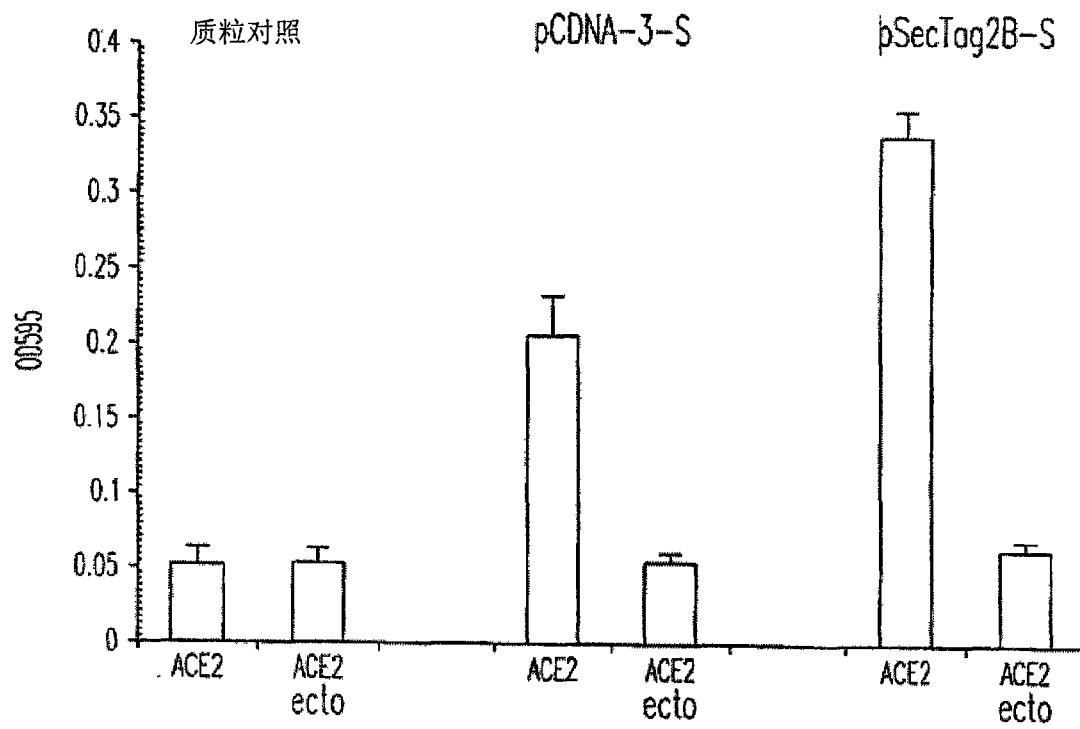


图7C

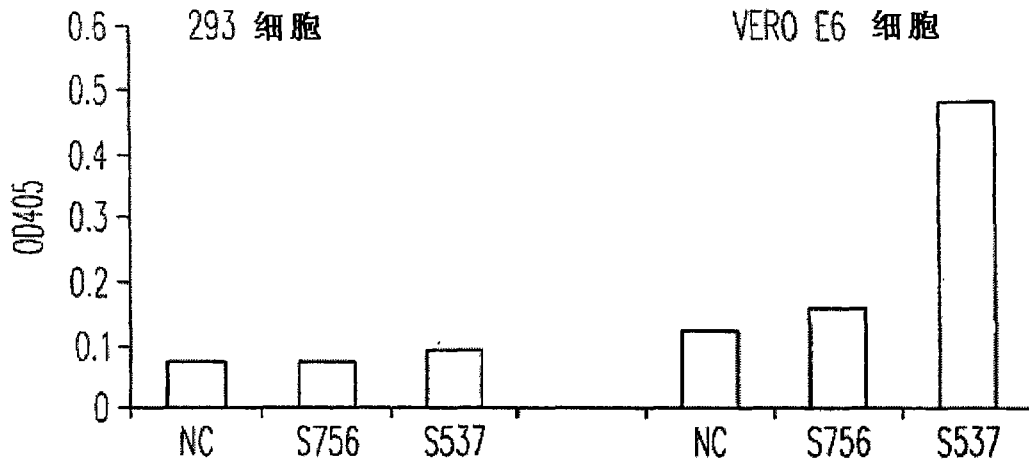


图8A

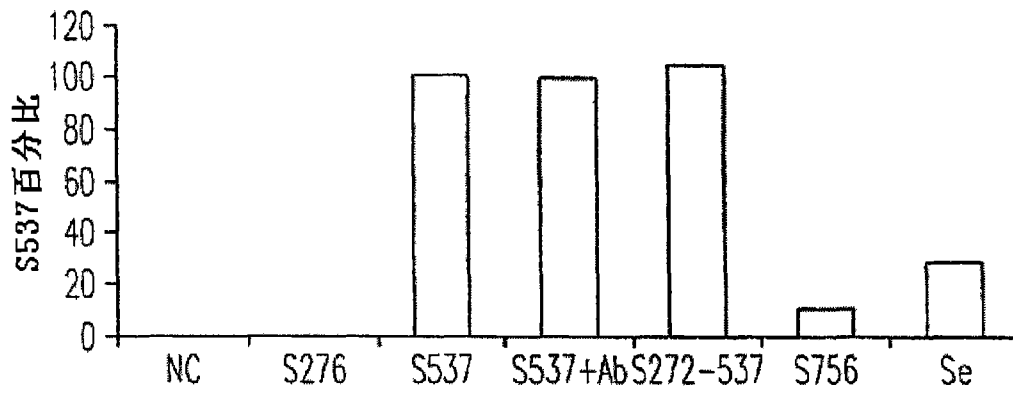


图8B

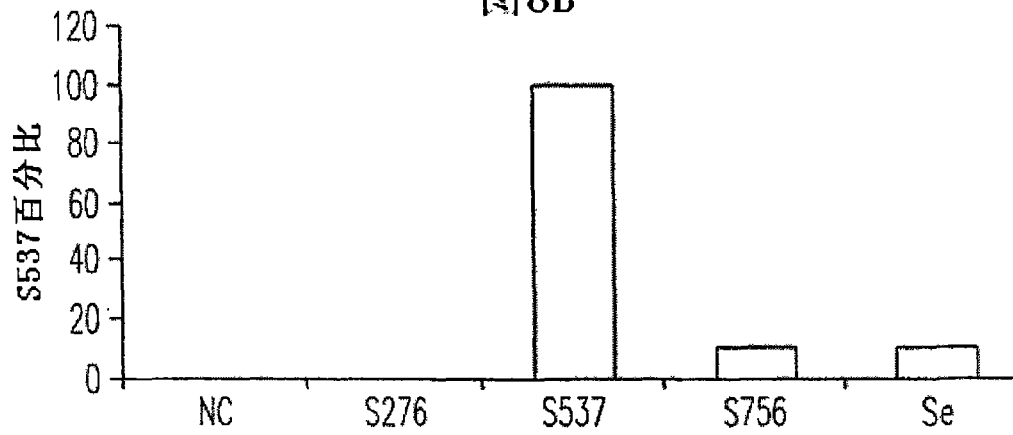


图8C

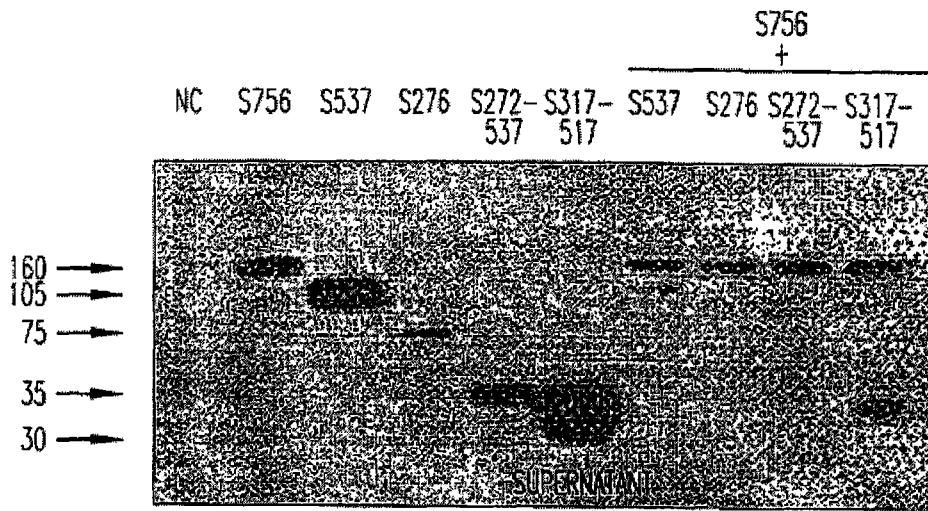


图9A

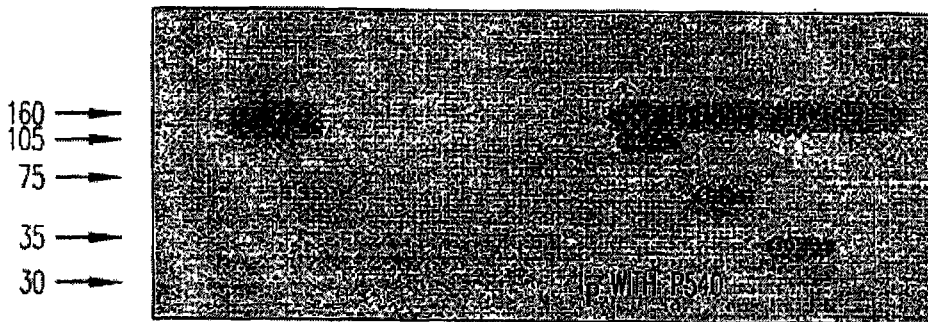


图9B

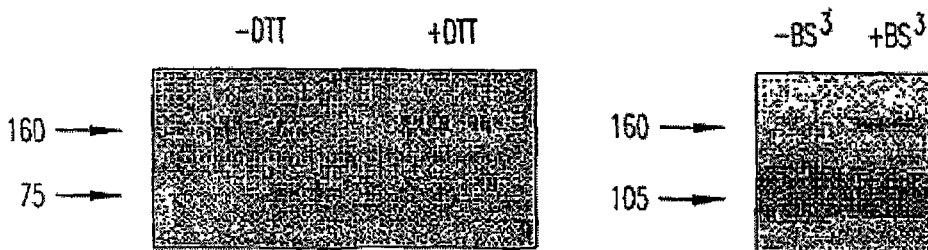


图9C

图9D

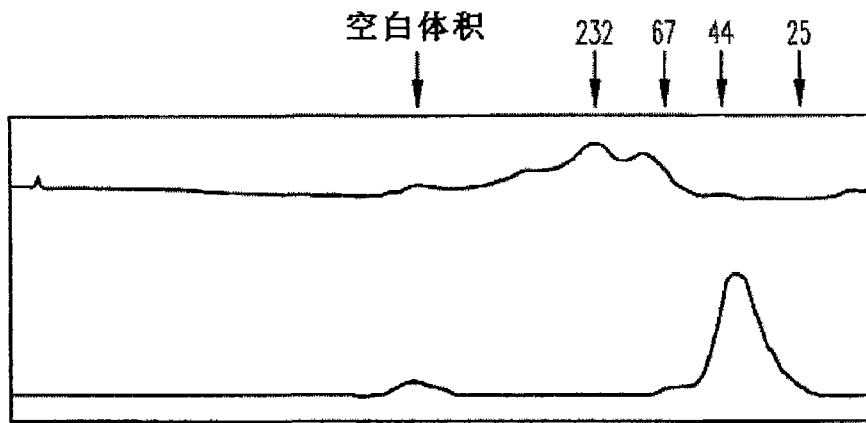


图10A

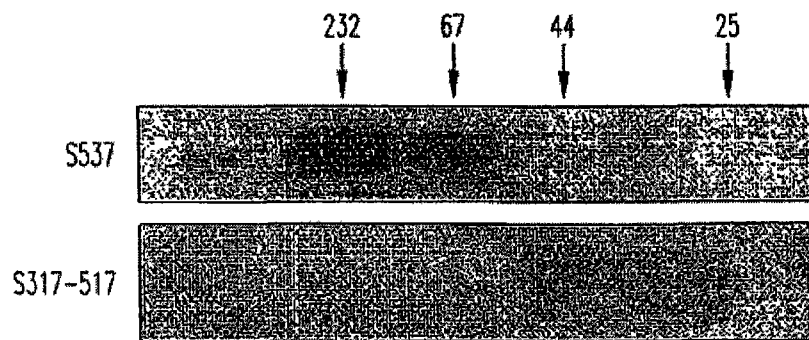


图10B

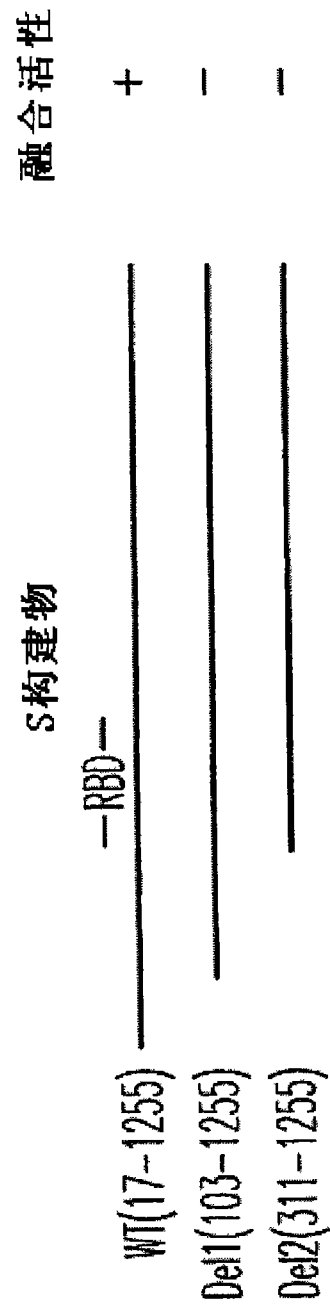


图 11A

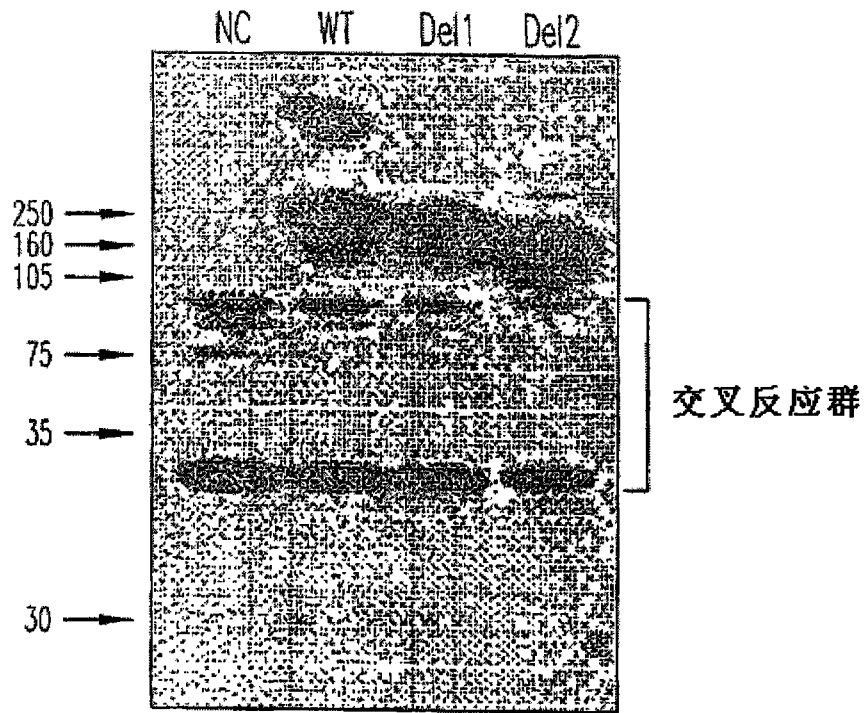


图11B

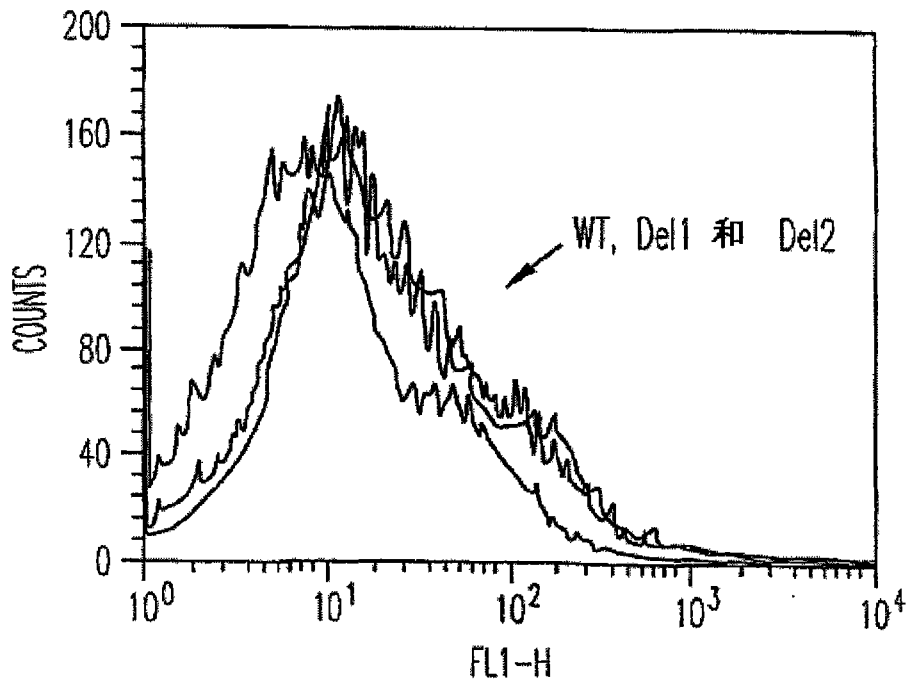


图11C

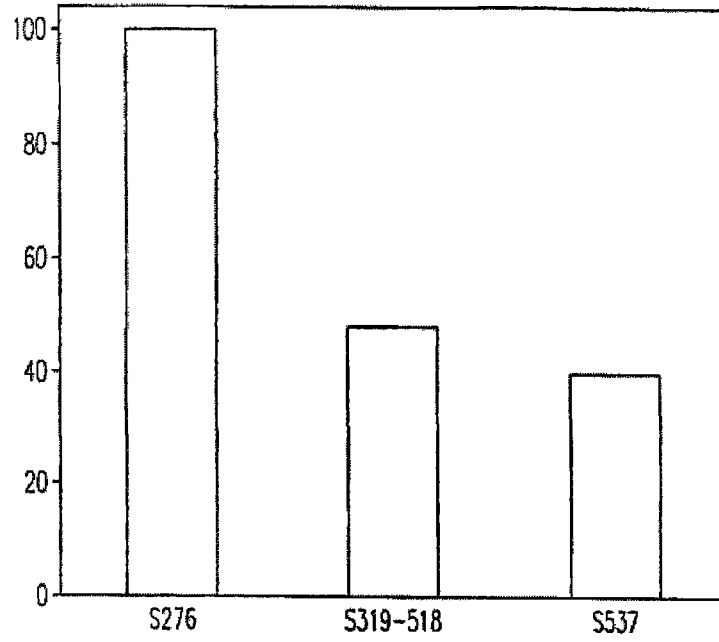


图12A

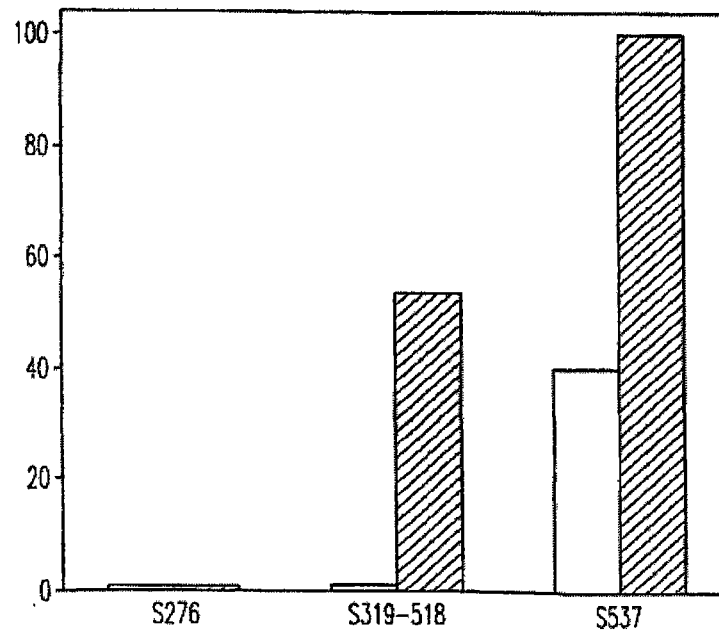


图12B

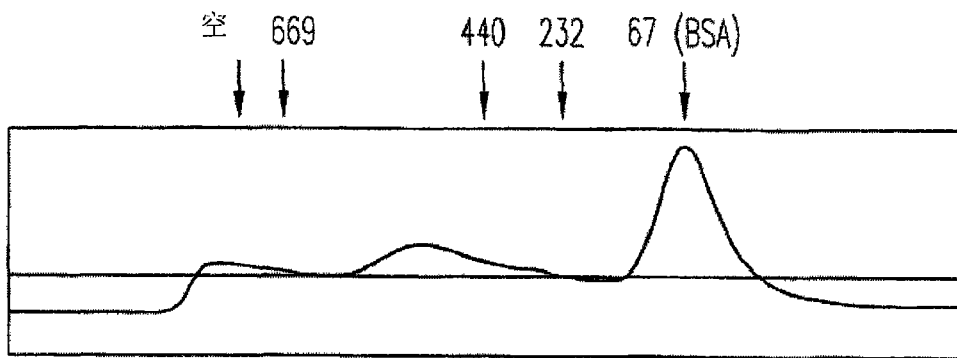


图13A

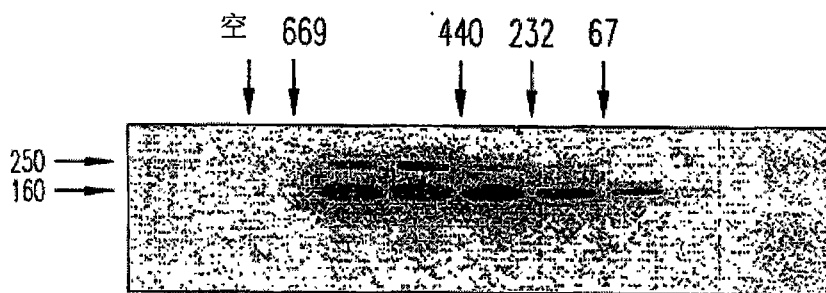


图13B

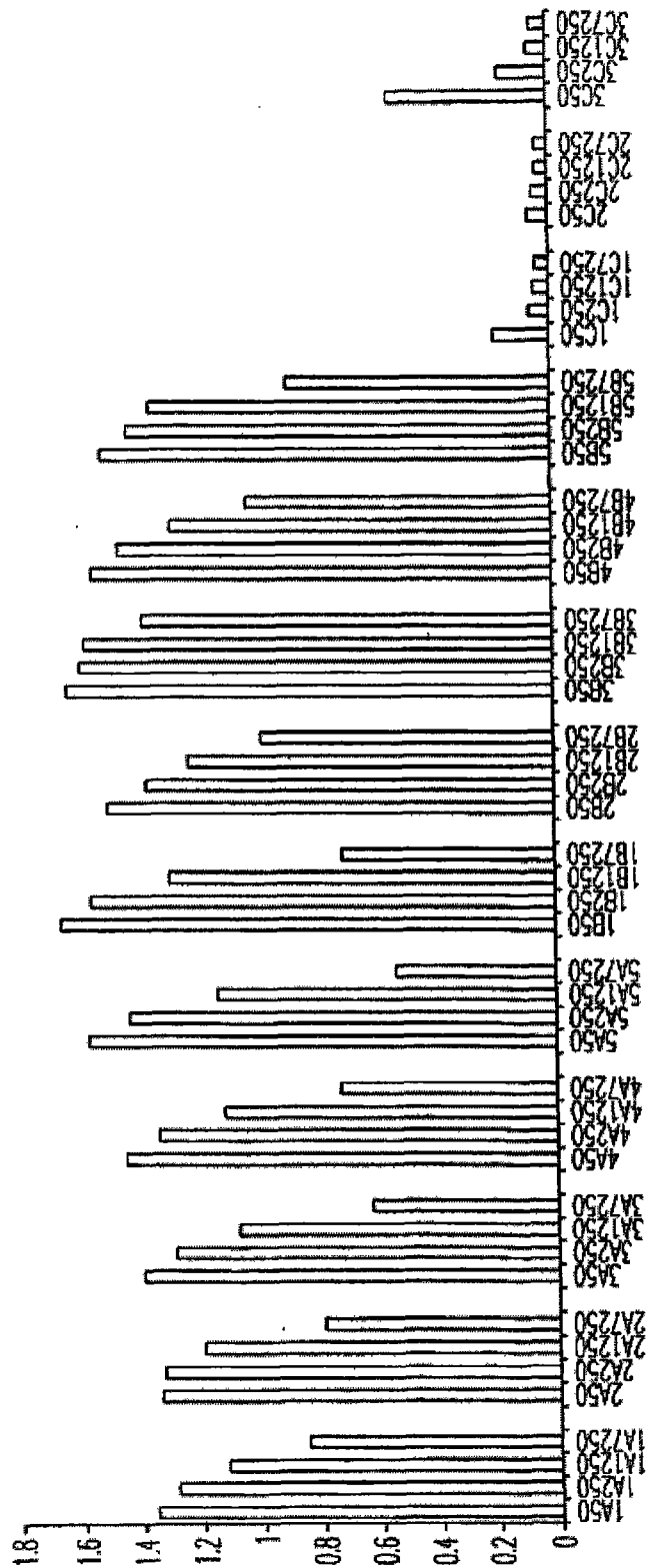


图 14A

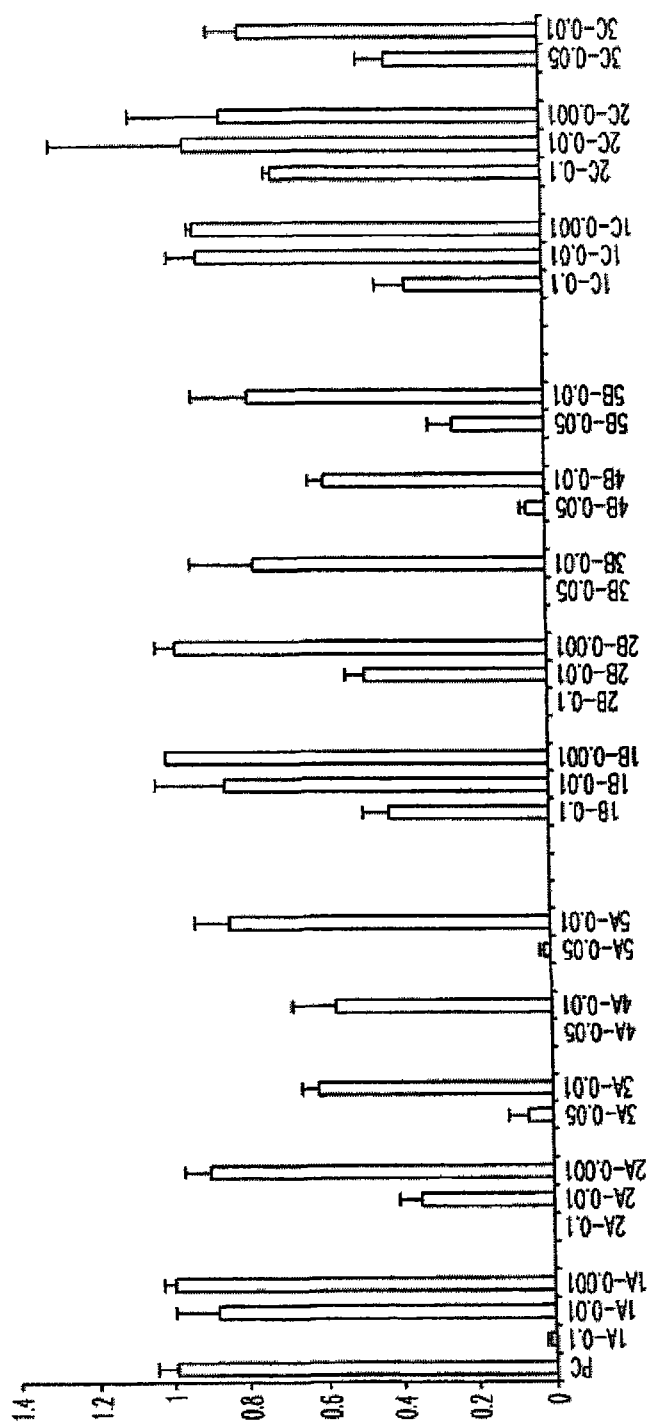


图 14B

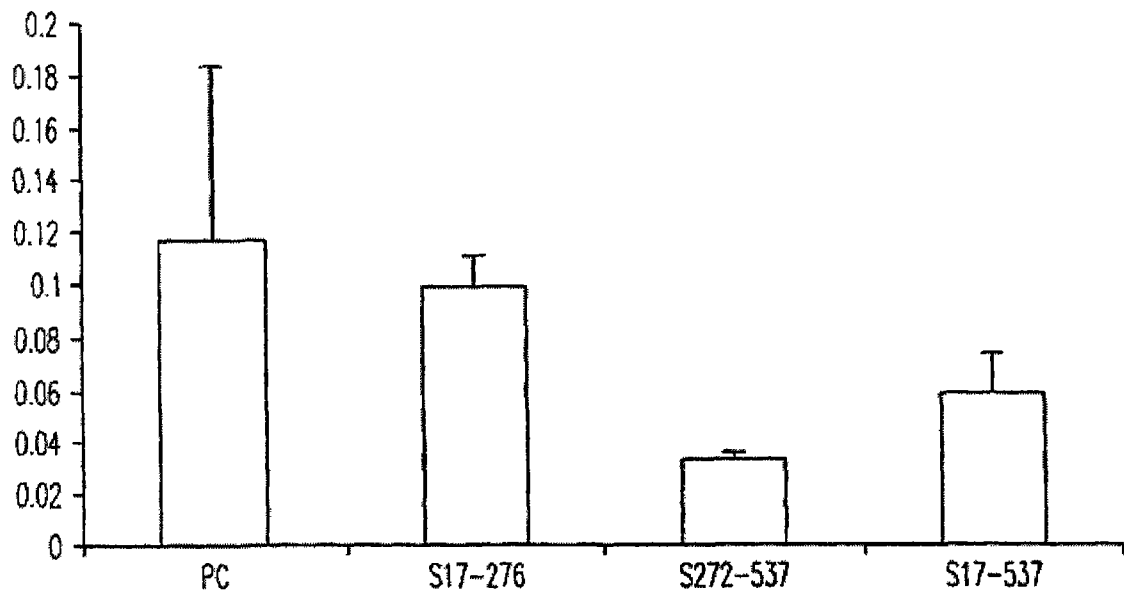


图 15