[19] 中华人民共和国国家知识产权局

[12] 发明专利说明书

专利号 ZL 99810706.9

[51] Int. Cl.
C12N 15/12 (2006.01)
C07K 14/705 (2006.01)
G01N 33/50 (2006.01)
A61K 38/17 (2006.01)
C07K 16/28 (2006.01)

[45] 授权公告日 2009 年 4 月 22 日

[11] 授权公告号 CN 100480387C

[22] 申请日 1999.7.22 [21] 申请号 99810706.9

[30] 优先权

[32] 1998. 7.22 [33] US [31] 60/093,843

[86] 国际申请 PCT/US1999/016676 1999.7.22

[87] 国际公布 WO2000/005375 英 2000.2.3

[85] 进入国家阶段日期 2001.3.8

[73] 专利权人 范德比尔特大学

地址 美国田纳西

[72] 发明人 C·G·赫勒奎斯特 傅长林

[56] 参考文献

WO9814603A1 1998.4.9

审查员 刘玉玲

[74] 专利代理机构 中国国际贸易促进委员会专利商标事务所

代理人 黄革生

权利要求书 7 页 说明书 82 页 附图 5 页

[54] 发明名称

GBS 毒素受体

[57] 摘要

本发明提供了新的 GBS 毒素受体，其制备和使用方法。还提供了 GBS 毒素受体多核苷酸和多肽，以及涉及这些多核苷酸和多肽的检测、筛选和治疗方法及药物组合物。

1. 包括 SEQ ID NO: 9 之核酸序列的分离的多核苷酸。

2. 包括与选自 SEQ ID NO: 1 的残基 61 至 1542、SEQ ID NO: 7 的残基 266 至 1870 和 SEQ ID NO: 3 的残基 87 至 1568 的核酸序列具有 100%相同性之核酸序列的分离的多核苷酸。

3. 一种分离的多核苷酸，包括(a)可与 SEQ ID NO: 1 的残基 61 至 1542、SEQ ID NO: 3 的残基 87 至 1568 或 SEQ ID NO: 7 的残基 266 至 1870 在高度严格条件下杂交，且(b)编码可与 GBS 毒素结合的多肽的核酸序列，其中所述高度严格条件如下：杂交条件为 50%甲酰胺、5×SSC 或 SSPE、1-5×Denhardt's 溶液、0.1-1%SDS、100-200μg 剪切的异源 DNA 或 tRNA、0-10%硫酸葡聚糖和 1×10^{5} 至 1×10^{7}cpm/ml 的比活性为 1×10^{8}cpm/μg 的变性探针，并于 42℃保温 6-36 小时；洗涤条件为 0.1×SSC、0.1-1%SDS 和 68℃，每 5-30 分钟换一次洗涤溶液。

4. 包括权利要求 3 之多核苷酸的载体。

5. 用权利要求 4 的载体转化的宿主细胞。

6. 生产哺乳动物 GBS 毒素受体或其片段的方法，包括在适当的培养基中培养权利要求 5 的宿主细胞。

7. 一种包括 GBS 毒素受体或其能够与 GBS 毒素 CM101 结合的区域的分离的多肽，其中该 GBS 毒素受体能够与 GBS 毒素 CM101 相结合，且由可在高度严格条件下与 SEQ ID NO: 1 的残基 61 至 1542、SEQ ID NO: 3 的残基 87 至 1568、或 SEQ ID NO: 7 的残基 266 至 1870 杂交的核酸序列编码，其中所述高度严格条件如下：杂交条件为 50%甲酰胺、5×SSC 或 SSPE、1-5×Denhardt's 溶液、0.1-1%SDS、100-200μg 剪切的异源 DNA 或 tRNA、0-10%硫酸葡聚糖和 1×10^{5} 至 1×10^{7}cpm/ml 的比活性为 1×10^{8}cpm/μg 的变性探针，并于 42℃保温 6-36 小时；洗涤条件为 0.1×SSC、0.1-1%SDS 和 68℃，每 5-30 分钟换一次洗涤溶液。

8. 权利要求 7 的多肽，其中受体与 SEQ ID NO: 2、SEQ ID NO: 4 或 SEQ ID NO: 8 的氨基酸序列有 100%相同性。

9. 权利要求 7 的多肽，其中受体与 SEQ ID NO: 4 或 SEQ ID NO: 8 之氨基酸序列中能够与 GBS 毒素 CM101 相结合的区域有 100%相同性。

10. 能够与 GBS 毒素 CM101 相结合的分离的多肽，其中包含与选自由 SEQ ID NO: 2、SEQ ID NO: 4 和 SEQ ID NO: 8 组成之组中的氨基酸序列只有不超过 1%氨基酸残基不同的氨基酸序列。

11. 权利要求 10 的分离的多肽，其中所述分离多肽的氨基酸序列中有一个氨基酸残基不同于选自所述组中的氨基酸序列。

12. 权利要求 10 的分离的多肽，其中所述不同的氨基酸残基是选自所述组中的氨基酸序列的相应残基的保守取代。

13. 包括有下列通式之氨基酸序列的分离的多肽:

AAl—AAn—AAm

其中:

AAl 不存在或者是 M,

AAn 是 SEQ ID NO: 8 的残基 2—42 的连续链，并且

AAm 是包括 AA43 至 AA536 的 494 个氨基酸的连续链，其中:

(1) AA43, AA47, AA51, AA52, AA57, AA58, AA65, AA66, AA72, AA85, AA87, AA93, AA94, AA96, AA115, AA116, AA122, AA123, AA125, AA134, AA143, AA173, AA174, AA178, AA185, AA186, AA189, AA190, AA196, AA200, AA204, AA206, AA207, AA220, AA253, AA260, AA276, AA277, AA280, AA283, AA287, AA294, AA295, AA298, AA300, AA301, AA312, AA324, AA326, AA360, AA365, AA373, AA374, AA379, AA396, AA403, AA407, AA418, AA480, AA483, AA486, AA491, AA494, AA502, AA528, AA529, AA532 和 AA536 是分别相当于下列（a）－（c）中任一所示的氨基酸残基:

(a) SEQ ID NO:8 的残基 43, 47, 51, 52, 57, 58, 65, 66, 72, 85, 87, 93, 94, 96, 115, 116, 122, 123, 125, 134, 143, 173, 174, 178, 185, 186, 189, 190, 196, 200, 204, 206, 207, 220, 253, 260, 276, 277, 280, 283, 287, 294, 295, 298, 300, 301, 312, 324, 326, 360, 365, 373, 374, 379, 396, 403, 407, 418, 480, 483, 486, 491, 494, 502, 528, 529, 532 和 536,

(b) SEQ ID NO:4 的残基 2, 6, 10, 11, 16, 17, 24, 25, 31, 44, 46, 52, 53, 55, 74, 75, 81, 82, 84, 93, 102, 132, 133, 137, 144, 145, 148, 149, 155, 159, 163, 165, 166, 179, 212, 219, 235, 236, 239, 242, 246, 253, 254, 257, 259, 260, 271, 283, 285, 319, 324, 332, 333, 338, 355, 362, 366, 377, 439, 442, 445, 450, 453, 461, 487, 488, 491 和 495, 或者

(c) 其保守取代;

(2) AA44-AA46, AA48-AA50, AA53-AA56, AA59-AA64, AA67-AA71, AA73-AA84, AA86, AA88-AA92, AA95, AA97-AA114, AA117-AA121, AA124, AA126-AA133, AA135-AA142, AA144-AA172, AA175-AA177, AA179-AA184, AA187-AA188, AA191-AA195, AA197-AA199, AA201-AA203, AA205, AA208-AA219, AA221-AA252, AA254-AA259, AA261-AA275, AA278-AA279, AA281-AA282, AA284-AA286, AA288-AA293, AA296-AA297, AA299, AA302-AA311, AA313-AA323, AA325, AA327-AA359, AA361-AA364, AA366-AA372, AA375-AA378, AA380-AA395, AA397-AA402, AA404-AA406, AA408-AA417, AA419-AA478, AA481-AA482, AA484-AA485, AA487-AA490, AA492-AA493, AA495-AA501, AA503-AA527, AA530-AA531 和 AA533-AA535 分别是:

(a) SEQ ID NO:8 的残基 44-46, 48-50, 53-56, 59-64, 67-71, 73-84, 86, 88-92, 95, 97-114, 117-121, 124, 126-133, 135-142, 144-172, 175-177, 179-184, 187-188, 191-195, 197-199, 201-203, 205, 208-219, 221-252, 254-259, 261-275, 278-279, 281-282, 284-286, 288-293, 296-297, 299, 302-311, 313-323, 325, 327-359, 361-364, 366-372, 375-378, 380-395, 397-402, 404-406, 408-417, 419-478, 481-482, 484-485, 487-490, 492-493, 495-501, 503-527, 530-531 和 533-535, 或者

(b) 其保守取代; 并且

(3) AA315-AA367 中的一个或多个任选不存在。

14. 识别权利要求 7-13 中任一项的多肽的抗体。

15. 包括与权利要求 7-13 中任一项的多肽相结合的 GBS 毒素的分离的复合体。

16. 形成复合体的方法, 包括:

使 GBS 毒素与权利要求 7-13 中任一项的多肽在允许 GBS 毒素与多

肽特异性结合的条件下接触，及

使复合体形成。

17. 纯化可与 GBS 毒素受体结合的化合物的方法，该方法包括:

提供权利要求 7-13 中任一项的多肽;

在允许化合物与多肽特异性结合的条件下，使所述多肽与含该化合物的样品接触，并

将结合的化合物与样品的其余物质分开。

18. 权利要求 7-13 中任一项的多肽在制备检测样品中是否存在 GBS 毒素的试剂中的用途。

19. 检测权利要求 7-13 中任一项的多肽存在的试剂在制备检测哺乳动物组织中的病理性脉管系统的检测试剂中的用途。

20. 鉴定可抑制 GBS 毒素与权利要求 7-13 中任一项的多肽结合的化合物的方法，其包括:

在含有 GBS 毒素的反应混合物中并在允许 GBS 毒素与权利要求 7-13 中任一项的多肽发生特异性结合的条件下，使试验化合物与包括所述 GBS 毒素受体或区域的多肽合并，并且

检测该化合物对 GBS 毒素与所述多肽之间的结合的抑制量。

21. GBS 毒素与权利要求 7-13 中任一项的多肽结合的抑制剂。

22. 鉴定可特异性结合权利要求 7-13 中任一项的多肽的化合物的方法，其包括:

在允许发生特异性结合的条件下，使试验化合物与权利要求 7-13 中任一项的多肽接触，并且

检测所述试验化合物与所述多肽间形成的复合体。

23. 检测试验嵌合化合物的细胞毒性的方法，该方法包括:

使在细胞表面上表达权利要求 7-13 中任一项的多肽的细胞，与包括偶联到所述 GBS 毒素上的细胞毒性剂的试验嵌合化合物接触，并且

检测细胞毒性信号。

24. 包括共价连接到可特异性结合权利要求 7-13 中任一项的多肽之分子上的细胞毒性剂的嵌合化合物。

25. 鉴定 GBS 毒素受体的抑制剂的方法，该方法包括:

在有或没有试验化合物存在下，在没有试验化合物存在时被保温的细胞即能够增殖或迁移的条件下，对试验细胞进行保温，其中试验细胞表达权利要求 7-13 中任一项的多肽；并且

将有试验化合物存在时保温的试验细胞的增殖或迁移与没有试验化合物存在时保温的试验细胞的增殖或迁移相比较，其中有试验化合物存在时增殖或迁移较少即指示试验化合物是 GBS 毒素受体的抑制剂。

26. 鉴定内皮细胞增殖或迁移的抑制剂的方法，该方法包括:

在有或没有试验化合物存在下，在没有试验化合物存在时被保温的细胞即能够增殖或迁移的条件下，对试验内皮细胞进行保温，其中试验细胞表达权利要求 7-13 中任一项的多肽；并且

将有试验化合物存在时保温的试验细胞的增殖或迁移与没有试验化合物存在时保温的试验细胞的增殖或迁移相比较，其中有试验化合物存在时增殖或迁移较少即指示试验化合物是内皮细胞增殖或迁移的抑制剂。

27. 鉴定用于治疗或预防以病理性血管生成或新血管形成为特征的医学状况的治疗化合物的方法，该方法包括:

在有或没有试验化合物存在下保温试验细胞，其中试验细胞表达权利要求 7-13 中任一项的多肽;

将有试验化合物存在时保温的试验细胞的增殖或迁移与没有试验化合物存在时保温的试验细胞的增殖或迁移相比较，其中有试验化合物存在时增殖或迁移较少即指示试验化合物是治疗或预防所述医学状况的候选治疗化合物。

28. 权利要求 27 的方法，其中医学状况是癌性肿瘤。

29. 权利要求 27 的方法，其中医学状况是重灌注损伤。

30. 权利要求 27 的方法，其中医学状况是伤口愈合期间的瘢痕形成。

31. 权利要求 27 的方法，其中医学状况是瘢痕疙瘩。

32. 权利要求 27 的方法，其中医学状况是慢性炎性疾病。

33. 权利要求 27 的方法，其中医学状况是神经损伤。

34. 鉴定可抑制 GBS 毒素与权利要求 7-13 中任一项的多肽结合的化合物的方法，其包括:

(a) 模拟并选择 GBS 毒素受体的最可能的构象;

(b) 设计基本上模拟该 GBS 毒素受体能量上最可能的三维结构的化学改性类似物;

(c) 化学合成该类似物，以及

(d) 评估该类似物的生物活性。

35. 鉴定可结合权利要求 7-13 中任一项的多肽的化合物的方法，其包括:

(a) 模拟并选择 GBS 毒素受体的最可能的构象，

(b) 推测多肽的最可能的结合结构域，

(c) 设计将与该多肽形成能量上最可能的复合体的化合物，

(d) 化学合成该化合物，并且

(e) 评估该化合物的生物活性。

36. GBS 毒素与权利要求 7-13 中任一项的多肽之间结合的抑制剂在制备可用于预防或治疗人新生儿的新生儿发作性疾病的药物中的用途。

37. 选自于下组中的分子在制备可用于抑制哺乳动物组织中病理性或缺氧驱使的内皮细胞增殖或迁移的药物中的用途:

当结合于权利要求 7-13 中任一项的多肽上时能够引发炎性反应的化合物;

包括偶联于可特异性结合权利要求 7-13 中任一项的多肽的化合物上的细胞毒性化合物的嵌合化合物;

权利要求 7-13 中任一项的多肽磷酸化的抑制剂; 以及

权利要求 7-13 中任一项的多肽活性的抑制剂。

38. 包括药学有效量的选自下列一组中的分子和药学可接受载体的药物组合物:

权利要求 7-13 中任一项的多肽;

权利要求 7-13 中任一项的多肽的抑制剂; 以及

包括偶联到可结合权利要求 7-13 中任一项的多肽的化合物上的细胞毒性剂的嵌合化合物。

39. 包括选自下列一组中的成分的试剂盒:

权利要求 7-13 中任一项的多肽;

用于检测权利要求 7-13 中任一项的多肽之存在的试剂; 以及

用于检测编码权利要求 7-13 中任一项的多肽之多核苷酸存在的试剂。

40. 权利要求 7-13 中任一项的多肽的抑制剂, 或 GBS 毒素与权利要求 7-13 中任一项的多肽之间结合的抑制剂, 在制备用于治疗以病理性或缺氧驱使之血管生成或新血管形成为特征的医学状况的药物中的用途。

41. 权利要求 7-13 中任一项的多肽在制备用于通过测定得自新生儿的样品中 GBS 毒素的有无诊断新生儿中早期发作性疾病的试剂中的用途。

42. 权利要求 7-13 中任一项的多肽的抑制剂, 或 GBS 毒素与权利要求 7-13 中任一项的多肽之间结合的抑制剂, 在制备用于治疗人新生儿的新生儿发作性疾病的药物中的用途。

GBS 毒素受体

技术领域

本发明提供与 GBS 毒素受体多核苷酸和多肽有关的组合物及方法。本发明涉及从细菌来源分离的多糖受体。

背景

B 组β溶血性链球菌(GBS)是普遍存在的微生物。除婴幼儿外，尚不知道 GBS 可引起人的什么有害感染。GBS 肺炎(也称为“早期发作性疾病”)与新生儿的高发病率和死亡率有关。

在 Carl G. Hellerqvist 博士及其同事(Vanderbilt 大学医学院，Nashville, Tennessee)进行的一系列研究工作中，已鉴定了一种多糖 GBS 毒素。确定该毒素为 GBS 肺炎的主要致病因子，并发现其可作治疗剂，通过抑制血管形成来对抗肿瘤(美国专利 5,010,062)。

另外，如美国专利 5,858,991 和国际专利 WO 98/32453 中所述，GBS 毒素可减小瘢痕形成并加速伤口愈合，同时减缓伤口相关肿瘤进展。

WO 98/32452 和 WO 98/32448 描述了使用 GBS 毒素作为治疗剂用于治疗慢性炎性疾病，如类风湿性关节炎和牛皮癣，并用于增强神经损伤的修复。

在本发明之前，尚没有鉴定出 GBS 毒素的受体。本发明人相信 GBS 毒素受体存在于上述状况中经受血管生成的组织的发展着的脉管系统细胞上，为此进行了一系列实验并完成了本发明。

发明概要

本发明第一次鉴定了 B 组β溶血性链球菌 GBS 毒素的新的受体(GBS 毒素受体)。本发明的一个方面提供包括 GBS 毒素受体的多肽或

其多肽片段。优选的具体实例包括哺乳动物 GBS 毒素受体。还提供了识别 GBS 毒素受体或其片段的抗体。可使用本发明多肽筛选可用于治疗或预防因病理性或缺氧造成的血管生成或新血管生成的状况，如肿瘤、慢性炎性疾病、伤口愈合期间的瘢痕生成、疤痕疙瘩、神经损伤及再灌注损伤。

本发明的另一方面是提供编码 GBS 毒素受体或其片段的多核苷酸，以及可与这些多核苷酸杂交的多核苷酸。优选的多核苷酸是长度至少 10 个碱基，并包括编码编码哺乳动物 GBS 毒素受体或其多肽片段的核酸序列或互补于这样的核酸序列。

本发明的第三个方面是包括与哺乳动物毒素受体或其片段结合之 GBS 毒素的复合体。还提供了形成这种复合体的方法。该方法包括在允许 GBS 毒素与多肽特异性结合并形成复合体的条件下，使 GBS 毒素与包括哺乳动物 GBS 毒素受体的多肽或其可与 GBS 毒素结合的片段接触。

本发明的再一个方面是纯化可与 GBS 毒素受体结合的化合物的方法。该方法包括提供含哺乳动物 GBS 毒素受体的多肽，或其可结合 GBS 毒素的片段，在允许化合物与多肽特异性结合的条件下使多肽与含有化合物的样品接触，并将结合的化合物与样品的残留物分离开。

本发明的再一个方面是确定样品中是否存在 GBS 毒素的方法。该方法包括在允许 GBS 毒素与 GBS 毒素受体特异性结合的条件下，使样品与包括哺乳动物 GBS 毒素受体的多肽或其可结合 GBS 毒素的片段接触，并确定是否已发生了 GBS 毒素的特异性结合。新生儿的样品中存在 GBS 毒素是早期发作性疾病的指征。

本发明的第六个方面是检测哺乳动物组织中病理性脉管系统的方法。可使用该方法检测或监测与血管生成或新血管形成有关的各种医学状况，如检测癌性肿瘤转移、或监测经受肿瘤治疗的哺乳动物体内的肿瘤界限。

本发明的再一个方面是提供鉴定用于治疗以病理性和/或缺氧促使的血管生成或新血管形成为特征的医学状况的候选药物的方法。一

个实例是鉴定可与哺乳动物 GBS 毒素受体特异性结合的化合物的方法。该方法包括在允许发生特异性结合的条件下，使试验化合物与哺乳动物 GBS 毒素受体或其可与 GBS 毒素结合的片段结合，然后检测试验化合物与多肽之间形成的复合体。另一个具体实例是确定试验嵌合化合物的细胞毒性的方法。该方法包括使表达哺乳动物 GBS 毒素受体或其可结合 GBS 毒素之片段的细胞，与包括偶联到 GBS 毒素上之细胞毒性剂的试验嵌合化合物接触，并检测毒性征记。再一个具体实例是鉴定 GBS 毒素受体抑制剂的方法，该方法包括在有及没有试验化合物存在下，并在没有试验化合物存在时培养的细胞能够增殖或迁移的条件下，保温表达 GBS 毒素受体或其片段的试验细胞，并比较有及没有试验化合物存在时保温的试验细胞的增殖或迁移，其中在有试验化合物存在时增殖或迁移更少即是试验化合物为 GBS 毒素受体抑制剂的指征。可用上述方法鉴定内皮细胞增殖或迁移抑制剂，其中当有试验化合物存在时试验细胞的增殖或迁移率较低，即是试验化合物为内皮细胞增殖或迁移抑制剂的指征。可以用上述方法鉴定用于治疗或预防以病理性血管生成或新血管形成为特征的医学状况的治疗化合物，其中当存在试验化合物时试验细胞的增殖或迁移率较低，即表明试验化合物可以作为用于治疗或预防所说的医学状况的候选治疗化合物。

本发明还提供鉴定可抑制 GBS 毒素与哺乳动物 GBS 毒素受体结合之化合物的方法。该方法包括模拟并选择哺乳动物 GBS 毒素受体的最可能的构象，设计实质上模拟该多肽的能量上最可能的三维结构的化学改性类似物，化学合成类似物，并估测类似物的生物活性。另外还提供鉴定可与哺乳动物 GBS 毒素受体结合之化合物的方法。该方法包括模拟并选择哺乳动物 GBS 毒素受体的最可能的构象，推测多肽的最可能的结合区，设计可与多肽形成能量上最可能之复合体的化合物，化学合成该化合物，并估计化合物的生物活性。

本发明的再一个方面涉及经投用 GBS 毒素与人 GBS 毒素受体结合的抑制剂，以预防或治疗新生儿的新生儿发作性疾病的方法。

本发明的再一个方面涉及抑制哺乳动物组织中病理性或缺氧引起

的内皮细胞增殖或迁移的方法。该方法包括使分子与组织中至少一个细胞的表面上存在的 GBS 毒素受体特异性结合，该分子选自下组：当与哺乳动物中的 GBS 毒素受体结合后即可引发炎性反应的化合物，包括与可特异性结合 GBS 毒素受体的化合物偶联之细胞毒性化合物的嵌合化合物，GBS 毒素受体磷酸化抑制剂、GBS 毒素受体活性抑制剂。

本发明还提供用于人或动物的治疗方法或用于治疗以病理性或缺氧驱使的血管生成或新血管生成为特征之医学状况的药物的制备中的 GBS 毒素受体或其片段、GBS 毒素受体抑制剂、GBS 毒素与 GBS 毒素受体结合的抑制剂。还提供用于治疗人或动物体之方法中的，包括与可结合 GBS 毒素受体的化合物偶联之细胞毒性剂的嵌合化合物。

本发明还提供包括 GBS 毒素受体抑制剂和/或嵌合化合物，以及药用载体的药物组合物，其中所述嵌合化合物包括与可结合 GBS 毒素受体的化合物偶联的细胞毒性剂。

本发明还提供包括 GBS 毒素受体或片段和/或用于检测 GBS 毒素受体或其多肽片段之存在的试剂，或检测其多核苷酸编码序列之存在的试剂盒。

附图简要说明

图 1 显示合理设计药物的方法步骤。

图 2A 和 2B 显示使用实施例 4 中所述的抗体 Pab55，分别对癌性和正常卵巢组织中的 GBS 毒素受体表达进行免疫组化分析的结果。

图 3A 和 3B 显示使用实施例 4 中所述的抗体 Pab57，分别对癌性和正常人卵巢组织中的 GBS 毒素受体表达进行免疫组化分析的结果。

图 4A—4C 显示嵌合化合物向实施例 6 中所述之癌性组织中表达的 GBS 毒素受体的定向送递。

发明的详细描述

一般说来，本文所使用的所有技术和科学术语均具有本发明所属技术领域普通技术人员所通常理解的意义。本文使用的命名法和细胞

培养、分子遗传学、核酸化学及下述杂交等方面的实验室方法都是本领域技术人员熟知的。用于核酸重组、多核苷酸合成、微生物培养及转化(如电穿孔、脂染法)的方法均为常规标准方法。生产商提供的酶促反应和纯化步骤都是按照生产商提供的说明书进行的。一般都按照本领域及各种常用参考文献(如Sambrook et al., Molecular Cloning: A laboratory Manual, 2nd Ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)中所述的常规方法完成各种技术和操作步骤。本文所使用的术语及下述的分析化学、有机合成化学及药物配制中的实验室方法均为本领域已知和常用的。可使用标准技术进行化学合成、化学分析、药物配制与送递，以及治疗病人。除另外指出者外，本文通编使用的术语均具有下述含义:

“GBS 毒素受体”是指能够与 B 组β溶血性链球菌，如 CM101 的毒素(GBS 毒素)结合的蛋白质分子。自然界中 GBS 毒素受体通常见于细胞表面上。可以用本领域已知的实验室技术生产膜结合和可溶性的重组 GBS 毒素受体。

术语“分离的多核苷酸”是指已经受过操作的多核苷酸，从而该分离的多核苷酸不再与其正常与之天然关联的染色体或细胞相联系。例如“分离的多核苷酸”是基因组、重组或合成来源的多核苷酸,或它们的某些组合。

术语“分离的蛋白质”是指不再与正常蛋白质天然与之联系的细胞相联系的蛋白质。例如(1) 至少没有某些同样来源之其他蛋白质的蛋白质; (2)由来自不同种的细胞表达的蛋白质; (3)非天然存在的蛋白质，(4)由 cDNA、重组 DNA 或合成来源或其组合方式产生的蛋白质。

术语“多肽”作为上位概念术语，是指天然蛋白质、片段或多肽序列的类似物。这里，天然蛋白质、片段及类似物均为多肽属的种。

术语“天然存在的”是指自然界发现的。例如，存在于天然生物体(包括病毒)中的，并且未在实验室中被人为修饰的多肽或多核苷酸序列即为天然存在的。

术语“可操作地连接的”是指其中所述的成分处于一种允许它们

以预期方式发挥功能的邻接关系。控制序列“可操作地连接到”编码序列上，是指两者的连接方式使得编码序列的表达可在与控制序列相配合的条件下实现。

术语“控制序列”是指对其所连接的编码序列的表达起必要作用的多核苷酸序列。这些控制序列的性质根据宿主生物体而有所不同；在原核生物中，这样的控制序列一般包括启动子、核糖体结合位点及转录终止序列；在真核生物中，这样的控制序列一般包括启动子和转录终止序列。术语“控制序列”倾向于至少包括表达所必要的所有成分，并可包括其存在带来益处的附加序列，例如前导序列和融合对象序列。

术语“多核苷酸”是指长度至少为 10 个碱基的聚合形式的核苷酸，可以是核糖核苷酸或脱氧核糖核苷酸或其修饰形式，且包括单链和双链形式的 DNA。

术语“寡核苷酸”是指通过天然存在的及非天然存在的寡核苷酸键连接在一起的天然存在的和修饰的核苷酸。寡核苷酸通常是长度为 200 个碱基或更短的多核苷酸。优选的寡核苷酸长度为 10 至 60 个碱基，最好为 15 至 35 个碱基。寡核苷酸通常为单链的，例如用作探针；但也可能是双链的，例如用于构建基因突变体。本发明的寡核苷酸可以是有义或反义寡核苷酸。术语“天然存在的核苷酸”是指脱氧核糖核苷酸和核糖核苷酸。“修饰的核苷酸”包括带有修饰的或取代的糖基等的核苷酸。术语“核苷酸键”包括寡核苷酸键如硫代磷酸酯、二硫代磷酸酯、硒代磷酸酯(phosphoroselenoate)、二硒代磷酸酯(phosphorodiselenoate)、氨苯基硫代磷酸酯(phosphoroanilothioate)、氨苯基磷酸酯(phosphoraniladate)、氨基磷酸酯等。需要的话，寡核苷酸可包括检测标记。

“互补”或“互补体”是指当腺嘌呤出现于第一核苷酸序列中时，“互补”序列中就出现胸腺嘧啶或尿嘧啶，反之亦然，并且当鸟嘌呤出现于第一核酸序列中时，“互补”序列中即可见有胞嘧啶，反之亦然。

术语“序列相同性”是指两个核酸序列之间的碱基匹配比例，或两个氨基酸序列间的氨基酸匹配比例，即两个序列间的相同性程度。当序列相同性以百分数表示时，例如 50%,即代表与某些其他序列相比，GBS 毒素受体序列的总序列中准确匹配的比例。可使用各种计算机比对程序确定序列相同性。在其最简单形式中，%相同性是将两个核酸序列间或两个氨基酸序列间的准确匹配数除以参考序列中的核苷酸或氨基酸总数而计算出来的。例如，长度为 400 个氨基酸的序列间有 300 个匹配，则序列就有 75%相同性。当将核糖核苷酸序列与脱氧核糖核苷酸序列比较时，认为尿嘧啶与胸腺嘧啶是同样的。

当用于多核苷酸时，术语“实质上相同”是指当最佳比对时，如使用程序 BLAST (Altschul et al., J. Mol. Biol. 215:403－410, 1990)分析时，两个核酸序列同有至少约 85%，较好至少约 90%，最好约 95%或更大的序列相同性。当使用计算机比对程序时，可在两序列中任何一条上造成缺口以达到最大匹配；通常使用 15 个碱基或更小的缺口，较好是 6 个碱基或更小的缺口，最好是 2 个碱基或更小的缺口。当使用寡核苷酸作为探针或用于治疗时，靶核酸和寡核苷酸之间的序列相同性一般是 20 个可能的碱基配对中有不少于 17 个靶碱基匹配(85%)，较好为 9/10 (90%)，最好为 19/20 (95%)个靶碱基匹配。

优选不相同的碱基为该氨基酸位置上编码同一氨基酸之简并密码子的一部分。或者，不相同的碱基是编码该氨基酸位置上保守氨基酸取代之简并密码子的一部分。

当用于多肽时，术语“实质上相同”是指两个肽序列当按 BLAST 计算机程序最优比对时，两个序列至少有约 80%，较好有至少约 86%，更好有至少约 95%，最好有至少约 99%，甚至为 100%的序列相同性。可允许两匹配序列中任何一条序列上存在缺口以利于最大匹配；缺口长度较好为 5 个氨基酸以下，最好为 2 个氨基酸以下。不相同的残基位置最好是因保守氨基酸取代造成的不同。保守氨基酸取代是指具有相似侧链的残基间的可交换性。例如，一组具有脂族侧链的氨基酸是甘氨酸、丙氨酸、缬氨酸、亮氨酸和异亮氨酸；一组具有脂族羟基侧

链的氨基酸是丝氨酸和苏氨酸；一组具有含酰胺侧链的氨基酸是天冬酰胺和谷氨酰胺；一组具有芳族侧链的氨基酸是苯丙氨酸、酪氨酸和色氨酸；一组具有含硫侧链的氨基酸是半胱氨酸和蛋氨酸，一组具有碱性侧链的氨基酸是赖氨酸、精氨酸和组氨酸。优选的组内保守氨基酸取代是：缬氨酸—亮氨酸—异亮氨酸、苯丙氨酸—酪氨酸、赖氨酸—精氨酸、丙氨酸—缬氨酸、谷氨酸—天冬氨酸，和天冬酰胺—谷氨酰胺。

术语“可在高度严格条件下杂交”是指能够在只有彼此间实质性相同性大于 95%的核酸序列才能杂交的条件下发生特异性结合。这些条件是本领域已知的并在本文中讨论。

术语“简并密码子”是指按照遗传密码编码所需氨基酸的任何核苷酸密码子三联体。可根据宿主生物体如大肠杆菌中已知的优选密码子用法来选择密码子。

术语“多肽片段”是指具有氨基末端和/或羧基末端缺失，但残留的氨基酸序列相同于例如从全长度 DNA 序列推断的天然存在之序列中的相应位置的多肽。片段一般至少为 3 个氨基酸长，较好为 5—10 个氨基酸长，更好为 10—50 个氨基酸长，最好为 50 个以上氨基酸长，并且包括至少一个 GBS 毒素受体的胞外结构域。最优选的片段包括 GBS 毒素受体的整个胞外结构域，还包括足以维持多肽片段在细胞表面、脂质膜、脂质体、微胶粒及其他亲脂结构上呈功能性立体化学构象的跨膜部分和胞内结构域。

术语“免疫学反应性”是指具有抗原性质或能够被可特异地结合 GBS 毒素受体的抗体特异性结合。当一种物质在本领域已知有利于产生可识别并结合特定抗原的抗体的条件下投用于动物时，如果能够产生单克隆或多克隆抗体，则该物质即具有抗原性质。

“异源多肽”是不同于特定细胞正常产生之多肽的多肽。例如，由正常情况下并不产生 GBS 毒素受体多肽或其片段的细胞重组产生的这种 GBS 毒素受体多肽或其片段即为异源多肽。如果连接到 GBS 毒素受体多肽或其片段上的第二种多肽并不是天然与之连接的，则这样的

连接多肽也是异源多肽。

术语“标记”或“标记的”是指例如掺入放射标记的氨基酸或连接到可用标记的抗生物素蛋白(例如含有可用光学或比色方法检测的荧光标记或酶促活性的链霉抗生物素蛋白)检测的带生物素基部分的多肽上，以掺入可检测标记。可使用本领域已知的各种标记多肽和糖蛋白的方法。多肽的标记物包括但不只限于：放射性同位素(如 ^{3}H、^{14}C、^{35}S、^{125}I、^{131}I)、荧光标记(如 FITC、罗丹明、镧系磷光体)、酶促标记(如辣根过氧化物酶、β 半乳糖苷酶、荧光素酶、碱性磷酸酶)、化学发光团、生物素基、由第二报导分子识别的预定的多肽表位(如亮氨酸拉链对序列、第二抗体结合位点、金属结合区、表位标签)。在某些具体实施中，由不同长度的间隔臂连接标记物，以减少可能的空间位阻。

术语“化合物”是指肽、拟肽(peptidomimetic)、有机或其他化学分子，以及核酸分子或其化学衍生物。化合物可以是一本发明多核苷酸或多肽或者可与其相互作用。

除特别指出者外，单数形式的“一个”、“一种”和“该”也包括多数事物。

下列表 1 中总结了本文所述核酸和氨基酸序列的 SEQ ID NOs:

表 1

核酸和氨基酸序列

SEQ ID NO:	序列类型	说 明
SEQ ID NO: 1	核酸	部分人 GBS 毒素受体（HP55）
SEQ ID NO: 2	氨基酸	部分人 GBS 毒素受体（HP55）
SEQ ID NO: 3	核酸	绵羊 GBS 毒素受体（SP55）
SEQ ID NO: 4	氨基酸	绵羊 GBS 毒素受体（SP55）
SEQ ID NO: 5	核酸	引物
SEQ ID NO: 6	核酸	引物
SEQ ID NO: 7	核酸	全长度人 GBS 毒素受体（HP59）
SEQ ID NO: 8	氨基酸	全长度人 GBS 毒素受体（HP59）
SEQ ID NO: 9	核酸	人/绵羊共有 GBS 毒素受体编码区（碱基码为 a、c、g、t、m、r、w、s、y、k）
SEQ ID NO: 10	氨基酸	人/绵羊共有 GBS 毒素受体编码区（SEQ ID NO: 9 的翻译）
SEQ ID NO: 11	核酸	人/绵羊共有 GBS 毒素受体编码区（碱基码为 a、c、g、t、n）
SEQ ID NO: 12	氨基酸	人/绵羊共有 GBS 毒素受体编码区（SEQ ID NO: 11 的翻译）

本文给出的题头描述所讨论的一般性课题，并且不排除其他部分中讨论的信息。有关信息、方法、组合物及其他方面可应用于不只一个发明的具体实施方案，并且可以相互组合。

引言

GBS 可与经受病理性的、缺氧驱使的，和胚胎学的血管生成或新血管形成的组织结合。本发明人至少鉴定了如本文讨论的两种哺乳动物 GBS 毒素受体。实施例 1 和 2 描述了某些 GBS 毒素受体的克隆和特征。本发明人已将 GBS 毒素受体分类为带有七个跨膜区的整合膜蛋白。表 7 中显示了预测的各片段。该蛋白质有几个 cAMP 依赖性激酶、蛋白激酶 C (PKC)和酪蛋白激酶Ⅱ (CK2)的推测的磷酸化位点。通常，这样的整合膜蛋白在结合分子(如配体或细胞外信使)后，经受一次有利于在如上文讨论的磷酸化位点磷酸化的构象改变。蛋白质在这些位点的磷酸化可能触发信号转导级联，从而常常使已暴露于结合分子的细胞发生增殖或其他核应答。血管生成或新血管形成包括有内皮细胞的增殖和迁移。如在实施例 4 和 5 中详细讨论的，GBS 毒素受体表达与涉及病理性的、缺氧驱使的、和胚胎学的血管生成或新血管形成的医学状况相关联。GBS 毒素受体多肽可用于各种不同的目的，其中包括筛选可抑制由 GBS 毒素受体介导的内皮细胞增殖和/或迁移的化合物、筛选可结合并破坏表达 GBS 毒素受体的细胞的细胞毒性嵌合化合物。另外，例如可使用 GBS 毒素受体多核苷酸设计反义多核苷酸，用以阻断编码 GBS 毒素受体的信使 RNA 的翻译，还可用于其它目的。

多核苷酸

本发明的一个方面涉及至少 10 个碱基长的分离的多核苷酸，其编码 GBS 毒素受体或其片段或者互补于其编码核酸序列。GBS 毒素受体是哺乳动物 GBS 毒素受体，尤其是绵羊、牛或猫 GBS 毒素受体，最好是人 GBS 毒素受体。分离的多核苷酸可以是天然存在的或非天然存在的。分离的多核苷酸可包括 DNA 序列或 RNA 序列(其中每个 T 均被 U 取

代)。为了确定相同性百分比，T 被视为等同于 U。多核苷酸最好包括绵羊、牛、猫或人 GBS 毒素受体的等位基因，并可包括其他哺乳动物 GBS 毒素受体的等位基因。可以按照下述方法，使用衍生于 SEQ ID NO: 1、3、7、9 和 11 的多核苷酸分离这些多核苷酸。

本发明的多核苷酸、寡核苷酸及片段可在与非特异性核酸发生最小量的可检测结合的杂交和洗涤条件下，选择性地与核酸链杂交。多核苷酸可在高度严格条件下与具有 SEQ ID NO: 1 或 SEQ ID NO: 3 的至少 20 个连续核苷酸，较好至少 30 个连续核苷酸之核酸序列的核酸分子，甚至含有 SEQ ID NO: 1、3、7、9 或 11 的核酸序列或它们的互补序列的核酸分子杂交。这样的多核苷酸可用于完成选择性的高度严格杂交，并特别适用于以聚合酶链反应(PCR)法扩增核酸以确定样品中是否存在 GBS 毒素受体，分离天然存在的编码 GBS 毒素受体的核酸(参见实施例 3)，以及作为反义分子以阻断 GBS 毒素受体 mRNA 的翻译。特别优选的是在高度严格条件下与具有包括 SEQ ID NO: 7 之核苷酸 266－1870(除起始密码子外的推测的全长度人 GBS 毒素受体编码区)、SEQ ID NO: 7 之核苷酸 266－1870(包括起始密码子的，推测的全长度人 GBS 毒素受体编码区)、SEQ ID NO: 1 之核苷酸 61－1542(除起始密码子外的人 GBS 毒素受体部分编码区)、SEQ ID NO: 1 之核苷酸 58－1542(包括起始密码子的人 GBS 毒素受体部分编码区)、SEQ ID NO: 3 之核苷酸 87－1568(除起始密码子外的绵羊 GBS 毒素受体编码区)、SEQ ID NO: 3 之核苷酸 84－1568(包括起始密码子的绵羊 GBS 毒素受体编码区)的核酸序列或其互补核酸序列的核酸序列的核酸分子可杂交的多核苷酸。

多核苷酸与 SEQ ID NO: 1、3 或 7 中相应区域的核酸序列或 SEQ ID NO: 1、3 或 7 相应区域的互补序列具有大约 85%至 100%，较好大于约 87%，最好大于约 95%，特别是大约 99%至 100%的相同性。特别优选的多核苷酸包括 SEQ ID NO: 7 中核苷酸 266－1870、或 SEQ ID NO: 3 中核苷酸 87－1568、SEQ ID NO: 9、SEQ ID NO: 11 的核酸序列或其互补核酸序列。

优选多核苷酸包括编码与 GBS 毒素受体片段的氨基酸序列有大约85—100%、较好大于 86%、更好大于 95%，最好为 99—100%相同性之多肽的核酸序列或者与这样的核酸序列互补的核酸序列。优选片段可结合 GBS 毒素。优选的片段包括 SEQ ID NO: 2 中残基 1—495 或 SEQ ID NO:8 中残基 1—536 的全部或一部分。特别优选的是包括编码与 SEQ ID NO: 4 之残基 1—495、SEQ ID NO: 2 之残基 1—495、SEQ ID NO: 8 之残基 1—536 的氨基酸序列有 100%相同性之多肽的核酸序列的多核苷酸。

可按本文所述方法从产生 GBS 毒素受体的不同种的各种组织来源或细胞培养物，例如肿瘤内皮细胞、类风湿性关节炎的滑膜组织，或丧失或限制血流的缺氧组织如再灌注损伤或受伤组织中分离编码天然存在之 GBS 毒素受体的多核苷酸。可以利用探针以杂交方法，或利用寡核苷酸的聚合酶链反应方法，以及借助本领域已知的其他分子生物学技术分离这样的多核苷酸。所利用的探针和寡核苷酸一般包括 SEQ ID NO: 1、3、7、9 或 11，特别是 SEQ ID NO: 1、3 或 7 所示序列的各个区域，或者编码 SEQ ID NO: 2、4、8、10 或 12，特别是 SEQ ID NO: 2、4 或 8 所示序列的各个区域。

可以通过同源蛋白质的氨基酸序列，确定可用于克隆编码各种生物体之 GBS 毒素受体的基因的多核苷酸(见表 4)。例如可以定位保守区，以合成用作杂交或核酸扩增之探针的寡核苷酸或简并寡核苷酸(详见下文和实施例 1)。可以改变严格度以达到选择性杂交条件，实现相互间有小于 95%相同性的核酸序列能彼此杂交。这些条件是本领域已知的，在本文中将进行讨论并提供实例。一般说来，本发明的多核苷酸、寡核苷酸和片段与感兴趣的核酸序列间的核酸序列相同性至少约为 85%，最好有渐增的相同性，即至少约为 90%、95%、99%和 100%。

如本领域所知,可在高度严格洗涤条件和相应的杂交条件下使用多核苷酸作为探针。本领域中,小的多核苷酸，例如 200 碱基或更短的多核苷酸常常被称为寡核苷酸。使用多核苷酸作为探针以检测同一或相关核酸序列的技术是本领域已知的(例如参见Sambrook等人的上述

文献，特别是第 11 章,该文献全文引入作为参考)。通常可由长度为 10—200 个碱基的多核苷酸制得探针。较好从长度为 10—60 个碱基，最好 12—40 个碱基的多核苷酸制得探针。可以按照 Pearson 和 Lipman (Proc. Natl. Acad. Sci. USA 85: 2444—2448, 1988)的方法，利用可显示相比较之序列间的相同性程度的核酸同源性计算机程序如 FASTA，基于序列比较得到的结果设计特异性探针。探针的大小取决于其与之杂交的基因的相应区域。探针的大小随着与不希望的核酸序列间的同源性程度增加而增加。可使用长度为 10—50 个核苷酸，较好 50 个以上核苷酸，最好 100 个以上核苷酸的探针。最好使用由 GBS 毒素受体的整个编码区制得的探针。为了减少假阳性的数目，最好使用两个探针鉴定出在杂交和洗涤条件下可与两个探针结合的克隆。可以按照生产商提供的使用说明书，在 Applied BioSystems 寡核苷酸合成仪上合成寡核苷酸。

通常按照常规杂交方法进行杂交和洗涤。筛选噬斑复制物的典型杂交条件(Benton and Davis, Science 196: 180, 1978)可以是: 50%甲酰胺、5×SSC (NaCl、柠檬酸钠)或 SSPE (NaCl、磷酸钠、EDTA)、1—5×Denhardt's 溶液、0.1—1%SDS、100—200μg 剪切的异源 DNA 或 tRNA、0—10%硫酸葡聚糖、1×10^5—1×10^7cpm/ml 变性探针(比活性约为 1×10^8cpm/μg)，并于 42℃保温大约 6—36 小时。除了不包括探针并且保温时间通常缩短外，预杂交的其他条件基本上与杂交条件相同。洗涤条件一般是 1—3×SSC、0.1—1%SDS、42—70℃，每约 5—30 分钟换一次洗涤溶液。可以按这种方式得到包括等位基因序列在内的同族细菌序列。对高度严格性杂交条件来说，可以改变各种参数，例如提高与标记探针的保温温度，以增加杂交的严格性。为使实验设计有更大的灵活性，最好可以在较低温度下如室温下与探针杂交，并可通过改变盐浓度和洗涤溶液的温度来修饰严格性。为达到高度严格条件，可使用大于或等于 42℃的洗涤温度(例如 68℃)，并使用盐浓度小于 3×SSC 例如 0.1×SSC 的洗涤缓冲液。在某些情况下，可使用 TMACL(特别是对于富含 G—C 碱基对的多核苷酸)，以减少非特异性结

合。可以使用低严格性洗涤条件，以使其有较低相同性的多核苷酸或长度小于60碱基对的多核苷酸杂交。对于低严格性洗涤，可在盐浓度大于或等于 2×SSC 的洗涤缓冲液中使用小于或等于 42℃的温度洗涤。

本发明包括扩增靶核酸的方法，例如聚合酶链反应("PCR")技术。可利用 PCR 技术鉴定各种生物体基因组中的相关序列，并检测可疑样品中的核苷酸序列，其中均使用彼此分隔开的、并且基于本文所示基因序列的寡核苷酸引物。引物互补于双链 DNA 分子的相反链，并且隔开大约 50—450 或更多个核苷酸(通常不超过 2000 个核苷酸)。该方法包括制备特异性寡核苷酸引物，然后重复靶 DNA 变性、引物结合及用 DNA 聚合酶延伸的循环，以得到具有基于引物间隔而定的预期长度的 DNA 片段。由一个引物产生的延伸产物作为另一个引物的附加靶序列。靶序列扩增的程度受所完成的循环次数控制并且可按简单公式 2n(其中 n 是循环次数)从理论上计算出来。如果每次循环的平均效率为大约 65%至 85%，25 次循环即产生 0.3—4.8×10^6个靶序列拷贝。许多出版物中都描述了 PCR 方法，例如可参见 Saiki 等人(Science 230: 1350—1354, 1985; Nature 324: 163—166, 1986)和 Scharf 等人(Science 233: 1076—1078, 1986)的文章，也可参见美国专利 4,683,194、4,683,195 和 4,683,203,各专利内容均引入本文作为参考。另外，PCR Technology: Principles and Applications for DNA Amplification, ed. HA Erlich, Freeman Press, New York, NY (1992); PCR Protocols: A Guide to Methods and Applications, eds. Innis, Gelfland, Snisky and White, Academic Press, San Diego, CA (1990); Mattila et al., Nucleic Acids Res. 19: 4967 (1991); Eckert, K. A. and Kunkel, T. A., PCR Methods and Applications 1:17 (1991); PCR eds. McPherson, Quirkes and Taylor, IRL Press, Oxford(各文献均引入本文作为参考)中还描述了 RCR 扩增的其他附加方法。

在另一个具体实例中，可给哺乳动物投用反义多核苷酸以治疗或预防涉及病理性和/或缺氧驱使的血管生成的医学状况。可用已知的化

学寡核苷酸合成方法合成本发明的反义寡核苷酸(例如参见 Winnacker, From Genes to Clones: Introduction to Gene Technology, VCH Verlagsgesellschaft mbH (H. Ibelgaufts trans. 1987))。可利用任何一种已知的寡核苷酸合成方法制备本发明的反义寡核苷酸。最好利用可从市场上购得的自动核酸合成仪制备反义寡核苷酸。用于制备本文所述寡核苷酸的设备 Applied Biosystems 380B 型 DNA 合成仪是利用氰乙基亚磷酰胺化学合成法。可使用本领域技术人员已知的寡核苷酸合成方法制备可与 mRNA 转录本的任何部分杂交的反义寡核苷酸。虽然本发明的实践中可利用任何长度寡核苷酸，但短于 12 个碱基的序列可能在与靶 GBS 毒素受体 mRNA 杂交中的特异性较小，而且可能更容易受到酶促消化而破坏。因此，最好是含有 12 个或更多个核苷酸的寡核苷酸。由于长度在 18—21 个核苷酸以上的序列被靶细胞摄入减少，所以有时在抑制 GBS 毒素受体翻译中效果较差。因此，在本发明的实践中优选 12—21 个核苷酸，特别是 12—18 个核苷酸的寡聚体。原则上，与 GBS 毒素受体 mRNA 转录本的任何部分互补并可与之杂交的寡核苷酸均可有效地抑制转录本的翻译，并能够诱导本文所述的作用。在起始密码子处或附近阻断 mRNA 可最为有效地抑制翻译过程。因此，优选互补于 GBS 毒素受体 mRNA 转录本之 5′ 区的寡核苷酸。在这个区域中可能干扰杂交的二级或三级结构是最小的。另外，在 3′ 方向距离起始位点太远的序列由于“通读”现象而可能使核糖体松解反义/有义双链从而允许信息翻译，所以其与 mRNA 转录本杂交的效率可能较差(参见 Shakin, J. Biochemistry 261, 16018(1986))。反义寡核苷酸最好是针对蛋白质合成的 ATG 起始密码子处位点或其附近。优选互补于 GBS 毒素受体 mRNA 中包括起始密码子的部分的寡核苷酸。虽然优选互补于 GBS 毒素受体转录本的 5′ 区域特别是包括起始密码子之区域的反义寡聚体，但应理解有用的反义寡聚体不只限于那些与 mRNA 转录本中翻译区域内的序列互补者，而且也包括与 5′ 和 3′ 非翻译区域内包含的或延伸到其中的核苷酸序列互补的寡聚体。因为 GBS 毒素受体的不适当表达可导致内皮细胞的过度

增殖，所以可将反义核苷酸或反义表达构建体用于治疗或预防与病理性的或缺氧驱使的血管生成及新血管形成有关的疾病。

在一个具体实例中，本发明的多核苷酸可能是线性形式的。在另一个具体实例中，多核苷酸可以以环形形式作为质粒的一部分存在。

在另一个具体实例中，探针或 PCR 引物包括一组在不同位置上含有不同的简并密码子的多核苷酸，这些多核苷酸编码全部或部分 GBS 毒素受体或互补于其编码序列。这样的多核苷酸可用于分离那些编码与绵羊或人 GBS 毒素受体的氨基酸序列有至少约 85%相同性之多肽(如其他生物体的 GBS 毒素受体)的核酸序列。可以使用程序化处理的合成仪，在某些位置上掺入特定组合的核酸残基，以化学合成这些多核苷酸。

下列表 2 中显示了典型常用的代码。

表 2

碱基代码

符号	含义
A	A;腺嘌呤
C	C;胞嘧啶
G	G;鸟嘌呤
T	T;胸腺嘧啶
U	U;尿嘧啶
M	A 或 C
R	A 或 G
W	A 或 T/U
S	C 或 G
Y	C 或 T/U
K	G 或 T/U
V	A 或 C 或 G;不是 T/U
H	A 或 C 或 T/U;不是 G
D	A 或 G 或 T/U;不是 C
B	C 或 G 或 T/U;不是 A
N	A 或 C 或 G 或 T/U

多肽

本发明的另一个方面提供下列多肽：(1) 全长度 GBS 毒素受体蛋白或其天然存在的等位基因变体，(2) 包括 SEQ ID NO: 2、4、8、10 或 12 所示氨基酸序列中至少 3 个氨基酸的片段，以及(3) 与 SEQ ID NO: 2、4 或 8 之相应区域的氨基酸序列有大约 80%至 100%氨基酸相同性的 GBS 毒素受体蛋白、多肽或多肽片段。优选 SEQ ID NO: 2、4、8、10 或 12 所示氨基酸序列的片段至少有 5、6、7、8 或 9 个氨基酸长并且是免疫学活性的，即免疫原性的，更优选长度至少 25 个氨基酸的片段，以及包括 SEQ ID NO: 2 的残基 181—419 或 SEQ ID NO: 4 的残基 1—240 的氨基酸序列。最优选的是能够与 GBS 毒素结合的片段。优选 GBS 毒素受体蛋白质、多肽或多肽片段与 SEQ ID NO: 2、4 或 8 之相应区域的氨基酸序列至少有约 86%、较好是至少约 95%、更好是至少约 99%相同性直至只有 1 个氨基酸差异，最好有 100%氨基酸相同性。优选的多肽与 SEQ ID NO: 2 中残基 181—419、SEQ ID NO: 4 中残基 1—495 构成的氨基酸序列具有至少约 89%、较好是至少约 95%、更好是至少约 99%相同性直至只有 1 个氨基酸差异，且最好有 100%相同性。优选全长度 GBS 毒素受体蛋白包括 SEQ ID NO: 2 中残基 1—495、SEQ ID NO: 4 中残基 1—495、或 SEQ ID NO: 8 中残基 1—536 的氨基酸序列或其等位基因变体。本发明的多肽除 GBS 毒素受体蛋白、多肽或多肽片段外还可包括其它氨基酸。这样的多肽一般包括与衍生于 GBS 毒素受体的第二种多肽连接的异源多肽。附加氨基酸最好共价连接到 GBS 毒素受体蛋白、多肽或多肽片段的氨基末端或羧基末端上。

可以按照已知方法制备 GBS 毒素受体的片段或类似物。片段或类似物的优选的氨基和羧基末端出现在功能区边界附近。例如，这样的功能区包括当 GBS 毒素与 GBS 毒素受体结合后提供诱导炎性反应性质的结构域。GBS 毒素被表达 GBS 毒素受体之内皮细胞的 C3 结合并受其调理。这样的结构域可包括全部或部分 GBS 毒素的结合位点，或对于 GBS 毒素受体结构和/或功能必需的结构域。最好使用计算机化比较方法鉴定出现于已知结构和/或功能的其他蛋白质中的序列基元或推断的蛋白质构象结构域。鉴定那些折迭成已知三维空间结构的蛋白质序

列的方法是已知的(Bowie et al., Science 253: 164, 1991)。可使用计算机化预测方法，例如 PC/GENE 中“Soap”程序所提供的亲水性图，鉴定推测的结构和功能性区域。本发明人已使用 Klein、Kanehisa 和 Delise (Biochim, Biophys. Acta 815: 468—476, 1985)的方法分类鉴定了一种可能具有七个跨膜区的整合膜蛋白，即绵羊 GBS 毒素受体 SP55。这种蛋白质也被俗称为“7-spanner”。下列表 3 中列示了预测的各区段。

表 3 SP55 推测的跨膜结构域

	内边界		外边界		区段	P:I odds*
No.	从	至	从	至	序列	
1	232	248	226	252	FFGIVGIIWFILWICLV (232-248， SEQ ID No. 4)	2.589323E-05
2	369	385	365	389	LIGMIGPAIFLVAAGFI (369-385，SEQ ID No. 4)	1.007311E-03
3	458	474	456	479	TVFCIAAAINVFGAIFF (458-474，SEQ ID No. 4)	2.482542E-03
4	137	153	135	157	LLLGFGIFATAIFTLFT (137-153，SEQ ID No. 4)	7.564906E-03
5	42	58	42	58	LAFLSFFGFFVLYSLRV (42-58，SEQ ID No. 4)	8.236557E-02
6	328	344	328	345	GFLSAVPYLGCWLCMI L (328-344， SEQ ID No. 4)	.1925022
7	390	406	390	407	SLAVAFLTISTTLGGFC (390-406，SEQ ID No. 4)	.8064944

* 相对于外周序列疏水性的整合序列疏水性。

有较高疏水性数值的整合序列很可能是跨膜结构域的一部分。

计算机化对比各种生物体 GBS 毒素受体的氨基酸序列为制备优选的片段提供了进一步的指南。例如参见表 4，其中比较了人 GBS 毒素受体(HP59) (SEQ ID NO: 8的残基42—536) 和绵羊GBS毒素受体(SP55)之残基 42—536 的氨基酸序列。

表 4 人和绵羊 GBS 毒素受体氨基酸序列的比对

```
SP55   MKSPVSDLAPSDGEEGSDRTPLLQRAPRAEPAPVCCSARYNLAFLSFFGF   50
HP55   MRSPVRDLARNDGEESTDRTPLLPGAPRAEAAPVCCSARYNLAILAFFGF   50

SP55   FVLYSLRVNLSVALVDMVDSNTTAKDNRTSYECAEHSAPIKVLHNQTGKK  100
HP55   FIVYALRVNLSVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHHNQTGKK  100

SP55   YRWDAETQGWILGSFFYGYIITQIPGGYVASRSGGKLLLGFGIFATAIFT  150
HP55   YQWDAETQGWILGSFFYGYIITQIPGGYVASKIGGKMLLGFGILGTAVLT  150

SP55   LFTPLAADFGVGALVALRALEGLGEGVTYPAMHAMWSSWAPPLERSKLLS  200
HP55   LFTPIAADLGVGPLIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLLS  200

SP55   ISYAGAQLGTVVSLPLSGVICYYMNWTYVFYFFGIVGIIWFILWICLVSD  250
HP55   ISYAGAQLGTVISLPLSGIICYYMNWTYVFYFFGTIGIFWFLLWIWLVSD  250

SP55   TPETHKTITPYEKEYILSSLKNQLSSQKSVPWIPMLKSLPLWAIVVAHFS  300
HP55   TPQKHKRISHYEKEYILSSLRNQLSSQKSVPWVPILKSLPLWAIVVAHFS  300

SP55   YNWTFYTLLTLLPTYMKEVLRFNIQENGFLSAVPYLGCWLCMILSGQAAD  350
HP55   YNWTFYTLLTLLPTYMKEILRFNVQENGFLSSLPYLGSWLCMILSGQAAD  350

SP55   NLRARWNFSTLWVRRVFSLIGMIGPAIFLVAAGFIGCDYSLAVAFLTIST  400
HP55   NLRAKWNFSTLCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFLTIST  400

SP55   TLGGFCSSGFSINHLDIAPSYAGILLGITNTFATIPGMIGPIIARSLTPE  450
HP55   TLGGFCSSGFSINHLDIAPSYAGILLGITNTFATIPGMVGPVIAKSLTPD  450

SP55   NTIGEWQTVFCIAAAINVFGAIFFTLFAKGEVQNWAISDHQGHRN       495
HP55   NTVGEWQTVFYIAAAINVFGAIFFTLFAKGEVQNWALNDHHGHRH       495
```

HP55 - SEQ ID NO: 2
SP55 - SEQ ID NO: 4

因此，前述实施例证明，本领域技术人员可以识别那些可用于确定 GBS 毒素受体序列中的结构和功能结构域的序列基元及结构构象。

虽然一类优选的片段是具有相当于功能区边缘附近之氨基酸位置的氨基和/或羧基末端的片段，但也可以制备另外可替代的片段。对这

些片段的氨基和羧基末端的选择将由实践者来决定，并且将取决于各种实验需要，如易于构建、对蛋白水解作用的稳定性、热稳定性、免疫反应性、氨基或羧基末端残基修饰等。多肽片段通常含有至少九个氨基酸并可含有任何数目的氨基酸，条件是肽片段与 SEQ ID NO: 2、SEQ ID NO: 4 或 SEQ ID NO: 8 的相应片段有至少约 80%相同性。与绵羊GBS毒素受体相比，人GBS毒素受体在N末端有41个附加氨基酸(SEQ ID NO: 4 和 SEQ ID NO: 8 相比较)。类似物可以包括这 41 个 N 末端氨基酸的全部或部分加入或缺失。也可包括可用于如纯化和/或抗体识别的 N 末端和 C 末端加入。例子包括组氨酸标签、FLAG(苯丙氨酸、亮氨酸、丙氨酸、鸟氨酸)表位、融合对象如谷胱甘肽 S 转移酶、氯霉素乙酰转移酶(CAT)、荧光素酶、β 半乳糖苷酶等。也可包括非保守氨基酸的缺失，条件是没有实质损害 GBS 毒素受体的结构完整性和/或结合特性。

类似物也可包括氨基酸取代，特别是保守取代。另外优选的是在种间具有较少相同性的区域中的保守和/或非保守取代。例如，GBS 毒素受体的变体可包括相当于 SEQ ID NO: 4 之残基

2, 6, 10, 11, 16, 17, 24, 31, 44,

46, 52, 53, 55, 74, 75, 81, 82, 84, 93, 102, 132, 133, 137, 144, 145, 148, 149, 155,

159, 163, 165, 166, 179, 212, 219. 235. 236, 239, 242, 246, 253. 254, 257, 259, 260,

271, 283, 285, 319, 324, 332, 333, 338, 355. 362, 366, 377, 439, 442, 445, 450, 453,

461, 487, 488, 491 和 495 的氨基酸的保守和/或非保守取代。最好是存在于 SEQ ID NO: 4 或 SEQ ID NO: 8 的相应位置中的氨基酸取代。例如参照图 4 中的序列比对图，相当于 SEQ ID NO: 4 之位置 152 的氨基酸可以是精氨酸(R)、谷氨酰胺(Q)或者 R 或 Q 的保守或非保守取代，并且最好是 R 或 Q。可根据氨基酸序列对比图表(如表 4 所示)鉴定这些区域。优选的氨基酸取代能：(1) 降低对蛋白水解作用的敏感性，(2) 降低对氧化作用的敏感性，(3) 改变对 GBS 毒素的结合亲和性；(4) 赋予或修饰这些类似物的其他物理化学或功能性质。类似物可包括天然肽序列以外的其他各种突变序列，如带有单一或多个氨基酸取代的序列。

保守氨基酸取代一般不会实质上改变亲代序列的结构特征(例如取代的氨基酸并不会破坏亲本序列中存在的螺旋结构，不会破坏亲本序列中特有的二硫键或其他类型的二级结构)。下列文献中描述了本领域熟知的多肽二级和三级结构的例子: Proteins, Structures and Molecular Principles (1984) Creighton (ed.), W. H. Freeman and Company, Nen York; Introduction to Protein Structure(1991), C. Branden and J. Tooze, Garland Publishing, New York, NY; Thornton et al., (1991) Nature 354: 105(引入本文作为参考)。保守取代是“多肽中的氨基酸被具有相似特征的氨基酸所取代”(McGraw-Hill Dictionary of Scientific and Technical Terms, Fifth Edition, 1994, Sybil P. Parker, Editor in Chief)。天然存在之氨基酸的结构和特征早已为人们所知(Biochemistry, 2nd. Eds., Albert L. Lehninger, p.71—76, 1975)。例如，因其疏水性R基而表现出相似性的氨基酸是丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、色氨酸和蛋氨酸。因其脂族R基而相似的是丙氨酸、缬氨酸、亮氨酸和异亮氨酸。因其芳族R基而相似的是苯丙氨酸和色氨酸。因其不带电荷的极性R基而相似的是甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺。因其小分子量而相似的是甘氨酸和丙氨酸。因其R基团中的羟基而相似的是丝氨酸和苏氨酸。天冬酰胺和谷氨酰胺不同只在于一个甲基。同样，天冬氨酸和谷氨酸不同也只是一个甲基，而且它们还因酸性R基团而相似。赖氨酸、精氨酸和组氨酸因其碱性R基团而相似。另外，赖氨酸和精氨酸因R基团中脂族链末端上的氨基而相似。酪氨酸和苯丙氨酸因芳族基团而相似。常见的氨基取代包括缬氨酸取代亮氨酸或异亮氨酸、丙氨酸取代甘氨酸、丝氨酸取代苏氨酸、天冬酰胺取代谷氨酰胺、天冬氨酸取代谷氨酸，赖氨酸取代精氨酸，酪氨酸取代苯丙氨酸，以及反方向替代。

一般本领域技术人员应尽量不改变那些GBS毒素受体中的保守氨基酸，而只对其进行合理地保守取代。例如，表4中指出了应避免在

相当于 SEQ ID NO: 4 下述残基的氨基酸处发生取代，特别是非保守取代: 1, 3-5, 7-9, 12-15, 18-23, 26-30, 32-43, 45, 47-51, 54, 56-73, 76-80, 83, 85-92, 94-101, 103-131, 134-136, 138-143, 146-147, 150-154, 156-158, 160-162, 164, 167-178, 180-211, 213-218, 220-234, 237-238, 240-241, 243-245, 247-252, 255-256, 258, 261-270, 272-282, 284, 286-318, 320-323, 325-331, 334-337, 339-354, 356-361, 363-365, 367-376, 378-438, 440-441, 443-444, 446-449, 451-452, 454-460, 462-486, 489-490 和 492-494，

所述残基在表 4 所示的 GBS 毒素受体之间是保守的。

表 5 和 6 分别列示了 HP59 和 SP55 内推测的酰胺化、N-糖基化、cAMP 磷酸化、CK2 磷酸化、十四烷基化(加入不饱和脂肪酸分子)及 PKC 磷酸化位点(Omega 1.1 序列分析程序)的序列。这些表中给出的信息为本领域技术人员设计 GBS 毒素受体变体及片段提供了指导。例如，当设计多肽变体时，可避免这些区域中的全部或部分取代。又如，当设计除免疫原性多肽片段以外的多肽片段时，可以选择地包括这些区域的全部或部分取代。

表 5
HP59 中推测的识别位点

位点	Seq. ID NO:8 残基:	序列
酰胺化	23-26	SGRR
酰胺化	138-141	TGKK
ASN-糖基化	100-103	NLSV
ASN-糖基化	112-115	NTTL
ASN-糖基化	118-121	NRTS
ASN-糖基化	136-139	NQTG
ASN-糖基化	266-269	NWTY
ASN-糖基化	343-346	NWTF
ASN-糖基化	398-401	NFST
CAMP-磷酸化位点	297-300	KRIS
CK2-磷酸化位点	113-116	TTLE
CK2-磷酸化位点	114-117	TLED
CK2-磷酸化位点	300-303	SHYE
CK2-磷酸化位点	493-496	TVGE
十四烷基化	66-71	GAPRAE
十四烷基化	167-172	GGYVAS
十四烷基化	183-188	GILGTA
十四烷基化	213-218	GLGEGV
十四烷基化	246-251	GAQLGT
十四烷基化	250-255	GTVISL
十四烷基化	378-383	GSWLCM
十四烷基化	427-432	GCDYSL
十四烷基化	444-449	GGFCSS
十四烷基化	464-469	GILLGI
十四烷基化	468-473	GITNTF
PKC-磷酸化位点	23-25	SGR
PKC-磷酸化位点	58-60	TDR
PKC-磷酸化位点	78-80	SAR
PKC-磷酸化位点	120-122	TSK
PKC-磷酸化位点	138-140	TGK
PKC-磷酸化位点	310-312	SLR
PKC-磷酸化位点	317-320	SQK

表 6
SP55 中推测的识别位点

位点	Seq. ID NO:4 残基:	序列
酰胺化	97-100	TGKK
ASN-糖基化	59-62	NLSV
ASN-糖基化	71-74	NTTA
ASN-糖基化	77-80	NRTS
ASN-糖基化	95-98	NQTG
ASN-糖基化	225-228	NWTY
ASN-糖基化	302-305	NWTF
ASN-糖基化	357-360	NFST
CK2-磷酸化位点	11-14	SDGE
CK2-磷酸化位点	73-76	TAKD
CK2-磷酸化位点	79-82	TSYE
CK2-磷酸化位点	259-262	TPYE
CK2-磷酸化位点	452-455	TIGE
十四烷基化	126-131	GGYVAS
十四烷基化	142-147	GIFATA
十四烷基化	162-167	GALVAL
十四烷基化	172-177	GLGEGV
十四烷基化	205-210	GAQLGT
十四烷基化	209-214	GTVVSL
十四烷基化	337-342	GCWLCM
十四烷基化	386-391	GCDYSL
十四烷基化	403-408	GGFCSS
十四烷基化	423-428	GILLGI
十四烷基化	427-432	GITNTF
PKC-磷酸化位点	17-19	SDR
PKC-磷酸化位点	37-39	SAR
PKC-磷酸化位点	55-57	SLR
PKC-磷酸化位点	73-75	TAK
PKC-磷酸化位点	97-99	TGK
PKC-磷酸化位点	254-256	THK
PKC-磷酸化位点	269-271	SLK
PKC-磷酸化位点	276-278	SQK

根据上述结果可见，优选的多肽包括下式的氨基酸序列：

$AA1—AAn—AA_m$

其中 AA1 不存在或者是 M;

AAn 是 SEQ ID NO: 8 中 0—100 个氨基酸，较好 0—41 个氨基酸，最好为残基 2—42 的连续链，并且

AAm 是包括 AA43 至 AA536 的 494 个氨基酸的连续链，其中：

(1)

AA43, AA47, AA51, AA52, AA57, AA58, AA65, AA66, AA72, AA85, AA87, AA93, AA94, AA96, AA115, AA116, AA122, AA123, AA125, AA134, AA143, AA173, AA174, AA178, AA185, AA186, AA189, AA190, AA196, AA200, AA204, AA206, AA207, AA220, AA253, AA260, AA276, AA277, AA280, AA283, AA287, AA294, AA295, AA298, AA300, AA301, AA312, AA324, AA326, AA360, AA365, AA373, AA374, AA379, AA396, AA403, AA407, AA418, AA480, AA483, AA486, AA491, AA494, AA502, AA528, AA529, AA532 和 AA536是基本氨基酸或修饰的氨基酸，并且最好是分别相当于：

(a) SEQ ID NO: 8 的残基

43, 47, 51, 52, 57, 58, 65, 66, 72, 85, 87, 93, 94, 96, 115, 116, 122, 123, 125, 134, 143, 173, 174, 178, 185, 186, 189, 190, 196, 200, 204, 206, 207, 220, 253, 260, 276, 277, 280, 283, 287, 294, 295, 298, 300, 301, 312, 324, 326, 360, 365, 373, 374, 379, 396, 403, 407, 418, 480, 483, 486, 491, 494, 502, 528, 529, 532 和 536,

(b) SEQ ID NO: 4 的残基

2, 6, 10, 11, 16, 17, 24, 25, 31, 44, 46, 52, 53, 55, 74, 75, 81, 82, 84, 93, 102, 132, 133, 137, 144, 145, 148, 149, 155, 159, 163, 165, 166, 179, 212, 219, 235, 236, 239, 242, 246, 253, 254, 257, 259, 260, 271, 283, 285, 319, 324, 332, 333, 338, 355, 362, 366, 377, 439, 442, 445, 450, 453, 461, 487, 488, 491 和 495,

(c) 其保守取代；

(2) AA44-AA46, AA48-AA50, AA53-AA56, AA59-AA64, AA67-AA71, AA73-AA84, AA86, AA88-AA92, AA95, AA97-AA114, AA117-AA121, AA124, AA126-AA133, AA135-AA142, AA144-AA172, AA175-AA177, AA179-AA184, AA187-AA188, AA191-AA195, AA197-AA199, AA201-AA203, AA205, AA208-AA219, AA221-AA252, AA254-AA259, AA261-AA275, AA278-AA279, AA281-AA282, AA284-AA286, AA288-AA293, AA296-AA297, AA299, AA302-AA311, AA313-AA323, AA325, AA327-AA359, AA361-AA364, AA366-AA372, AA375-AA378, AA380-AA395, AA397-AA402, AA404-AA406, AA408-AA417, AA419-AA478, AA481-AA482, AA484-AA485, AA487-AA490, AA492-AA493, AA495-AA501, AA503-AA527, AA530-AA531 和 AA533-AA535

分别是:

(a) SEQ ID NO: 8 的残基

44-46, 48-50, 53-56, 59-64, 67-71, 73-84, 86, 88-92, 95, 97-114, 117-121, 124, 126-133, 135-142, 144-172, 175-177, 179-184, 187-188, 191-195, 197-199, 201-203, 205, 208-219, 221-252, 254-259, 261-275, 278-279, 281-282, 284-286, 288-293, 296-297, 299, 302-311, 313-323, 325, 327-359, 361-364, 366-372, 375-378, 380-395, 397-402, 404-406, 408-417, 419-478, 481-482, 484-485, 487-490, 492-493, 495-501, 503-527, 530-531 和 533-535,

或者是

(b) 其保守取代; 以及

(3) 也可能没有 AA315 至 AA367。

优选的多肽包括 SEQ ID NO: 4、SEQ ID NO: 8 的氨基酸序列, 或只在绵羊 GBS 毒素受体(SEQ ID NO: 4)和人 GBS 毒素受体(SEQ ID NO: 8)之间非保守的特定残基上有所不同的氨基酸序列。在这些变异中, 最优选的变异是产生由 SEQ ID NO: 11 编码之多肽的变异。更为优选的变异是 SEQ ID NO: 4 之氨基酸序列相应位置上的氨基酸变异。特别优选的是包括与 SEQ ID NO: 2、SEQ ID NO: 4 或 SEQ ID NO: 8 的氨基酸序列不同不超过约 20%, 较好为不超过大约 10%、5%、1%氨基酸残基, 甚至只有 1 个至 0 个氨基酸差异之氨基酸序列的多肽。

除了定向改变特定氨基酸外, 还可以利用活性选择如噬菌体展示、定向进化、DNA 改组及原核或真核生物体之不同种基因的同源重组等技术(例如参见美国专利 5,605,793、5,830,721、5,811,238、5,837,458、5,093,257、5,223,409、5,403,484、5,571,698 和 5,837,500,其引入本文作为参考)制备 GBS 毒素受体类似物。

可以检测本文描述的人和绵羊 GBS 毒素受体的变体或片段是否可与 GBS 毒素结合, 以试验其必要的活性。

当在细胞特别是被稳定转染的细胞表面上表达时, 这些多肽可用

于各种体外试验，并可作为药物开发和纯化的试剂。例如，可使用结合试验法，根据化合物与 GBS 毒素受体结合的能力筛选多糖等试验化合物。可以使用这些方法鉴定某些可阻断 GBS 毒素与 GBS 毒素受体结合的候选药物。这些药物可用于预防和/或治疗新生儿的早期发作性疾病。某些多肽可用于竞争性抑制 GBS 毒素与 GBS 毒素受体的结合。

可使用本发明的多肽亲和纯化 GBS 毒素、GBS 毒素嵌合化合物，及其他可与 GBS 毒素受体结合的多糖或化合物。

也可使用多肽发展向表达 GBS 毒素受体的细胞定向送递细胞毒性剂的方法。例如，可将细胞毒性剂偶联到可结合 GBS 毒素受体的分子上，以使选择性地向生长着的肿瘤的新生脉管系统送递细胞毒性剂。这种送递系统可以高度浓缩地定位攻击正在生长的肿瘤，同时最大限度地减小一般化学治疗常常遇到的全身性副作用。在一例子中，细胞毒性剂可以是 GBS 毒素，当其与 GBS 毒素受体结合后，即可诱导炎性反应(参见 Hellerqvist et al., Angiogenesis: Molecular Biology, Clinical Aspects, M.E. Maragoudakis 等编, Plenum Press, New York 1994, pp. 265－269)。可以按相似方法，向表达 GBS 毒素受体的细胞选择性地送递治疗剂，以有利地治疗肿瘤、类风湿性关节炎或神经损伤，或促进伤口愈合。

也可使用本发明的多肽筛选和/或设计改善了治疗性质如改善了抑制缺氧诱发的新血管形成或血管生成能力的 GBS 毒素模拟药物。这样模拟化合物可用于治疗和预防缺氧诱发的新血管生成或血管形成相关的病理状况，如肿瘤生长、伤口愈合期间的斑痕形成、神经损伤修复期间的神经胶质增生、再灌注损伤、再狭窄、类风湿性关节炎、牛皮癣，以血管生成为特征的其他慢性炎性疾病等。可以通过提高生物学稳定性、对 GBS 毒素受体的亲和性、补体结合活性、降低抗原性等来改善治疗性质。

也可使用本发明的多肽产生用于各种治疗及研究目的的抗体。可使用本发明的多肽免疫兔、小鼠、山羊、鸡或本领域已知适于这种免疫的其它动物。一般优选单克隆抗体，但如果能够检测 GBS 毒素受体

抗体与 GBS 毒素受体的结合，则也可以使用多克隆抗体。生产非人单多克隆抗体如鼠类单克隆抗体的方法是已知的(如参见 Harlow et al., Antibodies A Laboratory Manual, Cold Spring Harbor Press, pp. 139—240，1989,将其引入本文作为参考)。因为可能难以产生抗人受体或结合区多肽的人单克隆抗体，所以可以用重组 DNA 技术将非人单克隆抗体的抗原结合区，如 $F(ab')_2$ 或鼠单克隆抗体的高变区转移到人恒定区(Fc)或构架区上，以产生实质上的人分子(参见美国专利 4,816,397 和 4,946,778，以及欧洲专利申请公开 EP 173,494 和 239,400。或者，可以按照 WO 90/14430 中介绍的一般方法筛选人 B 细胞的 DNA 文库，以分离编码可与受体蛋白特异性结合的人单克隆抗体或其部分的DNA序列，然后克隆并扩增编码具有所需特异性之抗体(或结合片段)的序列。

用于生产抗体的多肽通常是从推测的胞外结构域设计的全长度受体或受体片段，其中胞外结构域可用各种已知方法，包括基于其一级氨基酸序列预测多肽之二级或三级结构的计算机程序鉴定。设计抗原性肽的另一种方法是利用预测特定一级氨基酸序列内亲水性高点的计算机程序。例如使用 Happ 和 Woods (Proc. Natl. Acad. Sci. USA 78: 3824—3829, 1981)的方法，本发明人通过 PC/GENE 中的“抗原”程序鉴定了三个高亲水性区域(如表 7 所示)，并使用该结果设计用于制备抗 GBS 毒素受体之抗体的抗原性肽(参见实施例 3)。

表 7

SP55 中的亲水性高点

序号	平均亲水性	序列
1	2.05	Glu-Glu-Gly-Ser-Asp-Arg(SEQ ID NO.2 的 14-19)
2	1.52	Lys-Asp-Asn-Arg-Thr-Ser(SEQ ID NO.2 的 75-80)
3	1.33	Arg-Ala-Pro-Arg-Ala-Glu(SEQ ID NO.2 的 25-30)

可使用识别完整 GBS 毒素受体各个部分的抗体进一步研究受体的结构和功能。本发明的多肽可以引发产生能识别各种形式的 GBS 毒素

受体的抗体。这些受体形式包括但不只限于在细胞表面上表达的完整的 GBS 毒素受体、变性的或未变性的 GBS 毒素受体，以及从细胞成分中纯化的 GBS 毒素受体或包含在细胞溶解物中的 GBS 毒素受体。可使用 GBS 毒素受体抗体研究种间差异及各种类型细胞中的 GBS 毒素受体表达水平。

识别全部或部分胞外结构域的抗体特别适合作为诊断剂用于监测肿瘤生长和转移、检测或鉴别类风湿性关节炎或牛皮癣等慢性炎性状况，以及检测因缺氧驱使的血管生成引起的其他医学状况如再狭窄。这种抗体可应用于各种标准的研究与诊断技术如 Western 印迹、免疫沉淀、ELISA、放射免疫试验(RIA)、BIACOR®、酶联免疫试验(EIA)、免疫荧光、荧光激活的细胞分选法(FACS)，及磁共振成象(MRI)、核磁共振(NMR)、计算机化轴向断层显象摄影(CAT)扫描和方位发射断层显象(position emission tomography，PET)等体内诊断性显象系统。

另外，可使用能阻断 GBS 毒素与 GBS 毒素受体结合的抗体治疗或预防新生儿的早期发作性疾病。这些抗体可直接或间接封闭 GBS 毒素受体上的 GBS 毒素结合位点。

在一个具体实例中，GBS 毒素受体蛋白是天然存在的，并且可用本领域已知的蛋白质纯化技术如离子交换柱层析、高效液相层析(HPLC)、反相 HPLC 或利用可识别 GBS 毒素受体之抗体的亲和层析法从细胞提取物中分离之。

或者，可使用适于在哺乳动物细胞，也可以是适于在细菌、昆虫或酵母细胞中表达的表达载体，利用与适当的控制序列如启动子可操作地连接的编码本发明之多肽的多核苷酸表达分离的蛋白质和多肽。

通常可以在哺乳动物系统中表达 GBS 毒素受体多核苷酸或其片段。这种表达一般要依赖于哺乳动物启动子，即能够结合哺乳动物 RNA 聚合酶并启动下游编码序列(如结构基因)转录成 mRNA 的任何 DNA 序列。启动子具有通常位于编码序列之 5′ 末端附近的转录起始区。该转录起始区通常包括 RNA 聚合酶结合位点和转录起始位点。

适于在哺乳动物细胞中复制的载体是本领域已知的，并可包括病

毒复制子，或可确保编码 PAK65 的序列整合到宿主基因组中的序列。适用的载体可包括衍生于 SV40、逆转录病毒、牛乳头状病毒、痘苗病毒及腺病毒的载体。

例如一种适用的载体是衍生于痘苗病毒的载体。在这种情况下，异源 DNA 被插入到痘苗病毒基因组中。将外来 DNA 插入到痘苗病毒基因组中的技术是本领域已知的，例用可利用同源重组技术。一般将外源 DNA 插入到天然非必要的基因如可提供选择性标志的胸苷激酶基因(tk)中。已描述了极有利于重组病毒构建的质粒穿梭载体(如参见 Mackett et al. (1984); Chakrabarti et al. (1985); Moss (1987))。然后在已用活的重组痘苗病毒免疫的细胞或个体中可发生外源多肽的表达。

这样的合适哺乳动物表达载体通常含有一个或多个可促进在哺乳动物细胞中表达的真核转录单位。转录单位至少由介导外源 DNA 序列转录的启动子元件组成。哺乳动物细胞适用的启动子是本领域已知的，并包括如衍生于猿猴病毒 40(SV40) (Subramani et al., Mol. Cell. Biol. 1: 854—864, 1981)、巨细胞病毒(CMV) (Boshart et al., Cell 41: 521—530, 1985)、劳氏肉瘤病毒(RSV)、腺病毒(ADV) (Kaufman and Sharp, Mol. Cell. Biol. 2: 1304—1319, 1982)和牛乳头瘤病毒(BPV)的启动子，以及细胞启动子如小鼠金属硫蛋白 1 启动子(美国专利 No. 4,579,821)、小鼠 VK 启动子(Bergman et al., Proc. Natl. Acad. Sci. USA 81: 7041—7045, 1993; Grant et al., Nuc. Acids Res. 15: 5496, 1987)和小鼠 VH 启动子(Loh et al., Cell 33: 85—93, 1983)。

可以存在与本文所述的启动子元件组合的增强子元件(增强子)，借以提高表达水平。增强子是当与内源或外源启动子连接，在正常 mRNA 起始位点开始合成时可刺激转录效率提高达 1000 倍的任何调节性 DNA 序列。当其位于转录起始位点的上游或下游，处于正常或倒转方向，或距离启动子达 1000 多个核苷酸以上时，增强子仍是有活性的(Maniatis et al., Science 236: 1237, 1987; Alberts et al.,

Molecular Biology of the Cell, 2nd ed.)。因为病毒有较宽的宿主范围，所以病毒衍生的增强子元件可能是特别有用的。可用于哺乳动物细胞的增强子包括SV40早期基因增强子(Dijkema et al. EMBO J. 4: 761, 1985)和衍生于劳氏肉瘤病毒之长末端重复区(LTR)的增强子/启动子(Gorman et al., PNAS, 79:6777, 1982b)、衍生于人巨细胞病毒的增强子(Boshart et al., Cell 41:521, 1985)以及小鼠μ增强子(Gillies, Cell 33: 717—728, 1983)。另外，有些增强子是可以调节的，并且只有存在诱导剂如激素或金属离子时才成为有活性的(Sassone-Corsi and Borelli, Trends Genet., 2: 215, 1986; Maniatis et al., Science 236: 1237, 1987)。

此外，转录单位还可包括与GBS毒素受体编码序列可操作连接的终止序列和多聚腺苷酸化信号。多聚腺苷酸化信号包括但不只限于SV40的早期或晚期多聚腺苷酸化信号(Kaufman and Sharp)、腺病毒5 E1B区的多聚腺苷酸化信号及人生长激素基因终止子(DeNoto et al., Nuc. Acids Res. 9: 3719—3730, 1981)。

可引起基因扩增的序列也可能是必要的，编码可选择标志的序列可能也是需要的。适于哺乳动物细胞的可选择标志在本领域是众所周知的，包括胸苷激酶、二氢叶酸还原酶(与氨甲喋呤一起作为DHFR扩增剂)、氨基糖苷磷酸转移酶、潮霉素B磷酸转移酶、天冬酰胺合成酶、腺苷脱氨酶及抗生素(如新霉素)抗性基因。

可以在细胞表面上或以可溶或分泌形式表达GBS毒素受体或其片段。例如可以在编码所需受体片段和至少一个跨膜结构域的核酸序列上可操作连接分泌性前导序列，以实现在细胞表面上的表达。分泌前导序列可以是由GBS毒素受体基因编码的，或者可以是本领域通用的异源性前导序列，例如非洲粟酒裂殖酵母$phol^{+}$酸性磷酸酶的前导序列(Braspenning et al., Biochem. Biophys. Res. Commun. 245: 166—171, 1998)、人白介素2基因的前导序列(Sasada et al., Cell Struc. Funct. 13: 129—141, 1988)。例如，通过从基因构建体中切掉编码跨膜结构域的核酸序列即可完成可溶或分泌形式的表达。在某些情况

下，可以利用融合对象，例如易于在选择的宿主细胞中表达的氯霉素乙酰转移酶(CAT)、β半乳糖苷酶或其他基因等，实现可溶性和/或分泌。

可以使用编码 GBS 毒素受体的载体转化适当的哺乳动物宿主细胞。可以用已知的任何方法将多核苷酸导入宿主细胞中，包括在病毒中包装多核苷酸并用病毒转导或按已知方法转染(如参见美国专利4,399,216、4,912,040、4,740,461 和 4,959,455 等，将其引入本文作为参考)宿主细胞。使用的转化方法取决于待被转化的宿主细胞。将异源多核苷酸导入哺乳动物细胞的方法是已知的，其中包括葡聚糖介导的转染、磷酸钙沉淀、1,5-二甲基-1,5-二氮十一亚甲基聚甲溴化物介导的转染、原生质体融合、电穿孔、在脂质体中包裹多核苷酸及向核内直接显微注射 DNA。

可用作表达宿主的哺乳动物细胞系是本领域已知的，并包括可得自美国典型培养物保藏中心(ATCC)的许多无限增殖化的细胞系，其中包括但不只限于中国仓鼠卵巢细胞(CHO)、Hela 细胞、幼仓鼠肾(BHK)细胞、猴肾细胞(COS)、NIE-115 (Liles et al., J. Biol. Chem. 261: 5307－5313, 1986)、PC12 人肝细胞癌细胞(如 HepG2)及许多其他细胞系，如昆虫来源的细胞系 IF9 和 IF21。特别优选的细胞系是组成型结构表达重组 GBS 毒素受体构建体的细胞系、继后建立被转化细胞特征的细胞系，及优先在细胞表面上表达 GBS 毒素受体或片段的细胞系。特别优选的是 ECV 细胞(原来在科学文献中称为内皮细胞系的膀胱癌细胞系)、人脐静脉内皮细胞(HUVEC)，及牛、绵羊和人肾上腺髓质内皮细胞。

可以在适当的培养基中和适当的细胞培养条件下，培养表达 GBS 毒素受体或其片段的宿主细胞，以生产重组受体或其片段。可以根据特殊宿主细胞系的需要来改变培养基和培养条件，这在本领域是众所周知的。一般在有固定量 CO_2(通常为 5－10%)的细胞培养箱中 37℃培养细胞。

在另一个具体实例中，可以用本领域已知的技术，如固相肽合成

技术(Stewart et al., Solid Phase Peptide Synthesis, W. H. Freeman Co., San Francisco (1963); Merrifield, J. Am. Chem Soc. 85: 2149—2154 (1963))化学合成多肽片段。美国专利 3,862,925、3,842,067、3,972,859 和 4,105,602 中也举例描述了这些及其他一些肽合成方法。可以手工合成，或者使用 Applied Biosystems 430A 或 431A 型肽合成仪(Foster City, California)，按照生产商提供的说明书进行自动化合成。肽合成领域的普通技术人员可以理解到，在合成本发明类似化合物过程中构建的中间体本身是新的和有用的化合物，因此也包括在本发明范围内。

除了仅由天然氨基酸构成的多肽外，还提供了肽模拟物。肽类似物通常是作为与模板肽有相似性质的非肽药物用于医药工业的。这些类型的非肽化合物被称为“肽模拟物”(Fauchere, J., Adv. Drug Res. 15: 29, 1986; Veber and Freidinger, TINS p. 392, 1985; Evans et al., J. Med. Chem. 30: 1229, 1987,各文献均引入本文作为参考)，并且通常是借助于计算机化的分子建模而开发的。可使用结构上与治疗有用的肽相似的肽模拟物产生等同的治疗或预防效果。一般说来，肽模拟物在结构上与范式多肽(即具有生化特性或药学活性的多肽)相似，但其中有一个或多个肽键任选地被选自下列一组的键所取代：$-CH_2NH-$、$-CH_2S-$、$-CH_2CH_2-$、$-CH=CH-$(顺式和反式)、$-COCH_2-$、$-CH(OH)CH_2-$及$-CH_2SO-$(有关方法参见 Spatola, A. F., “Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins”, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983); Vol. 1, Issue 3, “Peptide Backbone Modifications”(一般评述); Morley, J. S., Trends, Pharm. Sci. (1980) pp. 463—468 (一般评述); Hudson, D. et al., Int. J. Pept. Prot. Res. (1979) 14: 177—185 ($-CH_2NH-$, $-CH_2CH_2-$); Spatola, A. F. et al., Life Sci. (1986) 38: 1243—1249 ($-CH_2-S-$); Hann, M. M., J. Chem. Soc. Perkin Trans. I (1982) 307-314 (-CH-CH-, 顺式和反式); Almquist, R. G. et al., J. Med. Chem. (1980) 23: 1392—1398 ($-COCH_2-$); Jennings-White, C. et

al., Tetrahedron Lett. (1982) 23: 2533 (-$COCH_2$-); Szelke, M. et al.,欧洲专利申请 EP 45665 (1982) CA: 97: 39405 (1982) (-$CH(OH)CH_2$-); Holladay, M. W. et al., Tetrahedron Lett. (1983) 24: 4401—4404 (-$C(OH)CH_2$-); Hruby, V. J., Life Sci. (182) 31: 189—199 (-CH_2-S-);各文献均引入本文作为参考)。特别优选的非肽键是-CH_2NH-。这些肽模拟物可比前述多肽有明显的优点，例如更易于经济化生产、有更高的化学稳定性、提高了药学性质(半寿期,吸附,潜力,效力等)、改变了特异性(例如有更宽的生物学活性谱)、降低了抗原性等。标记肽模拟物通常包括将一个或多个标记直接或通过间隔子(如酰胺基团)共价连接到肽模拟物的非干扰位置上，其中所说的位置是根据定量结构—活性数据和/或分子模拟推测的。这种非干扰位置一般是并不与GBS毒素形成直接接触的位置(例如不是GBS毒素受体之GBS毒素结合区中的接触点)。肽模拟物的衍生化(如标记)不应在实质上干扰肽模拟物的所需生物学或药学活性。

可以用同一类型的D-氨基酸系统取代共有序列中的一个或多个氨基酸(如D-赖氨酸代替L-赖氨酸)，以产生更为稳定的肽。另外，可用本领域已知的方法产生包括共有序列或实质上相同的共有序列变异的受限肽(参见Rizo and Gierasch, Ann. rev. Biochem. 61: 387, 1992,将其引入本文作为参考)，例如可以加入内部半胱氨酸残基，以形成可使肽分子环化的分子内二硫键。

本发明还提供包括与哺乳动物 GBS 毒素受体或其片段结合之 GBS 毒素的复合体。复合体最好包括与上述可结合 GBS 毒素的 GBS 毒素受体多肽结合的 GBS 毒素。复合体一般是在可使 GBS 毒素与多肽特异性结合的条件下，使 GBS 毒素与这种多肽接触而形成的。GBS 毒素可以是被标记或者未被标记的。多肽可以存在于细胞表面上，或固定于小孔内或微球上，或者是存在于溶液中。

检测方法

本发明的另一个方面提供检测或监测以病理性和/或缺氧驱使的

血管生成或新血管形成为特征的各种医学状况的方法。这些医学状况包括但不只限于新生儿早期发作性疾病和肿瘤进展。

可以通过检测病人体内是否存在 GBS 毒素来诊断早期发作性疾病。一种检测方法是竞争法，其中包括确定可疑样品对 GBS 毒素与 GBS 毒素受体或其片段间之复合体形成的影响。例如，该方法包括在有及没有疑含有 GBS 毒素的样品存在下和允许 GBS 毒素与多肽特异性结合的条件下，使对照 GBS 毒素与 GBS 毒素受体多肽接触，并将存在有可疑样品时复合体形成的量与没有可疑样品时复合体形成的量相比较。对照 GBS 毒素最好是实质上纯化的并已知浓度。对照 GBS 毒素最好还包括有标记。适用的标记包括但不只限于放射性同位素、发色团、荧光团、生物素、抗生物素蛋白及其他常用标记。另一种方法是直接检测可疑样品中存在的 GBS 毒素与 GBS 毒素受体多肽间复合体的形成，而不是经对照 GBS 毒素竞争进行测定。

可以通过检测肿瘤区域内是否存在 GBS 毒素受体来确定哺乳动物组织中的病理性脉管系统，存在 GBS 毒素受体即为病理性脉管系统的指征。可用该方法监测肿瘤生长或转移。一种检测方法包括使用可与 GBS 毒素受体，最好是 GBS 毒素受体之胞外区域特异性结合的分子，如抗体。该方法一般包括向哺乳动物组织如患有癌性肿瘤的哺乳动物体内投用可识别 GBS 毒素受体的抗体，并检测抗体的特异性结合。抗体一般是标记的抗体。观察的结果最好是定量的并且是可以看见的。

在外科手术期间，可使用本领域已知的及上述的显象技术观察肿瘤的边界。肿瘤的显象是通过检测标记的抗体或其他分子与肿瘤病理性脉管系统上的 GBS 毒素受体的结合而达到的。这种外科也称为虚拟外科，因为进行手术时，外科医生是在成象屏上间接观察肿瘤。

药物开发

本发明的第四个方面提供使用本发明的多肽鉴定用于治疗以缺氧驱使的血管生成或新血管形成为特征的医学状况的候选药物的方法。优选的化合物是 GBS 毒素与 GBS 毒素受体结合的竞争性抑制剂或抑制 GBS 毒素受体活性的化合物。特别优选的是抑制信号转导途径中第一

个磷酸化步骤的化合物。可使用本领域已知的各种随机药物设计方法生产这些化合物。所说的方法例如包括组合化学法(美国专利5,646,285、5,639,603)、肽文库(美国专利5,591,646、5,367,053、5,747,334)、噬菌体展示(美国专利5,403,484、5,223,409)、SELEX®(美国专利5,773,598、5,763,595、5,763,566)，以及组合糖类化学法(Hirschmann et al, J.Med. Chem. (1996) 39: 2441－2448; Hirschmann et al., J. Med. Chem. (1998) 41: 1382－1391; Sofia MJ, Mol. Divers (1998) 3: 75－94; 美国专利5, 780, 603和5, 756, 712)。

另外的方法是旨在生产改善了治疗性质的GBS毒素模拟物或GBS毒素受体模拟物的合理药物设计方法，包括X射线结晶法、核磁共振(NMR)相关光谱法(美国专利5,698,401)、计算机辅助的分子建模(美国专利5,579,250、5,612,895、5,680,331, Cooper et al., J. Comput.-Aided Mol. Design 3: 253－259 (1989); Brent et al., J. Comput.-Aided Mol. Design 2: 310－311 (1988))及本领域已知的其他合理药物设计方法。图1显示了本发明的某些合理药物设计方法中主要步骤的全面观。例如，合理药物设计的一种方法包括使用计算机程序如INSIGHTⅡ(可购自Biosym Technologies, 10065 Barnes Canyon Road, San Diego, California)，根据同源性建模来鉴定蛋白质中的活性位点。这种方法有利于使用已知结构的相似蛋白质绘制蛋白质模型。例如可从Day Light Information Systems公司(Irvine, California 92714)和Molecular Design有限公司(2132 Faralton Drive, San Leandro, California 94577)等供应商购得包含进行三维数据比较的检索算法的商业软件。

在一个具体实例中，化合物能够结合GBS毒素受体并以相似于GBS毒素与GBS毒素受体结合的方式诱导炎性反应。例如，可使用这些化合物作为药物将炎性反应控制在肿瘤中正在发展着的脉管系统中。

在另一个具体实例中，化合物可以在诱导或不诱导炎性反应，特别是不诱导炎性反应的情况下与GBS毒素受体结合。在一种情况下，

可使用化合物作为载体，以将用于治疗的细胞毒性剂导向病理性新脉管系统部位。例如，可以将细胞毒性剂化学偶联到化合物上，以形成嵌合药物。这些嵌合药物可用于治疗肿瘤、类风湿性关节炎、伤口愈合、脊髓损伤，和以缺氧驱使的血管生成或新血管形成为特征的其他状况。在另一个例子中，可直接使用化合物竞争性抑制 GBS 毒素与 GBS 毒素受体的结合。可使用这样的化合物治疗新生儿的早期发作性疾病。

在第三个具体实例中，化合物可与 GBS 毒素结合并可用于治疗新生儿的早期发作性疾病。

可在随机诱变系统如噬菌体展示或酵母双杂交系统中表达本发明的多核苷酸，以合成并鉴定可与 GBS 毒素结合的突变肽 GBS 毒素受体多肽。或者，可使用本发明的固相化或可溶性 GBS 毒素受体片段筛选组合肽文库和组合化学文库、非随机的重组体和合成肽及其他化合物（如非肽分子）与 GBS 毒素受体结合的情况。然后可在功能试验中进一步定性可与 GBS 毒素或 GBS 毒素受体结合之化合物的任何上述活性，以鉴定可用于治疗涉及血管生成或新血管形成之医学状况的候选药物。

可以在允许 GBS 毒素与 GBS 毒素受体或片段结合的条件下，使试验化合物与能够结合 GBS 毒素的哺乳动物 GBS 毒素受体或其片段合并，并确定该化合物对 GBS 毒素与 GBS 毒素受体或片段结合的抑制水平，以鉴定能抑制 GBS 毒素与 GBS 毒素受体结合的化合物。

在一个优选实例中，由细胞特别是在细胞表面上表达 GBS 毒素受体或其片段。在有或没有试验化合物存在下，使细胞与标记的 GBS 毒素接触。然后确定 GBS 毒素与 GBS 毒素受体结合的改变。或者，GBS 毒素是未标记的，而可识别 GBS 毒素的抗体是被标记的。使用标记的抗体检测化合物对 GBS 毒素与 GBS 毒素受体或其片段之结合的抑制作用。在另一个实例中，GBS 毒素受体或其片段与细胞没有联系，而是偶联到基质如微量滴定板的小孔或微球上。其他适用的固相支持物包括胶乳、聚苯乙烯小珠(Interfacial Dynamics Corp. Portland,

Oreg.)、磁性颗粒(Advanced Magnetics, Cambridge, Mass.)及尼龙球(Hendry et al., J. Immunological Meth. 35: 285—296, 1980)。可以将受体或其片段直接偶联到基质上，或者通过可与受体片段结合的抗体间接偶联到基质上。在第三个具体实例中，GBS 毒素受体或片段是可溶的，并可以用识别受体或片段的抗体免疫沉淀之。

鉴定可结合哺乳动物 GBS 毒素受体之化合物的优选方法包括以下步骤：(1) 使试验化合物在得以发生特异性结合的条件下与 GBS 毒素受体或其片段合并，(2) 检测试验化合物与 GBS 毒素受体或片段间形成的复合体。一种优选的方法是竞争试验，其中检测试验化合物竞争结合 GBS 毒素受体或其片段的能力。在这一试验中，在有或没有试验化合物存在下，使 GBS 毒素与 GBS 毒素受体或其片段合并。有试验化合物时比没有时降低了 GBS 毒素的特异性结合，即表明试验化合物有能力与哺乳动物 GBS 毒素受体结合。另一种方法包括在如试验化合物相同的条件下使对照化合物与 GBS 毒素受体或片段合并，并比较试验化合物或对照化合物与 GBS 毒素受体或其片段之间形成的复合体的量。试验化合物和/或对照化合物最好是被标记的。试验化合物可以是任何一类的化合物，例如小的有机分子(如用于组合化学法的和以组合化学方法得到的)、多糖、多肽、RNA、抗体及单链抗体。在一优选实例中，多肽是由细胞表达的，最好是由稳定转染的细胞表达的，且最好在细胞表面上表达。这样一个系统特别适于试验包括细胞毒性剂之嵌合化合物的效力。可以通过使在细胞表面上表达 GBS 毒素受体的细胞与试验嵌合化合物接触并检测细胞毒性征象，以确定化合物的细胞毒活性。可以通过细胞存活性染色、检测细胞凋亡(例如 DNA 梯度试验或 TUNEL™(其与断裂的 DNA 结合)染色)、检测细胞内氚标记的胸苷掺入、定量分析激酶依赖性磷酸化(如使用磷酸抗体或各种磷酸显象技术)等方法检测这些征象。

在另一个具体实例中，本发明提供鉴定 GBS 毒素受体抑制剂的方法。该方法包括在有或没有试验化合物存在下保温试验细胞。试验细胞表达 GBS 毒素受体或其具有 GBS 毒素受体活性的片段 (如相对于并

不表达这些片段的相同细胞类型的对照细胞而言，可增加表达细胞的增殖或迁移的片断)。在没有试验化合物时被保温的细胞也能增殖或迁移的条件下保温试验细胞。并不表达 GBS 毒素受体或其片段的对照细胞的增殖或迁移比表达 GBS 毒素受体或片段的细胞更小。比较有或没有试验化合物存在时保温的检测细胞的增殖或迁移(本文中也称“迁移率”)。有试验化合物时比没有试验化合物时的增殖或迁移小，表明试验化合物是 GBS 毒素受体的抑制剂。为了确定试验化合物是否特异地抑制 GBS 毒素受体，还比较有或没有试验化合物时对照细胞的增殖和迁移。在有或没有试验化合物时被保温的对照细胞的增殖或迁移没有差别的情况下，与试验化合物接触的试验细胞比不接触试验化合物的试验细胞的增殖或迁移率低，表明试验化合物特异性地抑制 GBS 毒素受体。很显然，该方法的对照部分不需要与方法的试验部分同时完成。例如，可以将对照细胞与一组试验化合物保温，以便在试验细胞与试验化合物保温之前确定试验化合物的细胞效应。可以检测细胞在培养皿上的运动来确定细胞的游动和迁率。可以用许多方法检测细胞增殖，例如检测氚标记的胸苷掺入、细胞计数、凋亡试验及存活性试验。优选的细胞包括 GBS 毒素受体转染的细胞，特别是 GBS 毒素受体转染的内皮细胞，最好是血管内皮细胞或微血管内皮细胞。表达 GBS 毒素受体的原代细胞，例如已在细胞培养基中传代不超过 8 或 9 次的，汇合的内皮细胞也是优选的。一类优选的试验化合物包括激酶抑制剂，优选 cAMP 依赖型激酶抑制剂、PKC 抑制剂及 CK2 抑制剂，它们可作为开发更特异的 GBS 毒素受体抑制剂之起始点。另一类化合物包括 GBS 毒素受体特异性抗体。特别优选的是单链抗体，特别是识别 GBS 毒素受体上各种表位的单链抗体。其次是对 GBS 毒素受体配体的结合位点特异的双价抗体，因为它们在二聚化后可激发信号转导级联。

本发明的另一个具体实例是鉴定作为血管生成之基本成分的内皮细胞增殖或迁移抑制剂的方法。该方法基本上包括前示段落中所述的步骤并使用内皮细胞。

本发明的再一个具体实例是鉴定用于治疗或预防以病理性或缺氧

驱使的血管生成或新血管形成为特征的医学状况的治疗化合物的方法。该方法基本上包括上述步骤并使用衍生于受上述医学状况影响之哺乳动物组织的细胞，或作为受影响组织之模型的细胞。

设计可抑制 GBS 毒素与哺乳动物 GBS 毒素受体结合之化合物的优选方法包括以下步骤：(1) 模拟并选择 GBS 毒素受体或其片段的最可能的构象，(2) 设计实质上模拟 GBS 毒素受体或其片段能量上最可能的三维结构的化学改性类似物，(3) 以化学方法合成该类似物，(4) 估测该类似物的生物活性。步骤(a)和(b)最好是借助计算机程序完成的。

设计与哺乳动物 GBS 毒素受体结合之化合物的优选方法包括下列步骤：(1) 模拟并选择 GBS 毒素受体或其片段的最可能的构象，(2) 推测受体或片段的最可能的结合区，(3) 设计可与受体或片段形成能量上最可能的复合体的化合物，(4) 化学合成该化合物，(5) 估测该化合物的生物活性。步骤(a)-(c)最好是借助计算机程序来完成。

用于上述筛选试验的多肽最好是与 SEQ ID NO: 2 或 SEQ ID NO: 4 的氨基酸序列或其具有 GBS 毒素受体活性的片段至少有约 85%，较好有至少约 95%，最好有大于约 99%的序列相同性的多肽。最优选的是具有 SEQ ID NO: 2、4 或 8 的氨基酸序列的多肽或其具有 GBS 毒素受体活性的片段。

纯化方法

本发明的另一方面是纯化能结合 GBS 毒素受体的化合物，例如天然配体、其他多糖或 GBS 毒素受体特异性抗体的方法。该方法包括：提供包括哺乳动物 GBS 毒素受体或其可结合 GBS 毒素的片段的多肽，使该多肽与含有化合物的样品在允许化合物与多肽特异性结合的条件下接触，并从样品的残留物中分离出结合的化合物。多肽可以是水溶性的，但最好是固定在基质上，如小珠、膜或细胞表面上，特别是固定在稳定转染的细胞的表面上。

治疗方法

可使用 GBS 毒素受体多肽和干扰 GBS 毒素结合的抗体治疗人或动物体。例如，可给病人投用这种 GBS 毒素结合抑制剂，以治疗或预防涉及 GBS 毒素与 GBS 毒素受体的结合的医学状况，如新生儿的早期发作性疾病。

可使用可结合和/或抑制 GBS 毒素受体的 GBS 毒素模拟物或其他化合物(有些可以用本发明的药物发现方法鉴定)治疗人或动物体，或者可用于制造治疗或预防涉及病理性和/或缺氧驱使之血管生成的医学状况，如癌性肿瘤、慢性炎性疾病、伤口愈合或神经损伤修复期间的瘢痕形成等的药物。

在一优选实例中，这些化合物如 GBS 毒素一样，通过结合 GBS 毒素受体并引发炎性反应而发挥其治疗作用。这些化合物最好包括可用于结合补体 C3 的硫氢基、羟基或氨基。

在另一个优选实例中，化合物是 GBS 毒素活性的抑制剂。优选的抑制剂包括但不只限于激酶抑制剂、GBS 毒素受体特异性单链抗体，及在高度严格条件下可与 GBS 毒素受体核酸序列，如 SEQ ID NO: 1、SEQ ID NO: 3 或 SEQ ID NO: 7 的核酸序列特异性杂交的反义多核苷酸。

在另一优选实例中，化合物发挥其治疗作用而不引发炎性反应。可使用化合物向接近于表达 GBS 毒素受体之细胞的组织，如经受血管生成的肿瘤送递细胞毒性剂。化合物最好与细胞毒性剂共价连接，并可以与细胞毒性剂非共价结合，如结合到脂质体、微胶粒或其他亲脂性药物包裹结构的外表面上。优选的细胞毒性剂包括本领域已知的抗肿瘤剂,如氮芥(mechlorethamine)、苯丁酸氮芥、环磷酰胺、苯丙氨酸氮芥、异环磷酰胺及其他烷基化剂、氨甲喋呤及其他叶酸拮抗剂、6-巯基嘌呤及其他嘌呤拮抗剂、5-氟尿嘧啶及其他嘧啶拮抗剂、阿糖胞苷、硫酸长春碱(ovinblastine)、硫酸长春新碱(vincustine)及其他长春花碱(vincas)、依托伯苷及其他鬼臼毒素、多柔比星(doxorubicin)、博来霉素、丝裂霉素及其他抗生素、卡莫司汀、洛莫

司汀及其他亚硝基脲、顺铂、干扰素、天冬酰胺酶(asparaginase)、他莫昔芬、氟他胺及紫杉酚。其他优选的生物剂包括衍生于特异性肿瘤启动子或抑制基因的有义和/或反义 RNA 或 DNA 序列，所述特异性肿瘤启动子或抑制基因例如是 p53 和 TGF 基因家族，信号转导蛋白家族成员如 ras 和 myc，以及生长因子受体激酶如 flt2 和 flk1、Tai1、Tai2 和 neuropholin，及参与癌性疾病和其他病理性血管生成相关疾病的其他基因。

在另一个具体实例中，可作为诱饵给病人投用 GBS 毒素受体多肽或其片段，以减低存在于患病组织(如肿瘤组织)中的 GBS 毒素受体的受刺激作用，从而降低导致受影响组织中细胞增殖和迁移的细胞反应。最好以可溶形式，特别是无跨膜区的形式投用 GBS 毒素受体多肽或其片段。

药物组合物

可以使用具有预期的生物可利用性、稳定性及体内药物动力学方面的其他重要参数的，包括 GBS 毒素结合所必需之结构域的本发明多肽作为 GBS 毒素结合的竞争性抑制剂，用于治疗其中 GBS 毒素结合不合乎需要的医学状况，例如新生儿早期发作性疾病。适用的多肽可以包括具有 SEQ ID NO: 2 或 SEQ ID NO: 4 所示之部分或全部氨基酸序列的片段或其类似物。

可以使用借助本发明的多肽确定的能结合 GBS 毒素受体和/或诱导炎性反应，并具有预期的药物动力学性质的化合物，治疗其中 GBS 毒素结合可能有治疗作用的医学状况，例如涉及病理性或缺氧驱使的血管生成或新血管形成的医学状况。

本发明的药物组合物包括可含有各种成分的药学上可接受的载体，其中所说的成分可提供调节药物浓度、调节可溶性、化学稳定化、调节粘度、增强吸收率、调节 pH 等功能。例如，水溶性配方中的药物组合物优选包括缓冲液如磷酸缓冲液或 pH 在约 7—8 之间的其他有机酸盐。其他成分可包括抗氧化剂如抗坏血酸、亲水性聚合物如单糖、

二糖及包括纤维素或其衍生物、糊精在内的其他碳水化合物、螯合剂如 EDTA，以及医药科学领域中已知的其他成分(参见 Remington's Phamaceutical Science，最新版，Mack Publishing Company，Easton，PA)。

如 GBS 毒素受体多肽、模拟物或类似物，或者 GBS 毒素模拟物或类似物等活性化合物在特定应用中的有效量取决于几个方面的因素，包括多肽、模拟物或类似物的化学性质、被治疗的疾病、给药方法等。为达到有效量，优选多肽或模拟物在细胞表面上靶 GBS 毒素受体处的浓度约为 0.0001 至 100μM，较好低于 10μM，最好低于 1μM.

可以将活性化合物与各种药用可接受的惰性载体合并制成不同的剂型，如根据选用的给药途径可制成片剂、胶囊剂、锭剂、糖锭、硬糖果、粉末、喷雾剂、酏剂、糖浆、可注射或滴眼溶液等口服或胃肠道外给药制剂。胃肠道外给药包括下列途径：静脉内、肌肉内、皮下、眼内、滑膜内、经皮(包括透皮、眼内、舌下、颊部)、局部(包括眼、皮肤、眼内、直肠、经吹入和气雾剂鼻吸入)及直肠系统。

可将活性化合物与例如惰性稀释剂或可食用载体混合口服给药，可将其装入硬或软壳明胶胶囊、压成片剂，或直接掺入到食物中。为了口服给药，可以将活性化合物掺入赋形剂，并以可摄入的片剂、含片、锭剂、胶囊、酏剂、悬浮剂、糖浆、糯米纸囊剂等剂型使用。这些组合物和制剂应至少含有 0.1%的活性化合物。当然，组合物和制剂的百分比例可以有所变化，并可以在相应剂型单位的大约 2—6%重量之间。这些治疗有用的组合物中活性化合物的量应能保证病人获得适当的剂量。制备本发明的优选组合物或制剂时，应使口服剂量单位含有大约 1—1000mg 活性化合物。

片剂、锭剂、丸剂、胶囊等还可含有下列成分：聚乙烯吡咯烷酮、黄耆胶、阿拉伯胶(acacia)、蔗糖、玉米淀粉或明胶等粘结剂，磷酸钙、柠檬酸钠和碳酸钙等赋形剂，玉米淀粉、马铃薯淀粉、木薯淀粉、甘些复合硅酸盐、海藻酸等崩解剂，十二烷基硫酸钠、滑石粉和硬脂酸镁等润滑剂，蔗糖、乳糖或糖精等甜味剂，薄荷、冬青油或樱桃香

精等香味剂。也可以使用相似类型的固体组合物作为软和硬明胶胶囊的填充剂；这方面的优选材料包括乳糖或高分子量聚乙二醇。当剂量单位为胶囊时，其可另外含有液体载体。可以作为包衣而存在各种其他材料，或借以修饰剂量单位的物理形态。例如，可以用虫胶、糖或两者包被片剂、丸剂或胶囊。糖浆或酏剂可以含有活性化合物、作为甜味剂的蔗糖、作为防腐剂的羟苯甲酸甲酯或丙酯、染料、樱桃或桔橙香味剂、乳化剂和/或悬浮剂、以及水、乙醇、丙二醇、甘油等稀释剂，以及各种组合。当然，用于制备任何剂型的任何材料都应是药品纯的，并且在所使用的量下基本上是无害的。此外，可以将活性化合物掺入到缓释制剂或配方中。

活性化合物也可以经胃肠道外或腹膜内给药。为了经胃肠道外给药，可以使用在芝麻油或花生油，或含水丙二醇中制成的溶液，以及相应水溶性碱金属或碱土金属盐的无菌水溶液。必要时，这些水溶液应是经过适当缓冲的，并用足够的盐水或葡萄糖将液体稀释剂制成等渗液。活性化合物作为游离碱或药用可接受盐的溶液，可在水中与适当的表面活性剂如羟丙基纤维素混合制成。也可以在甘油、液态聚乙二醇及其混合物和油中制成分散剂。在常规储存和使用条件下，这些制剂可含有防止微生物生长的防腐剂。这些特殊的含水溶液特别适于静脉内、肌肉内、皮下和腹膜内注射。在这方面，所使用的无菌含水介质都可用本领域已知的标准技术容易地得到。

适于注射使用的药物剂型包括无菌水溶液或分散剂，或用于临时配制无菌可注射溶液或分散剂的无菌粉末。各种剂型都必须是无菌的，并且必须以易于注射的流体存在。其在制造和储存条件下必须是稳定的，并必须防腐保存以防止细菌和真菌等微生物的污染。载体可以是含有水、乙醇、多元醇(例如甘油、丙二醇、液态聚乙二醇等)、其适当的混合物及植物油的溶剂或分散介质。可以使用卵磷脂等涂层、保持分散剂中颗粒的所需大小及使用表面活性剂来维持适当的流动性。可以用各种抗细菌和抗真菌剂，例如对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸、硫柳汞等防止微生物污染。在许多情况下，优选应包

括等渗剂例如糖或氯化钠。可以使用延迟吸收的药剂，如一硬脂酸铝和明胶来延长可注射组合物的吸收。

将必需量的活性成分掺入到含上述各种其他成分的适当溶剂中，然后过滤除菌即制得无菌注射液。一般说来，将无菌的活性成分掺入含有基本分散介质和上述其他必要成分的无菌载体中，即制得分散剂。为了得到用于配制无菌注射液的无菌粉末，优选用真空干燥或冷冻干燥技术由前述无菌过滤溶液得到活性成分加上所需附加成分的粉末。

为了局部给药，可以在适于滴眼给药的容器内制成相似于上述非胃肠道给药溶液的无菌稀水溶液(通常浓度约 0.1—5%)。本发明的化合物可以单独或与药用载体合并用于哺乳动物。如上所述，活性化合物和载体的相对比例是根据化合物的溶解度和化学性质、选择的给药途径、选用的特定化合物及被治疗病人的生理特征等因素确定的。

试剂盒

本发明的再一个方面是用于完成上述方法的试剂盒。优选的试剂盒包括 GBS 毒素受体或其片段。受体或其片段最好是固相化的。可使用优选的试剂盒鉴定可与 GBS 毒素受体结合的化合物，并至少包括一种表达 GBS 毒素受体的细胞。

另一个具体实例是用于监测肿瘤生长或转移的，包括检测 GBS 毒素受体表达之试剂的试剂盒。这种试剂的例子包括但不只限于可与 GBS 毒素受体核酸序列杂交的多核苷酸探针，及可与 GBS 毒素受体结合的化合物，例如可特异性识别 GBS 毒素受体的抗体、GBS 毒素、GBS 毒素模拟物，或用上述筛选方法鉴定到的其他化合物。

第三个具体实例是纯化可与 GBS 毒素受体结合之化合物的试剂盒，其中包括可与该化合物结合的 GBS 毒素受体或其片段。优选的化合物包括 GBS 毒素、GBS 毒素模拟物、可特异性结合 GBS 毒素受体的抗体，以及用上述筛选方法鉴定到的其他化合物。

试剂盒的附加成分包括但不只限于检测所必需的附加试剂、参考

标准、使用说明书等。适用的参考标准包括阳性对照、阴性对照、这些对照的照片、这些对照的表格或图解数据等。试剂盒中还可包括完成上述方法的说明书，最好是印刷的说明书。

实施例

实施例 1: 绵羊 GBS 毒素受体的克隆

绵羊肺内皮细胞的原代培养物

将从 7 周龄绵羊身上切下的小块原始肺组织放入含有 10mM HEPES 缓冲液(Life Technology)、1%青霉素/链霉素和 0.1%庆大霉素的 Hank's 平衡盐溶液(HBSS)中，并于 37℃在绵羊肺完全培养基(Life Technology)中培养。培养 1 周后，根据 Cobblestone 形态学鉴定绵羊肺内皮细胞的克隆并使用克隆环收获到 24 孔组织培养板(Falcon)中。当细胞生长至连片时，用胰酶制剂离散细胞并转移到 60mm 组织培养皿或 T25 组织培养瓶(Falcon)中。当细胞再次生长至成片时，分散细胞并转入少数几个 100mm 组织培养板(Falcon)中培养。每次分散细胞都认为是一代。重复该方法直到有足够分离 mRNA 的细胞(约 10^8 个)为止。

mRNA 的分离及 cDNA 文库的构建

使用标准方法(Pharmacia)从 9.2×10^7 个绵羊肺内皮细胞(8 和 9 代)中分离 poly(A)$^+$RNA。总共得到 16μg poly (A)$^+$RNA。使用 2.5μg mRNA 构建 cDNA 文库。poly (A)$^+$ RNA 是用 Oligo (dT)引发(带有 Not I 限制性位点)并转化成双链 cDNA。加入 BstX I /EcoR I 衔接子后，将 cDNA 单方向克隆到 pCDNA3.1(+) (Invitrogen)的 BstX I 和 Not I 位点中。使用大肠杆菌 Top10F′ (Invitrogen)作为扩增的宿主菌株。产生 5.38×10^6 个原始克隆应是合适的。将细胞铺在 50 个含有氨苄青霉素(100μg/ml)的大 LB 琼脂板上以扩增文库。刮取平板上的克隆并分成若干等份，以使每个等份代表 10 个平板。用 Qiagen Max 柱(Qiagen)纯化 DNA。

从 cDNA 文库中筛选编码 GBS 毒素受体的基因

使用独特的比色方法筛选 cDNA 文库中编码 GBS 毒素受体的基因。使用各个 cDNA 文库池的 5μg 质粒 DNA 转染 COS7 细胞。在 4—8 个 96 孔组织培养板(Falcon)上培养被转染的细胞以进行瞬时表达。每个孔于 DMEM 培养基(Life Technology)中含有约 20,000 个被转染细胞。使用 pCDNA3.1(+)转染的 COS7 细胞作为对照。表达 3 天之后，小心去掉培养基。各个孔用含有 Mg^{2+} 和 Ca^{2+} 的 HPSS 缓冲液(洗涤缓冲液)(Life Technology)淋洗 3 次。

然后将细胞与生物素化的毒素(50μl/孔；1—1.5μg/ml)在室温下保温 1 小时。保温后，弃去生物素化的毒素并用洗涤缓冲液将各孔洗 3 次。将细胞与链霉抗生物素蛋白-β-gal 溶液温育，并用洗涤缓冲液洗孔 3 次。然后加 PNPG (50μl/孔；在底物缓冲液中 1mg/ml) 37℃保温细胞。分别在 1 和 20 小时时用 ELISA 读数器测出 405nm 吸光值。收获给出最高 OD 值的细胞。用 Hirt 提取法分离质粒 DNA。在大肠杆菌中扩增质粒 DNA 以用于下一步的转染(富集)。

以此比色法富集 8 次。在后几次富集中加入各孔中的被转染细胞的数目逐渐减少，并向各孔中加入未被转染的细胞以使每孔的细胞总数为 20,000 个，以利细胞成片生长并减小 3 天表达后的本底。最后一次富集时，各孔只有 1—10 个被转染的细胞。收获给出最高 OD 值的细胞。分离 DNA 并在大肠杆菌中扩增。

用此比色方法个别检测许多分离的克隆。测定与 CM101 有较高结合率之克隆的序列。

序列分析

对带有 2.1kb 插入片段的克隆 pFU102 进行 DNA 序列分析，揭示出编码部分内在糖蛋白的序列。用 5′ RACE 方法(Life Technology)得到 N 末端序列，并将全长度基因定名为 SP55。三段连接得到含有 SP55 之整个编码区的 pCD55。

SP55 的 mRNA 有 2844 个核苷酸，编码预测分子量 55KDa 的长 495 个氨基酸的蛋白质(SP55)。用 Klein 等人(Klein et al., Biochim. Biophys. Acta 815: 468—476, 1985)的方法分析证明，SP55 为具有

七个跨膜区的膜内在蛋白。SP55 具有 N 糖基化和激酶磷酸化位点。对 SP55 进行 Swiss-Prot. 检索，未发现与已知人蛋白质有高度同源性。然而，SP55 与人、兔、小鼠和大鼠的肾钠依赖性磷酸转运蛋白有一定的相同性(约 30%)。此外，SP55 与 C. elegans 的假设蛋白(HYP50 和 HYP63)有一定的相同性(约 30—39%)。

实施例 2: 人 GBS 毒素受体的克隆

绵羊 GBS 毒素受体序列与 C. elegans 的两种假设蛋白质 HYP50 和 HYP63 分别有大约 37%和 33%的相同性。在相当于 SEQ ID NO: 2 的氨基酸残基 180—186 和 443—449 的区域中，七氨基酸序列段中的 5 个氨基酸在这三种蛋白质之间是绝对保守的。

设计第一个简并寡核苷酸 CMR3-S: 5′ -CGGGATCCCGCCNGCNATGCAYRSHRTSTGG-3′ (SEQ ID NO: 5)，以包括编码 SP55、HYP50 和 HYP63 之 180—186 区域中氨基酸序列的所有可能的密码子。设计第二个简并寡核苷酸 CMR4-AS2: 5′ -GGAATTCCDGGDGCRATKTCNARRTRRTT-3′ (SEQ ID NO: 6)，以包括编码 SP55、HYP50 和 HYP63 之 443—449 区域中氨基酸序列的所有可能密码子的互补序列。

使用这些寡核苷酸并以人胚肺 cDNA 文库作为模板进行聚合酶链反应(PCR)。反应产生了三个长度约 400bp，包括 SEQ ID NO: 3 之部分核酸序列的交迭序列。然后使用这些序列作为探针，克隆基因的其余部分，得到 HP59 (SEQ ID NO: 7)。

实施例 3: 抗 GBS 毒素受体之抗体的制备

用表 8 中所示的合成肽免疫兔。在 0.01M 磷酸盐缓冲液中制备肽加 KLH 的 1mg/ml 溶液。为了进行第一次免疫，充分乳化 200μg 肽加 KLH(200μl)和等体积弗氏完全佐剂，然后在兔的颈和肩部背侧面的 3—4 点注射。两周后，第二次免疫(加强免疫)使用同样浓度的，但在弗氏不完全佐剂中乳化的免疫原。加强免疫的注射部位与第一次免疫相同。两周后，采血并用 ELISA 方法检测动物对无 KLH 之肽的反应。必要时可进行再次加强注射，以进一步提高抗体滴度。

表 8

免疫原性肽

肽	氨基酸序列	长度	序列对照
p56a	APSDGEEGSDRTPLLQRAPRAEPAPVC	27aa	SEQ ID NO:4 的残基 8-35
p55a	LAPSDGEEGSDRTPL	15aa	SEQ ID NO:4 的残基 7-22
p57a	NTTAKDNRTSYECA	14aa	SEQ ID NO:4 的残基 71-84

肽 p55 是 GBS 毒素受体之胞外结构域的片段。肽 p57a 是 GBS 毒素受体之胞内结构域的片段。用这些肽免疫动物分别产生多克隆抗体 Pab55 和 Pab57。

实施例 4: 检测 GBS 毒素受体在肿瘤细胞中的表达

该实施例显示可以在肿瘤细胞中检测到 GBS 毒素受体。使用兔多克隆抗体 Pab55 和 Pab57 对成双的正常或肿瘤来源的人和小鼠组织进行免疫组织化学分析。

在 10%中性福尔马林中固定小鼠和人肿瘤组织。然后将组织脱水、石蜡包埋并切成 10—20×8 微米切片以进行免疫组化染色。

按照生产商推荐的方法（Ventana, Tucson, Arizona），使用自动 Ventana 免疫组化染色器进行免疫组化分析。用二甲苯体对切片脱蜡。然后用在 30%含水甲醇中制备的 1%过氧化氢将切片室温下处理 20 分钟，以淬灭内源性过氧化物酶活性。然后用 PBS 洗切片，用溶于 PBS 的 5%BSA 和 5%山羊血清封闭，再次洗涤后与适当稀释的(1：100)抗体 37℃保温 30 分钟。使用辣根过氧化物酶标记的山羊抗兔 IgG 作为二抗。加 DAB/H_2O_2 保温切片以便观察结果。最后将切片与铜增强剂(Ventana)保温 4 分钟，洗涤并用苏木精复染，然后在无甲苯的封固剂中封固。照相并保存影象以备分析。结果如表 9 中所示。数字是指载玻片号。

表 9

肿瘤和正常组织的免疫组织化学

人组织：

	抗体	放大倍数	信号
1. 卵巢癌（95-02V016）高度乳头状癌	Pab 55	400x	+
2. 卵巢癌（95-02V016）高度乳头状癌	Pab 55	400x	+
3. 正常卵巢（96-08Z008）对照组织	Pab 55	400x	-
4. 卵巢癌（95-02V016）高度乳头状癌	Pab 57	400x	+
5. 卵巢癌（95-02V016）高度乳头状癌	Pab 57	400x	+
6. 卵巢癌（95-02V016）高度乳头状癌	Pab 57	400x	+
7. 正常卵巢（96-08Z008）对照组织	Pab 57	400x	-
8. 结肠癌（95-14664）分化差的腺癌	Pab 55	400x	+
9. 正常结肠（9708V008）对照	Pab 55	400x	-
10. 结肠癌（95-14664）分化差的腺癌	Pab 57	400x	+
11. 结肠癌（95-14664）分化差的腺癌	Pab 57	400x	+
12. 正常结肠（9708V008）对照	Pab 57	400x	-
13. 女性乳腺肿瘤（97-10V03a）	Pab 55	400x	+
侵入性乳腺癌			
14. 男性乳腺肿瘤（无编号）乳腺癌	Pab 55	400x	+
15. 正常女性乳腺（97-12V020-3）对照	Pab 55	400x	-
16. 女性乳腺肿瘤（97-10V03a）	Pab 57	400x	+
侵入性乳腺癌			
17. 男性乳腺肿瘤（无编号）乳腺癌	Pab 57	400x	+
18. 正常女性乳腺（97-12V020-3）对照	Pab 57	400x	-
19. 肺癌（97-1 0V022-5）分化差的 NOJ	Pab 55	400x	+
小细胞癌			
20. 正常肺（98-01V011）对照	Pab 55		-
21. 肺癌（97-1 0V022-5）分化差的 NOJ	Pab 57	400x	+
小细胞癌			

22. 肺癌（97-1 0V022-5）分化差的 NOJ 小细胞癌	Pab 57	400x	+
23. 正常肺（98-01V011）对照	Pab 57		-

小鼠组织:

	抗体	放大倍数	信号
24. 未用 CM101 治疗的 Madison 肺癌（MLT）	Pab 55		+
25. 未用 CM101 治疗的 MLT	Pab 55		+
26. 正常小鼠肺	Pab 55		-
27. 未用 CM101 治疗的 MLT	Pab 57		+
28. 正常小鼠肺	Pab 57		-

Pab55 抗体着染人卵巢癌组织切片中的血管内层细胞，但这种着染在正常人卵巢组织则不明显(见图 2A 和 2B)。用 Pab57 抗体染色可得到相似的结果(见图 3A 和 3B)。如上列表中和图 2A—3B 中所示，抗 GBS 毒素受体片段的抗体可特异地结合肿瘤组织，但不与正常组织结合，提示 GBS 毒素受体是在肿瘤细胞而不是正常细胞中表达的。

实施例 5: 检测 GBS 毒素受体在患类风湿性关节炎小鼠体内的表达

本实施例显示可在类风湿性关节炎(RA)哺乳动物模型的细胞中检测到 GBS 毒素受体。用 CM101 或载体治疗患有胶原蛋白诱发之关节炎的小鼠。CM101 可逆转炎性损伤并抑制血管翳形成。杀死用 CM101 治疗的 8 号和 15 号小鼠及两只对照小鼠(未用 CM101 治疗的)以进行免疫组化分析。

表 10

类风湿性关节炎小鼠的免疫组织化学

29. 未用 CM101	Pab55	+
30. 小鼠 8-5′ （血管）	Pab55	+
31. 未用 CM101	Pab57	+
32. 小鼠 15-5′ （血管）	Pab57	+
33. 小鼠 8-5′ （关节间）	Pab57	+
34. 小鼠 15-5′	Pab57	+
35. 未用 CM101（骨髓）	Pab57	+
36. 小鼠 15-5′ （骨髓）	Pab57	+

如上所示，Pab55 和 Pab57 特异地结合血管翳中的病理性新脉管系统，提示患有类风湿性关节炎的小鼠体内表达 GBS 毒素受体。关节炎小鼠关节的生长板中正常新脉管系统内未见有 CM101 结合。

实施例 6: 向表达 GBS 毒素受体的组织定向送递嵌合化合物

本实施例显示向表达 GBS 毒素受体的组织定向送递嵌合化合物。嵌合化合物是 CM101-生物素缀合物。给患 Madison 肺癌(MLT)的小鼠静脉内灌注生物素化的 CM101。

用肼基化的生物素处理 CM101 以在多糖 CM101 的还原性末端形成生物素腙。简单地说，将 25μg 冻干的 CM101 溶解在 250μl 标记缓冲液（100mM 乙酸钠、0.02%叠氮钠）中。加入含水偏高碘酸(30mM, 125μl)并在暗处室温下氧化 30 分钟。向溶液中加入 80mM Na_2SO_3 以终止反应。使所得到的醛与 125μl 5mM NHS-LC-生物素(分子量 556.58)于室温下反应 1 小时以形成生物素化的 CM101。于 4℃对 1 升 PBS 透析 4 次，以除去过量的生物素。在 Ultrahydrogel 1000HPLC 柱上凝胶过滤纯化产物，冻干并于-70℃储存备用。

输注 CM101 后 5 分钟回收组织并进行免疫组化分析。用小鼠抗 CM101 mAb(7A3)，再用第二 mAb-HRP 缀合物(图 4B 中称之为 MLT CM101-Biot.5′ +McAb)，或与 HRP 缀合的抗生物素蛋白(其特异性地结

合生物素）（图 4A 中称之为 MLT CM101-Biot. 5′ +Strep.HRP)分析肿瘤和正常小鼠组织切片的 CM101 结合。

图 4A—4C 显示取自同一肿瘤的，并包括靠近图中心之同一血管纵切面的不同的切片。图 4A 中的深染色显示血管被覆细胞中生物素成分的定位。同样，图 4B 显示 CM101 成分在血管被覆细胞中的定位。图 4C 为未接触 CM101 的阴性对照组织。分析结果明确表明，7A3 和抗生物素蛋白与肿瘤组织中的同一血管结合。因此，生物素已借助其与可结合 GBS 毒素受体之化合物(CM101)的物理结合被送递到肿瘤组织的血管上。

这些研究表明，嵌合化合物可被送递到经受病理性和/或缺氧驱使的血管生成或新血管形成的组织。作为嵌合化合物的一部分，细胞毒性分子可被导向这些组织，如肿瘤组织。可以将细胞毒性分子直接偶联到可与 GBS 毒素受体结合的分子如 GBS 毒素上。或者，可以将可结合 GBS 毒素受体的分子偶联到生物素上，将细胞毒性分子偶联到抗生物素蛋白上。

实施例 7：提高表达 GBS 毒素受体之细胞对 GBS 毒素依赖性细胞毒性的敏感性

本实施例显示相对于对照组细胞，提高了 GBS 毒素受体转染的细胞对 GBS 毒素依赖性细胞毒性的敏感性。虽然不拘泥于理论，但本发明人相信，补体可结合细胞上的 GBS 毒素受体结合的 GBS 毒素，从而使细胞被白细胞(WBC)杀死。

将人膀胱癌细胞(ECV 细胞)用人 GBS 毒素受体基因稳定转染。所得到的细胞系是 ECV711。只用载体稳定转染的细胞称为 V23。将 ECV711 和 V23 按每孔 5,000 个细胞接种于 96 孔平板中。

按下述方法从健康供体身上收集白细胞。用标准静脉切开放血法将血液收集到肝素化试管(30U/ml)中并以 2000rpm 离心 20 分钟。将界面层小心转移到新的试管中并用培养基(RPMI-1640)离心洗涤两次。将细胞重新悬浮在加有 5%胎牛血清(FBS)和 γ 干扰素(IFN) (100U/ml)的 RPMI 1640 中，并在 5%CO_2 培养箱中 37℃保温过夜。然后将细胞重新

悬浮在加有5%FBS的新鲜培养基中。

将WBC制剂的5,000个细胞加到含有被转染细胞的各小孔中。向小孔内加入终浓度为1μg/ml的CM101及适配人供体的人血清。将细胞37℃保温6小时。

使用Promega公司的Cyto Tox96非放射性检测试剂盒检测乳酸脱氢酶以确定细胞毒性(Nachlas et al., Anal. Biochem. 1:317, 1960; Korzeniewski et al., J. Immunol. Methods 64:313, 1983; Decker et al., J. Immunol. Methods 115:61; Brander et al., Eur. J. Immunology 23:3217, 1993; Behl et al., Cell 77:817, 1994; Lappalainen et al., Pharm. Research 11: 1127, 1994; Allen et al., Promega Notes 45:7, 1994; Sinensky et al., Toxicol. Letters 75:02, 1995; Moravec, Promega Notes 45: 11, 1994)。按生产厂商推荐的方法计算细胞毒性百分比。结果如表11中所示。

表11

细胞毒性	ECV711	V23
WBC, IFN, C3, -CM101	29.1%	27.5%
WBC, IFN, C3, +CM101	40.45%	22.46%

当ECV711细胞与CM101、WBC和人血清(C3来源)一起保温时，细胞毒性比细胞未加CM101保温时增加39%。用单独载体转染的对照细胞V23没有显示出细胞毒性的CM101依赖性增加。

本说明书中提到的所有出版物和专利申请除特别指出者外，均单独地全文列为本文参考文献。

现已全面描述了本发明，但本领域技术人员应能理解到，可以在不背离本发明待批权利要求的精神和范围的前提下，对本发明作出许多改动和改进。

序列表

<110> Hellerqvist, Carl
Fu, Changlin

<120> GBS毒素受体

<130> CARB-008/01WO

<140>
<141>

<150> 60-093,043
<151> 1998-07-22

<160> 12

<170> PatentIn Ver. 2.0

<210> 1
<211> 3602
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (58)..(1542)

<400> 1
tcgggccggc gctcccttct ctgccaggtg gcgagtacac ctgctcacgt aggcgtc 57

atg agg tct ccg gtt cga gac ctg gcc cgg aac gat ggc gag gag agc 105
Met Arg Ser Pro Val Arg Asp Leu Ala Arg Asn Asp Gly Glu Glu Ser
1 5 10 15

acg gac cgc acg cct ctt cta ccg ggc gcc cca cgg gcc gaa gcc gct 153
Thr Asp Arg Thr Pro Leu Leu Pro Gly Ala Pro Arg Ala Glu Ala Ala
20 25 30

cca gtg tgc tgc tct gct cgt tac aac tta gca att ttg gcc ttt ttt 201
Pro Val Cys Cys Ser Ala Arg Tyr Asn Leu Ala Ile Leu Ala Phe Phe
35 40 45

ggt ttc ttc att gtg tat gca tta cgt gtg aat ctg agt gtt gcg tta 249
Gly Phe Phe Ile Val Tyr Ala Leu Arg Val Asn Leu Ser Val Ala Leu
50 55 60

gtg gat atg gta gat tca aat aca act tta gaa gat aat aga act tcc 297

```
Val Asp Met Val Asp Ser Asn Thr Thr Leu Glu Asp Asn Arg Thr Ser
 65                  70                  75                  80

aag gcg tgt cca gag cat tct gct ccc ata aaa gtt cat cat aat caa   345
Lys Ala Cys Pro Glu His Ser Ala Pro Ile Lys Val His His Asn Gln
                 85                  90                  95

acg ggt aag aag tac caa tgg gat gca gaa act caa gga tgg att ctc   393
Thr Gly Lys Lys Tyr Gln Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
            100                 105                 110

ggt tcc ttt ttt tat ggc tac atc atc aca cag att cct gga gga tat   441
Gly Ser Phe Phe Tyr Gly Tyr Ile Ile Thr Gln Ile Pro Gly Gly Tyr
        115                 120                 125

gtt gcc agc aaa ata ggg ggg aaa atg ctg cta gga ttt ggg atc ctt   489
Val Ala Ser Lys Ile Gly Gly Lys Met Leu Leu Gly Phe Gly Ile Leu
    130                 135                 140

ggc act gct gtc ctc acc ctg ttc act ccc att gct gca gat tta gga   537
Gly Thr Ala Val Leu Thr Leu Phe Thr Pro Ile Ala Ala Asp Leu Gly
145                 150                 155                 160

gtt gga cca ctc att gta ctc aga gca cta gaa gga cta gga gag ggt   585
Val Gly Pro Leu Ile Val Leu Arg Ala Leu Glu Gly Leu Gly Glu Gly
                165                 170                 175

gtt aca ttt cca gcc atg cat gcc atg tgg tct tct tgg gct ccc cct   633
Val Thr Phe Pro Ala Met His Ala Met Trp Ser Ser Trp Ala Pro Pro
            180                 185                 190

ctt gaa aga agc aaa ctt ctt agc att tcg tat gca gga gca cag ctt   681
Leu Glu Arg Ser Lys Leu Leu Ser Ile Ser Tyr Ala Gly Ala Gln Leu
        195                 200                 205

ggg aca gta att tct ctt cct ctt tct gga ata att tgc tac tat atg   729
Gly Thr Val Ile Ser Leu Pro Leu Ser Gly Ile Ile Cys Tyr Tyr Met
    210                 215                 220

aat tgg act tat gtc ttc tac ttt ttt ggt act att gga ata ttt tgg   777
Asn Trp Thr Tyr Val Phe Tyr Phe Phe Gly Thr Ile Gly Ile Phe Trp
225                 230                 235                 240

ttt ctt ttg tgg atc tgg tta gtt agt gac aca cca caa aaa cac aag   825
Phe Leu Leu Trp Ile Trp Leu Val Ser Asp Thr Pro Gln Lys His Lys
                245                 250                 255

aga att tcc cat tat gaa aag gaa tac att ctt tca tca tta aga aat   873
```

```
Arg Ile Ser His Tyr Glu Lys Glu Tyr Ile Leu Ser Ser Leu Arg Asn
            260                 265                 270

cag ctt tct tca cag aag tca gtg ccg tgg gta ccc att tta aaa tcc   921
Gln Leu Ser Ser Gln Lys Ser Val Pro Trp Val Pro Ile Leu Lys Ser
        275                 280                 285

ctg cca ctt tgg gct atc gta gtt gca cac ttt tct tac aac tgg act   969
Leu Pro Leu Trp Ala Ile Val Val Ala His Phe Ser Tyr Asn Trp Thr
    290                 295                 300

ttt tat act tta ttg aca tta ttg cct act tat atg aag gag atc cta   1017
Phe Tyr Thr Leu Leu Thr Leu Leu Pro Thr Tyr Met Lys Glu Ile Leu
305                 310                 315                 320

agg ttc aat gtt caa gag aat ggg ttt tta tct tca ttg cct tat tta   1065
Arg Phe Asn Val Gln Glu Asn Gly Phe Leu Ser Ser Leu Pro Tyr Leu
                325                 330                 335

ggc tct tgg tta tgt atg atc ctg tct ggt caa gct gct gac aat tta   1113
Gly Ser Trp Leu Cys Met Ile Leu Ser Gly Gln Ala Ala Asp Asn Leu
            340                 345                 350

agg gca aaa tgg aat ttt tca act tta tgt gtt cgc aga att ttt agc   1161
Arg Ala Lys Trp Asn Phe Ser Thr Leu Cys Val Arg Arg Ile Phe Ser
        355                 360                 365

ctt ata gga atg att gga cct gca gta ttc ctg gta gct gct ggc ttc   1209
Leu Ile Gly Met Ile Gly Pro Ala Val Phe Leu Val Ala Ala Gly Phe
    370                 375                 380

att ggc tgt gat tat tct ttg gcc gtt gct ttc cta act ata tca aca   1257
Ile Gly Cys Asp Tyr Ser Leu Ala Val Ala Phe Leu Thr Ile Ser Thr
385                 390                 395                 400

aca ctg gga ggc ttt tgc tct tct gga ttt agc atc aac cat ctg gat   1305
Thr Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Asp
                405                 410                 415

att gct cct tcg tat gct ggt atc ctc ctg ggc atc aca aat aca ttt   1353
Ile Ala Pro Ser Tyr Ala Gly Ile Leu Leu Gly Ile Thr Asn Thr Phe
            420                 425                 430

gcc act att cca gga atg gtt ggg ccc gtc att gct aaa agt ctg acc   1401
Ala Thr Ile Pro Gly Met Val Gly Pro Val Ile Ala Lys Ser Leu Thr
        435                 440                 445

cct gat aac act gtt gga gaa tgg caa acc gtg ttc tat att gct gct   1449
```

```
Pro Asp Asn Thr Val Gly Glu Trp Gln Thr Val Phe Tyr Ile Ala Ala
    450                 455                 460

gct att aat gtt ttt ggt gcc att ttc ttt aca cta ttc gcc aaa ggt   1497
Ala Ile Asn Val Phe Gly Ala Ile Phe Phe Thr Leu Phe Ala Lys Gly
465                 470                 475                 480

gaa gta caa aac tgg gct ctc aat gat cac cat gga cac aga cac       1542
Glu Val Gln Asn Trp Ala Leu Asn Asp His His Gly His Arg His
                485                 490                 495

tgaaggaacc aataaataat cctgcctcta ttaatgtatt tttatttatc atgtaacctc 1602

aaagtgcctt ctgtattgtg taagcattct atgtcttttt ttaattgtac ttgtattaga 1662

tttttaaggc ctataatcat gaaatatcac tagttgccag aataataaaa tgaactgtgt 1722

ttaattatga ataatatgta agctaggact tctactttag gttcacatac ctgcctgcta 1782

gtcgggcaac atgaagtagg acagttctgt tgatttttta gggccatact aaagggaatg 1842

agctgaaaca gacctcctga taccttttgct taattaaact agatgataat tctcaggtac 1902

tgataaacac ctgttgttgt tcactttcct cataaaaatt gtcagcttct tctgacactt 1962

agacctcaaa ctttagcatc tctgtggagc tgccatccac tgtataattt cgcctggcaa 2022

ctggactgag gggagtgtgc ccaggcagct gccaagcact ccctccctgg cttcagggtc 2082

agagtgccca gcgtttatca gaggcagcat ccaagcccag agccagtgtc gactcttcgg 2142

ctggtgcctt tcctctgagg ggctatcaat gtgtagataa agccctgagt aggcaagagc 2202

agtgagatcc actgctatgg tcttgataca tcctcaaact ttcccttccc agcacagagg 2262

aatattggct ggcatgcaac ctgcaaaaga aaaatgcgaa gcggccgggc acggtggctc 2322

atgcctgtaa tcccagcact ttgggggggct gaggtgggcg aatcatgaga tcaggagttc 2382

gagaccagcc tggccagcat ggtgaaaccc catctctact aaaaatacaa aaaattagct 2442

gggcgtggtg acgggcgcct gtaatcccag atactcagga ggctgaggta ggagaatcac 2502

ttgaacctgg gaggtggaag ttgcagtgaa ccaagatcac gccactgcac tccagcctgg 2562

gcgatggagc gagactccaa ctcaaaaaaa aaaaaaaaaa                       2602
```

<210> 2
<211> 495
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Arg Ser Pro Val Arg Asp Leu Ala Arg Asn Asp Gly Glu Glu Ser
 1               5                  10                  15

Thr Asp Arg Thr Pro Leu Leu Pro Gly Ala Pro Arg Ala Glu Ala Ala
            20                  25                  30

Pro Val Cys Cys Ser Ala Arg Tyr Asn Leu Ala Ile Leu Ala Phe Phe
        35                  40                  45

Gly Phe Phe Ile Val Tyr Ala Leu Arg Val Asn Leu Ser Val Ala Leu
    50                  55                  60

Val Asp Met Val Asp Ser Asn Thr Thr Leu Glu Asp Asn Arg Thr Ser
65                  70                  75                  80

Lys Ala Cys Pro Glu His Ser Ala Pro Ile Lys Val His His Asn Gln
                85                  90                  95

Thr Gly Lys Lys Tyr Gln Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
            100                 105                 110

Gly Ser Phe Phe Tyr Gly Tyr Ile Ile Thr Gln Ile Pro Gly Gly Tyr
        115                 120                 125

Val Ala Ser Lys Ile Gly Gly Lys Met Leu Leu Gly Phe Gly Ile Leu
    130                 135                 140

Gly Thr Ala Val Leu Thr Leu Phe Thr Pro Ile Ala Ala Asp Leu Gly
145                 150                 155                 160

Val Gly Pro Leu Ile Val Leu Arg Ala Leu Glu Gly Leu Gly Glu Gly
                165                 170                 175

Val Thr Phe Pro Ala Met His Ala Met Trp Ser Ser Trp Ala Pro Pro
            180                 185                 190

Leu Glu Arg Ser Lys Leu Leu Ser Ile Ser Tyr Ala Gly Ala Gln Leu
        195                 200                 205

Gly Thr Val Ile Ser Leu Pro Leu Ser Gly Ile Ile Cys Tyr Tyr Met
    210                 215                 220
```

```
Asn Trp Thr Tyr Val Phe Tyr Phe Phe Gly Thr Ile Gly Ile Phe Trp
225                 230                 235                 240

Phe Leu Leu Trp Ile Trp Leu Val Ser Asp Thr Pro Gln Lys His Lys
                245                 250                 255

Arg Ile Ser His Tyr Glu Lys Glu Tyr Ile Leu Ser Ser Leu Arg Asn
            260                 265                 270

Gln Leu Ser Ser Gln Lys Ser Val Pro Trp Val Pro Ile Leu Lys Ser
        275                 280                 285

Leu Pro Leu Trp Ala Ile Val Val Ala His Phe Ser Tyr Asn Trp Thr
    290                 295                 300

Phe Tyr Thr Leu Leu Thr Leu Leu Pro Thr Tyr Met Lys Glu Ile Leu
305                 310                 315                 320

Arg Phe Asn Val Gln Glu Asn Gly Phe Leu Ser Ser Leu Pro Tyr Leu
                325                 330                 335

Gly Ser Trp Leu Cys Met Ile Leu Ser Gly Gln Ala Ala Asp Asn Leu
            340                 345                 350

Arg Ala Lys Trp Asn Phe Ser Thr Leu Cys Val Arg Arg Ile Phe Ser
        355                 360                 365

Leu Ile Gly Met Ile Gly Pro Ala Val Phe Leu Val Ala Ala Gly Phe
    370                 375                 380

Ile Gly Cys Asp Tyr Ser Leu Ala Val Ala Phe Leu Thr Ile Ser Thr
385                 390                 395                 400

Thr Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Asp
                405                 410                 415

Ile Ala Pro Ser Tyr Ala Gly Ile Leu Leu Gly Ile Thr Asn Thr Phe
            420                 425                 430

Ala Thr Ile Pro Gly Met Val Gly Pro Val Ile Ala Lys Ser Leu Thr
        435                 440                 445

Pro Asp Asn Thr Val Gly Glu Trp Gln Thr Val Phe Tyr Ile Ala Ala
    450                 455                 460

Ala Ile Asn Val Phe Gly Ala Ile Phe Phe Thr Leu Phe Ala Lys Gly
465                 470                 475                 480
```

```
Glu Val Gln Asn Trp Ala Leu Asn Asp His His Gly His Arg His
                485                 490                 495


<210> 3
<211> 2844
<212> DNA
<213> Ovis sp.

<220>
<221> CDS
<222> (84)..(1568)

<400> 3
ccccgggggcg gggggcttcg gcggtcccgc tggagctctc ttttccgcgg agcaggtttg 60

cgccgtagct ccctgaaggc atc atg aag tcc ccg gtt tcg gac tta gcc ccg 113
                          Met Lys Ser Pro Val Ser Asp Leu Ala Pro
                            1               5                  10

agc gac ggc gag gag ggc tcg gac cgc aca ccg ctc ctg cag cgc gcc  161
Ser Asp Gly Glu Glu Gly Ser Asp Arg Thr Pro Leu Leu Gln Arg Ala
                  15                  20                  25

ccg cgg gcg gaa ccc gct cca gta tgc tgc tct gct cgt tac aac cta  209
Pro Arg Ala Glu Pro Ala Pro Val Cys Cys Ser Ala Arg Tyr Asn Leu
              30                  35                  40

gca ttt ttg tcc ttt ttt ggt ttc ttc gtt ctc tat tca tta cgg gtg  257
Ala Phe Leu Ser Phe Phe Gly Phe Phe Val Leu Tyr Ser Leu Arg Val
          45                  50                  55

aat ctg agc gtt gca cta gtg gac atg gtg gat tca aac aca act gcc  305
Asn Leu Ser Val Ala Leu Val Asp Met Val Asp Ser Asn Thr Thr Ala
      60                  65                  70

aaa gat aat aga acg tcc tac gag tgt gca gag cat tct gct ccc ata  353
Lys Asp Asn Arg Thr Ser Tyr Glu Cys Ala Glu His Ser Ala Pro Ile
  75                  80                  85                  90

aaa gtt ctt cac aac caa acg ggt aaa aag tac cgg tgg gat gca gaa  401
Lys Val Leu His Asn Gln Thr Gly Lys Lys Tyr Arg Trp Asp Ala Glu
                  95                 100                 105

act caa gga tgg att ctc gga tct ttt ttc tat ggc tac atc atc aca  449
Thr Gln Gly Trp Ile Leu Gly Ser Phe Phe Tyr Gly Tyr Ile Ile Thr
              110                 115                 120
```

```
caa att cct gga gga tat gtt gcc agc aga agt ggg ggg aag ctg ttg   497
Gln Ile Pro Gly Gly Tyr Val Ala Ser Arg Ser Gly Gly Lys Leu Leu
        125                 130                 135

cta gga ttg ggg atc ttt gct aca gct atc ttc acc ctg ttc act ccc   545
Leu Gly Phe Gly Ile Phe Ala Thr Ala Ile Phe Thr Leu Phe Thr Pro
    140                 145                 150

ctc gct gca gat ttc gga gtc gga gcc ctt gtt gca ctc agg gca cta   593
Leu Ala Ala Asp Phe Gly Val Gly Ala Leu Val Ala Leu Arg Ala Leu
155                 160                 165                 170

gaa ggg cta gga gag ggt gtc aca tat cca gcc atg cat gcc atg tgg   641
Glu Gly Leu Gly Glu Gly Val Thr Tyr Pro Ala Met His Ala Met Trp
                175                 180                 185

tct tca tgg gct ccc cct ctt gaa aga agc aag ctt ctg agt att tca   689
Ser Ser Trp Ala Pro Pro Leu Glu Arg Ser Lys Leu Leu Ser Ile Ser
            190                 195                 200

tat gca gga gca caa ctt ggg aca gta gtt tct ctt cct ctt tct gga   737
Tyr Ala Gly Ala Gln Leu Gly Thr Val Val Ser Leu Pro Leu Ser Gly
        205                 210                 215

gta att tgc tac tat atg aat tgg act tat gtc ttc tat ttc ttt ggc   785
Val Ile Cys Tyr Tyr Met Asn Trp Thr Tyr Val Phe Tyr Phe Phe Gly
    220                 225                 230

att gtt gga atc atc tgg ttt att tta tgg atc tgc tta gtt agt gat   833
Ile Val Gly Ile Ile Trp Phe Ile Leu Trp Ile Cys Leu Val Ser Asp
235                 240                 245                 250

aca cca gaa act cac aag aca atc act ccc tat gaa aag gag tat att   881
Thr Pro Glu Thr His Lys Thr Ile Thr Pro Tyr Glu Lys Glu Tyr Ile
                255                 260                 265

ctt tca tca tta aaa aat cag ctc tct tca cag aag tca gtg ccg tgg   929
Leu Ser Ser Leu Lys Asn Gln Leu Ser Ser Gln Lys Ser Val Pro Trp
            270                 275                 280

ata cct atg ctg aaa tca ctg cca ctt tgg gct att gtc gtt gca cat   977
Ile Pro Met Leu Lys Ser Leu Pro Leu Trp Ala Ile Val Val Ala His
        285                 290                 295

ttt tct tac aac tgg act ttt tat act ttg ttg acc tta ttg cct act  1025
Phe Ser Tyr Asn Trp Thr Phe Tyr Thr Leu Leu Thr Leu Leu Pro Thr
    300                 305                 310
```

```
tac atg aag gaa gtc cta agg ttc aat att caa gag aat ggg ttt tta   1073
Tyr Met Lys Glu Val Leu Arg Phe Asn Ile Gln Glu Asn Gly Phe Leu
315                 320                 325                 330

tct gca gtc cct tat tta ggt tgt tgg tta tgt atg atc ctg tcg ggt   1121
Ser Ala Val Pro Tyr Leu Gly Cys Trp Leu Cys Met Ile Leu Ser Gly
                335                 340                 345

caa gct gct gac aat tta agg gca aga tgg aat ttt tca act ctg tgg   1169
Gln Ala Ala Asp Asn Leu Arg Ala Arg Trp Asn Phe Ser Thr Leu Trp
            350                 355                 360

gtt cga aga gtt ttt agc ctt ata ggg atg att gga cct gcg ata ttc   1217
Val Arg Arg Val Phe Ser Leu Ile Gly Met Ile Gly Pro Ala Ile Phe
        365                 370                 375

ctg gtt gcc gca gga ttt ata ggc tgt gat tat tcc ttg gct gtt gca   1265
Leu Val Ala Ala Gly Phe Ile Gly Cys Asp Tyr Ser Leu Ala Val Ala
    380                 385                 390

ttc cta acc ata tca aca acc ctg gga ggc ttt tgc tct tct gga ttt   1313
Phe Leu Thr Ile Ser Thr Thr Leu Gly Gly Phe Cys Ser Ser Gly Phe
395                 400                 405                 410

agc atc aac cat ctg gac att gct cct tcg tat gct ggt att ctc ctg   1361
Ser Ile Asn His Leu Asp Ile Ala Pro Ser Tyr Ala Gly Ile Leu Leu
                415                 420                 425

ggc atc aca aat acc ttt gcc act att cct gga atg att ggg ccc atc   1409
Gly Ile Thr Asn Thr Phe Ala Thr Ile Pro Gly Met Ile Gly Pro Ile
            430                 435                 440

att gcc aga agt ctt acc cct gag aac act att gga gaa tgg caa act   1457
Ile Ala Arg Ser Leu Thr Pro Glu Asn Thr Ile Gly Glu Trp Gln Thr
        445                 450                 455

gtt ttc tgc atc gct gct gct atc aat gta ttt ggt gcc att ttc ttc   1505
Val Phe Cys Ile Ala Ala Ala Ile Asn Val Phe Gly Ala Ile Phe Phe
    460                 465                 470

aca cta ttc gcc aaa ggt gaa gtg caa aac tgg gcc atc agt gat cac   1553
Thr Leu Phe Ala Lys Gly Glu Val Gln Asn Trp Ala Ile Ser Asp His
475                 480                 485                 490

caa gga cac aga aac tgaaggaacc aataaataat cctgtctcta ttaatgtatc   1608
Gln Gly His Arg Asn
                495
```

tttgtttatc atgtaaccta aaagtgcctt tgatatttta atgtgtaagc aatctatata 1668

caagataaaa ttgtactaga aaaattgtgt tagatttgta aggcttgtaa tcatgaaatg 1728

tcactagttg ccatataagc aaaattagct atttttaatt attattaacc cgtttgctgg 1788

aacttacaat tcagggtcac atatctggct gcaagtcagg caacccacaa taggggagtt 1848

ctatttattt ataagaccat acctaaagag atgagctgaa atagaccctt ctatacctttt 1908

gcttaattaa ggtggataat aattctcagg tcttgttaaa catctgtttt tgtacacctt 1968

cctcaaaaaa ttatttgtca tcagcaatcc ctgacatgta ggtctcaaac tttagcctct 2028

ccacggagct ggcagccact gtatcattca gcctggcaac ttcactgagg gaagcatgcc 2088

caggcagctg ccacatgtcc cctctctggc ttcagggaca gtgcccagca cttaggcagc 2148

atccaagacc agggtcagcg ccaaggcttt ggacggtatt ctttcccctgg ggctgttaat 2208

gtgtggatga agccctgagc caacagggac agcgcgatcc acagtcatgg tttccatgca 2268

ccctctccct tcccttccca gcacactgga gtattgcctg gcatgtaacc tgcaaaagaa 2328

agtgtgatgc ctaattagcc acatataaca tcatccttga tgatcctacc ttcacatgga 2388

tcagagtata aatcttcaag tcctgtgttc taggagctac accagaataa ttaaaatata 2448

aaaagaaaca aaacattttt tctgtctgac acctaagtgt ctggttgcag ttcaaggtta 2508

aagtgacttc tacttcacat aacctgcaac cggtggtgta atcatctttа gtgttggttt 2568

cttaaatctt atttttccag tttttcctgg accatcttcc agtggttttg agcatgcttt 2628

gagggcattt atgtgattta gaacttgatt aatgtttcac tgtgtatgtt caacactacc 2688

tgtaatattt taactaaagc tatttaatgt aatatgatgt gtatacattc tgtaaattaa 2748

tttttaaatc tgtaaatagc tttaagttgc tatggtgata tttcttttac aaatcaaaat 2808

aaatcttttt ggaatgataa aaaaaaaaaa aaaaaa 2844

<210> 4
<211> 495
<212> PRT
<213> Ovis sp.

<400> 4

```
Met Lys Ser Pro Val Ser Asp Leu Ala Pro Ser Asp Gly Glu Glu Gly
 1               5                  10                  15

Ser Asp Arg Thr Pro Leu Leu Gln Arg Ala Pro Arg Ala Glu Pro Ala
            20                  25                  30

Pro Val Cys Cys Ser Ala Arg Tyr Asn Leu Ala Phe Leu Ser Phe Phe
        35                  40                  45

Gly Phe Phe Val Leu Tyr Ser Leu Arg Val Asn Leu Ser Val Ala Leu
    50                  55                  60

Val Asp Met Val Asp Ser Asn Thr Thr Ala Lys Asp Asn Arg Thr Ser
65                  70                  75                  80

Tyr Glu Cys Ala Glu His Ser Ala Pro Ile Lys Val Leu His Asn Gln
                85                  90                  95

Thr Gly Lys Lys Tyr Arg Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
            100                 105                 110

Gly Ser Phe Phe Tyr Gly Tyr Ile Ile Thr Gln Ile Pro Gly Gly Tyr
        115                 120                 125

Val Ala Ser Arg Ser Gly Gly Lys Leu Leu Leu Gly Phe Gly Ile Phe
    130                 135                 140

Ala Thr Ala Ile Phe Thr Leu Phe Thr Pro Leu Ala Ala Asp Phe Gly
145                 150                 155                 160

Val Gly Ala Leu Val Ala Leu Arg Ala Leu Glu Gly Leu Gly Glu Gly
                165                 170                 175

Val Thr Tyr Pro Ala Met His Ala Met Trp Ser Ser Trp Ala Pro Pro
            180                 185                 190

Leu Glu Arg Ser Lys Leu Leu Ser Ile Ser Tyr Ala Gly Ala Gln Leu
        195                 200                 205

Gly Thr Val Val Ser Leu Pro Leu Ser Gly Val Ile Cys Tyr Tyr Met
    210                 215                 220

Asn Trp Thr Tyr Val Phe Tyr Phe Phe Gly Ile Val Gly Ile Ile Trp
225                 230                 235                 240

Phe Ile Leu Trp Ile Cys Leu Val Ser Asp Thr Pro Glu Thr His Lys
                245                 250                 255
```

Thr Ile Thr Pro Tyr Glu Lys Glu Tyr Ile Leu Ser Ser Leu Lys Asn
260 265 270

Gln Leu Ser Ser Gln Lys Ser Val Pro Trp Ile Pro Met Leu Lys Ser
275 280 285

Leu Pro Leu Trp Ala Ile Val Val Ala His Phe Ser Tyr Asn Trp Thr
290 295 300

Phe Tyr Thr Leu Leu Thr Leu Leu Pro Thr Tyr Met Lys Glu Val Leu
305 310 315 320

Arg Phe Asn Ile Gln Glu Asn Gly Phe Leu Ser Ala Val Pro Tyr Leu
325 330 335

Gly Cys Trp Leu Cys Met Ile Leu Ser Gly Gln Ala Ala Asp Asn Leu
340 345 350

Arg Ala Arg Trp Asn Phe Ser Thr Leu Trp Val Arg Arg Val Phe Ser
355 360 365

Leu Ile Gly Met Ile Gly Pro Ala Ile Phe Leu Val Ala Ala Gly Phe
370 375 380

Ile Gly Cys Asp Tyr Ser Leu Ala Val Ala Phe Leu Thr Ile Ser Thr
385 390 395 400

Thr Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Asp
405 410 415

Ile Ala Pro Ser Tyr Ala Gly Ile Leu Leu Gly Ile Thr Asn Thr Phe
420 425 430

Ala Thr Ile Pro Gly Met Ile Gly Pro Ile Ile Ala Arg Ser Leu Thr
435 440 445

Pro Glu Asn Thr Ile Gly Glu Trp Gln Thr Val Phe Cys Ile Ala Ala
450 455 460

Ala Ile Asn Val Phe Gly Ala Ile Phe Phe Thr Leu Phe Ala Lys Gly
465 470 475 480

Glu Val Gln Asn Trp Ala Ile Ser Asp His Gln Gly His Arg Asn
485 490 495

<210> 5

<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述: PCR 引物

<400> 5
cgggatcccg ccngcnatgc ayrshrtstg g 31

<210> 6
<211> 29
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述: PCR 引物

<400> 6
ggaattccdg gdgcratktc narrtrrtt 29

<210> 7
<211> 2930
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (263)..(1870)

<400> 7
gttcggtcga agccctcccc ttaattatgt gcaattcaag tccccactgc ccgcccgcaa 60

gcccccactc atcctcgctg cgggcagggt ggcccctgca ctttacaagg gggtgcagga 120

gcgggagacg gtcgtccgaa caccggctcc ccggcatgcg tagaccggcg ggcggagcgg 180

gctcactttg cgccaatcct acgagaactc ccagaactcc gcttccctag tccaacccaa 240

gccagagttg cccacaccta ag atg gcg gcg ggg gcg atg aca ccg ccc cgc 292
Met Ala Ala Gly Ala Met Thr Pro Pro Arg
1 5 10

ccg gtc cag cca gct cgg ccc ggg ggc ttc ggg ctg tcg ggc cgg cgc 340
Pro Val Gln Pro Ala Arg Pro Gly Gly Phe Gly Leu Ser Gly Arg Arg
15 20 25

```
tcc ctt ctc tgc cag gtg gcg agt aca cct gct cac gta ggc gtc atg   388
Ser Leu Leu Cys Gln Val Ala Ser Thr Pro Ala His Val Gly Val Met
             30                  35                  40

agg tct ccg gtt cga gac ctg gcc cgg aac gat ggc gag gag agc acg   436
Arg Ser Pro Val Arg Asp Leu Ala Arg Asn Asp Gly Glu Glu Ser Thr
         45                  50                  55

gac cgc acg cct ctt cta ccg ggc gcc cca cgg gcc gaa gcc gct cca   484
Asp Arg Thr Pro Leu Leu Pro Gly Ala Pro Arg Ala Glu Ala Ala Pro
     60                  65                  70

gtg tgc tgc tct gct cgt tac aac tta gca att ttg gcc ttt ttt ggt   532
Val Cys Cys Ser Ala Arg Tyr Asn Leu Ala Ile Leu Ala Phe Phe Gly
 75                  80                  85                  90

ttc ttc att gtg tat gca tta cgt gtg aat ctg agt gtt gcg tta gtg   580
Phe Phe Ile Val Tyr Ala Leu Arg Val Asn Leu Ser Val Ala Leu Val
                 95                 100                 105

gat atg gta gat tca aat aca act tta gaa gat aat aga act tcc aag   628
Asp Met Val Asp Ser Asn Thr Thr Leu Glu Asp Asn Arg Thr Ser Lys
            110                 115                 120

gcg tgt cca gag cat tct gct ccc ata aaa gtt cat cat aat caa acg   676
Ala Cys Pro Glu His Ser Ala Pro Ile Lys Val His His Asn Gln Thr
        125                 130                 135

ggt aag aag tac caa tgg gat gca gaa act caa gga tgg att ctc ggt   724
Gly Lys Lys Tyr Gln Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu Gly
    140                 145                 150

tcc ttt ttt tat ggc tac atc atc aca cag att cct gga gga tat gtt   772
Ser Phe Phe Tyr Gly Tyr Ile Ile Thr Gln Ile Pro Gly Gly Tyr Val
155                 160                 165                 170

gcc agc aaa ata ggg ggg aaa atg ctg cta gga ttt ggg atc ctt ggc   820
Ala Ser Lys Ile Gly Gly Lys Met Leu Leu Gly Phe Gly Ile Leu Gly
                175                 180                 185

act gct gtc ctc acc ctg ttc act ccc att gct gca gat tta gga gtt   868
Thr Ala Val Leu Thr Leu Phe Thr Pro Ile Ala Ala Asp Leu Gly Val
            190                 195                 200

gga cca ctc att gta ctc aga gca cta gaa gga cta gga gag ggt gtt   916
Gly Pro Leu Ile Val Leu Arg Ala Leu Glu Gly Leu Gly Glu Gly Val
        205                 210                 215
```

```
aca ttt cca gcc atg cat gcc atg tgg tct tct tgg gct ccc cct ctt   964
Thr Phe Pro Ala Met His Ala Met Trp Ser Ser Trp Ala Pro Pro Leu
    220                 225                 230

gaa aga agc aaa ctt ctt agc att tcg tat gca gga gca cag ctt ggg   1012
Glu Arg Ser Lys Leu Leu Ser Ile Ser Tyr Ala Gly Ala Gln Leu Gly
235                 240                 245                 250

aca gta att tct ctt cct ctt tct gga ata att tgc tac tat atg aat   1060
Thr Val Ile Ser Leu Pro Leu Ser Gly Ile Ile Cys Tyr Tyr Met Asn
                255                 260                 265

tgg act tat gtc ttc tac ttt ttt ggt act att gga ata ttt tgg ttt   1108
Trp Thr Tyr Val Phe Tyr Phe Phe Gly Thr Ile Gly Ile Phe Trp Phe
            270                 275                 280

ctt ttg tgg atc tgg tta gtt agt gac aca cca caa aaa cac aag aga   1156
Leu Leu Trp Ile Trp Leu Val Ser Asp Thr Pro Gln Lys His Lys Arg
        285                 290                 295

att tcc cat tat gaa aag gaa tac att ctt tca tca tta aga aat cag   1204
Ile Ser His Tyr Glu Lys Glu Tyr Ile Leu Ser Ser Leu Arg Asn Gln
    300                 305                 310

ctt tct tca cag aag tca gtg ccg tgg gta ccc att tta aaa tcc ctg   1252
Leu Ser Ser Gln Lys Ser Val Pro Trp Val Pro Ile Leu Lys Ser Leu
315                 320                 325                 330

cca ctt tgg gct atc gta gtt gca cac ttt tct tac aac tgg act ttt   1300
Pro Leu Trp Ala Ile Val Val Ala His Phe Ser Tyr Asn Trp Thr Phe
                335                 340                 345

tat act tta ttg aca tta ttg cct act tat atg aag gag atc cta agg   1348
Tyr Thr Leu Leu Thr Leu Leu Pro Thr Tyr Met Lys Glu Ile Leu Arg
            350                 355                 360

ttc aat gtt caa gag aat ggg ttt tta tct tca ttg cct tat tta ggc   1396
Phe Asn Val Gln Glu Asn Gly Phe Leu Ser Ser Leu Pro Tyr Leu Gly
        365                 370                 375

tct tgg tta tgt atg atc ctg tct ggt caa gct gct gac aat tta agg   1444
Ser Trp Leu Cys Met Ile Leu Ser Gly Gln Ala Ala Asp Asn Leu Arg
    380                 385                 390

gca aaa tgg aat ttt tca act tta tgt gtt cgc aga att ttt agc ctt   1492
Ala Lys Trp Asn Phe Ser Thr Leu Cys Val Arg Arg Ile Phe Ser Leu
395                 400                 405                 410
```

```
ata gga atg att gga cct gca gta ttc ctg gta gct gct ggc ttc att   1540
Ile Gly Met Ile Gly Pro Ala Val Phe Leu Val Ala Ala Gly Phe Ile
                415                 420                 425

ggc tgt gat tat tct ttg gcc gtt gct ttc cta act ata tca aca aca   1588
Gly Cys Asp Tyr Ser Leu Ala Val Ala Phe Leu Thr Ile Ser Thr Thr
            430                 435                 440

ctg gga ggc ttt tgc tct tct gga ttt agc atc aac cat ctg gat att   1636
Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Asp Ile
        445                 450                 455

gct cct tcg tat gct ggt atc ctc ctg ggc atc aca aat aca ttt gcc   1684
Ala Pro Ser Tyr Ala Gly Ile Leu Leu Gly Ile Thr Asn Thr Phe Ala
    460                 465                 470

act att cca gga atg gtt ggg ccc gtc att gct aaa agt ctg acc cct   1732
Thr Ile Pro Gly Met Val Gly Pro Val Ile Ala Lys Ser Leu Thr Pro
475                 480                 485                 490

gat aac act gtt gga gaa tgg caa acc gtg ttc tat att gct gct gct   1780
Asp Asn Thr Val Gly Glu Trp Gln Thr Val Phe Tyr Ile Ala Ala Ala
                495                 500                 505

att aat gtt ttt ggt gcc att ttc ttt aca cta ttc gcc aaa ggt gaa   1828
Ile Asn Val Phe Gly Ala Ile Phe Phe Thr Leu Phe Ala Lys Gly Glu
            510                 515                 520

gta caa aac tgg gct ctc aat gat cac cat gga cac aga cac           1870
Val Gln Asn Trp Ala Leu Asn Asp His His Gly His Arg His
        525                 530                 535

tgaaggaacc aataaataat cctgcctcta ttaatgtatt tttatttatc atgtaacctc 1930

aaagtgcctt ctgtattgtg taagcattct atgtcttttt ttaattgtac ttgtattaga 1990

tttttaaggc ctataatcat gaaatatcac tagttgccag aataataaaa tgaactgtgt 2050

ttaattatga ataatatgta agctaggact tctactttag gttcacatac ctgcctgcta 2110

gtcgggcaac atgaagtagg acagttctgt tgatttttta gggccatact aaagggaatg 2170

agctgaaaca gacctcctga tacctttgct taattaaact agatgataat tctcaggtac 2230

tgataaacac ctgttgttgt tcactttcct cataaaaatt gtcagctctc tctgacactt 2290

agacctcaaa ctttagcatc tctgtggagc tgccatccac tgtataattt cgcctggcaa 2350
```

ctggactgag gggagtgtgc ccaggcagct gccaagcact ccctccctgg cttcagggtc 2410

agagtgccca gcgtttatca gaggcagcat ccaagcccag agccagtgtc gactcttcgg 2470

ctggtgcctt tcctctgagg ggctatcaat gtgtagataa agccctgagt aggcaagagc 2530

agtgagatcc actgctatgg tcttgataca tcctcaaact ttcccttccc agcacagagg 2590

aatattggct ggcatgcaac ctgcaaaaga aaaatgcgaa gcggccgggc acggtggctc 2650

atgcctgtaa tcccagcact ttgggggggct gaggtgggcg aatcatgaga tcaggagttc 2710

gagaccagcc tggccagcat ggtgaaaccc catctctact aaaaatacaa aaaattagct 2770

gggcgtggtg acgggcgcct gtaatcccag atacttcagga ggctgaggta ggagaatcac 2830

ttgaacctgg gaggtggaag ttgcagtgaa ccaagatcac gccactgcac tccagcctgg 2890

gcgatggagc gagactccaa ctcaaaaaaa aaaaaaaaaa 2930

<210> 8
<211> 536
<212> PRT
<213> Homo sapiens

<400> 8

Met Ala Ala Gly Ala Met Thr Pro Pro Arg Pro Val Gln Pro Ala Arg
1 5 10 15

Pro Gly Gly Phe Gly Leu Ser Gly Arg Arg Ser Leu Leu Cys Gln Val
20 25 30

Ala Ser Thr Pro Ala His Val Gly Val Met Arg Ser Pro Val Arg Asp
35 40 45

Leu Ala Arg Asn Asp Gly Glu Glu Ser Thr Asp Arg Thr Pro Leu Leu
50 55 60

Pro Gly Ala Pro Arg Ala Glu Ala Ala Pro Val Cys Cys Ser Ala Arg
65 70 75 80

Tyr Asn Leu Ala Ile Leu Ala Phe Phe Gly Phe Phe Ile Val Tyr Ala
85 90 95

Leu Arg Val Asn Leu Ser Val Ala Leu Val Asp Met Val Asp Ser Asn
100 105 110

```
Thr Thr Leu Glu Asp Asn Arg Thr Ser Lys Ala Cys Pro Glu His Ser
        115                 120                 125

Ala Pro Ile Lys Val His His Asn Gln Thr Gly Lys Lys Tyr Gln Trp
    130                 135                 140

Asp Ala Glu Thr Gln Gly Trp Ile Leu Gly Ser Phe Phe Tyr Gly Tyr
145                 150                 155                 160

Ile Ile Thr Gln Ile Pro Gly Gly Tyr Val Ala Ser Lys Ile Gly Gly
                165                 170                 175

Lys Met Leu Leu Gly Phe Gly Ile Leu Gly Thr Ala Val Leu Thr Leu
            180                 185                 190

Phe Thr Pro Ile Ala Ala Asp Leu Gly Val Gly Pro Leu Ile Val Leu
        195                 200                 205

Arg Ala Leu Glu Gly Leu Gly Glu Gly Val Thr Phe Pro Ala Met His
    210                 215                 220

Ala Met Trp Ser Ser Trp Ala Pro Pro Leu Glu Arg Ser Lys Leu Leu
225                 230                 235                 240

Ser Ile Ser Tyr Ala Gly Ala Gln Leu Gly Thr Val Ile Ser Leu Pro
                245                 250                 255

Leu Ser Gly Ile Ile Cys Tyr Tyr Met Asn Trp Thr Tyr Val Phe Tyr
            260                 265                 270

Phe Phe Gly Thr Ile Gly Ile Phe Trp Phe Leu Leu Trp Ile Trp Leu
        275                 280                 285

Val Ser Asp Thr Pro Gln Lys His Lys Arg Ile Ser His Tyr Glu Lys
    290                 295                 300

Glu Tyr Ile Leu Ser Ser Leu Arg Asn Gln Leu Ser Ser Gln Lys Ser
305                 310                 315                 320

Val Pro Trp Val Pro Ile Leu Lys Ser Leu Pro Leu Trp Ala Ile Val
                325                 330                 335

Val Ala His Phe Ser Tyr Asn Trp Thr Phe Tyr Thr Leu Leu Thr Leu
            340                 345                 350

Leu Pro Thr Tyr Met Lys Glu Ile Leu Arg Phe Asn Val Gln Glu Asn
        355                 360                 365
```

```
Gly Phe Leu Ser Ser Leu Pro Tyr Leu Gly Ser Trp Leu Cys Met Ile
    370                 375                 380

Leu Ser Gly Gln Ala Ala Asp Asn Leu Arg Ala Lys Trp Asn Phe Ser
385                 390                 395                 400

Thr Leu Cys Val Arg Arg Ile Phe Ser Leu Ile Gly Met Ile Gly Pro
                405                 410                 415

Ala Val Phe Leu Val Ala Ala Gly Phe Ile Gly Cys Asp Tyr Ser Leu
            420                 425                 430

Ala Val Ala Phe Leu Thr Ile Ser Thr Thr Leu Gly Gly Phe Cys Ser
        435                 440                 445

Ser Gly Phe Ser Ile Asn His Leu Asp Ile Ala Pro Ser Tyr Ala Gly
    450                 455                 460

Ile Leu Leu Gly Ile Thr Asn Thr Phe Ala Thr Ile Pro Gly Met Val
465                 470                 475                 480

Gly Pro Val Ile Ala Lys Ser Leu Thr Pro Asp Asn Thr Val Gly Glu
                485                 490                 495

Trp Gln Thr Val Phe Tyr Ile Ala Ala Ala Ile Asn Val Phe Gly Ala
            500                 505                 510

Ile Phe Phe Thr Leu Phe Ala Lys Gly Glu Val Gln Asn Trp Ala Leu
        515                 520                 525

Asn Asp His His Gly His Arg His
    530                 535
```

<210> 9
<211> 1485
<212> DNA
<213> 人工序列
<220>
<223> 人工序列的描述：人/绵羊共有序列

<220>
<221> CDS
<222> (1)..(1485)

<400> 9

```
atg arg tcy ccg gtt ysr gac ytr gcc csg arc gay ggc gag gag rgc   48
Met Xaa Xaa Pro Val Xaa Asp Xaa Ala Xaa Xaa Xaa Gly Glu Glu Xaa
  1               5                  10                  15

wcg gac cgc acr cck cty ctr cmg sgc gcc ccr cgg gcs gaa scc gct   96
Xaa Asp Arg Xaa Xaa Xaa Xaa Xaa Xaa Ala Xaa Arg Xaa Glu Xaa Ala
            20                  25                  30

cca gtr tgc tgc tct gct cgt tac aac yta gca wtt ttg kcc ttt ttt  144
Pro Xaa Cys Cys Ser Ala Arg Tyr Asn Xaa Ala Xaa Leu Xaa Phe Phe
        35                  40                  45

ggt ttc ttc rtt sts tat kca tta cgk gtg aat ctg agy gtt gcr yta  192
Gly Phe Phe Xaa Xaa Tyr Xaa Leu Xaa Val Asn Leu Xaa Val Xaa Xaa
    50                  55                  60

gtg gay atg gtr gat tca aay aca act kym raa gat aat aga ack tcc  240
Val Xaa Met Xaa Asp Ser Xaa Thr Thr Xaa Xaa Asp Asn Arg Xaa Ser
65                  70                  75                  80

was gmg tgt sca gag cat tct gct ccc ata aaa gtt cwt cay aay caa  288
Xaa Xaa Cys Xaa Glu His Ser Ala Pro Ile Lys Val Xaa Xaa Xaa Gln
                85                  90                  95

acg ggt aar aag tac crr tgg gat gca gaa act caa gga tgg att ctc  336
Thr Gly Xaa Lys Tyr Xaa Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
            100                 105                 110

ggw tcy ttt tty tat ggc tac atc atc aca car att cct gga gga tat  384
Xaa Xaa Phe Xaa Tyr Gly Tyr Ile Ile Thr Xaa Ile Pro Gly Gly Tyr
        115                 120                 125

gtt gcc agc ara akw ggg ggg aar mtg ytg cta gga tty ggg atc ytt  432
Val Ala Ser Xaa Xaa Gly Gly Xaa Xaa Xaa Leu Gly Xaa Gly Ile Xaa
    130                 135                 140

gsy acw gct rtc ytc acc ctg ttc act ccc mty gct gca gat ttm gga  480
Xaa Xaa Ala Xaa Xaa Thr Leu Phe Thr Pro Xaa Ala Ala Asp Xaa Gly
145                 150                 155                 160

gty gga scm cty rtt gya ctc agr gca cta gaa ggr cta gga gag ggt  528
Xaa Gly Xaa Xaa Xaa Xaa Leu Xaa Ala Leu Glu Xaa Leu Gly Glu Gly
                165                 170                 175

gty aca twt cca gcc atg cat gcc atg tgg tct tcw tgg gct ccc cct  576
Xaa Thr Xaa Pro Ala Met His Ala Met Trp Ser Xaa Trp Ala Pro Pro
            180                 185                 190
```

```
ctt gaa aga agc aar ctt ctk agy att tcr tat gca gga gca car ctt   624
Leu Glu Arg Ser Xaa Leu Xaa Xaa Ile Xaa Tyr Ala Gly Ala Xaa Leu
        195                 200                 205

ggg aca gta rtt tct ctt cct ctt tct gga rta att tgc tac tat atg   672
Gly Thr Val Xaa Ser Leu Pro Leu Ser Gly Xaa Ile Cys Tyr Tyr Met
    210                 215                 220

aat tgg act tat gtc ttc tay tty ttt ggy ayt rtt gga atm wty tgg   720
Asn Trp Thr Tyr Val Phe Xaa Xaa Phe Xaa Xaa Xaa Gly Xaa Xaa Trp
225                 230                 235                 240

ttt mtt ttr tgg atc tgs tta gtt agt gay aca cca saa amw cac aag   768
Phe Xaa Xaa Trp Ile Xaa Leu Val Ser Xaa Thr Pro Xaa Xaa His Lys
                245                 250                 255

asa aty wcy cmy tat gaa aag gar tay att ctt tca tca tta ara aat   816
Xaa Xaa Xaa Xaa Tyr Glu Lys Xaa Xaa Ile Leu Ser Ser Leu Xaa Asn
            260                 265                 270

cag cty tct tca cag aag tca gtg ccg tgg rta ccy atk ytr aaa tcm   864
Gln Xaa Ser Ser Gln Lys Ser Val Pro Trp Xaa Xaa Xaa Xaa Lys Xaa
        275                 280                 285

ctg cca ctt tgg gct aty gtm gtt gca cay ttt tct tac aac tgg act   912
Leu Pro Leu Trp Ala Xaa Xaa Val Ala Xaa Phe Ser Tyr Asn Trp Thr
    290                 295                 300

ttt tat act ttr ttg acm tta ttg cct act tay atg aag gar rtc cta   960
Phe Tyr Thr Xaa Leu Xaa Leu Leu Pro Thr Xaa Met Lys Xaa Xaa Leu
305                 310                 315                 320

agg ttc aat rtt caa gag aat ggg ttt tta tct kca kts cct tat tta  1008
Arg Phe Asn Xaa Gln Glu Asn Gly Phe Leu Ser Xaa Xaa Pro Tyr Leu
                325                 330                 335

ggy tst tgg tta tgt atg atc ctg tck ggt caa gct gct gac aat tta  1056
Xaa Xaa Trp Leu Cys Met Ile Leu Xaa Gly Gln Ala Ala Asp Asn Leu
            340                 345                 350

agg gca ara tgg aat ttt tca act ytr tgk gtt cgm aga rtt ttt agc  1104
Arg Ala Xaa Trp Asn Phe Ser Thr Xaa Xaa Val Xaa Arg Xaa Phe Ser
        355                 360                 365

ctt ata ggr atg att gga cct gcr rta ttc ctg gtw gcy gcw ggm tty  1152
Leu Ile Xaa Met Ile Gly Pro Xaa Xaa Phe Leu Xaa Xaa Xaa Xaa Xaa
    370                 375                 380
```

```
atw ggc tgt gat tat tcy ttg gcy gtt gcw ttc cta acy ata tca aca   1200
Xaa Gly Cys Asp Tyr Xaa Leu Xaa Val Xaa Phe Leu Xaa Ile Ser Thr
385                 390                 395                 400

acm ctg gga ggc ttt tgc tct tct gga ttt agc atc aac cat ctg gay   1248
Xaa Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Xaa
                405                 410                 415

att gct cct tcg tat gct ggt aty ctc ctg ggc atc aca aat acm ttt   1296
Ile Ala Pro Ser Tyr Ala Gly Xaa Leu Leu Gly Ile Thr Asn Xaa Phe
            420                 425                 430

gcc act att ccw gga atg rtt ggg ccc rtc att gcy ara agt ctk acc   1344
Ala Thr Ile Xaa Gly Met Xaa Gly Pro Xaa Ile Xaa Xaa Ser Xaa Thr
        435                 440                 445

cct gak aac act rtt gga gaa tgg caa acy gtk ttc try aty gct gct   1392
Pro Xaa Asn Thr Xaa Gly Glu Trp Gln Xaa Xaa Phe Xaa Xaa Ala Ala
    450                 455                 460

gct aty aat gtw ttt ggt gcc att ttc tty aca cta ttc gcc aaa ggt   1440
Ala Xaa Asn Xaa Phe Gly Ala Ile Phe Xaa Thr Leu Phe Ala Lys Gly
465                 470                 475                 480

gaa gtr caa aac tgg gcy mtc art gat cac caw gga cac aga mac       1485
Glu Xaa Gln Asn Trp Xaa Xaa Xaa Asp His Xaa Gly His Arg Xaa
                485                 490                 495
```

<210> 10
<211> 495
<212> PRT
<213> 人工序列

```
<400> 10
Met Xaa Xaa Pro Val Xaa Asp Xaa Ala Xaa Xaa Xaa Gly Glu Glu Xaa
  1               5                  10                  15

Xaa Asp Arg Xaa Xaa Xaa Xaa Xaa Xaa Ala Xaa Arg Xaa Glu Xaa Ala
            20                  25                  30

Pro Xaa Cys Cys Ser Ala Arg Tyr Asn Xaa Ala Xaa Leu Xaa Phe Phe
        35                  40                  45

Gly Phe Phe Xaa Xaa Tyr Xaa Leu Xaa Val Asn Leu Xaa Val Xaa Xaa
    50                  55                  60
```

Val Xaa Met Xaa Asp Ser Xaa Thr Thr Xaa Xaa Asp Asn Arg Xaa Ser
65 70 75 80

Xaa Xaa Cys Xaa Glu His Ser Ala Pro Ile Lys Val Xaa Xaa Xaa Gln
85 90 95

Thr Gly Xaa Lys Tyr Xaa Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
100 105 110

Xaa Xaa Phe Xaa Tyr Gly Tyr Ile Ile Thr Xaa Ile Pro Gly Gly Tyr
115 120 125

Val Ala Ser Xaa Xaa Gly Gly Xaa Xaa Xaa Leu Gly Xaa Gly Ile Xaa
130 135 140

Xaa Xaa Ala Xaa Xaa Thr Leu Phe Thr Pro Xaa Ala Ala Asp Xaa Gly
145 150 155 160

Xaa Gly Xaa Xaa Xaa Xaa Leu Xaa Ala Leu Glu Xaa Leu Gly Glu Gly
165 170 175

Xaa Thr Xaa Pro Ala Met His Ala Met Trp Ser Xaa Trp Ala Pro Pro
180 185 190

Leu Glu Arg Ser Xaa Leu Xaa Xaa Ile Xaa Tyr Ala Gly Ala Xaa Leu
195 200 205

Gly Thr Val Xaa Ser Leu Pro Leu Ser Gly Xaa Ile Cys Tyr Tyr Met
210 215 220

Asn Trp Thr Tyr Val Phe Xaa Xaa Phe Xaa Xaa Xaa Gly Xaa Xaa Trp
225 230 235 240

Phe Xaa Xaa Trp Ile Xaa Leu Val Ser Xaa Thr Pro Xaa Xaa His Lys
245 250 255

Xaa Xaa Xaa Xaa Tyr Glu Lys Xaa Xaa Ile Leu Ser Ser Leu Xaa Asn
260 265 270

Gln Xaa Ser Ser Gln Lys Ser Val Pro Trp Xaa Xaa Xaa Xaa Lys Xaa
275 280 285

Leu Pro Leu Trp Ala Xaa Xaa Val Ala Xaa Phe Ser Tyr Asn Trp Thr
290 295 300

Phe Tyr Thr Xaa Leu Xaa Leu Leu Pro Thr Xaa Met Lys Xaa Xaa Leu
305 310 315 320

```
Arg Phe Asn Xaa Gln Glu Asn Gly Phe Leu Ser Xaa Xaa Pro Tyr Leu
                325                 330                 335

Xaa Xaa Trp Leu Cys Met Ile Leu Xaa Gly Gln Ala Ala Asp Asn Leu
            340                 345                 350

Arg Ala Xaa Trp Asn Phe Ser Thr Xaa Xaa Val Xaa Arg Xaa Phe Ser
        355                 360                 365

Leu Ile Xaa Met Ile Gly Pro Xaa Xaa Phe Leu Xaa Xaa Xaa Xaa Xaa
    370                 375                 380

Xaa Gly Cys Asp Tyr Xaa Leu Xaa Val Xaa Phe Leu Xaa Ile Ser Thr
385                 390                 395                 400

Xaa Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Xaa
                405                 410                 415

Ile Ala Pro Ser Tyr Ala Gly Xaa Leu Leu Gly Ile Thr Asn Xaa Phe
            420                 425                 430

Ala Thr Ile Xaa Gly Met Xaa Gly Pro Xaa Ile Xaa Xaa Ser Xaa Thr
        435                 440                 445

Pro Xaa Asn Thr Xaa Gly Glu Trp Gln Xaa Xaa Phe Xaa Xaa Ala Ala
    450                 455                 460

Ala Xaa Asn Xaa Phe Gly Ala Ile Phe Xaa Thr Leu Phe Ala Lys Gly
465                 470                 475                 480

Glu Xaa Gln Asn Trp Xaa Xaa Xaa Asp His Xaa Gly His Arg Xaa
                485                 490                 495
```

<210> 11
<211> 1485
<212> DNA
〈213〉 人工序列
〈220〉
〈223〉 人工序列的描述：人/绵羊共有序列

<220>
<221> CDS
<222> (1)..(1485)

<400> 11

```
atg ang tcn ccg gtt nnn gac ntn gcc cng anc gan ggc gag gag ngc   48
Met Xaa Xaa Pro Val Xaa Asp Xaa Ala Xaa Xaa Xaa Gly Glu Glu Xaa
 1               5                  10                  15

ncg gac cgc acn ccn ctn ctn cng ngc gcc ccn cgg gcn gaa ncc gct   96
Xaa Asp Arg Xaa Xaa Xaa Xaa Xaa Xaa Ala Xaa Arg Xaa Glu Xaa Ala
            20                  25                  30

cca gtn tgc tgc tct gct cgt tac aac nta gca ntt ttg ncc ttt ttt  144
Pro Xaa Cys Cys Ser Ala Arg Tyr Asn Xaa Ala Xaa Leu Xaa Phe Phe
        35                  40                  45

ggt ttc ttc ntt ntn tat nca tta cgn gtg aat ctg agn gtt gcn nta  192
Gly Phe Phe Xaa Xaa Tyr Xaa Leu Xaa Val Asn Leu Xaa Val Xaa Xaa
    50                  55                  60

gtg gan atg gtn gat tca aan aca act nnn naa gat aat aga acn tcc  240
Val Xaa Met Xaa Asp Ser Xaa Thr Thr Xaa Xaa Asp Asn Arg Xaa Ser
65                  70                  75                  80

nan gng tgt nca gag cat tct gct ccc ata aaa gtt cnt can aan caa  288
Xaa Xaa Cys Xaa Glu His Ser Ala Pro Ile Lys Val Xaa Xaa Xaa Gln
                85                  90                  95

acg ggt aan aag tac cnn tgg gat gca gaa act caa gga tgg att ctc  336
Thr Gly Xaa Lys Tyr Xaa Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
            100                 105                 110

ggn tcn ttt ttn tat ggc tac atc atc aca can att cct gga gga tat  384
Xaa Xaa Phe Xaa Tyr Gly Tyr Ile Ile Thr Xaa Ile Pro Gly Gly Tyr
        115                 120                 125

gtt gcc agc ana ann ggg ggg aan ntg ntg cta gga ttn ggg atc ntt  432
Val Ala Ser Xaa Xaa Gly Gly Xaa Xaa Xaa Leu Gly Xaa Gly Ile Xaa
    130                 135                 140

gnn acn gct ntc ntc acc ctg ttc act ccc ntn gct gca gat ttn gga  480
Xaa Xaa Ala Xaa Xaa Thr Leu Phe Thr Pro Xaa Ala Ala Asp Xaa Gly
145                 150                 155                 160

gtn gga ncn ctn ntt gna ctc agn gca cta gaa ggn cta gga gag ggt  528
Xaa Gly Xaa Xaa Xaa Xaa Leu Xaa Ala Leu Glu Xaa Leu Gly Glu Gly
                165                 170                 175

gtn aca tnt cca gcc atg cat gcc atg tgg tct tcn tgg gct ccc cct  576
Xaa Thr Xaa Pro Ala Met His Ala Met Trp Ser Xaa Trp Ala Pro Pro
            180                 185                 190
```

```
ctt gaa aga agc aan ctt ctn agn att tcn tat gca gga gca can ctt   624
Leu Glu Arg Ser Xaa Leu Xaa Xaa Ile Xaa Tyr Ala Gly Ala Xaa Leu
        195                 200                 205

ggg aca gta ntt tct ctt cct ctt tct gga nta att tgc tac tat atg   672
Gly Thr Val Xaa Ser Leu Pro Leu Ser Gly Xaa Ile Cys Tyr Tyr Met
    210                 215                 220

aat tgg act tat gtc ttc tan ttn ttt ggn ant ntt gga atn ntn tgg   720
Asn Trp Thr Tyr Val Phe Xaa Xaa Phe Xaa Xaa Xaa Gly Xaa Xaa Trp
225                 230                 235                 240

ttt ntt ttn tgg atc tgn tta gtt agt gan aca cca naa ann cac aag   768
Phe Xaa Xaa Trp Ile Xaa Leu Val Ser Xaa Thr Pro Xaa Xaa His Lys
                245                 250                 255

ana atn ncn cnn tat gaa aag gan tan att ctt tca tca tta ana aat   816
Xaa Xaa Xaa Xaa Tyr Glu Lys Xaa Xaa Ile Leu Ser Ser Leu Xaa Asn
            260                 265                 270

cag ctn tct tca cag aag tca gtg ccg tgg nta ccn atn ntn aaa tcn   864
Gln Xaa Ser Ser Gln Lys Ser Val Pro Trp Xaa Xaa Xaa Xaa Lys Xaa
        275                 280                 285

ctg cca ctt tgg gct atn gtn gtt gca can ttt tct tac aac tgg act   912
Leu Pro Leu Trp Ala Xaa Xaa Val Ala Xaa Phe Ser Tyr Asn Trp Thr
    290                 295                 300

ttt tat act ttn ttg acn tta ttg cct act tan atg aag gan ntc cta   960
Phe Tyr Thr Xaa Leu Xaa Leu Leu Pro Thr Xaa Met Lys Xaa Xaa Leu
305                 310                 315                 320

agg ttc aat ntt caa gag aat ggg ttt tta tct nca ntn cct tat tta  1008
Arg Phe Asn Xaa Gln Glu Asn Gly Phe Leu Ser Xaa Xaa Pro Tyr Leu
                325                 330                 335

ggn tnt tgg tta tgt atg atc ctg tcn ggt caa gct gct gac aat tta  1056
Xaa Xaa Trp Leu Cys Met Ile Leu Xaa Gly Gln Ala Ala Asp Asn Leu
            340                 345                 350

agg gca ana tgg aat ttt tca act ntn tgn gtt cgn aga ntt ttt agc  1104
Arg Ala Xaa Trp Asn Phe Ser Thr Xaa Xaa Val Xaa Arg Xaa Phe Ser
        355                 360                 365

ctt ata ggn atg att gga cct gcn nta ttc ctg gtn gcn gcn ggn ttn  1152
Leu Ile Xaa Met Ile Gly Pro Xaa Xaa Phe Leu Xaa Xaa Xaa Xaa Xaa
    370                 375                 380
```

```
atn ggc tgt gat tat tcn ttg gcn gtt gcn ttc cta acn aca tca aca   1200
Xaa Gly Cys Asp Tyr Xaa Leu Xaa Val Xaa Phe Leu Xaa Ile Ser Thr
385                 390                 395                 400

acn ctg gga ggc ttt tgc tct tct gga ttt agc atc aac cat ctg gan   1248
Xaa Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Xaa
                405                 410                 415

att gct cct tcg tat gct ggt atn ctc ctg ggc atc aca aat acn ttt   1296
Ile Ala Pro Ser Tyr Ala Gly Xaa Leu Leu Gly Ile Thr Asn Xaa Phe
            420                 425                 430

gcc act att ccn gga atg ntt ggg ccc ntc att gcn ana agt ctn acc   1344
Ala Thr Ile Xaa Gly Met Xaa Gly Pro Xaa Ile Xaa Xaa Ser Xaa Thr
        435                 440                 445

cct gan aac act ntt gga gaa tgg caa acn gtn ttc tnn atn gct gct   1392
Pro Xaa Asn Thr Xaa Gly Glu Trp Gln Xaa Xaa Phe Xaa Xaa Ala Ala
    450                 455                 460

gct atn aat gtn ttt ggt gcc att ttc ttn aca cta ttc gcc aaa ggt   1440
Ala Xaa Asn Xaa Phe Gly Ala Ile Phe Xaa Thr Leu Phe Ala Lys Gly
465                 470                 475                 480

gaa gtn caa aac tgg gcn ntc ant gat cac can gga cac aga nac       1485
Glu Xaa Gln Asn Trp Xaa Xaa Xaa Asp His Xaa Gly His Arg Xaa
                485                 490                 495
```

<210> 12
<211> 495
<212> PRT
<213> 人工序列

<400> 12

```
Met Xaa Xaa Pro Val Xaa Asp Xaa Ala Xaa Xaa Xaa Gly Glu Glu Xaa
  1               5                  10                  15

Xaa Asp Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ala Xaa Arg Xaa Glu Xaa Ala
            20                  25                  30

Pro Xaa Cys Cys Ser Ala Arg Tyr Asn Xaa Ala Xaa Leu Xaa Phe Phe
        35                  40                  45

Gly Phe Phe Xaa Xaa Tyr Xaa Leu Xaa Val Asn Leu Xaa Val Xaa Xaa
    50                  55                  60

Val Xaa Met Xaa Asp Ser Xaa Thr Thr Xaa Xaa Asp Asn Arg Xaa Ser
```

65 70 75 80

Xaa Xaa Cys Xaa Glu His Ser Ala Pro Ile Lys Val Xaa Xaa Xaa Gln
85 90 95

Thr Gly Xaa Lys Tyr Xaa Trp Asp Ala Glu Thr Gln Gly Trp Ile Leu
100 105 110

Xaa Xaa Phe Xaa Tyr Gly Tyr Ile Ile Thr Xaa Ile Pro Gly Gly Tyr
115 120 125

Val Ala Ser Xaa Xaa Gly Gly Xaa Xaa Xaa Leu Gly Xaa Gly Ile Xaa
130 135 140

Xaa Xaa Ala Xaa Xaa Thr Leu Phe Thr Pro Xaa Ala Ala Asp Xaa Gly
145 150 155 160

Xaa Gly Xaa Xaa Xaa Xaa Leu Xaa Ala Leu Glu Xaa Leu Gly Glu Gly
165 170 175

Xaa Thr Xaa Pro Ala Met His Ala Met Trp Ser Xaa Trp Ala Pro Pro
180 185 190

Leu Glu Arg Ser Xaa Leu Xaa Xaa Ile Xaa Tyr Ala Gly Ala Xaa Leu
195 200 205

Gly Thr Val Xaa Ser Leu Pro Leu Ser Gly Xaa Ile Cys Tyr Tyr Met
210 215 220

Asn Trp Thr Tyr Val Phe Xaa Xaa Phe Xaa Xaa Xaa Gly Xaa Xaa Trp
225 230 235 240

Phe Xaa Xaa Trp Ile Xaa Leu Val Ser Xaa Thr Pro Xaa Xaa His Lys
245 250 255

Xaa Xaa Xaa Xaa Tyr Glu Lys Xaa Xaa Ile Leu Ser Ser Leu Xaa Asn
260 265 270

Gln Xaa Ser Ser Gln Lys Ser Val Pro Trp Xaa Xaa Xaa Xaa Lys Xaa
275 280 285

Leu Pro Leu Trp Ala Xaa Xaa Val Ala Xaa Phe Ser Tyr Asn Trp Thr
290 295 300

Phe Tyr Thr Xaa Leu Xaa Leu Leu Pro Thr Xaa Met Lys Xaa Xaa Leu
305 310 315 320

Arg Phe Asn Xaa Gln Glu Asn Gly Phe Leu Ser Xaa Xaa Pro Tyr Leu

325 330 335

Xaa Xaa Trp Leu Cys Met Ile Leu Xaa Gly Gln Ala Ala Asp Asn Leu
340 345 350

Arg Ala Xaa Trp Asn Phe Ser Thr Xaa Xaa Val Xaa Arg Xaa Phe Ser
355 360 365

Leu Ile Xaa Met Ile Gly Pro Xaa Xaa Phe Leu Xaa Xaa Xaa Xaa Xaa
370 375 380

Xaa Gly Cys Asp Tyr Xaa Leu Xaa Val Xaa Phe Leu Xaa Ile Ser Thr
385 390 395 400

Xaa Leu Gly Gly Phe Cys Ser Ser Gly Phe Ser Ile Asn His Leu Xaa
405 410 415

Ile Ala Pro Ser Tyr Ala Gly Xaa Leu Leu Gly Ile Thr Asn Xaa Phe
420 425 430

Ala Thr Ile Xaa Gly Met Xaa Gly Pro Xaa Ile Xaa Xaa Ser Xaa Thr
435 440 445

Pro Xaa Asn Thr Xaa Gly Glu Trp Gln Xaa Xaa Phe Xaa Xaa Ala Ala
450 455 460

Ala Xaa Asn Xaa Phe Gly Ala Ile Phe Xaa Thr Leu Phe Ala Lys Gly
465 470 475 480

Glu Xaa Gln Asn Trp Xaa Xaa Xaa Asp His Xaa Gly His Arg Xaa
485 490 495

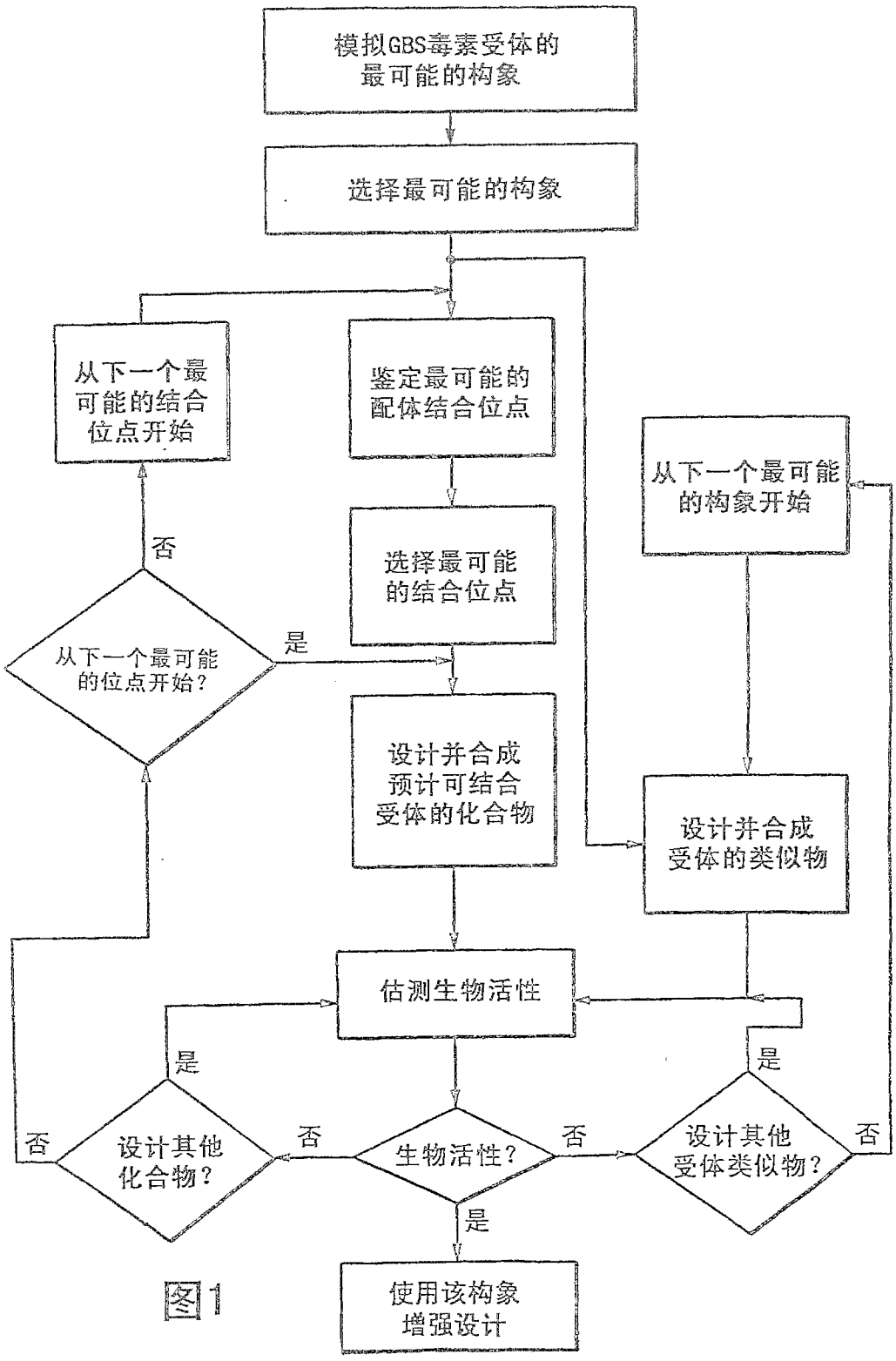
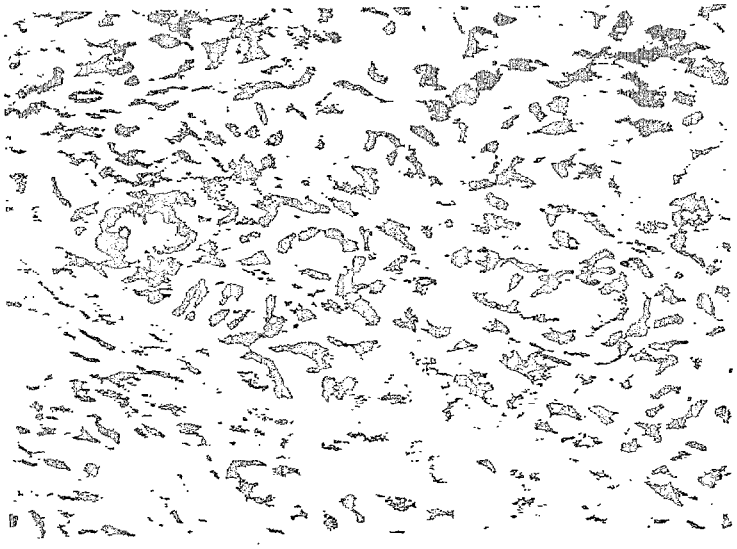

模拟GBS毒素受体的最可能的构象
选择最可能的构象
从下一个最可能的结合位点开始
鉴定最可能的配体结合位点
从下一个最可能的构象开始
否
选择最可能的结合位点
从下一个最可能的位点开始？
是
设计并合成预计可结合受体的化合物
设计并合成受体的类似物
估测生物活性
是
是
否
设计其他化合物？
否
生物活性？
否
设计其他受体类似物？
否
是
使用该构象增强设计

图1



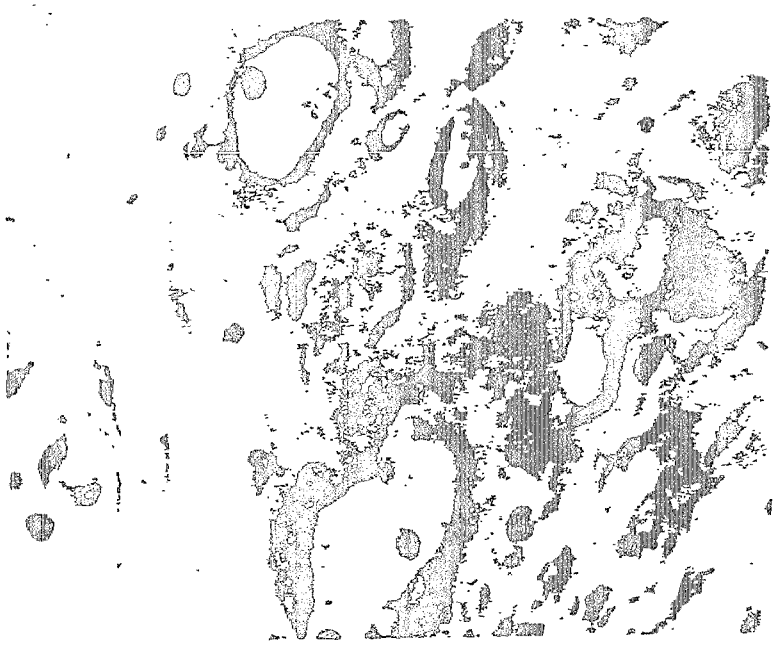
人卵巢癌+Pab 55

图 2A



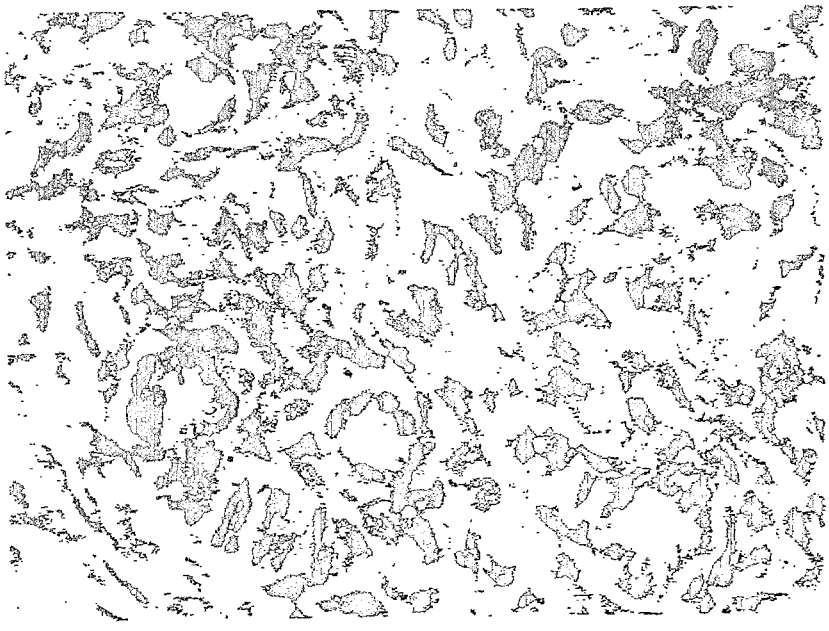
正常人卵巢+Pab 55

图 2B



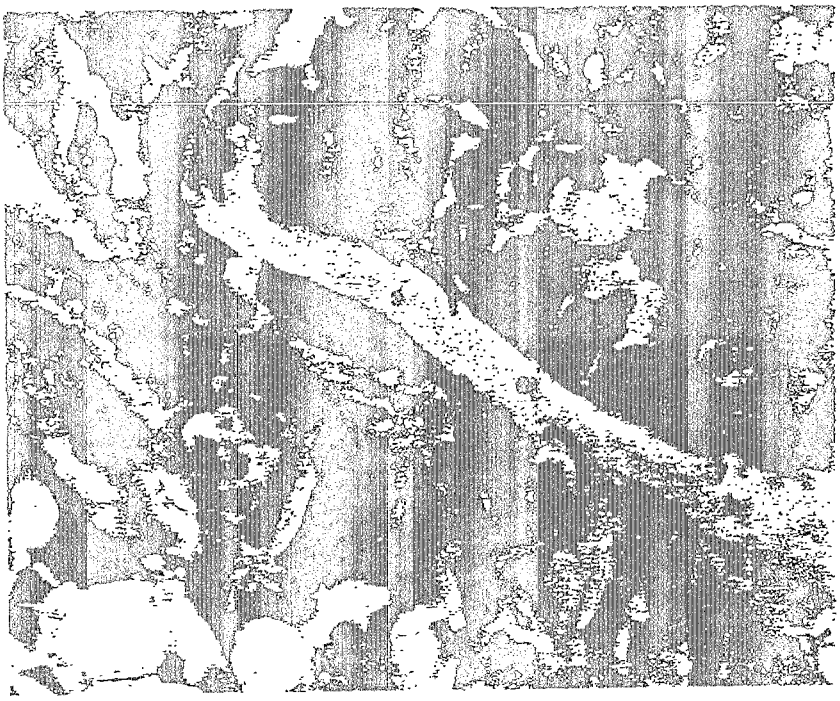
人卵巢癌+Pab 57

图 3A



正常人卵巢+Pab 57

图 3B



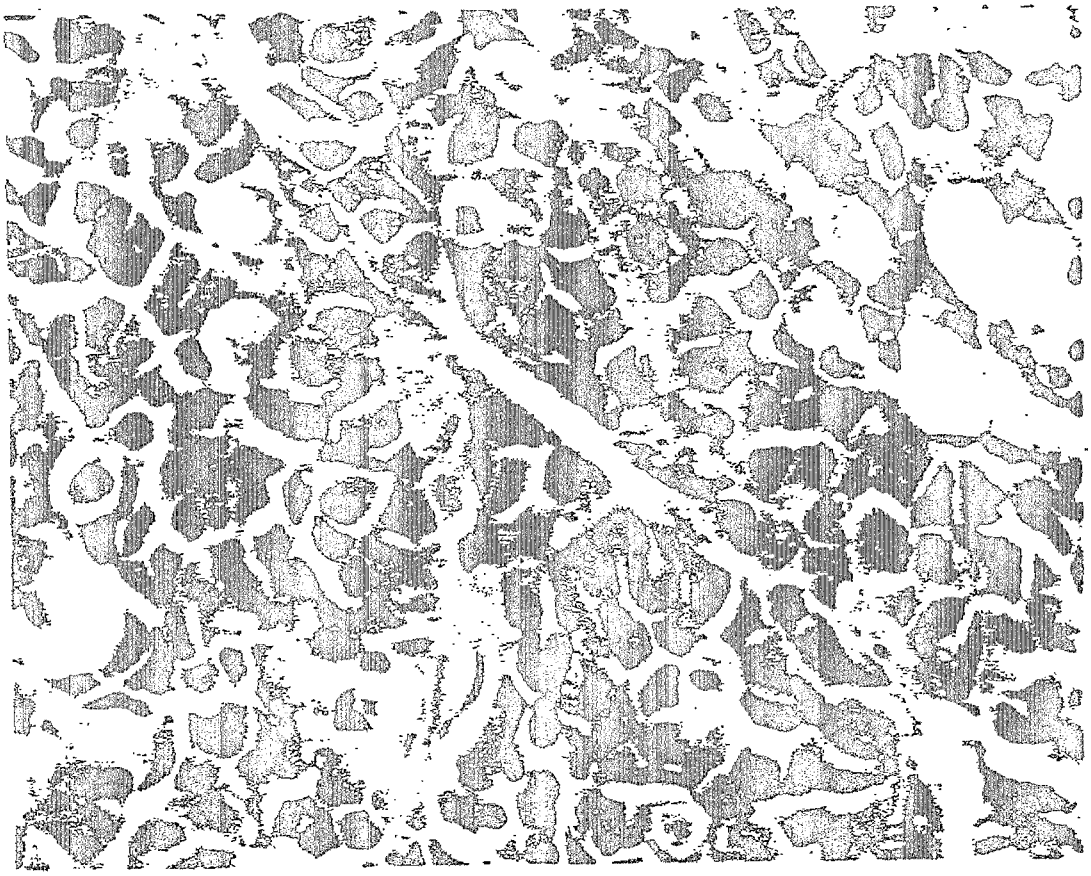
MLT CM101 - Biot.5'+ Strep.HRP

图 4A



MLT CM101 - Biot.5' + mAb

图 4B



MLT - PBS 5' + Streptavidin - HRP

图 4C