

[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 200380108789.2

[51] Int. Cl.

A61K 39/00 (2006.01)

A61K 39/395 (2006.01)

C07K 14/00 (2006.01)

A61P 7/00 (2006.01)

[45] 授权公告日 2010年3月17日

[11] 授权公告号 CN 100594037C

[22] 申请日 2003.11.13

[21] 申请号 200380108789.2

[30] 优先权

[32] 2002.11.15 [33] US [31] 60/426,676

[86] 国际申请 PCT/US2003/036459 2003.11.13

[87] 国际公布 WO2004/045520 英 2004.6.3

[85] 进入国家阶段日期 2005.7.14

[73] 专利权人 MUSC 研究发展基金会

地址 美国南卡罗来纳州

共同专利权人 科罗拉多大学董事会

[72] 发明人 S·汤姆林森 M·V·霍勒斯

[56] 参考文献

US6458360B1 2002.10.1

审查员 苏 林

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 隗永良

权利要求书 2 页 说明书 133 页 附图 31 页

[54] 发明名称

通过补体受体 2 定向的补体调节剂

[57] 摘要

对补体系统的调节是与补体激活有关的大量病理疾病的一种治疗形式。在制备靶向补体激活和疾病位点的补体抑制剂的策略中，公开了这样的组合物，该组合物包含与补体受体(CR)2连接的补体抑制剂。所公开的该组合物可以通过调节补体系统用于治疗病原性疾病和炎性疾病的方法中。

- 1、一种包含构建体的组合物，其中该构建体包含：
  - (a) CR2 或其片段，其中该片段包含 CR2 蛋白质的至少前两个 N-末端 SCR 结构域；和
  - (b) 补体活性调节剂，其中该补体活性调节剂包含补体抑制剂。
- 2、权利要求 1 的组合物，其中 CR2 或其片段包含 CR2 蛋白质的前四个 N-末端 SCR 结构域。
- 3、权利要求 1 的组合物，其中 CR2 包含全长 CR2 蛋白质。
- 4、如权利要求 1-3 之任一项所述的组合物，其中该构建体是融合蛋白。
- 5、权利要求 4 的组合物，其中 CR2 或其片段与补体活性抑制剂的 N 端或 C 端融合。
- 6、如权利要求 5 所述的组合物，其中补体抑制剂包含衰变加速因子 (DAF) 的前四个 SCR 结构域、人 CD59 的细胞外区域、鼠 CD59 的细胞外区域、Crry 的前 5 个 SCR 结构域、MCP 的前 4 个 SCR 结构域、补体受体 1 (CR1) 的胞外域、CR1 的前 4 个 SCR 结构域、CR1 的 SCR 结构域 8-11、或 CR1 的 SCR 结构域 15-18。
- 7、如权利要求 5 所述的组合物，其中补体抑制剂选自 DAF、人 CD59、鼠 CD59、Crry、MCP 和 CR1。
- 8、如权利要求 5-7 之任一项所述的组合物，其中该组合物包含 SEQ ID NO: 10 (DAF-CR2)、SEQ ID NO:6 (CR2-DAF)、SEQ ID NO: 12 (人 CD59-CR2)、SEQ ID NO: 8 (CR2-人 CD59)、SEQ ID NO: 17 (Crry)、SEQ ID NO: 14 (CR1)、或 SEQ ID NO: 16 (MCP)。
- 9、如权利要求 1 所述的组合物，其中该构建体是 CR2 或其片段与补体抑制剂的免疫缀合物。
- 10、权利要求 1-9 中任一项的组合物在制造用于治疗对象中病毒感染、真菌感染或炎症疾病的药物中的用途。

---

11、如权利要求 10 所述的用途，其中该炎性疾病是中风或局部缺血/再灌注损伤。

12、权利要求 1-9 中任一项的组合物在制造用于减少对象中补体介导的损伤的药物中的用途。

13、核苷酸序列，其编码权利要求 2 的融合蛋白。

## 通过补体受体 2 定向的补体调节剂

本申请要求 2002 年 11 月 15 日提交的美国临时申请号 60/426,676 的权益，这里将此临时申请的整体并入为参考文献。

### 背景技术

补体是一系列血液蛋白的总称，并且是免疫系统的主要效应子机制。补体激活和其在靶结构上的沉积可以引起直接的补体介导的细胞裂解，或者可以由于有效的炎症调节剂的产生和免疫效应细胞的募集和激活而间接地引起细胞或组织破坏。在补体途径中的各个点产生介导组织损害的补体激活产物。宿主组织上不适当的补体激活在许多自身免疫病和炎症疾病的病理学中起重要作用，并且也是造成与例如心肺炎症和移植排斥后的生物不相容性有关的许多病状的原因。补体抑制是治疗这些免疫介导的疾病和病状的潜在治疗方式。已经证明了系统地抑制补体的补体抑制蛋白在各种疾病的动物模型中（和在一些临床试验中）是有效的，但是在安全性和有效性方面，靶向疾病和补体激活位点的补体抑制剂提供了显著潜在的优点。

在健康个体中，通过在细胞表面表达补体抑制蛋白阻止宿主细胞膜上的补体沉积。这些补体抑制蛋白也常常以增加的水平在肿瘤细胞表面上表达，并且被认为是肿瘤细胞抵抗单克隆抗体（靶向肿瘤细胞和激活补体的单克隆抗体）介导的免疫治疗的重要起作用因素。

补体系统包含约 30 种蛋白质，并且是免疫系统的主要效应子机制之一。补体级联主要通过经典（通常是依赖于抗体的）或旁路（通常是非依赖于抗体的）途径激活。通过两种途径中任何一种途径的激活都引起了 C3 转化酶的产生，该 C3 转化酶是该级联的中心酶复合物。C3 转化酶将血清 C3 切割成 C3a 和 C3b，其中 C3b 共价地结合激活位点，并且引起 C3 转

化酶的进一步产生（放大循环）。激活产物 C3b（和只通过经典途径产生的 C4b）和其分解产物是重要的调理素，并且参与促进细胞介导的靶细胞裂解（通过吞噬细胞和 NK 细胞）及免疫复合物转运和溶解。C3/C4 激活产物和它们在免疫系统的各种细胞上的受体在调节细胞免疫应答方面也是重要的。C3 转化酶参与 C5 转化酶形成，C5 转化酶是一种复合物，切割 C5 以产生 C5a 和 C5b。C5a 具有有效的促炎和趋化特性，并且可以募集和激活免疫效应细胞。C5b 的形成起始了终末补体途径，引起了补体蛋白质 C6, C7, C8 和 (C9)<sub>n</sub> 的顺序组装以形成膜攻击复合物 (MAC 或 C5b-9)。靶细胞膜中 MAC 的形成可以引起直接的细胞裂解，但是也可以引起细胞激活和各种炎性调节剂的表达/释放。

存在两大类的膜补体抑制剂；补体激活途径的抑制剂（抑制 C3 转化酶形成），和终末补体途径的抑制剂（抑制 MAC 形成）。补体激活的膜抑制剂包括补体受体 1 (CR1)，衰变加速因子 (DAF) 和膜辅因子蛋白 (MCP)。它们都具有如下蛋白质结构，该蛋白质结构由可变数量的称为短共有重复 (SCR) 的约 60-70 个氨基酸的重复单元组成，该短共有重复 (SCR) 是 C3/C4 结合蛋白的共同特征。已经鉴定了人补体激活抑制剂的啮齿类动物的同系物。啮齿类动物蛋白质 Crry 是广泛分布的补体激活抑制剂，其作用与 DAF 和 MCP 相似。尽管 Crry 看起来是啮齿类动物中功能上最重要的补体激活调节剂，但是啮齿类动物也表达 DAF 和 MCP。尽管在人类中没有发现 Crry 同系物，但是在动物模型中 Crry 的研究和其用途具有临床相关性。

终末补体途径和宿主细胞膜上 MAC 的形成主要通过 CD59 活性来控制，CD59 是广泛分布的 20kD 糖蛋白，其通过葡糖基磷脂酰肌醇 (GPI) 锚附着质膜。在组装 MAC 时 CD59 结合 C8 和 C9，并且阻止膜插入。

目前正在研究用于治疗与生物不相容性相关的炎症疾病和病状的各种类型的补体抑制蛋白。两种最具治疗特征的人补体抑制剂是可溶形式的补体受体 1 (sCR1) 和抗 C5 单克隆抗体。这些系统活性的抑制蛋白在各种疾病的动物模型中以及最近在临床试验中已经显示了效力 (1-5, 6: #

1037)。抗 C5 mAb 抑制 C5a 和 MAC 的产生，而 sCR1 是补体激活的抑制剂，并且也抑制 C3 激活产物的产生。在炎性和生物不相容性动物模型中，也已经证明可溶形式的人衰变加速因子(DAF)和膜辅因子蛋白(MCP)（补体激活的膜抑制剂）具有保护性（7-11）。CD59 是补体的膜抑制剂，其阻断 MAC 的装配，但是不引起补体调理素或 C3a 和 C5a 的产生。已经产生了可溶形式的 CD59，但是在体外，特别是在血清存在的情况下，其低的功能活性表明 sCD59 几乎没有或根本没有治疗效力（12-15）。

将补体抑制剂靶向补体激活和疾病位点可能改进它们的效力。因为补体在宿主防御和免疫复合物分解代谢中起重要的作用，所以定向的补体抑制剂也可以减少特别是长期补体抑制带来的潜在严重的副作用。最近，制备了装饰有唾液酸-Lewis x (sLex) 的修饰形式的 sCR1，并且表明其结合表达 P 和 E 选择蛋白的内皮细胞。在炎症疾病的啮齿类动物模型中，证明 sCR1sLex 是比 sCR1 更有效的治疗剂（16,17）。在体外可行性研究中，证明与保护未靶向细胞相比，抗体-DAF（18）和抗体-CD59（19）融合蛋白更有效地保护靶向的细胞免受补体的破坏。通过偶联抑制剂与膜插入肽，也已经获得了重组补体抑制剂的非特异性膜靶向（20，21）。

C3 激活片段是在补体激活位点发现的丰富的补体调理素，并且它们做为各种 C3 受体的配体。这样的一种受体是补体受体 2 (CR2)，一种跨膜蛋白，它通过主要在成熟 B 细胞和滤泡树突细胞上的表达而在体液免疫中起重要的作用（22，23）。CR2 是 C3 结合蛋白家族的成员，并且由 15-16 个短共有重复 (SCR) 结构域（这些蛋白质的特征性结构单元）组成，C3 结合位点包含在 2 个 N 末端 SCR 中（24，25）。不同于补体激活抑制剂 (DAF,MCP,CR1 和 Crry)，CR2 不是补体抑制剂，并且它不结合 C3b。CR2 的天然配体是 iC3b，C3dg 和 C3d，其是结合 CR2 两个 N 末端 SCR 结构域的 C3b 的细胞结合裂解片段（26，27）。C3 的裂解最初引起 C3b 的产生及其在激活细胞表面上的沉积。C3b 片段参与酶复合物的产生，该酶复合物扩大补体级联。在细胞表面上，特别是当在含有补体激活调节剂的宿主表面（即，大多数宿主组织）上沉积 C3b 时，C3b 快速转化为失活

的 iC3b。即时在不存在膜结合的补体调节剂的情况下，也形成相当高水平的 iC3b。随后，iC3b 被血清蛋白酶消化为膜结合片段 C3dg 和 C3d，但是这个过程相对慢（28，29）。因此，CR2 的 C3 配体一旦产生后就相对长寿，并且以高浓度存在于补体激活位点。

## 发明内容

根据这里所体现和充分描述的本发明目的，一方面，本发明涉及通过 CR2 定向的补体活性调节剂。

本发明的其它优点部分地在下面的说明书中阐述，并且部分可以从说明书中显而易见，或者通过实施本发明可以获知。本发明的优点特别可以通过所附权利要求中指出的要素和组合来实现和获得。应当理解，前述一般性描述和下面详细的描述都是只是示例性和解释性的，并不限制本发明。

## 附图说明

包含在本说明书并且组成本说明书的一部分的附图与说明书一起示例了本发明的几种实施方式，用于解释本发明的原理。

图 1 表示 CR2-补体抑制剂融合蛋白例子的示图。

图 2 表示纯化的重组融合蛋白和可溶补体抑制剂的 SDS-PAGE 和 Western 印迹分析。用考马斯亮蓝染色凝胶（10% 丙烯酰胺）。利用补体抑制剂的抗体做为初级抗体显影 Western 印迹。

图 3 表示重组融合蛋白与 C3 调理的 CHO 细胞的结合。在缺乏 C6 的血清中温育抗体敏化的 CHO 细胞，洗涤并且与 20 $\mu$ g/ml 可溶补体抑制剂（黑色痕迹）或者在 N 末端（浅灰痕迹）或 C 末端（深灰痕迹）具有 CR2 的融合蛋白质温育。利用抗 DAF 或抗 CD59 mAbs，通过流式细胞术检测重组蛋白的细胞结合。CHO 细胞与 PBS，而不是补体抑制剂的温育产生了与 sDAF 和 sCD59 相似的荧光谱。代表 3 次单独的试验。

图 4 表示通过表面等离子体共振（surface plasmon resonance）对 CR2 融合蛋白和 C3d 之间相互作用的分析。实线表示不同浓度的 CR2 融合蛋

白。虚线表示拟合 1:1 Langmuir 结合模型的曲线。

图 5 表示通过重组 sDAF 和 DAF 融合蛋白对补体介导的裂解的抑制。抗体敏化的 CHO 细胞 (图 a) 或绵羊红细胞 (图 b) 与重组蛋白和 10% 人血清 (CHO 细胞) 或 0.33% 人血清 (红细胞) 温育。这些浓度引起了约 90% 未保护细胞的裂解。在 37°C 温育 45 分钟后测定裂解。通过在加热灭活的血清中温育细胞测定的背景裂解小于 5%，并且该背景裂解被扣除。平均值  $\pm$  SD, n = 4。

图 6 表示通过重组 sCD59 和 CD59 融合蛋白对补体介导的裂解的抑制。抗体敏化的 CHO 细胞 (图 a) 或绵羊红细胞 (图 b) 与重组蛋白和 10% 人血清 (CHO 细胞) 或 0.33% 人血清 (红细胞) 温育。这些浓度引起了约 90% 未保护细胞的裂解。在 37°C 温育 45 分钟后测定裂解。通过在加热灭活的血清中温育细胞测定的背景裂解小于 5%，并且该背景裂解被扣除。平均值  $\pm$  SD, n = 4。

图 7 表示重组融合蛋白对 U937 细胞粘附的影响。用 IgM 抗体敏化绵羊红细胞，并且在缺乏 C6 的血清中温育。在 500nM 重组融合蛋白或 PBS 存在的情况下，C3 调理的红细胞与 U937 细胞共温育。温育后，通过显微镜检测每个红细胞结合的平均 U937 细胞数。平均值  $\pm$  SD, n = 3。

图 8 表示成熟人 CR2-DAF 的核苷酸和预测的氨基酸序列。下划线的氨基酸表示 CR2 和 DAF 之间的连接序列。

图 9 表示成熟人 CR2-CD59 的核苷酸和预测的氨基酸序列。下划线的氨基酸表示 CR2 和 CD59 之间的连接序列。

图 10 表示成熟人 DAF-CR2 的核苷酸和预测的氨基酸序列。下划线的氨基酸表示 DAF 和 CR2 之间的连接序列。

图 11 表示成熟人 CD59-CR2 的核苷酸和预测的氨基酸序列。下划线的氨基酸表示 CD59 和 CR2 之间的连接序列。

图 12 表示含有 CR2 的融合蛋白靶向 C3 包被的 CHO 细胞。通过在 10% 抗 CHO 抗血清和 10% C6 耗竭的人血清中温育细胞 (以阻止膜攻击复合物的形成和细胞裂解)，在 CHO 细胞上产生 C3 配体。洗细胞，并且与融

合蛋白温育 (20 $\mu$ g/ml, 4 $^{\circ}$ C, 30 分钟)。利用抗适当补体抑制剂 (DAF 或 CD59) 的抗体, 通过流式细胞术分析检测结合。黑线: 对照 (无融合蛋白); 浅灰: 在 C 末端的 CR2; 深灰: 在 N 末端的 CR2。

图 13 表示通过表面等离子体共振对 CR2-DAF 结合 C3dg 的分析。

图 14 表示通过表面等离子体共振对 CR2-CD59 结合 C3dg 的分析。

图 15 表示通过表面等离子体共振对 DAF-CR2 结合 C3dg 的分析。

图 16 表示通过表面等离子体共振对 CD59-CR2 结合 C3dg 的分析。

图 17 表示定向和未定向的 DAF 对补体介导的 CHO 细胞裂解的影响。用抗 CHO 抗血清 (10% 浓度, 4 $^{\circ}$ C, 30 分钟) 敏化 CHO 细胞使其对补体敏感, 随后在不同浓度的补体抑制蛋白存在的情况下, 与 10% 的正常人血清 (NHS) 温育 (37 $^{\circ}$ C, 60 分钟)。然后通过台盼蓝排除测定法检测细胞裂解。代表性试验显示平均值  $\pm$  SD ( $n=3$ )。利用不同的融合蛋白制备物进行 3 个单独的试验。

图 18 表示定向和未定向的 CD59 对补体介导的 CHO 细胞裂解的影响。如图 17 的图注中所述进行测定。代表性试验显示平均值  $\pm$  SD ( $n=3$ )。利用不同的融合蛋白制备物进行 3 个单独的试验。

图 19 表示定向和未定向的 DAF 对补体介导的溶血的影响。用抗绵羊 E 抗体敏化绵羊红细胞 (E), 随后在不同浓度的补体抑制蛋白存在的情况下, 与 1/300 稀释度的 NHS 温育 (37 $^{\circ}$ C, 60 分钟)。通过测定释放的血红蛋白 (在 412nm 的吸光度) 检测细胞裂解。代表性试验显示平均值  $\pm$  SD ( $n=3$ )。利用不同的融合蛋白制备物进行 2 个单独的试验。

图 20 表示定向和未定向的 CD59 对补体介导的溶血的影响。如图 19 的图注中所述进行测定。代表性试验表示平均值  $\pm$  SD ( $n=3$ )。利用不同的融合蛋白制备物进行 2 个单独的试验。

图 21 表示成熟人 CR2-人 IgG1 Fc 的核苷酸和预测的氨基酸序列。下划线的氨基酸表示 CR2 和 Fc 区之间的连接序列。表达质粒含有基因组 Fc 区 (铰链-内含子-CH2-内含子-CH3)。

图 22 表示 CR2-Fc 融合蛋白的 SDS-PAGE 分析。在非还原 (道 1) 或

还原（道2）条件下电泳纯化的 CR2-Fc。用考马斯亮蓝染色凝胶。（道3为分子量标记，参见图2）。

图23表示 CR2-Fc 靶向 C3 包被的 CHO 细胞。如所述（图12的图注），产生 C3 配体。洗细胞并且与 CR2-Fc 温育（20 $\mu$ g/ml，4 $^{\circ}$ C，30分钟）。利用缀合 FITC 的抗人 Fc 抗体，通过流式细胞术分析检测结合。上方的图表示 CR2-Fc 与 C3 包被的 CHO 细胞温育的结果，下方的图表示 CR2-Fc 与对照 CHO 细胞温育的结果。

图24表示表面等离子体共振的感应曲线（sensorgram），该曲线显示 CR2-Fc 与固定在芯片上的 C3d 配体的结合。

图25表示在34周龄 NZB/W F1 小鼠中  $^{125}$ I-CR2-DAF 和  $^{125}$ I-sDAF 的生物分布。将放射性标记的蛋白质注射到尾部静脉中，并且24小时后检测放射性标记的生物分布。将每种蛋白质注射到2只小鼠中。

图26表示结合24周龄 MRL/lpr 小鼠肾小球的 CR2-DAF 的成像。每种蛋白质尾静脉注射后24小时分析 CR2-DAF (a) 和 sDAF (b) 的肾小球结合。该图表示肾截面的免疫荧光染色。

图27表示单链抗体 CD59-Crry 构建体。该图表示该构建体包含来源于 K9/9mAb 的可变轻链（VL）和可变重链（VH）。在酵母表达载体 pPICZ $\alpha$ (Invitrogen)中制备该构建体。

图28表示大鼠中补体抑制剂和 K9/9 单链 Ab 的生物分布。PAN 处理后4天施用碘化重组蛋白质，48小时后测定器官中放射性。

图29表示在用 PAN 处理和接受所示治疗的大鼠中肌酐的清除(n=4, +/- SD)。

图30表示 PAS 染色的肾皮质。图30A表示无 PAN 对照，图30B: PAN 加 PBS 处理，图30C: PAN 加定向的 K9/9 Crry 处理，和图30D: PAN 加 sCrry 处理。

图31表示重组蛋白质给药后血清中补体抑制活性。通过敏化的绵羊红细胞的裂解来测量。所示是相对于来源于对照大鼠的血清的百分比抑制活性。

## 发明详述

通过参考下面本发明优选实施方式的详细描述及本文所包含的实施例，和参考附图和附图前后的描述可以更容易地理解本发明。

在公开和描述本发明的化合物、组合物、物品、装置和/或方法以前，应该理解除非另外的说明，本发明不限于特定的合成方法、特定的重组生物技术方法，或者特定的试剂，因为这些是当然可以改变的。也应该理解这里所用的术语仅仅旨在描述特定实施方式，而不意在限制。

### A.定义

除了上下文另外明确地规定以外，说明书和所附权利要求中所用的单数形式“a/an”及“the”包括复数。因此，例如提到“a”药物载体包括两种或多种此类载体的混合物，等等。

这里可以将范围表示为从“约”一个特定数值，和/或到“约”另一个特定的数值。当表示这种范围时，另一种实施方式包括从一个特定的数值和/或到另一个特定的数值。相似地，通过利用先行词“约”将数值表示为近似值时，应理解特定的数值形成了另一个实施方式。也应理解每个范围的端点不论是相对于另一端点而言还是独立于另一端点而言都是有意义的。

在本说明书和随后权利要求书中，将提及大量的术语，所述术语被定义具有下面的含义：

“可选的”或“可选地”意指随后描述的事件或情形可以发生或可以不发生，并且该描述包括所述事件或情形发生的情况和所述事件或情形不发生的情况。

“治疗”意指向患有疾病的对象施用组合物，其中该疾病可以是任何病原性疾病，自身免疫病，癌症或炎性疾病。向对象施用组合物的效果可以是但不限于：减少疾病的症状、降低疾病的严重性或者完全消除疾病。

这里“抑制”意指减少活性。应理解抑制可以意指轻微的减少活性到完全去除所有活性。“抑制剂”可以是减少活性的任何物质。

这里“活化”或“激活”意指增加活性。应理解活化可以意指增加已有的活性及诱导新活性。“激活剂”可以是增加活性的任何物质。

#### B. 补体抑制和激活构建体

本发明公开了包含如下构建体的组合物，其中该构建体包含 CR2 和补体活性调节剂。

CR2 具有由 15 或 16 个称为短共有重复 (SCR) 的重复单元组成的细胞外部分。氨基酸第 1-20 位包含前导肽，氨基酸 23-82 包含 SCR1，氨基酸 91-146 包含 SCR2，氨基酸 154-210 包含 SCR3，氨基酸 215-271 包含 SCR4。活性位点 (C3dg 结合位点) 位于 SCR1-2 (头 2 个 N-末端 SCR) 中。SCR 单元由作为间隔区的可变长度短序列分隔。应理解可以使用含有活性位点的任何数量的 SCR。在一个实施方式中，构建体含有 4 个 N-末端 SCR 单元。在另一个实施方式中，构建体含有头 2 个 N-末端 SCR。在另一个实施方式中，构建体含有头 3 个 N-末端 SCR。

可以理解对于公开的肽、多肽、蛋白质、蛋白质片段和组合物，存在物种和株系变异。特别公开所公开的肽、多肽、蛋白质、蛋白质片段和组合物的所有物种和株系变异。

也公开了组合物，其中构建体是融合蛋白。

这里的“融合蛋白”意指可操作地连接的包含肽、多肽或蛋白质的两个或多个成分。CR2 可以通过氨基酸连接序列连接补体抑制剂或激活剂。接头的例子是本领域所熟知的。接头的例子可以包括但不限于  $(\text{Gly}_4\text{Ser})_3(\text{G4S})$ ,  $(\text{Gly}_3\text{Ser})_4(\text{G3S})$ ,  $\text{SerGly}_4$  和  $\text{SerGly}_4\text{SerGly}_4$ 。连接序列也可以由在人类 (或小鼠) 蛋白质中 SCR 单元之间发现的“天然”连接序列，例如 VSVFPLE，人 CR2 的 SCR2 和 3 之间的连接序列组成。也可以构建无连接序列的融合蛋白。

也公开了本发明的组合物，其中融合蛋白抑制补体。

也公开了本发明的组合物，其中补体活性调节剂包含补体抑制剂。

也公开了本发明的组合物，例如，其中补体抑制剂是衰变加速因子 (DAF) SEQ ID NO:1 (核苷酸) 和 SEQ ID NO:2 (氨基酸)。例如，DAF

可以是没有糖基磷脂酰锚和丝氨酸-苏氨酸富集区的包含 4 个 SCR 结构域的可溶人 DAF。DAF 也可以是包含 4 个 SCR 结构域和丝氨酸-苏氨酸富集区但是没有糖基磷脂酰锚的可溶人 DAF。

DAF 细胞外区由 N-末端的 4 个 SCR 单元及接着的丝氨酸-苏氨酸富集区组成。氨基酸 1-34 包含前导肽, 氨基酸 35-95 包含 SCR1, 氨基酸 97-159 包含 SCR2, 氨基酸 162-221 包含 SCR3, 氨基酸 224-284 包含 SCR4, 氨基酸 287-356 包含 S/T 区。在本发明的一个实施方式中, 本发明的组合物包含所有 4 个 SCR 单元。在本发明的另一个实施方式中, 组合物包含 DAF 的 SCR2-4。

公开了本发明的组合物, 其中补体抑制剂包含 CD59 和另一补体抑制剂的融合蛋白, 该另一补体抑制剂选自 DAF, MCP, Crry 和 CR1。也公开了本发明的组合物, 其中补体抑制剂是两种或多种补体抑制剂的融合蛋白。

也公开了本发明的组合物, 其中融合蛋白包含 CR2-DAF (SEQ ID NO:6)。也公开了本发明的组合物, 其中由包含 SEQ ID NO:5 的核苷酸编码融合蛋白。

也公开了本发明的组合物, 其中融合蛋白包含 DAF-CR2 (SEQ ID NO:10)。也公开了本发明的组合物, 其中由包含 SEQ ID NO:9 的核苷酸编码融合蛋白。

也公开了本发明的组合物, 其中补体抑制剂是人 CD59 (SEQ ID NO:3 (核苷酸) 和 SEQ ID NO:4 (氨基酸))。人 CD59 可以是可溶人 CD59, 其包含没有糖基磷脂酰锚的成熟蛋白。

也公开了本发明的组合物, 其中融合蛋白包含 CR2-人 CD59 (SEQ ID NO:8)。也公开了本发明的组合物, 其中由包含 SEQ ID NO:7 的核苷酸编码融合蛋白。

也公开了本发明的组合物, 其中融合蛋白包含人 CD59-CR2 (SEQ ID NO:12)。也公开了本发明的组合物, 其中由包含 SEQ ID NO:10 的核苷酸编码融合蛋白。

也公开了本发明的组合物，其中补体抑制剂是抗 C5 的抗体。也公开了本发明的组合物，其中融合蛋白包含 CR2-抗-C5 抗体。

也公开了本发明的组合物，其中补体抑制剂是 CR1 (SEQ ID NO:13 (核苷酸) 和 SEQ ID NO:14 (氨基酸))。CR1 的细胞外区域可以包含 30 个 SCR 单元。组合物可以包含 CR1 的整个细胞外区，这是本发明的一个实施方式。在本发明的另一个实施方式中，组合物包含 CR1 的一个或多个活性位点。CR1 的活性位点是包含前导肽的氨基酸 1-46，包含 SCR1-4 的氨基酸 47-300 (C4b 结合位点，对 C3b 亲和性较低)，包含 SCR8-11 的氨基酸 497-750 (C3b 结合位点，对 C4b 亲和性较低)，包含 SCR15-18 的氨基酸 947-1200 (C3b 结合位点，对 C4b 亲和性较低) 和包含 C1q 结合位点的氨基酸 1400-1851。在本发明的另一个实施方式中，本发明的组合物可以包含 CR1 的任何一个活性位点或其活性位点的任意组合，或者所有的活性位点。

也公开了本发明的组合物，其中补体抑制剂包含 CR1 活性位点，并且其中此一个或多个活性位点进一步包括含有氨基酸 6-46 的前导肽，含有 SCR1-4 的氨基酸 47-300 (C4b 结合位点，对 C3b 亲和性较低)，含有 SCR8-11 的氨基酸 497-750 (C3b 结合位点，对 C4b 亲和性较低)，含有 SCR15-18 的氨基酸 947-1200 (C3b 结合位点，对 C4b 亲和性较低) 和含有 C1q 结合位点的氨基酸 1400-1851。在本发明的另一个实施方式中，本发明的组合物可以包含 CR1 的任何一个活性位点或其活性位点的任意组合，或者所有的活性位点。

也公开了本发明的组合物，其中补体抑制剂是 MCP (SEQ ID NO:15 (核苷酸) 和 SEQ ID NO:16 (氨基酸))。细胞外区域由 4 个 SCR 单元及接着的 ser/thr 区组成。氨基酸 1-34 包含前导肽，氨基酸 35-95 包含 SCR1，氨基酸 96-158 包含 SCR2，氨基酸 159-224 包含 SCR3，氨基酸 225-285 包含 SCR4，并且氨基酸 286-314 包含 S/T 区。

也公开了本发明的组合物，其中补体抑制剂是 Crry (SEQ ID NO:17)。Crry 可以是可溶的小鼠 Crry，其包含 5 个 N-末端 SCR 结构域且没有跨膜

区。

也公开了本发明的组合物，其中补体抑制剂是鼠 CD59。鼠 CD59 可以是可溶的鼠 CD59，其包含没有糖基磷脂酰锚的成熟蛋白。

也公开了本发明的组合物，其中融合蛋白激活补体。

因此，也公开了本发明的组合物，其中补体活性的调节剂包含补体激活剂。

也公开了本发明的组合物，其中补体激活剂是人 IgG1 Fc (SEQ ID NO:18)。

也公开了本发明的组合物，其中补体激活剂包含 CR2-人 IgG1 Fc (SEQ ID NO:20)。也公开了本发明的组合物，其中由包含 SEQ ID NO:21 的核苷酸编码融合蛋白。

公开了本发明的组合物，其中融合蛋白是人 IgM (SEQ ID NO:19)。

也公开了本发明的组合物，其中融合蛋白包含 CR2-人 IgM Fc。

公开了本发明的组合物，其中补体激活剂是小鼠 IgG3 (SEQ ID NO:22)。

也公开了本发明的组合物，其中融合蛋白包含 CR2-鼠 IgG3 Fc。

也公开了本发明的组合物，其中融合蛋白包含 CR2-鼠 IgM Fc。

特别可以想到通过组合物的 Fc 部分，补体激活剂也可以增加依赖于抗体的细胞介导的细胞毒性 (ADCC)。ADCC 指通过 NK 细胞上的 Fc 受体对 Fc 区的识别和接触，由天然杀伤 (NK) 细胞破坏靶细胞。这可以是 IgG1 Fc 或 IgG3 Fc 的 FcγRIII 识别。Fc 受体与 Fc 接触后，NK 细胞通过使用穿孔素和粒酶裂解靶细胞。该机制在控制肿瘤生长方面是重要的。

公开了本发明的组合物，其中 CR2-Fc 融合蛋白不具有免疫原性。可以理解受试者自身的免疫应答不太可能攻击和失活不具免疫原性 (即，不激发免疫应答) 的组合物。预期 CR2-Fc 缺少免疫原性也使之潜在地优于抗 C3d 的抗体 (即使抗体被人源化)。本发明的一个实施方式中，融合 CR2 的 Fc 区可以来源于任何人类或小鼠的 IgG 同种型，人类或小鼠的 IgM，或者含有 mu-尾片 (mu-tailpiece) 的任何人类或小鼠的 IgG 同种型。

Mu-尾片 (mu-tailpiece) 是来源于 IgM 的 18 个氨基酸的 C-末端区, 当向 IgG Fc 序列 C-末端添加时, 引起了聚合形式 IgG 的产生 (与 IgM 相似), 该聚合形式 IgG 有效地激活补体并且对 Fc 受体具有增强的亲和力。融合可以在组合物 Fc 部分的铰链区发生。

含有具 mu-尾片的 IgM 或 IgG Fc 区的 CR2 融合蛋白可以优于 CR2-IgG Fc 融合蛋白。IgM 或 IgG-mu Fc 区将导致产生具有多达 6 个 Fc 和 12 个 CR2 位点的聚合融合蛋白。这些构建体对于 C3 配体具有增强的亲和力并具有增强的效应子功能 (补体激活和 Fc 受体结合)。

也公开了本发明的组合物, 其中补体激活剂是 CVF (SEQ ID NO:23 (核苷酸) 和 SEQ ID NO:24 (氨基酸))。

在本发明的一个实施方式中, CVF 可以偶联可溶的 CR2。可以理解 CVF 结合因子 B 并且通过形成 CVFBb 激活补体的旁路途径, CVFBb 是一种不被补体抑制蛋白失活的 C3/C5 转化酶。CVFBb 的半衰期是约 7 小时, 相比较, 对于生理旁路途径的转化酶 C3bBb 而言半衰期是约 1 分钟。

CVF 可以化学偶联可溶的 CR2 是本发明的实施方式。

公开了本发明的组合物, 其中构建体位于载体中。

公开了包含本发明载体的细胞。

也公开了组合物, 其中构建体是免疫缀合物。这里“免疫缀合物”意指通过化学交联剂可操作地连接的两种或多种包含肽、多肽或蛋白质的成分。可以在位于该成分上的反应性基团上发生免疫缀合物成分的连接。可以利用交联剂靶向的反应性基团包括伯胺、巯基、羧基、碳水化合物和羧酸, 或者可以向蛋白质添加活性基团。化学交联剂的例子是本领域所熟知的, 并且可以包括但不限于双马来酰亚氨基己烷, 间马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯, NHS-酯-马来酰亚胺交联剂如 MBS, 磺基-MBS, SMPB, 磺基-SMPB, GMBS, 磺基-GMBS, EMCS, 磺基-EMCS; 亚氨基酯交联剂如 DMA, DMP, DMS, DTBP; EDC[1-乙基-3-(3-二甲基氨基丙基)碳二亚胺盐酸盐], [2-(4-羧基苯基)乙基]-4-马来酰亚氨基甲基-环己烷-1-甲酰胺, DTME: 二硫代-双-马来酰亚氨基乙烷, DMA (二甲基己

二亚氨酸酯·2HCl, Dimethyl adipimidate), DMP (二甲基庚二亚氨酸酯·2HCl, Dimethyl pimelimidate·2HCl), DMS (二甲基辛二亚氨酸酯·2HCl, Dimethyl suberimidate·2HCl), DTBP (二甲基 3,3'-二硫代双丙亚氨酸·2HCl), MBS (间-马来酰亚氨基苯甲酰基-N-羟基琥珀酰亚胺酯), 磺基-MBS(间-马来酰亚氨基苯甲酰基-N-羟基琥珀酰亚胺酯), 磺基-SMPB(磺基琥珀酰亚氨基 4-[对马来酰亚氨基苯基]丁酸酯), GMBS(N-[间-马来酰亚氨基丁酰氧基]琥珀酰亚胺酯), EMCS(N-[间-马来酰亚氨基己酰氧基]琥珀酰亚胺酯), 和磺基-EMCS(N-[间-马来酰亚氨基己酰氧基]磺基琥珀酰亚胺酯).

### C.使用组合物的方法

目前正在研究用于治疗与生物不相容性相关的炎症疾病和病状的各种类型的补体抑制蛋白。两种最具治疗特征的人补体抑制剂是可溶形式的补体受体 1 (sCR1) 和抗 C5 单克隆抗体。这些系统活性的抑制蛋白在各种疾病的动物模型中以及最近在临床试验中已经显示了效力(1-5, 6: #1037, 这里将这些文献并入作为有关体内效力和临床结果的教导的参考)。

公开了在对象中治疗受补体影响的疾病的方法, 该方法包括向对象施用本发明的组合物。可以理解向对象施用组合物可以具有但是不限于下面的作用: 减少疾病的症状, 降低疾病严重性或完全消除疾病。

#### 1、使用组合物抑制补体的方法

公开了在对象中治疗受补体影响的疾病的方法, 该方法包括向对象施用本发明的组合物, 其中该组合物抑制补体活性。可以理解向对象施用组合物可以具有但是不限于下面的作用: 减少疾病的症状, 降低疾病严重性或完全消除疾病。

公开了减少补体介导的损害的方法, 该方法包括向对象施用抑制补体的本发明的组合物。

公开了本发明的方法, 其中所治疗的疾病是炎症疾病。也公开了本发明的方法, 其中炎症疾病可以选自哮喘、系统性红斑狼疮、类风湿性关节炎

炎、反应性关节炎、脊椎关节炎、系统性血管炎、胰岛素依赖型糖尿病、多发性硬化、实验性变应性脑脊髓炎、舍格伦综合征、移植物抗宿主病、炎性肠病包括局限性回肠炎、溃疡性结肠炎、局部缺血再灌注损伤、心肌梗塞、阿尔茨海默氏病、移植排斥（同种异体和异种）、烧伤、任何免疫复合物诱导的炎症、肾小球肾炎、重症肌无力、脑狼疮、吉-巴综合征、血管炎、系统性硬化、过敏症、导管反应(catheter reactions)、粉瘤(atheroma)、不育、甲状腺炎、ARDS、post-bypass 综合症、血液透析、幼年型类风湿、Behcets 综合症、溶血性贫血、天疱疮、大疱性类天疱疮、中风、动脉粥样硬化和硬皮病。

也公开了本发明的方法，其中疾病是病毒感染。也公开了本发明的方法，其中病毒感染可以选自下面的病毒：甲型流感病毒、乙型流感病毒、呼吸道合胞病毒、登革病毒、黄热病毒、埃博拉病毒、马尔堡病毒、拉沙热病毒、东方马脑炎病毒、日本脑炎病毒、圣路易脑炎病毒、Murray 河谷脑炎病毒、西尼罗病毒、立夫特山谷热病毒、轮状病毒 A、轮状病毒 B、轮状病毒 C、辛德毕斯病毒和汉坦病毒。

公开了本发明的方法，其中该疾病是对病毒载体的炎性反应。病毒载体可以选自下面的病毒：腺病毒、痘苗病毒、腺相关病毒、修饰的痘苗 ancara 病毒和巨细胞病毒。可以理解其它病毒载体可以用于疫苗递送。具体公开的是本领域已知的每种病毒载体。

本领域技术人员理解念珠菌属(*Candida*)表达 CR3 样蛋白质，该 CR3 样蛋白质具有与 CR2 相似的结合特性。CR3 样蛋白质看起来参与发病机理。因此，本发明的一个实施方式是治疗真菌感染的患者，其中该治疗阻断了真菌“CR3”功能及抑制了补体，该方法包括给患者施予本发明的组合物。

公开了本发明的方法，其中补体抑制剂可以增强基于细胞凋亡的治疗（例如，用表达 Fas 配体的腺病毒进行基因治疗）的结果。

正常发育期间出现的细胞凋亡是非炎性的，并且参与了免疫耐受性的诱导。尽管根据如何和在什么类型的细胞中激活细胞凋亡，细胞凋亡也可

以是炎性的（例如，连接 Fas 的治疗剂能诱导炎症），但是坏死的细胞死亡能引起持久有效的炎性应答，其中由释放的细胞内含物和受刺激的吞噬细胞释放的促炎细胞因子介导该炎性应答。正常地，通过吞噬细胞清除调亡的细胞和小泡，由此阻止细胞裂解的促炎后果。在本发明上下文中，已经证明调亡细胞和调亡小体直接固定补体，并且补体可以通过调理作用以及增强的对调亡细胞的吞噬作用来维持抗炎性应答。

炎症涉及免疫细胞的非特异性募集，其可以影响天然和适应性免疫应答。在基于细胞调亡的肿瘤治疗期间调节补体的激活以抑制调亡细胞/小体的吞噬细胞摄取，将增强肿瘤环境中的炎性/天然免疫应答。此外，调亡的细胞可以是免疫原性自身抗原的来源，并且未清除的调亡小体可以引起自身免疫。除了产生增强的免疫刺激环境外，在诱导肿瘤细胞发生细胞调亡的位点调节补体，还可以扩大或者触发特异性免疫，该特异性免疫对抗正常在宿主中具有耐受性的肿瘤。

本发明公开的组合物可以用作 CR2 和 CR3 拮抗剂。公开了通过抑制 CR2 来抑制补体活性的方法，该方法包括向对象施用本发明组合物。也公开了通过抑制 CR3 来抑制补体活性的方法，该方法包括向对象施用本发明组合物。正如 CR2 拮抗剂可以调节免疫应答，CR3 拮抗剂可以具有第二抗炎作用机理，因为 CR3 是结合内皮 ICAM1 的整联蛋白。ICAM1 在炎症位点表达，并且参与白细胞粘着和血细胞渗出。此外，补体激活产物可以上调 ICAM1 表达。

## 2、使用组合物激活补体的方法

公开了在对象中治疗受补体影响的疾病的方法，该方法包括向对象施用本发明的组合物，其中该组合物将激活补体。可以理解向对象给予组合物可以具有但是不限于下面的作用：减少疾病的症状，降低疾病严重性或完全去除疾病。

公开了增强补体介导的损害的方法，该方法包括向对象施用激活补体的本发明组合物。

也公开了本发明的方法，其中该疾病是癌症。癌症可以选自（霍奇金

和非霍奇金)淋巴瘤, B 细胞淋巴瘤、T 细胞淋巴瘤、髓性白血病、白血病、蕈样肉芽肿病、癌、实体组织癌、鳞状细胞癌、腺癌、肉瘤、神经胶质瘤、胚细胞瘤(blastoma)、神经母细胞瘤、浆细胞瘤、组织细胞瘤、黑素瘤、腺瘤、缺氧肿瘤(hypoxic tumour)、骨髓瘤、AIDS 相关淋巴瘤或肉瘤、转移性癌症、膀胱癌、脑癌、神经系统癌、头和颈的鳞状细胞癌、神经母细胞瘤/成胶质细胞瘤、卵巢癌、皮肤癌、肝癌、黑素瘤、口腔、喉、咽和肺的鳞状细胞癌、结肠癌、宫颈恶性肿瘤(cervical cancer)、宫颈癌(cervical carcinoma)、乳腺癌、上皮癌、肾癌、泌尿生殖器癌症、肺癌、食管癌、头和颈癌、造血性癌症(hematopoietic cancer)、睾丸癌症、结肠-直肠癌、前列腺癌或胰腺癌。

在本发明的一个实施方式中, 由于施用的抗肿瘤抗体或者由于正常无效的体液免疫应答, CR2 可以靶向肿瘤细胞上沉积的补体。

因此, 可以联合抗肿瘤抗体施用本发明的补体激活组合物。这种抗肿瘤抗体的例子是本领域熟知的, 并且可以包括抗 PSMA 单克隆抗体 J591, PEQ226.5 和 PM2P079.1 (Fracasso, G.等, (2002) *Prostate* 53(1): 9-23); 抗 Her2 抗体 hu4D5 (Gerstner, R. B.,等, (2002) *J. Mol. Biol.* 321(5): 851-62); 抗 disialosyl Gb5 单克隆抗体 5F3, 其可以用做抗肾细胞癌的抗体(Ito A.等, (2001) *Glycoconj. J.* 18 (6): 475-485); 抗 MAGE 单克隆抗体 57B (Antonescu, C. R.等, (2002) *Hum. Pathol.* 33 (2): 225-9); 抗癌单克隆抗体 CLN-Ig (Kubo, O.等, (2002) *Nippon Rinsho.* 60 (3): 497-503); 抗 Dalton 淋巴瘤相关抗原(DLAA)单克隆抗体 DLAB(Subbiah, K.等, (2001) *Indian J. Exp. Biol.* 39 (10): 993-7)。只要当抗体存在于肿瘤部位时本发明组合物也可以存在于肿瘤部位, 则本发明组合物可以在抗肿瘤抗体给药之前, 同时或之后给药。

所公开的组合物可以用于治疗的代表性, 但非限制性癌症如下: 淋巴瘤、B 细胞淋巴瘤、T 细胞淋巴瘤、蕈样肉芽肿病、多发性骨髓瘤、霍奇金病、髓性白血病、膀胱癌、脑癌、神经系统癌、头和颈癌、头和颈的鳞状细胞癌、肾癌、肺癌诸如小细胞肺癌和非小细胞肺癌、尿道上皮癌、腺

癌、肉瘤、神经胶质瘤、高级神经胶质瘤、胚细胞瘤、神经母细胞瘤、浆细胞瘤、组织细胞瘤、腺瘤、缺氧肿瘤、骨髓瘤、AIDS 相关淋巴瘤或肉瘤、转移性癌症、神经母细胞瘤/成胶质细胞瘤、卵巢癌、胰腺癌、前列腺癌、皮肤癌、肝癌、黑素瘤、口腔、喉、咽和肺的鳞状细胞癌、结肠癌、宫颈恶性肿瘤、宫颈癌、乳癌和上皮癌、肾癌、泌尿生殖器癌症、肺癌、食管癌、头和颈癌、大肠癌、造血性癌症、睾丸癌症、结肠直肠癌、胃癌、前列腺癌、瓦尔登斯特伦病或胰腺癌。

这里公开的补体激活组合物也可以用于治疗癌前疾病诸如宫颈和肛门发育异常、其它发育异常、严重发育异常、增生、不典型增生和瘤形成。公开了本发明的方法，其中疾病是癌前疾病。可以理解组合物将识别在癌前细胞表面上超表达的抗原。

也公开了利用本发明的补体激活组合物治疗病毒感染的方法。病毒感染可以选自下面的病毒：单纯疱疹病毒 1 型，单纯疱疹病毒 2 型，巨细胞病毒、EB 病毒、水痘-带状疱疹病毒、人疱疹病毒 6、人疱疹病毒 7、人疱疹病毒 8、天花病毒、水泡性口炎病毒、甲型肝炎病毒、乙型肝炎病毒、丙型肝炎病毒、丁型肝炎病毒、戊型肝炎病毒、鼻病毒、冠状病毒、甲型流感病毒、乙型流感病毒、麻疹病毒、多瘤病毒、人乳头瘤病毒、呼吸道合胞病毒、腺病毒、柯萨奇病毒、登革病毒、腮腺炎病毒、脊髓灰质炎病毒、狂犬病病毒、Rous 肉瘤病毒、黄热病毒、埃博拉病毒、马尔堡病毒、拉沙热病毒、东方马脑炎病毒、日本脑炎病毒、圣路易脑炎病毒、Murray 河谷脑炎病毒、西尼罗病毒、立夫特山谷热病毒、轮状病毒 A、轮状病毒 B、轮状病毒 C、辛德毕斯病毒、猿猴免疫缺陷病毒、人 T 细胞白血病病毒 1 型，汉坦病毒、风疹病毒、猿免疫缺陷病毒、人免疫缺陷病毒 1 型和人免疫缺陷病毒 2 型。

也公开了利用本发明的补体激活组合物治疗细菌感染的方法。也公开了本发明的方法，其中细菌感染可以选自下面的细菌：结核分枝杆菌 (*M. tuberculosis*)，牛分枝杆菌 (*M. bovis*)，牛分枝杆菌株 BCG，BCG 亚株，鸟分枝杆菌 (*M. avium*)，胞内分枝杆菌 (*M. intracellulare*)，非洲分枝

杆菌 (*M. africanum*), 堪萨斯分枝杆菌 (*M. kansasii*), 海分枝杆菌 (*M. marinum*), 溃疡分枝杆菌 (*M. ulcerans*), 鸟分枝杆菌副结核亚种 (*M. avium subspecies paratuberculosis*), 星状诺卡氏菌 (*Nocardia asteroides*), 其它诺卡氏菌属 (*Nocardia*) 物种, 嗜肺军团菌 (*Legionella pneumophila*), 其它军团菌属 (*Legionella*) 物种, 伤寒沙门氏菌 (*Salmonella typhi*), 其它沙门氏菌属 (*Salmonella*) 物种, 志贺氏菌属 (*Shigella*) 物种, 鼠疫耶尔森氏菌 (*Yersinia pestis*), 溶血巴斯德氏菌 (*Pasteurella haemolytica*), 多杀巴斯德氏菌 (*Pasteurella multocida*), 其它巴斯德氏菌属 (*Pasteurella*) 物种, 胸膜肺炎放线杆菌 (*Actinobacillus pleuropneumoniae*), 单核细胞增生李斯特氏菌 (*Listeria monocytogenes*), 伊氏李斯特氏菌 (*Listeria ivanovii*), 流产布鲁氏菌 (*Brucella abortus*), 其它布鲁氏菌属 (*Brucella*) 物种, 反刍类考德里氏体 (*Cowdria ruminantium*), 肺炎衣原体 (*Chlamydia pneumoniae*), 砂眼衣原体 (*Chlamydia trachomatis*), 鸚鵡热衣原体 (*Chlamydia psittaci*), 伯氏考克斯氏体 (*Coxiella burnetti*), 其它立克次氏体属 (*Rickettsial*) 物种, 埃里希氏体属 (*Ehrlichia*) 物种, 金黄色葡萄球菌 (*Staphylococcus aureus*), 表皮葡萄球菌 (*Staphylococcus epidermidis*), 酿脓链球菌 (*Streptococcus pyogenes*), 无乳链球菌 (*Streptococcus agalactiae*), 炭疽芽孢杆菌 (*Bacillus anthracis*), 大肠杆菌 (*Escherichia coli*), 霍乱弧菌 (*Vibrio cholerae*), 弯曲杆菌属 (*Campylobacter*) 物种, 脑膜炎奈瑟氏球菌 (*Neisseria meningitidis*), 淋病奈瑟氏球菌 (*Neisseria gonorrhoea*), 铜绿假单胞菌 (*Pseudomonas aeruginosa*), 其它假单胞菌属 (*Pseudomonas*) 物种, 流感嗜血菌 (*Haemophilus influenzae*), 杜氏嗜血菌 (*Haemophilus ducreyi*), 其它嗜血菌属 (*Hemophilus*) 物种, 破伤风梭菌 (*Clostridium tetani*), 其它梭菌属 (*Clostridium*) 物种, 小肠结肠炎耶尔森氏菌 (*Yersinia enterocolitica*) 和其它耶尔森氏菌属 (*Yersinia*) 物种。

也公开了利用本发明的补体激活组合物治疗寄生虫感染的方法。也公开了本发明的方法, 其中寄生虫感染可以选自鼠弓形体 (*Toxoplasma*

*gondii*), 恶性疟原虫 (*Plasmodium falciparum*), 间日疟原虫 (*Plasmodium vivax*), 三日疟原虫 (*Plasmodium malariae*), 其它疟原虫属 (*Plasmodium*) 物种, 布氏锥虫 (*Trypanosoma brucei*), 克氏锥虫 (*Trypanosoma cruzi*), 硕大利什曼原虫 (*Leishmania major*), 其它利什曼原虫属 (*Leishmania*) 物种, 曼氏血吸虫 (*Schistosoma mansoni*), 其它血吸虫属 (*Schistosoma*) 物种和痢疾阿米巴 (*Entamoeba histolytica*)。

也公开了利用本发明的补体激活组合物治疗真菌感染的方法。也公开了本发明的方法, 其中真菌感染可以选自: 白色念珠菌 (*Candida albicans*), 新型隐球酵母 (*Cryptococcus neoformans*), 荚膜组织胞浆菌 (*Histoplasma capsulatum*), 烟曲霉 (*Aspergillus fumigatus*), 粗球孢子菌 (*Coccidioides immitis*), 巴西类球孢子菌 (*Paracoccidioides brasiliensis*), 皮炎芽生菌 (*Blastomyces dermatidis*), 卡氏肺孢子虫 (*Pneumocystis carinii*), *Penicillium marneffi* 和链格孢 (*Alternaria alternata*)。在本发明的方法中, 对象可以是哺乳动物。例如哺乳动物可以是人类、非人灵长类、小鼠、大鼠、猪、狗、猫、猴子、牛或马。

### 3、使用组合物做为研究工具的方法

可以用各种方式, 将公开的组合物用做研究工具。例如, 公开的组合物可以用于研究补体激活的抑制剂。

公开的组合物可以用做涉及与补体激活有关的疾病, 诸如癌症、病毒感染、细菌感染、寄生虫感染和真菌感染的诊断工具。CR2-融合蛋白将靶向补体激活位点, 并且标记的 CR2 融合蛋白可以诊断与补体激活有关的疾病。例如, 肿瘤反应性抗体将在肿瘤细胞上激活补体, CR2 将靶向该补体。然后, 标记的 CR2-Fc 可以在抗体打靶后扩大信号。

### D、组合物

公开了用于制备公开的组合物成分及用在这里所公开的方法中的组合物本身。这里公开了这些和其它材料, 并且可以理解当公开这些材料的组合、子集、相互作用、群组等时, 尽管可能不明确地一一具体提到这些化合物的各种单独和集合的组合及排列中的每种情况, 但这里具体地考

虑和描述了每种情况。例如，除非有特别相反的说明外，如果描述了特定的 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, 和/或公开和讨论了其特定的组合，和/或讨论了可以对多种分子（包括 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF）和/或其组合进行的多种修饰，则具体地考虑的是可能的 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF 或者其组合的每种组合和排列和可能的修饰。因此，如果公开了一类分子 A, B 和 C 及公开了一类分子 D, E 和 F 和组合分子的例子，A-D，那么即使没有单独列举每种情况，也单独和集合地考虑了每种情况，即意味着组合 A-E, A-F, B-D, B-E, B-F, C-D, C-E 和 C-F 均被认为公开了。同样，也公开了这些组合的任何子集或组合。因此，例如，也认为公开了 A-E, B-F 和 C-E 的亚群组。这个概念适用于本申请的所有方面，包括但不限于制备和使用所公开的组合物中的步骤。因此，如果存在可以进行的各种附加步骤，可以理解这些附加步骤之每个步骤均可以在所公开的方法的任何特定实施方式或实施方式的组合中实施。

### 1、序列相似性

可以理解，如这里所述术语同源性和同一性的应用意指与相似性相同的含义。因此，例如，如果在两个非天然序列之间使用术语同源性，可以理解这不一定表明这两个序列之间的进化关系，而是考虑它们核酸序列之间的相似性或相关性。测定两个进化相关分子之间同源性的许多方法常规地适用于任何两个或多个核酸或蛋白质以测定序列相似性，而不管它们是否是进化相关还是不相关的。

一般地，可以理解定义这里所公开的基因和蛋白质的任何已知的变异体和衍生物或可能产生的变异体和衍生物的一种方式是以相对于具体公开的序列的同源性来定义这些变异体和衍生物。也在这里以外的其它地方讨论了这里公开的具体序列的这种同一性。例如，SEQ ID NO:25 给出了 CR2 的一个特定序列，SEQ ID NO:26 给出了 SEQ ID NO:25 所编码的蛋白质，即 CR2 蛋白质的一个特定序列。具有地，公开了这里所公开的这些和其它基因和蛋白质的变异体，该变异体与所述序列具有至少 70、

71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99 百分比同源性。本领域技术人员容易理解如何确定两个蛋白质或核酸诸如基因的同源性。例如，可以在为使同源性在其最高水平而比对两个序列后，计算同源性。

通过公开的算法可以进行另一种计算同源性的方法。通过 Smith 和 Waterman *Adv. Appl. Math.* 2: 482(1981)的局部同源性算法，通过 Needleman 和 Wunsch, *J. Mol Biol.* 48: 443 (1970)的同源性比对算法，通过 Pearson 和 Lipman, *Proc. Natl. Acad. Sci. U. S. A.* 85: 2444 (1988)的相似性检索方法，通过计算机化的这些算法的执行程序(GAP, BESTFIT, FASTA,和 TFASTA, Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI)或者通过视察可以进行所比较的序列的最佳比对 (alignment)。

例如，通过 Zuker, *M. Science* 244: 48-52,1989, Jaeger 等. *Proc.Natl. Acad. Sci. USA* 86: 7706-7710,1989, Jaeger 等. *Methods Enzymol.* 183: 281-306,1989 中公开的算法可以获得核酸的相同类型同源性，这里将这些文献并入至少作为核酸比对相关材料的参考。可以理解通常地可以使用任何方法，并且在一些情况下，这些不同方法的结果可以不同，但是本领域的技术人员理解如果用这些方法中至少一种方法发现了同一性，那么就称这些序列具有所述的同一性，并且在这里公开。

例如，如这里所述，所述与另一个序列具有特定百分比同源性的序列指这些序列具有通过上述一种或多种计算方法所计算的所述同源性。例如，如果利用 Zuker 计算方法计算得到第一序列与第二序列具有 80% 同源性，即使用任何其它计算方法计算第一序列与第二序列没有 80% 同源性，那么如这里所定义，第一序列与第二序列也具有 80% 同源性。再如，如果利用 Zuker 计算方法及 Pearson 和 Lipman 计算方法计算得到第一序列与第二序列具有 80% 同源性，即使用 Smith 和 Waterman 计算方法、Needleman 和 Wunsch 计算方法、Jaeger 计算方法或任何其它计算方法计算第一序列与第二序列没有 80% 同源性，那么如这里所定义，第一序列

与第二序列也具有 80% 同源性。再如，如果利用任何一种计算方法计算都得到第一序列与第二序列具有 80% 同源性，那么如这里所定义，第一序列与第二序列具有 80% 同源性（尽管实际上，不同的方法将常常产生不同的同源性百分数计算结果）。

## 2、核酸

这里公开了基于核酸的各种分子，例如，包括编码例如 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-抗-C5, CR2-IgG1 Fc(人), CR2-IgM Fc, CR2-IgG3 Fc(鼠)或 CR2-CVF 的核酸以及各种功能性核酸。公开的核酸由例如核苷酸、核苷酸类似物或核苷酸替代物组成。这里讨论了这些和其它分子的非限制性例子。可以理解例如，当在细胞中表达载体时，表达的 mRNA 通常由 A, C, G 和 U 组成。同样地，可以理解例如，将反义分子通过例如外源递送方式引入到细胞或细胞环境中时，有利地，反义分子由核苷酸类似物组成，其中该核苷酸类似物降低了反义分子在细胞环境中的降解。

### a) 核苷酸和相关的分子

核苷酸是含有碱基部分、糖部分和磷酸部分的分子。核苷酸可以通过它们的磷酸部分和糖部分产生核苷酸间的键而连接在一起。核苷酸的碱基部分可以是腺嘌呤-9-基(A), 胞嘧啶-1-基(C), 鸟嘌呤-9-基(G), 尿嘧啶-1-基(U), 和胸腺嘧啶-1-基(T)。核苷酸的糖部分是核糖或脱氧核糖。核苷酸的磷酸部分是五价磷酸。核苷酸的非限制性例子是 3'-AMP (3'-腺苷一磷酸) 或 5'-GMP (5'-鸟苷一磷酸)。

核苷酸类似物是含有对碱基、糖或磷酸部分的某种修饰的核苷酸。对碱基部分的修饰包括 A, C, G 和 T/U 及不同的嘌呤或嘧啶碱基诸如尿嘧啶-5-基(.psi.)，次黄嘌呤-9-基(I) 和 2-氨基腺嘌呤-9-基的天然和合成修饰。修饰的碱基包括，但不限于 5-甲基胞嘧啶(5-me-C)，5-羟甲基胞嘧啶，黄嘌呤，次黄嘌呤，2-氨基腺嘌呤，腺嘌呤和鸟嘌呤的 6-甲基和其它烷基

衍生物,腺嘌呤和鸟嘌呤的 2-丙基和其它烷基衍生物, 2-硫尿嘧啶, 2-硫胸腺嘧啶和 2-硫胞嘧啶, 5-卤代尿嘧啶和胞嘧啶, 5-丙炔基尿嘧啶和胞嘧啶, 6-氮杂尿嘧啶(6-azouracil)、胞嘧啶和胸腺嘧啶, 5-尿嘧啶(假尿嘧啶), 4-硫尿嘧啶, 8-卤代, 8-氨基, 8-巯基, 8-烷硫基, 8-羟基和其它 8-取代的腺嘌呤和鸟嘌呤, 5-卤代, 特别是 5-溴代, 5-三氟甲基和其它 5-取代的尿嘧啶和胞嘧啶, 7-甲基鸟嘌呤和 7-甲基腺嘌呤, 8-氮杂鸟嘌呤和 8-氮杂腺嘌呤, 7-脱氮杂鸟嘌呤和 7-脱氮杂腺嘌呤, 3-脱氮杂鸟嘌呤和 3-脱氮杂腺嘌呤。其它碱基修饰可以例如,在美国专利号 3,687,808, Englisch 等, *Angewandte Chemie, International Edition*, 1991,30, 613, 和 Sanghvi, Y. S., 15 章, *Antisense Research and Applications*, 289-302 页, Crooke, S. T. 和 Lebleu, B. 编辑, CRC Press, 1993 中发现。一些核苷酸类似物, 诸如 5-取代的嘧啶, 6-氮杂嘧啶和 N-2, N-6 和 O-6 取代的嘌呤, 包括 2-氨基丙基腺嘌呤, 5-丙炔基尿嘧啶和 5-丙炔基胞嘧啶。5-甲基胞嘧啶可以增加双螺旋形成的稳定性。常常, 碱基修饰可以与例如糖修饰, 诸如 2'-O-甲氧基乙基联合以获得独特的特性诸如增加的双螺旋稳定性。有大量的美国专利诸如 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121; 5,596,091; 5,614,617; 和 5,681,941 详细地描述了一系列的碱基修饰。这里将这些专利中每篇专利并入为参考文献。

核苷酸类似物也可以包括糖部分的修饰。糖部分的修饰包括核糖和脱氧核糖的天然修饰及合成修饰。糖部分的修饰包括但不限于下面的在 2' 位置的修饰: OH; F; O-, S-, 或 N-烷基; O-, S-, 或 N-烯基; O-, S- 或 N-炔基; 或 O-烷基-O-烷基, 其中烷基、烯基和炔基可以是取代的或未取代的 C<sub>1</sub> 到 C<sub>10</sub> 烷基或 C<sub>2</sub> 到 C<sub>10</sub> 烯基和炔基。2' 糖修饰也包括但不限于 -O[(CH<sub>2</sub>)<sub>n</sub>O]<sub>m</sub>CH<sub>3</sub>, -O(CH<sub>2</sub>)<sub>n</sub>OCH<sub>3</sub>, -O(CH<sub>2</sub>)<sub>n</sub>NH<sub>2</sub>, -O(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>, -O(CH<sub>2</sub>)<sub>n</sub>-ONH<sub>2</sub>, 和 -O(CH<sub>2</sub>)<sub>n</sub>ON[(CH<sub>2</sub>)<sub>n</sub>CH<sub>3</sub>]<sub>2</sub>, 其中 n 和 m 是 1 到约 10。

在 2' 位置的其它修饰包括但不限于: C<sub>1</sub> 到 C<sub>10</sub> 低级烷基, 取代的低级烷基、烷芳基、芳烷基、O-烷芳基或 O-芳烷基、SH、SCH<sub>3</sub>、OCN、Cl、

Br、CN、CF<sub>3</sub>、OCF<sub>3</sub>、SOCH<sub>3</sub>、SO<sub>2</sub>CH<sub>3</sub>、ONO<sub>2</sub>、NO<sub>2</sub>、N<sub>3</sub>、NH<sub>2</sub>，杂环烷基、杂环烷基芳基、氨基烷基氨基、多烷基氨基(polyalkylamino)、取代的甲硅烷基、RNA切割基团(RNA cleaving group)、报道基团、嵌入剂、改进寡核苷酸药代动力学特征的基团，或者改进寡核苷酸药物动力学特征的基团和具有相似特征的其他取代基。也可以在糖上其它位置，特别地在3'末端核苷酸或2'-5'连接的寡核苷酸糖中的3'位置及5'末端核苷酸的5'位置进行相似的修饰。修饰的糖也包括在桥环氧位置含有修饰，诸如CH<sub>2</sub>和S的糖。核苷酸糖类似物也可以具有糖模拟物，诸如代替呋喃戊糖基糖的环丁基部分。有大量的美国专利教导了这种修饰糖结构的制备，诸如4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633;和5,700,920，这里将每篇专利整体并入为参考文献。

也可以在磷酸部分修饰核苷酸类似物。修饰的磷酸部分包括但不限于可以修饰磷酸部分使两个核苷酸之间的键含有硫代磷酸酯、手性硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基磷酸三酯、甲基和其它烷基磷酸酯，包括3'-亚烷基磷酸酯和手性磷酸酯，次磷酸酯，氨基磷酸酯，包括3'-氨基氨基磷酸酯和氨基烷基氨基磷酸酯，硫羰基氨基磷酸酯，硫羰基烷基磷酸酯，硫羰基烷基磷酸三酯和boranophosphates。可以理解两个核苷酸之间的这些磷酸酯或修饰的磷酸酯键可以是3'-5'键或2'-5'键，并且该键可以含有反转极性诸如3'-5'到5'-3'或者2'-5'到5'-2'。也包括各种盐、混合的盐和自由酸形式。大量的美国专利教导了如何制备和使用含有修饰磷酸的核苷酸，这些专利包括但不限于3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361;和5,625,050，这里将每篇专利并入为参考文献。

可以理解核苷酸类似物只需要含有单一修饰，但是也可以在核苷酸的

一个部分中或不同部分之间含有多个修饰。

核苷酸替代物是与核苷酸具有相似功能特性的分子,但是该分子不含有磷酸部分,诸如肽核酸(PNA)。核苷酸替代物是可以以 Watson-Crick 或 Hoogsteen 方式识别核酸但是通过非磷酸部分的其它部分连接在一起的分子。当核苷酸替代物与适合的靶核酸相互作用时,其能遵循双螺旋类结构。

核苷酸替代物是磷酸部分和/或糖部分被置换的核苷酸或核苷酸类似物。核苷酸替代物不含有标准磷原子。例如,短链烷基或环烷基核苷间键、混合的杂原子和烷基或环烷基核苷间键、或者一个或多个短链杂原子的或杂环的核苷间键可以替代磷酸。这包括具有(部分地由核苷的糖部分形成的)吗啉键;硅氧烷骨架;硫醚、亚砷和砷骨架;formacetyl和thioformacetyl骨架;亚甲基formacetyl和thioformacetyl骨架;含有烯烃的骨架;氨基磺酸酯骨架;亚氨基亚甲基(methyleneimino)和胍基亚甲基(methylenehydrazino)骨架;磺酸酯和氨磺酰骨架;酰胺骨架的那些;和具有混合的N,O,S和CH<sub>2</sub>组成部分的那些。大量的美国专利公开了如何制备和使用这些类型的磷酸置换,这些专利包括但不限于5,034,506;5,166,315;5,185,444;5,214,134;5,216,141;5,235,033;5,264,562;5,264,564;5,405,938;5,434,257;5,466,677;5,470,967;5,489,677;5,541,307;5,561,225;5,596,086;5,602,240;5,610,289;5,602,240;5,608,046;5,610,289;5,618,704;5,623,070;5,663,312;5,633,360;5,677,437;和5,677,439,这里将每篇专利并入为参考文献。

也可以理解在核苷酸替代物中,核苷酸的糖和磷酸部分都可以被替代,例如被酰胺类型的键(氨基乙基甘氨酸)(PNA)替代。美国专利5,539,082;5,714,331;和5,719,262教导了如何制备和使用PNA分子,这里将每篇专利并入为参考文献。(也参见Nielsen等,Science,1991,254,1497-1500)。

也可以连接其它类型分子(缀合物)和核苷酸或核苷酸类似物以增强,例如细胞摄取。缀合物可以化学地连接核苷酸或核苷酸类似物。这种缀合

物包括但不限于脂质部分诸如胆固醇部分(Letsinger 等, Proc. Natl. Acad. Sci. USA, 1989,86, 6553-6556), 胆酸(Manoharan 等, Bioorg. Med. Chem. Let. , 1994,4, 1053-1060), 硫醚, 例如己基-S-三苯甲基硫醇(Manoharan 等, Ann. N. Y. Acad. Sci., 1992,660, 306-309; Manoharan 等, Bioorg. Med. Chem. Let. , 1993,3, 2765-2770), 硫代胆固醇(Oberhauser 等, Nucl. Acids Res. , 1992,20, 533-538), 脂族链, 例如十二烷二醇或十一烷基基团(Saison-Behmoaras 等, EMBO J. , 1991,10,1111-1118 ; Kabanov 等, FEBS Lett., 1990,259, 327-330; Svinarchuk 等, Biochimie, 1993,75, 49-54), 磷脂, 例如二-十六烷基-rac-甘油或三乙基铵 1,2-二-O-十六烷基-rac-甘油-3-H-膦酸盐/酯 (Manoharan 等, Tetrahedron Lett., 1995,36, 3651-3654; Shea 等, Nucl. Acids Res. , 1990,18, 3777-3783), 聚胺或聚乙二醇链 (Manoharan 等, Nucleosides & Nucleotides, 1995,14, 969-973), 或金刚烷醋酸 (Manoharan 等, Tetrahedron Lett. , 1995,36,3651-3654), 棕榈基部分(Mishra 等, Biochim. Biophys. Acta, 1995,1264, 229-237), 或者十八胺或己基氨基-羰基-氧基胆固醇部分(Crooke 等, J. Pharmacol. Exp. Ther. , 1996,277, 923-937。大量的美国专利教导了这些缀合物的制备, 这些专利包括但不限于美国专利号 4,828, 979; 4,948, 882; 5,218, 105; 5,525, 465; 5,541, 313; 5,545, 730; 5,552, 538; 5,578, 717; 5, 580,731 ; 5,580, 731; 5,591, 584; 5,109, 124 ; 5,118, 802; 5,138, 045; 5,414, 077; 5,486, 603; 5,512, 439; 5,578, 718; 5,608, 046; 4,587, 044; 4,605, 735; 4,667, 025; 4,762, 779; 4,789, 737; 4,824, 941; 4,835, 263; 4,876, 335; 4,904, 582; 4,958, 013; 5,082, 830; 5,112, 963; 5,214, 136; 5,082, 830; 5,112, 963; 5,214, 136; 5,245, 022; 5,254, 469; 5,258, 506; 5,262, 536; 5,272, 250; 5,292, 873; 5,317, 098; 5,371, 241; 5, 391,723 ; 5,416, 203; 5, 451,463 ; 5,510, 475; 5,512, 667; 5,514, 785; 5,565, 552;5,567, 810; 5,574, 142; 5,585, 481; 5,587, 371.; 5,595, 726; 5,597, 696; 5,599, 923; 5,599, 928 和 5,688, 941, 这里将每篇专利并入为参考文献。

Watson-Crick 相互作用是至少与核苷酸、核苷酸类似物或核苷酸替

代物的 Watson-Crick 面的一种相互作用。核苷酸、核苷酸类似物或核苷酸替代物的 Watson-Crick 面包括基于嘌呤的核苷酸、核苷酸类似物或核苷酸替代物的 C2、N1 和 C6 位置和基于嘧啶的核苷酸、核苷酸类似物或核苷酸替代物的 C2、N3、C4 位置。

Hoogsteen 相互作用是在核苷酸或核苷酸类似物的 Hoogsteen 面上发生的相互作用，该面暴露在双螺旋 DNA 的大沟中。Hoogsteen 面包括嘌呤核苷酸的 N7 位置和 C6 位置的活性基团 (NH<sub>2</sub> 或 O)。

#### b) 序列

存在各种与 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠) 或 CR2-CVF 基因相关的序列，包括例如，这里所公开的序列或文献中可以获得的序列。这里将这些序列和其它序列整体及其中含有的各单个亚序列并入为参考。

这里使用 SEQ ID NO:25 中所示一个特定序列做为例子以示例公开的组合物和方法。应理解除非特别地另外说明外，涉及该序列的描述可以应用于与 CR2, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠) 或 CR2-CVF 相关的任何序列。本领域的技术人员理解如何解决序列差异和不同并调节与特定序列相关的组合物和方法以适应于其它相关序列 (即, CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠) 或 CR2-CVF 的序列)。根据这里所公开和本领域已知的信息，可以设计用于任何 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc,

CR2-IgG3 Fc (鼠)或 CR2-CVF 序列的引物和/或探针。

### 3、将组合物递送给细胞

有多种组合物和方法可以在体外或体内用于向细胞递送本发明融合蛋白组合物，免疫缀合物组合物和核酸组合物。优选地，可以在药物学上可接受的载体中向对象给药本发明组合物。在 *Remington: The Science and Practice of Pharmacy* (第 19 版), A. R. Gennaro 编辑, Mack Publishing Company, Easton, PA 1995 中描述了适宜的载体和它们的制剂。通常地，在制剂中使用适当量的药物学上可接受的盐以使得制剂变得等渗。药物学上可接受的载体的例子包括但不限于盐水，水：油乳剂，油：水乳剂，水：油：水乳剂，和林格溶液和葡萄糖溶液。优选地，溶液的 pH 从约 5 到约 8，更优选地从约 7 到约 7.5。其它载体包括缓释制剂，诸如含有抗体的固体疏水性聚合物半透性基质，该基质是有形物体的形式，例如薄膜、脂质体或微粒。对本领域技术人员显而易见的是，例如，根据给药的途径和施用的抗体浓度，有些载体是更优选的。

可以通过注射（例如，静脉内、腹膜内、皮下、肌内），或者通过可以确保以有效的形式向血流递送组合物的其它方法诸如输注向对象、患者或细胞施用本发明的组合物。优选局部或静脉内注射。

可以凭经验确定施用本发明组合物的有效剂量和方案，并且做出这种决定是本领域技术人员技术范围之内的事情。本领域的技术人员将理解必须给药的本发明组合物的剂量将根据例如，将接受组合物的对象、给药的途径、所用的特定类型组合物和正在施用的其它药物而变化。根据上面提到的因素，单独使用时本发明组合物的典型日剂量范围从每天约 1 $\mu$ g/kg 直到 100mg/kg 体重，或者更多。

#### a) 基于核酸的递送系统

有可以在体外或体内用于向细胞递送核酸的大量组合物和方法。大致可以将这些方法和组合物分成 2 类：基于病毒的递送系统和基于非病毒的递送系统。例如，可以通过多种直接递送系统，诸如电穿孔、脂质转染法、磷酸钙沉淀、质粒、病毒载体、病毒核酸、噬菌体核酸、噬菌体、粘粒或

通过在细胞或运载体诸如阳离子脂质体中转移遗传物质来递送核酸。例如，Wolff, J. A., 等, *Science*, 247,1465-1468, (1990);和 Wolff, J. A. *Nature*, 352, 815-818,(1991)描述了转染的适宜方法，包括病毒载体、化学转染子或物理机理的方法诸如电穿孔和 DNA 的直接扩散。这些方法是本领域所熟知的，并且容易适用于与这里所述组合物和方法一起使用。在一些情况下，将修改这些方法以特异地用于大 DNA 分子。而且，通过利用运载体的靶向特征，这些方法可以用于靶向一些疾病和细胞群。

转移载体可以是用于将基因递送到细胞中的任何核苷酸构建体(例如，质粒)，或者是递送基因的一般策略的一部分，例如重组逆转录病毒或腺病毒的一部分(Ram 等. *Cancer Res.* 53: 83-88, (1993))。

这里所用的质粒或病毒载体是将公开的核酸，诸如 SEQ ID NO:25 转运到细胞中而不出现降解的试剂，其包括在基因所递送至的细胞中造成基因表达的启动子。在一些实施方式中，CR2, DAF, CD59,CR1, MCP, Crry,IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1,CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠)或 CR2-CVF 来源于病毒或逆转录病毒。例如，病毒载体是腺病毒、腺相关病毒、疱疹病毒、痘苗病毒、脊髓灰质炎病毒、AIDS 病毒、神经元向性病毒、辛德毕斯病毒和其它 RNA 病毒，包括具有 HIV 骨架的病毒。也优选的是具有这些病毒的如下特征的任何病毒家族，所述特征为使这些病毒适于用做载体的特征。逆转录病毒包括 Maloney 鼠白血病病毒，MMLV，和做为载体表达 MMLV 的期望特性的逆转录病毒。逆转录病毒载体与其它病毒载体相比能携带更大的遗传有效负载，即转基因或标记基因，并且因为这个原因，它们是通常使用的载体。然而，在非增殖细胞中不能使用它们。腺病毒载体相对稳定，并且容易操作，具有高的滴度，可以以气溶胶制剂形式被递送，可以转染非分裂的细胞。痘病毒载体是大的病毒载体，有几个用于插入基因的位点，它们是热稳定的，并且可以在室温下贮藏。优选的实施方式是被改造以致抑制了病毒抗原激发的宿主生物体免疫应答的病毒载

体。优选的这种类型载体将携带白细胞介素 8 或 10 的编码区。

病毒载体可以比化学或物理方法具有更高的将基因引入到细胞中的反式作用能力（引入基因的能力）。通常，病毒载体含有非结构性早期基因，结构晚期基因，RNA 聚合酶 III 转录物，复制和衣壳化所需的末端反向重复序列，和控制病毒基因组转录和复制的启动子。当改造做为载体时，病毒通常有一个或多个早期基因被移除，并且基因或基因/启动子盒插入到病毒基因组中置换移除的病毒 DNA。这种类型的构建体通常携带有达到约 8kb 的外源遗传物质。通常由已经被改造以反式表达早期基因产物的细胞系提供移除的早期基因的必需功能。

### (1) 逆转录病毒载体

逆转录病毒是属于逆转录病毒科 (Retroviridae) 的动物病毒，包括任何类型、亚科、属或向性。一般地，Verma, I. M., *Retroviral vectors for gene transfer. In Microbiology-1985, American Society for Microbiology, 229-232 页, Washington, (1985)* 描述了逆转录病毒载体，这里将这篇文献并入为参考文献。在美国专利号 4,868, 116 和 4,980, 286; PCT 申请 WO 90/02806 和 WO89/07136 ;和 Mulligan, (*Science 260: 926-932 (1993)*) 中描述了使用逆转录病毒载体用于基因治疗的方法的例子；这里将其中所教导的内容并入为参考文献。

逆转录病毒实质上是已经将核酸载货包装在其中的包装物。核酸载货上带有包装信号，这确保了复制的子代分子在包装外壳中被有效地包装。除了包装信号外，还有对于病毒复制和复制的病毒的包装而言顺式所需的大量分子。通常，逆转录病毒基因组含有 gag, pol 和 env 基因，这些基因参与蛋白质外壳的制备。待要转移到靶细胞中的外源 DNA 通常置换 gag, pol 和 env 基因。逆转录病毒载体通常含有用于掺入到包装衣壳中的包装信号，指示 gag 转录单元起始的信号序列，逆转录必需的元件，包括结合逆转录 tRNA 引物的引物结合位点，在 DNA 合成期间引导 RNA 链转换的末端重复序列，位于 3'LTR 的 5'端做为第二条链 DNA 合成的合成引发位点的嘌呤富集序列和能使逆转录病毒 DNA 状态的插入片段插入到宿主基因

组中的 LTRs 末端附近的特异性序列。gag, pol 和 env 基因的移除允许约 8kb 的外源序列被插入到病毒基因组中, 并且可以逆转录, 而且一旦复制后可以包装成新的逆转录病毒颗粒。根据每种转录物的大小, 这一核酸量足以使得人们递送一到多个基因。优选地, 在插入片段中包含阳性或阴性选择性标记及其它基因。

因为已经移除了大多数逆转录病毒载体中的复制机器和包装蛋白 (gag, pol 和 env), 所以通常通过放置在包装细胞系中来产生载体。包装细胞系是用逆转录病毒转染或转化的细胞系, 该逆转录病毒含有复制和包装机器但是缺少包装信号。当带有所选择的 DNA 的载体被转染到这些细胞系中时, 通过辅助细胞顺式提供的机器, 含有目的基因的载体被复制和包装成新的逆转录病毒颗粒。该机器的基因组不被包装, 因为它们缺少必需的信号。

## (2) 腺病毒载体

复制缺陷的腺病毒的构建已有描述(Berkner 等, *J. Virology* 61: 1213-1220 (1987); Massie 等, *Mol. Cell. Biol.* 6: 2872-2883 (1986); Haj-Ahmad 等, *J. Virology* 57: 267-274 (1986); Davidson 等, *J. Virology* 61: 1226-1239 (1987); Zhang "Generation and identification of recombinant adenovirus by liposome-mediated transfection and PCR analysis" *BioTechniques* 15: 868-872 (1993) )。使用这些病毒做为载体的益处是它们向其它细胞类型的传播是受限的, 因为它们可以在最初感染的细胞中复制, 但是不能形成新的感染性病毒颗粒。已经证明, 将重组腺病毒体内直接递送到气道上皮, 肝细胞, 血管内皮, CNS 实质和多种其它组织位点后, 可以获得高效率基因转移(Morsy, *J. Clin. Invest.* 92: 1580-1586 (1993); Kirshenbaum, *J. Clin. Invest.* 92: 381-387 (1993); Roessler, *J. Clin. Invest.* 92: 1085-1092(1993) ; Moullier, *Nature Genetics* 4: 154-159 (1993); La Salle, *Science* 259: 988-990 (1993); Gomez-Foix, *J. Biol. Chem.* 267: 25129-25134 (1992); Rich, *Human Gene Therapy* 4 : 461-476 (1993); Zabner, *Nature Genetics* 6: 75-83 (1994); Guzman, *Circulation Research*

73: 1201-1207 (1993); Bout, *Human Gene Therapy* 5:3-10 (1994); Zabner, *Cell* 75: 207-216 (1993); Caillaud, *Eur. J. Neuroscience* 5: 1287-1291 (1993); 和 Ragot, *J. Gen. Virology* 74: 501-507 (1993) )。重组腺病毒通过结合特异性细胞表面受体, 随后通过受体介导的胞吞作用以野生型和复制缺陷型腺病毒相同的方式内化而获得基因转导(Chardonnet 和 Dales, *Virology* 40: 462-477 (1970); Brown 和 Burlingham, *J. Virology* 12: 386-396 (1973); Svensson 和 Persson, *J. Virology* 55: 442-449 (1985); Seth 等, *J. Virol.* 51: 650-655 (1984); Seth 等, *Mol. Cell. Biol.* 4: 1528-1533 (1984); Varga 等, *J. Virology* 65: 6061-6070 (1991); Wickham 等, *Cell* 73 : 309- 319(1993))。

病毒载体也可以是基于移除了 E1 基因的腺病毒的病毒载体, 并且在细胞系诸如人类 293 细胞系中产生这些病毒体。在另一个优选的实施方式中, 从腺病毒基因组移除了 E1 和 E3 基因。

### (3) 腺相关病毒载体

另一种类型的病毒载体基于腺相关病毒 (AAV)。这种缺陷的细小病毒是优选的载体, 因为它可以感染许多细胞类型并且对人类是非致病的。AAV 类型载体可以转运约 4 到 5kb, 并且已知野生型 AAV 稳定地插入到染色体 19 中。优选含有这种位点特异性整合特性的载体。这种类型载体的特别优选的实施方式是 Avigen, San Francisco, CA 产生的 P4.1C 载体, 其可以含有单纯疱疹病毒胸苷激酶基因, HSV-tk 和/或标记基因, 诸如编码绿色荧光蛋白 GFP 的基因。

在另一种类型 AAV 病毒中, AAV 含有一对末端反向重复序列 (ITRs), 该 ITR 位于至少一个含有启动子的盒子的侧翼, 该启动子引导与之可操作地连接的异源基因的细胞特异性表达。在上下文中异源指不是 AAV 和 B19 细小病毒天然的任何核苷酸序列或基因。

通常, 缺失 AAV 和 B19 编码区以产生安全的非细胞毒性的载体。AAV ITRs 或其修饰形式赋予了感染性和位点特异性整合, 但是不导致细胞毒性, 并且该启动子引导细胞特异性表达。这里将涉及 AAV 载体相关材料的美国专利号 6,261, 834 并入为参考文献。

因此,本发明的载体提供了 DNA 分子,该 DNA 分子能整合到哺乳动物染色体中而没有实质的毒性。

病毒和逆转录病毒中插入的基因通常含有启动子和/或增强子以帮助控制期望基因产物的表达。启动子通常是一个或多个 DNA 序列,当位于相对于转录起始位点而言相对固定的位置时,其发挥作用。启动子含有 RNA 聚合酶和转录因子基本相互作用所需的核心元件,并且可以含有上游元件和应答元件。

#### (4) 大负载的病毒载体

使用大的人疱疹病毒的分子遗传试验提供了在允许用疱疹病毒感染的细胞中克隆、繁殖和建立大的异源 DNA 片段的方法(Sun 等, *Nature genetics* 8: 33-41,1994 ; Cotter and Robertson, . *Curr Opin Mol Ther* 5: 633- 644,1999)。这些大的 DNA 病毒(单纯疱疹病毒(HSV)和 EB 病毒(EBV))具有潜力向特定细胞递送大于 150kb 的人异源 DNA 片段。EBV 重组体可以在感染的 B 细胞中以游离型 DNA 形式维持大片段的 DNA。各单独的克隆可以带有看似遗传稳定的多达 330kb 的人基因组插入片段。这些游离体的维持需要在 EBV 感染期间组成型表达的特异的 EBV 核蛋白, EBNA1。此外,这些载体可以用于转染,从而可以体外瞬时产生大量蛋白质。疱疹病毒扩增子系统也用于包装大于 220kb 的 DNA 片段和用于感染可以稳定地以游离体形式维持 DNA 的细胞。

其它有用的系统包括例如,复制和宿主限制的非复制痘苗病毒载体。

#### b) 基于非核酸的系统

可以用各种方法将公开的组合物递送给靶细胞。例如,可以通过电穿孔或者通过脂质转染,或者通过磷酸钙沉淀递送组合物。选择的递送机理将部分地依赖于靶向的细胞类型和是否在例如体内或体外发生递送。

因此,组合物除了包含例如公开的 CR2, DAF, CD59,CR1, MCP, Crry,IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2,CR2-CR1,CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2,CR2-IgG1 Fc (人),CR2-IgM Fc, CR2-IgG3 Fc (鼠)或 CR2-CVF

或载体外，还可以包含脂质诸如脂质体，诸如阳离子脂质体（例如，DOTMA, DOPE, DC-胆固醇）或阴离子脂质体。如果希望，脂质体可以进一步包含有利于靶向特定细胞的蛋白质。包含化合物和阳离子脂质体的组合物的给药方式可以是向输入靶器官的血液施用或者吸入呼吸道从而达到呼吸道的靶细胞。关于脂质体，参见，例如 Brigham 等. *Am. J. Resp. Cell. Mol. Biol.* 1: 95-100 (1989); Felgner 等. *Proc. Natl. Acad. Sci USA* 84: 7413-7417(1987); 美国专利号 4,897, 355。而且，化合物可以做为微囊的成分给药，该微囊可以靶向特定的细胞类型，诸如巨噬细胞，或者可以设计以特定速率或剂量从微囊扩散或递送化合物。

在上述公开的包括给药和将外源 DNA 摄取到对象的细胞中（即，基因转导或转染）的方法中，可以通过各种机制向细胞递送组合物。做为一个例子，利用商购脂质体制剂诸如 LIPOFECTIN, LIPOFECTAMINE (GIBCO-BRL, Inc., Gaithersburg, MD), SUPERFECT (Qiagen, Inc. Hilden, Germany) 和 TRANSFECTAM (Promega Biotec, Inc., Madison, WI) 及根据本领域标准程序开发的其它脂质体，可以通过脂质体进行递送。此外，可以通过电穿孔，可从 Genetronics, Inc. (San Diego, CA) 获得的技术，及通过 SONOPORATION 机器 (ImaRx Pharmaceutical Corp., Tucson, AZ) 体内递送本发明的核酸或载体。

材料可以在溶液、悬浮液中（例如，掺入微颗粒、脂质体或细胞中）。通过抗体、受体或受体配体，这些材料可以靶向特定的细胞类型。下面的参考文献是应用这种技术将特定蛋白质靶向肿瘤组织的例子 (Senter 等, *Bioconjugate Chem.*, 2: 447-451, (1991); Bagshawe, K. D., *Br. J. Cancer*, 60: 275-281, (1989); Bagshawe 等, *Br. J. Cancer*, 58: 700-703, (1988); Senter 等, *Bioconjugate Chem.*, 4: 3-9, (1993); Battelli 等, *Cancer Immunol. Immunother.*, 35: 421-425, (1992); Pietersz 和 McKenzie, *Immunolog. Reviews*, 129: 57-80, (1992); 和 Roffler 等, *Biochem. Pharmacol.*, 42: 2062-2065, (1991))。这些技术可以用于各种其它特定的细胞类型。载体诸如“隐形” (stealth) 和其它抗体缀合的脂质体（包括靶向结肠癌的脂

质介导的药物)，借助细胞特异性配体而介导 DNA 靶向的受体，淋巴细胞指导的肿瘤靶向，高度特异性治疗性逆转录病毒对体内鼠神经胶质瘤细胞的靶向。下面的参考文献是这项技术用于将特定蛋白质靶向肿瘤组织的例子（Hughes 等, Cancer Research, 49:6214-6220, (1989)；Litzinger 和 Huang, Biochimica et Biophysica Acta,1104:179-187, (1992)）。一般地，受体参与组成性或配体诱导的胞吞作用途径。这些受体簇集在披有网格蛋白的小窝中，通过披有网格蛋白的小泡进入细胞，通过酸化的内体（endosome），在内体中分拣受体，然后受体或者重新回到细胞表面、或者细胞内贮藏，或者在溶酶体中降解。内在化途径具有各种功能，诸如营养物质摄取，活化蛋白的移除，大分子的清除，病毒和毒素的机会进入，配体的解离和降解，及受体水平的调节。根据细胞类型、受体浓度、配体类型、配体化合价和配体浓度，许多受体遵循一条以上的细胞内途径。已经综述了受体介导的胞吞作用的分子和细胞机制（Brown 和 Greene, DNA and Cell Biology 10:6, 399-409 (1991)）。

通常，递送给细胞的将要整合到宿主细胞基因组中的核酸含有整合序列。特别是使用基于病毒的系统时，这些序列常常是病毒相关序列。也可以将这些病毒整合系统掺入到待递送的核酸中，其中利用基于非核酸的递送系统，诸如脂质体递送所述核酸，这样，递送系统中含有的核酸可以整合到宿主基因组中。

用于整合到宿主基因组中的其它常规技术包括，例如设计的用于促进与宿主基因组同源重组的系统。通常，这些系统依赖于位于待要表达的核酸的侧翼的序列，该序列与宿主细胞基因组中靶序列具有足够的同源性，这样就发生了载体核酸和靶核酸之间的重组，引起递送的核酸整合到宿主基因组中。本领域技术人员已知促进同源重组所必需的这些系统和方法。

### c) 体内/离体

如上述，可以在药物学可接受的载体中施用组合物，并且可以用本领域已知的各种机制（例如，裸 DNA 的摄取、脂质体融合、通过基因枪的 DNA 肌肉注射，胞吞作用等等），体内和/或离体地将该组合物递送给对

象细胞。

如果使用离体的方法，可以移出细胞或组织，并且根据本领域熟知的标准方法在体外维持。可以通过任何基因转移机制诸如，例如磷酸钙介导的基因递送、电穿孔、微注射或脂蛋白体将组合物引入到细胞中。然后，根据用于该细胞或组织类型的标准方法，可以将转导的细胞输注（例如，在药理学可接受的载体中）或等位地移植回对象中。已知用于将各种细胞移植或输注到对象中的标准方法。

#### 4、表达系统

递送给细胞的核酸通常含有表达控制系统。例如，病毒和逆转录病毒系统中的插入基因通常含有启动子和/或增强子以帮助控制期望基因产物的表达。启动子通常是一个或多个 DNA 序列，当位于相对于转录起始位点而言相对固定位置时，其起作用。启动子含有 RNA 聚合酶和转录因子基本相互作用所需的核心元件，并且可以含有上游元件和应答元件。

##### a) 病毒启动子和增强子

可以从各种来源，例如病毒诸如：多瘤病毒、猿猴病毒 40 (SV40)、腺病毒、逆转录病毒、乙型肝炎病毒和最优选的巨细胞病毒的基因组，或者从异源哺乳动物启动子，例如  $\beta$ -肌动蛋白启动子获得在哺乳动物宿主细胞中控制载体转录的优选的启动子。可以方便地以 SV40 限制片段获得 SV40 病毒的早期和晚期启动子，所述 SV40 限制片段也含有 SV40 病毒复制起点(Fiers 等, Nature, 273: 113 (1978))。可以方便地以 HindIII E 限制片段获得人巨细胞病毒立即早期启动子(Greenway, P. J.等, Gene 18: 355-360 (1982))。当然，这里也可以使用来源于宿主细胞或相关物种的启动子。

增强子通常指在距离转录起始位点非固定距离处起作用的 DNA 序列，并且其可以在转录单元的 5' (Laimins, L.等, Proc. Natl. Acad. Sci. 78: 993 (1981))或 3' (Lusky, M. L.,等, Mol. Cell Bio. 3: 1108 (1983))。而且，增强子可以位于内含子中(Banerji, J. L.等, Cell 33: 729 (1983))及编码序列本身中(Osborne, T. F., et al., Mol. Cell Bio. 4: 1293 (1984))。它们的长度通常是 10 和 300bp 之间，并且它们顺式发挥作用。增强子的功能是增强从附近启

动子的转录。增强子也常常含有介导转录调节的应答元件。启动子也可以含有介导转录调节的应答元件。增强子常常决定基因的表达调节。尽管现在已知许多来源于哺乳动物基因（珠蛋白、弹性蛋白酶、白蛋白、胎蛋白和胰岛素）的增强子序列，但是通常人们使用来源于真核生物细胞病毒的增强子用于一般表达。优选的例子是 SV40 复制起点晚期侧上的增强子（bp100-270），巨细胞病毒早期启动子增强子，多瘤病毒复制起点晚期侧上的增强子和腺病毒增强子。

启动子和/或增强子可以由触发它们功能的光或特异性化学事件特异性地激活。可以用试剂诸如四环素和地塞米松调节系统。还有各种可以通过暴露于辐射，诸如  $\gamma$  辐射，或者烷基化化学治疗药物来增强病毒载体基因表达的方法。

在一些实施方式中，启动子和/或增强子区可以是组成型启动子和/或增强子，从而最大化待要转录的转录单元区的表达。在一些构建体中，启动子和/或增强子区在所有真核生物细胞类型中均具有活性，即使其只在特定的时间于特定的细胞类型中表达。这种类型的优选启动子是 CMV 启动子（650 个碱基）。其它优选的启动子是 SV40 启动子，巨细胞病毒（全长启动子）和逆转录病毒载体 LTR。

已经证明可以克隆所有的特异性调节元件，并且该特异性调节元件可以用于构建在特异性细胞类型，诸如黑素瘤细胞中选择性表达的表达载体。胶质原纤维酸性蛋白（GFAP）启动子已经用于选择性地胶质来源细胞中表达基因。

真核宿主细胞（酵母、真菌、昆虫、植物、动物、人类或有核的细胞）中使用的表达载体也可以含有转录终止必需的序列，该序列可以影响 mRNA 表达。在编码组织因子蛋白的 mRNA 的非翻译部分中，这些区域转录为聚腺苷酸化片段。3'非翻译区也包含转录终止位点。优选地，转录单元也含有聚腺苷酸化区。这种区的一个益处是它增加了转录单元象 mRNA 一样被加工和转运的可能性。已经充分确立了聚腺苷酸化信号的鉴定及其在表达构建体中的用途。优选地，在转基因构建体中使用同源的聚

腺苷酸化信号。在一些转录单元中，聚腺苷酸化区来源于 SV40 早期聚腺苷酸化信号，并且由约 400 个碱基组成。也优选地，转录单元含有单独的或者联合上述序列的其它标准序列，以改进构建体的表达或稳定性。

#### b) 标记

病毒载体可以包含编码标记产物的核酸序列。这种标记产物用于确定是否基因已经递送给细胞，以及基因递送给细胞后是否表达了该基因。优选的标记基因是编码  $\beta$ -半乳糖苷酶的大肠杆菌 lacZ 基因和绿色荧光蛋白。

在一些实施方式中，标记可以是选择性标记。哺乳动物细胞的适宜选择性标记的例子是二氢叶酸还原酶 (DHFR)、胸苷激酶、新霉素、新霉素类似物 G418、潮霉素和嘌呤霉素。当将这种选择性标记成功地转移到哺乳动物宿主细胞中时，如果放置在选择性压力下，转化的哺乳动物宿主细胞可以存活。有两种广泛使用的不同类型的选择性方案。第一种方案基于细胞代谢和突变细胞系的使用，该突变细胞系缺少独立于补料培养基而生长的能力。两个例子是 CHO DHFR 细胞和小鼠 LTK 细胞。在没有添加营养物诸如胸苷或次黄嘌呤的情况下，这些细胞缺少生长的能力。因为这些细胞缺少完整核苷酸合成途径所必需的一些基因，所以只有在补料培养基中提供了缺少的核苷酸时它们才能存活。对补料培养基的一个替代方案是向缺少 DHFR 或 TK 基因的细胞中引入相应完整的基因，由此改变它们的生长要求。没有被 DHFR 或 TK 基因转化的细胞个体将不能在非补料培养基中存活。

第二种方案是显性选择，所述显性选择指可以在任何细胞类型上使用的选择方案，不需要使用突变细胞系。这些方案通常使用药物阻止宿主细胞的生长。具有新基因的这些细胞将表达赋予药物抗性的蛋白质，并且在筛选中可以存活。这种显性筛选的例子使用药物新霉素 (Southern P. 和 Berg, P., J. Molec. Appl. Genet. 1: 327 (1982) ), 霉酚酸 (Mulligan, R. C. 和 Berg, P. Science 209: 1422 (1980) ) 或潮霉素 (Sugden, B. 等, Mol. Cell. Biol. 5: 410-413 (1985) )。这 3 个例子使用在真核生物控制下的细菌基因以分别赋予对适当药物 G418 或新霉素 (geneticin), xgpt (霉酚酸) 或潮霉素的

抗性。其它药物包括新霉素类似物 G418 和嘌呤霉素。

## 5、肽

### a) 蛋白质变异体

如这里所述, 存在大量 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠) 和 CR2-CVF 蛋白质的变异体, 这些变异体是已知的且是本文所考虑的。此外, 除了已知的功能性 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠) 和 CR2-CVF 的株系变异体, 尚有也可以在本公开的方法和组合物中起作用的 CR2、DAF、CD59、CR1、MCP、Crry、IgG1、IgM、IgG3、CVF、CR2-DAF、DAF-CR2、CR2-CD59、CD59-CR2、CR2-CR1、CR1-CR2、CR2-MCP、MCP-CR2、CR2-Crry、Crry-CR2、CR2-IgG1 Fc (人)、CR2-IgM Fc、CR2-IgG3 Fc (鼠) 和 CR2-CVF 蛋白质的衍生物。本领域技术人员充分理解蛋白质变异体和衍生物, 并且蛋白质变异体和衍生物可以包含氨基酸序列修饰。例如, 氨基酸序列修饰通常分为 3 类修饰: 置换、插入或缺失变异体。插入包括单个或多个氨基酸残基的氨基和/或羧基末端融合及序列内插入。插入通常是比氨基或羧基末端融合要小的插入, 例如约 1 到 4 个残基。通过体外交联或者通过利用以编码融合蛋白的 DNA 转化的重组细胞培养物, 可以将靶序列融合至足够大以致可以赋予免疫原性的多肽上, 制备诸如实施例中所述的免疫原性融合蛋白衍生物。缺失的特征是从蛋白质序列中移除一个或多个氨基酸残基。通常, 在蛋白质分子中任何一个位点缺失不多于约 2 到 6 个残基。通常, 通过编码蛋白质的 DNA 中核苷酸的位点特异性诱变, 由此产生编码变异体的 DNA, 并且此后在重组细胞培养物中表达该 DNA 而制备这些变异体。在具有已知序列的 DNA 中于预定位点进行置换诱变的技术是熟知的, 例如 M13 引物诱变和 PCR 诱变。通常, 氨基酸置换是单个残基置换, 但是也可以一次出现在多个不同的位置; 插入通常在约 1 到 10 个氨基酸残基的级别上; 缺失的范围是约

1 到 30 个残基。优选地，在相邻碱基对中进行缺失或插入，即，2 个残基的缺失或 2 个残基的插入。可以联合置换、缺失、插入或它们的任何组合以获得最终的构建体。突变不应当导致序列脱离阅读框，并且优选地，突变不导致可以产生二级 mRNA 结构的互补区。置换变异体是移除至少一个残基并且在该位置插入了不同残基的变异体。通常，根据下表 1 和 2 进行这种置换，并且这种置换称为保守性置换。

表 1 氨基酸缩写

氨基酸	缩写
丙氨酸	Ala A
allosoleucine	Alle
精氨酸	Arg R
天冬酰胺	Asn N
天冬氨酸	Asp D
半胱氨酸	Cys C
谷氨酸	Glu E
谷氨酰胺	Gln Q
甘氨酸	Gly G
组氨酸	His H
异亮氨酸	Ile I
亮氨酸	Leu L
赖氨酸	Lys K
苯丙氨酸	Phe F
脯氨酸	Pro P
焦谷氨酸	pGlu
丝氨酸	Ser S
苏氨酸	Thr T
酪氨酸	Tyr Y
色氨酸	Trp W
缬氨酸	Val V

表 2 氨基酸置换

原始残基示例性保守置换，其它置换是本领域已知的

Ala ; Ser
Arg; Lys; Gln
Asn;Gln ; His
Asp; Glu
Cys; Ser
Gln ; Asn, Lys
Glu; Asp
Gly; Pro
His; Asn; Gln
Ile ; Leu; Val
Leu; Ile ; Val
Lys; Arg;Gln ;
Met ; Leu ; Ile
Phe; Met; Leu ; Tyr
Ser; Thr
Thr; Ser
Trp ; Tyr
Tyr; Trp; Phe
Val; Ile ; Leu

通过选择没有表 2 中的置换保守的那些置换，即选择在维持 (a) 置换区中多肽骨架的结构，例如片层或螺旋构象，(b) 靶位点处分子的电荷或疏水性或 (c) 侧链的大小方面具有更显著不同作用的残基，可以造成功能或免疫学身份的实质性改变。通常预期会在蛋白质特性上造成最大改变的置换是 (a) 亲水残基，例如丝氨酸或苏氨酸置换 (或被置换为) 疏水

残基, 例如亮氨酸、异亮氨酸、苯丙氨酸、缬氨酸或丙氨酸;  
(b) 半胱氨酸或脯氨酸置换 (或被置换为) 任何其它的残基; (c) 具有正电侧链的残基, 例如赖氨酸、精氨酸或组氨酸置换 (或被置换为) 负电残基; 例如谷氨酸或天冬氨酸; 或 (d) 具有大侧链的残基, 例如苯丙氨酸置换 (或被置换为) 没有侧链的残基, 例如甘氨酸, 在本情况中, (e) 增加硫酸化和/或糖基化位点的数量。

例如, 一个氨基酸残基与另一个生物学和/或化学相似的氨基酸残基的替代是本领域技术人员已知的保守性置换。例如, 保守性置换将是用一个疏水性残基替代另一个疏水性残基, 或者用另一个极性残基置换另一个极性残基。这些置换包括组合诸如, Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; 和 Phe, Tyr。这里提供的嵌合多肽中包括每个明确公开序列的这种保守性置换变异。

可以利用置换或缺失诱变以插入 N-糖基化位点 (Asn-X-Thr/Ser) 或 O-糖基化位点 (Ser 或 Thr)。也可能希望缺失半胱氨酸或其它易变残基。缺失或置换潜在的蛋白酶解位点, 例如 Arg, 可以例如, 通过缺失或用谷氨酰胺或组氨酸置换其中一个碱性残基而完成。

一些翻译后衍生化是重组宿主细胞对表达的多肽的作用结果。谷氨酰胺和天冬酰胺残基常常在翻译后被脱酰胺成为相应的谷氨酸和天冬氨酸残基。做为可替代的方案, 可以在温和酸性条件下将这些残基脱酰胺。其它的翻译后修饰包括脯氨酸和赖氨酸的羟基化, 丝氨酸或苏氨酸残基的羟基基团的磷酸化, 赖氨酸、精氨酸和组氨酸侧链的 O-氨基的甲基化 (T. E. Creighton, *Proteins: Structure and Molecular Properties*, W. H. Freeman & Co., San Francisco 79-86 页 [1983]), N-末端氨基的乙酰化, 和在一些情况下, C-末端羧基的酰胺化。

可以理解定义这里公开的蛋白质的变异体和衍生物的一个方法是根据与具体已知序列的同源性/同一性来定义变异体和衍生物。例如, SEQ ID NO: 26 给出了 CR2 的一个特定序列, SEQ ID NO: 2 给出了 DAF 蛋白质的一个特定序列。而具体公开了这里公开的这些和其它蛋白质的变异体,

该变异体与所述序列具有至少 70%或 75%或 80%或 85%或 90%或 95%的同源性。本领域技术人员易于理解如何确定两种蛋白质的同源性。例如，可以在使同源性处于其最高水平而比对两个序列后计算同源性。

通过公开的算法可以进行另一种计算同源性的方法。通过 Smith 和 Waterman *Adv. Appl. Math.* 2: 482(1981)的局部同源性算法，通过 Needleman 和 Wunsch, *J. Mol Biol.* 48: 443 (1970)的同源性比对算法，通过 Pearson 和 Lipman, *Proc. Natl. Acad. Sci. U. S. A.* 85: 2444 (1988)的相似性检索方法，通过这些算法的计算机化实施(GAP, BESTFIT, FASTA, 和 TFASTA, Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI)或者通过视察可以进行所比较的序列的最佳比对。

例如，通过 Zuker, *M. Science* 244: 48-52,1989, Jaeger 等. *Proc.Natl. Acad. Sci. USA* 86: 7706-7710,1989, Jaeger 等. *Methods Enzymol.* 183: 281-306,1989 中公开的算法可以获得核酸的相同类型同源性，这里将该文献并入至少作为核酸比对相关材料的参考。

可以理解保守性突变和同源性的描述可以以任何组合方式联合，诸如与特定序列具有至少 70%同源性的实施方式，其中变异是保守性突变。

当本说明书讨论各种蛋白质和蛋白质序列时，应理解也公开了可以编码这些蛋白质序列的核酸。这将包括与特定蛋白质序列相关的所有简并序列，即，具有编码一个特定蛋白质序列的序列的所有核酸及编码公开的蛋白质序列的变异体和衍生物的所有核酸，包括简并核酸。因此，尽管不可能在这里写出每个特定的核酸序列，但是应理解实际上通过公开的蛋白质序列已在这里公开和描述了每个序列。例如，在 SEQ ID NO: 25 中给出了可以编码 SEQ ID NO: 26 中所示蛋白质序列的许多核酸序列之一。此外，例如，SEQ ID NO: 29 中示出了公开的 SEQ ID NO: 26 的保守性衍生物，其中在位置 9 的异亮氨酸 (I) 被改变为缬氨酸 (V)。应理解对于此突变，也公开了编码任何公开序列的这种特定衍生物的所有核酸序列。也应理解尽管氨基酸序列不表示怎样的具体 DNA 序列在生物体中编码该蛋白质，

但当这里描述所公开蛋白质的特定变异体时，也已知且在这里描述和公开了编码该蛋白质的起源蛋白质的已知核酸序列。

## 6、抗体

### a) 抗体概述

如这里所述，这里所用的术语“抗体”具有广泛的意义，并且包括多克隆和单克隆抗体。除了完整的免疫球蛋白分子，术语“抗体”也包括这些免疫球蛋白分子的片段或聚合物，和人或人源化形式的免疫球蛋白分子或其片段。可以利用这里所述的体外测定法，或者通过类似的方法检测抗体的期望活性，随后，根据已知的临床试验方法检测它们体内的治疗和/或预防活性。

这里所用的术语“抗体”包括但不限于任何类型的完整免疫球蛋白(即，完整的抗体)。天然抗体通常是异源四聚体糖蛋白，由2条相同的轻(L)链和2条相同的重(H)链组成。通常，每条轻链通过一个共价二硫键连接重链，而在不同免疫球蛋白同种型的重链之间二硫键的数量不同。每条重和轻链也具有规则间隔的链内二硫键。每条重链的一个末端具有可变结构域(V(H))，接着是多个恒定结构域。每条轻链的一个末端具有可变结构域(V(L))，另一端具有恒定结构域；轻链的恒定结构域与重链的第一恒定结构域对齐，并且轻链的可变结构域与重链的可变结构域对齐。据认为特定的氨基酸残基形成轻链和重链可变结构域之间的界面。来源于任何脊椎动物物种的抗体轻链均可以根据它们恒定结构域的氨基酸序列被归到称为 $\kappa$ 和 $\lambda$ 的两个明显不同类型中的一种类型。根据免疫球蛋白重链恒定结构域的氨基酸序列，可以将免疫球蛋白归为不同的类别。有5大类人免疫球蛋白：IgA, IgD, IgE, IgG 和 IgM，并且一些类别的免疫球蛋白可以进一步分为亚类(同种型)，例如 IgG-1, IgG-2, IgG-3, 和 IgG-4; IgA-1 和 IgA-2。本领域技术人员知道小鼠的相应类别。对应于不同类免疫球蛋白的重链恒定结构域分别被称为 $\alpha$ 、 $\delta$ 、 $\epsilon$ 、 $\gamma$  和  $\mu$ 。

这里使用术语“可变”描述可变结构域中的一些部分，其在抗体之间具有不同的序列并在每种特定抗体对其特定抗原的结合和特异性中具有用

途。然而，可变性通常不是平均分布在整个抗体的可变结构域上。通常集中在轻链和重链可变结构域中称为互补决定区(CDRs)或超变区的3个片段中。可变结构域中较高保守性的部分称为构架(FR)。天然重链和轻链的可变结构域都各包含4个FR区，大体上采用了通过3个CDRs连接的 $\beta$ 片层构型，该3个CDRs形成环连接该 $\beta$ 片层结构并在一些情况下形成 $\beta$ 片层结构的部分。每条链上CDRs通过FR区而紧靠在一起，并且与来源于另一链的CDRs一起促成抗体的抗原结合位点的形成(参见，Kabat E. A.等,"Sequences of Proteins of Immunological Interest,"National Institutes of Health, Bethesda, Md.(1987))。恒定结构域不直接地参与抗体和抗原的结合，但是显示各种效应子功能，诸如依赖于抗体的细胞毒性中抗体的参与。

这里所用的术语“抗体或其片段”包括具有双或多个抗原或表位特异性的嵌合抗体和杂合抗体，和诸如 scFv, sFv, F(ab')<sub>2</sub>, Fab', Fab 等片段，包括杂合片段。因此，提供了保留结合抗体的特异性抗原的能力的抗体片段。例如，术语“抗体或其片段”的含义中包括了维持补体结合活性的抗体片段。可以用本领域已知的技术制备这种抗体和片段，并且可以根据实施例所述的方法和用于产生抗体和筛选抗体特异性和活性的一般方法来筛选特异性和活性(参见，Harlow 和 Lane. *Antibodies, A Laboratory Manual*. Cold Spring Harbor Publications, New York, (1988))。

“抗体或其片段”的含义中也包括抗体片段和抗原结合蛋白质(单链抗体)的缀合物，例如，如美国专利号 4,704, 692 中所述，因此将其内容并入为参考文献。

这里所用的术语“抗体”也可以指人抗体和/或人源化抗体。许多非人抗体(例如，来源于小鼠、大鼠或兔子的抗体)在人类中具有天然抗原性，因此当向人施用这些抗体时，可以产生不期望的免疫应答。因此，本发明方法中应用人或人源化抗体可以减少由施用给人的抗体激发不期望的免疫应答的可能性。

可选地，在其它物种中产生抗体，并且经过“人源化”后在人类中施

用该抗体。非人（例如鼠的）抗体的人源化形式是嵌合免疫球蛋白、免疫球蛋白链或其片段（诸如，Fc, scFv, sFv, Fv, Fab, Fab', F(ab')<sub>2</sub>, 或者抗体的其它抗原结合亚序列），其含有最少的来源于非人免疫球蛋白的序列。人源化抗体包括人免疫球蛋白（受体抗体），其中用具有期望特异性、亲和性和能力的来源于非人物种（供体抗体）诸如小鼠、大鼠或兔子的 CDR 的残基替代受体互补决定区（CDR）的残基。在一些情况下，用相应的非人残基替代人免疫球蛋白的 Fv 构架残基。人源化抗体也可以包含在受体抗体中及在输入的 CDR 或构架序列中均不存在的残基。一般地，人源化抗体将基本包含至少一个，和通常 2 个可变结构域的全部，其中所有或基本所有的 CDR 区对应于非人免疫球蛋白的 CDR 区，并且所有或基本所有的 FR 区是具有人免疫球蛋白共有序列的 FR 区。可选地，人源化抗体也将包含免疫球蛋白恒定区的至少一部分（Fc），通常是人免疫球蛋白的部分(Jones 等, *Nature*, 321: 522-525 (1986); Riechmann 等, *Nature*, 332: 323-327 (1988);和 Presta, *Curr. Op. Struct. Biol.*, 2: 593-596 (1992))。

人源化非人抗体的方法是本领域已知的。一般地，人源化抗体具有从非人来源引入的一个或多个氨基酸残基。这些非人氨基酸残基常常被称为“输入”残基，通常从“输入”可变结构域获得该“输入”残基。可以通过将人抗体的相应序列置换为啮齿动物的 CDR 或 CDR 序列，基本根据 Winter 和其同事的方法进行人源化(Jones 等, *Nature*, 321: 522-525 (1986); Riechmann 等, *Nature*, 332: 323-327 (1988); Verhoeyen 等, *Science*, 239: 1534-1536 (1988) )。因此，这种“人源化”抗体是嵌合抗体（美国专利号 4,816,567），其中用来源于非人物种的相应序列置换了实质上不足全部的人可变结构域。实际上，人源化抗体通常是人抗体，其中一些 CDR 残基和可能的一些 FR 残基被来源于啮齿动物抗体中类似位点的残基置换。

为了减少抗原性，选择用于制备人源化抗体的人轻链和重链可变结构域是非常重要的。根据“最佳匹配方法”（best-fit method），相对于已知人可变结构域序列的完整文库，筛选啮齿动物抗体的可变结构域序列。然后，接受最接近啮齿动物序列的人序列做为人源化抗体的人构架（FR）

(Sims 等, *J. Immunol.*, 151: 2296 (1993)和 Chothia 等, *J. Mol. Biol.*, 196: 901 (1987))。另一种方法使用来源于特定轻链或重链亚型的所有人抗体的共有序列的特定构架。相同的构架可以用于几种不同的人源化抗体(Carter 等, *Proc. Natl. Acad. Sci. USA*, 89: 4285 (1992); Presta 等, *J. Immunol.*, 151: 2623 (1993))。

人源化抗体保留对抗原的高亲合力和其它有利的生物特性也是重要的。为了达到该目的, 根据优选的方法, 利用亲本和人源化序列的三维模型, 通过分析亲本序列和各种概念人源化产物而制备人源化抗体。本领域技术人员普遍可以获得并且熟悉三维免疫球蛋白模型。现有可以使用的计算机程序, 该计算机程序可以说明和显示选定的候选免疫球蛋白序列的可能三维构象结构。检查这些显示的结构可以分析残基在候选免疫球蛋白序列的功能中的可能作用, 即分析影响候选免疫球蛋白结合其抗原的能力的残基。以这种方式, 可以从共有和输入序列选择和组合 FR 残基, 以便获得期望的抗体特性诸如对靶抗原增加的亲合性。一般地, CDR 残基直接地和最实质性地参与影响抗原结合(参见, 1994年3月3日公开的 WO 94/04679)。

抗体消化中产生的 Fab 片段也含有轻链的恒定结构域和重链的第一恒定结构域。Fab'片段由于在重链结构域羧基末端添加几个残基, 包括来源于抗体铰链区的一个或多个半胱氨酸, 而不同于 Fab 片段。F(ab')<sub>2</sub>片段是二价片段, 其包含在铰链区通过二硫键连接的 2 个 Fab'片段。这里 Fab'-SH 表示其中恒定结构域的半胱氨酸残基具有自由巯基的 Fab'。最初, 产生 Fab'片段对形式的抗体片段, 其中该 Fab'片段对之间具有铰链半胱氨酸。抗体片段的其它化学偶联也是已知的。

也提供了分离的免疫原性特异的抗体互补位或片段。通过化学或机械地破坏抗体分子, 可以从整个抗体分离抗体的特异性免疫原性表位。通过这里所教导的方法, 检测由此获得的纯化片段的免疫原性和特异性。可选地, 可以直接合成抗体的免疫反应性互补位。免疫反应性片段定义为具有来源于抗体氨基酸序列的至少约 2 到 5 个连续氨基酸的氨基酸序列。

产生包含本发明抗体的蛋白质的一种方法是通过蛋白质化学技术将 2 或多个肽或多肽连接在一起。例如,利用目前可用的实验室设备,利用 Fmoc (9-芴基甲氧羰基)或 Boc(叔丁氧羰基)化学可以化学合成肽或多肽 (Applied Biosystems, Inc., Foster City, CA)。本领域技术人员容易理解可以例如,通过标准化学反应合成对应于本发明抗体的肽或多肽。例如,可以合成肽或多肽,并且不从其合成树脂上切割下来,而可以合成抗体的其它片段并且随后将其从树脂上切割下来,由此暴露了在该其它片段上被功能性封闭的端基。通过肽缩合反应,这两个片段可以分别在它们的羧基和氨基末端通过肽键共价连接以形成抗体或其片段(Grant GA (1992) *Synthetic Peptides : A User Guide*. W. H. Freeman 和 Co., N. Y. (1992);Bodansky M 和 Trost B.,编辑(1993) *Principles of Peptide Synthesis*. Springer-Verlag Inc., NY。做为替代的方案,如上所述在体内独立合成肽或多肽。一旦分离这些独立的肽或多肽,就可以通过相似的肽缩合反应,连接它们以形成抗体或其片段。

例如,克隆或合成的肽片段的酶促连接允许相对短的肽片段被连接以产生更大的肽片段、多肽或整个蛋白质结构域(Abrahmsen L 等, *Biochemistry*, 30: 4151 (1991) )。做为可替代的方案,合成肽的天然化学连接可以用于从较短的肽片段合成地构建大的肽或多肽。这种方法包括 2 步化学反应(Dawson 等. *Synthesis of Proteins by Native Chemical Ligation*. *Science*, 266: 776-779 (1994) )。第一步是未保护的合成肽- $\alpha$ -硫酯与另一含有氨基末端 Cys 残基的未保护肽片段的化学选择性反应,以产生硫酯连接的中间体作为最初的共价产物。不改变反应条件,该中间体经自发的快速分子内反应在连接位点形成天然的肽键。通过制备人白细胞介素 8 (IL-8) 示例了这种天然的化学连接方法在蛋白质分子的全合成中的应用 (Baggiolini M 等. (1992) *FEBS Lett.* 307: 97-101;Clark-Lewis I 等, *J. Biol. Chem.*, 269: 16075 (1994); Clark-Lewis I 等, *Biochemistry*, 30 :3128 (1991) );Rajarathnam K 等, *Biochemistry* 33: 6623-30 (1994))。

做为可替代的方案,可以化学连接未保护的肽片段,其中由于化学连

接在肽片段之间形成的键是非天然(非肽)键(Schnolzer, M 等. *Science*, 256: 221 (1992) )。这种技术已经用于合成蛋白质结构域类似物及大量具有全部生物活性的相对纯的蛋白质(deLisle Milton RC 等, *Techniques in Protein Chemistry IV*. Academic Press, New York, 257-267 页(1992))。

本发明也提供了具有生物活性的抗体片段。本发明的多肽片段可以是重组蛋白质,其中通过将编码多肽的核酸克隆在能产生其多肽片段的表达系统,诸如腺病毒或杆状病毒表达系统中获得所述重组蛋白质。例如,可以从特异性杂交瘤确定抗体的活性结构域,其中该杂交瘤可以引起与抗体和 Fc 受体的相互作用相关的生物学效果。例如,发现对抗体活性或结合特异性或亲合性不起作用的氨基酸可以被缺失,而不损失相应活性。例如,在各种实施方式中,可以从天然或修饰的非免疫球蛋白分子或免疫球蛋白分子顺序地移除氨基或羧基末端氨基酸并在许多可获得的测定法之一中测定各自的活性。在另一个例子中,抗体片段包括修饰的抗体,其中在特定的位置至少一个氨基酸已经置换了天然出现的氨基酸并且抗体的一部分氨基末端或羧基末端氨基酸或者甚至内部区域已被能利于修饰抗体纯化的多肽片段或其它部分,诸如生物素所替代。

只要与未修饰的抗体或抗体片段比较,没有显著改变或损害片段的活性,则无论是连接其它序列还是没连接其它序列的片段均可以包含特定区域或特定氨基酸残基的插入、缺失、置换或其它选定的修饰。这些修饰可以提供一些附加的特性,诸如移除或添加能形成二硫键的氨基酸、增加生物寿命,改变分泌特征等。在任何情况下,片段必须具有生物活性特性,诸如结合活性、对结合结构域的结合的调节等。通过蛋白质特异区的诱变,接着表达和检测表达的多肽,可以鉴定抗体的功能或活性区。这些方法对本领域技术人员显而易见,并且可以包括编码抗原的核酸的位点特异性诱变(Zoller MJ 等. *Nucl. Acids Res.* 10: 6487-500 (1982))。

可以使用各种免疫测定法以筛选选择性结合特定蛋白质,变异体或片段的抗体。例如,常规地使用固相 ELISA 免疫测定法以筛选与蛋白质、蛋白质变异体或其片段选择性发生免疫反应的抗体。参见, Harlow 和 Lane.

**Antibodies, A Laboratory Manual. Cold Spring Harbor Publications, New York, (1988)**, 描述了可以用于检测选择性结合的免疫测定形式和条件。例如, 通过 Scatchard 分析 (Munson 等, *Anal. Biochem.*, 107: 220 (1980)) 可以检测单克隆抗体的结合亲和力。

也提供了具有抗体试剂的试剂盒, 该试剂盒包含容器, 该容器装有本发明的单克隆抗体或其片段和用于检测抗体或其片段与 Fc 受体分子结合的一种或多种试剂。例如, 试剂可以包括荧光标签、酶标签或其它标签。试剂也可以包括二级或三级抗体或者用于酶促反应的试剂, 其中酶促反应产生可以观察的产物。

#### b) 人抗体

可以利用任何技术制备本发明的人抗体。产生人单克隆抗体的技术例子包括 Cole 等. (*Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, 77 页, 1985) 和 Boerner 等. (*J Immunol.*, 147 (1) : 86-95, 1991) 描述的技术。利用噬菌体展示文库也可以产生本发明的人抗体 (和其片段) (Hoogenboom 等, *J. Mol. Biol.*, 227: 381, 1991 ; Marks 等, *J. Mol. Biol.*, 222: 581, 1991)。

也可以从转基因动物获得本发明的人抗体。例如, 已描述了能应答免疫接种而产生人抗体全部组成成分的转基因突变小鼠 (参见, 例如 Jakobovits 等, *Proc. Natl. Acad. Sci. USA*, 90: 2551-2555 (1993); Jakobovits 等, *Nature*, 362: 255-258 (1993); Bruggermann 等, *Year in Immunol.*, 7: 33 (1993))。具体地, 在这些嵌合的种系突变小鼠中抗体重链连接区 (*J(H)*) 基因的纯合缺失使内源抗体产生受到完全抑制, 而一旦抗原攻击, 成功地转移到这种种系突变小鼠中的一系列人种系抗体基因将引起人抗体的产生。如这里所述, 可以利用 Env-CD4-共受体复合物, 筛选具有期望活性的抗体。

#### d) 抗体的施用

优选地, 在药理学可接受的载体中向对象施用本发明的抗体。在 *Remington : The Science and Practice of Pharmacy* (第 19 版) A. R. Gennaro 编辑, Mack Publishing Company, Easton, PA 1995 中描述了适合的载体和

它们的制剂。通常地，在制剂中使用适当量的药物学可接受的盐以使制剂等渗。药物学可接受的载体例子包括但不限于盐水、林格溶液和葡萄糖溶液。溶液的 pH 优选地从约 5 到约 8，更优选地从约 7 到约 7.5。其它的载体包括缓释制剂诸如含有抗体的固体疏水性聚合物的半透性基质，该基质是有形物体的形式，例如薄膜、脂质体或微颗粒。对本领域技术人员显而易见，根据例如，给药的途径和抗体的给药浓度，有些载体可能是更优选的。

可以通过注射（例如，静脉内、腹膜内、皮下、肌内），或者通过其它方法诸如输注向对象、患者或细胞施用抗体，其中所述其它方法将确保以有效的形式向血流递送抗体。优选局部或静脉内注射。

可以凭经验确定给予抗体的有效剂量和方案，并且做出这种决定是本领域技术人员技术范围之内的事情。本领域的技术人员将理解必须给予的抗体剂量将根据例如，将接受抗体的对象、给药的途径、所用的特定类型抗体和给予的其它药物而变化。在关于抗体治疗用途的文献中可以发现选择适合抗体剂量的指导，例如 *Handbook of Monoclonal Antibodies*, Ferrone 等编辑, Noyes Publications, Park Ridge, N. J., (1985) 22 章和 303-357 页; Smith 等, *Antibodies in Human Diagnosis and Therapy*, Haber 等编辑, Raven Press, New York (1977), 365-389 页。根据上面提到的因素，单独使用的抗体的典型日剂量范围可以从每天约 1 $\mu$ g/kg 直到 100mg/kg 体重，或者更多。

给予用于治疗、抑制或预防 HIV 感染的抗体后，可以用本领域技术人员熟知的各种方法估测治疗抗体的效力。例如，本领域普通技术人员将理解，当观察抗体减少病毒负载或阻止病毒负载的进一步增加时，本发明抗体有效地治疗或抑制对象中的 HIV 感染。可以用本领域已知的方法测量病毒负载，例如在来自于对象或患者的样品（例如，不限于血液）中利用聚合酶链式反应试验检测 HIV 核酸的存在或利用抗体试验检测 HIV 蛋白质的存在，或者测定患者中循环抗 HIV 抗体水平。通过测定 HIV 感染的对象中 CD4<sup>+</sup>T 细胞的数量，也可以测定抗体治疗的效力。抑制 HIV 阳性对

象或患者中 CD4<sup>+</sup>T 细胞的最初或进一步减少或者引起 HIV 阳性对象或患者中 CD4<sup>+</sup>T 细胞数量增加的抗体治疗是有效的抗体治疗。

#### d) 用于递送抗体的核酸方法

本发明组合物也可以以编码抗体或抗体片段的核酸制备物（例如，DNA 或 RNA）形式施用给患者或对象，这样患者或对象自己的细胞将摄取核酸，并且产生和分泌编码的组合物（例如，CR2-DAF、 DAF-CR2、 CR2-CD59， CD59-CR2、 CR2-CR1、 CR1-CR2、 CR2-MCP、 MCP-CR2、 CR2-Crry、 Crry-CR2、 CR2-IgG1 Fc (人)、 CR2-IgM Fc、 CR2-IgG3 Fc (鼠)或 CR2-CVF）。

#### e) 核酸递送

在上述包括外源 DNA 施用和摄取到对象细胞中（即，基因转导或转染）的方法中，本发明的核酸可以是裸 DNA 或 RNA 形式，或者该核酸可以在用于向细胞递送核酸的载体中，由此编码抗体的 DNA 片段受启动子转录调节控制，这是本领域普通技术人员将充分理解的。载体可以是商购的制备物，诸如腺病毒载体(Quantum Biotechnologies, Inc. (Laval, Quebec, Canada)。可以通过各种机制向细胞递送核酸或载体。做为一个例子，利用商购脂质体制备物诸如 LIPOFECTIN, LIPOFECTAMINE (GIBCO-BRL, Inc., Gaithersburg, MD), SUPERFECT(Qiagen, Inc. Hilden, Germany)和 TRANSFECTAM(Promega Biotec, Inc., Madison, WI), 及根据本领域标准程序开发的其它脂质体，可以通过脂质体进行递送。此外，可以通过电穿孔及通过 SONOPORATION 机器(ImaRx Pharmaceutical Corp., Tucson, AZ)体内递送本发明的核酸或载体，其中可以从 Genetronics, Inc. (San Diego, CA)获得电穿孔技术。

做为一个例子，可以通过病毒系统，诸如可以包装重组逆转录病毒基因组的逆转录病毒载体系统进行载体递送(参见, 例如 Pastan 等, *Proc. Natl. Acad. Sci. U. S. A.* 85: 4486, 1988 ; Miller 等, *Mol. Cell. Biol.* 6: 2895, 1986)。然后，重组逆转录病毒可以用于感染，并且向感染的细胞递送编码本发明中和抗体（或其活性片段）的核酸。当然，向哺乳动物细胞中引入改变的

核酸的确切方法不限于逆转录病毒载体的应用。广泛地可获得用于这个方法的其它技术，包括腺病毒载体(Mitani 等, *Hum. Gene Ther.* 5: 941-948,1994)，腺相关病毒(AAV)载体(Goodman 等, *Blood* 84: 1492-1500,1994)，慢病毒载体(Naidini 等, *Science* 272: 263-267,1996)，假型逆转录病毒载体(Agrawal 等, *Exper. Hemato.* 24: 738-747,1996)的应用。也可以使用物理转导技术，诸如脂质体递送及受体介导的和其它的胞吞机制(参见，例如 Schwartzberger 等., *Blood* 87: 472-478,1996)。本发明可以联合这些或其它常用的基因递送方法中任一种方法。

做为一个例子，如果用腺病毒载体向对象细胞递送编码补体调节构建体的本发明核酸，向人类给药腺病毒的剂量范围可以是约  $10^7$  到  $10^9$  噬斑形成单位(pfu)/每次注射，但也可以高达  $10^{12}$  pfu/每次注射(Crystal, *Hum. Gene Ther.* 8: 985-1001,1997 ; Alvarez 和 Curiel, *Hune. Gene Ther.* 8: 597-613,1997)。对象可以接受单次注射，或者如果需要另外的注射，可以每隔 6 个月(或者本领域医师确定的其它适合的时间间隔)重复注射，为期不限和/或直到已经确立治疗效力为止。

如果使用本发明核酸或载体的非肠胃外给药，一般地其特征在于注射。可以制备常规形式的注射剂，例如液体溶液或悬浮液，适于注射前在液体中制成溶液或悬浮液的固体形式或者乳液。最近修改的用于非肠道给药的方法包括使用缓释或持续释放系统，以维持恒定剂量。参见，例如美国专利号 3,610,795，这里将这篇专利并入为参考文献。对于适合的制剂和治疗化合物的各种给药途径的其它讨论，参见，例如 *Remington: The Science and Practice of Pharmacy* (19 版) A. R. Gennaro 编辑, Mack Publishing Company, Easton, PA 1995。

## 7、药理学载体/药理学产品的递送

如上所述，也可以在药理学可接受载体中体内施用组合物。“药理学可接受”意指不是生物学上或其它不期望的物质，即，该物质可以与核酸或载体一起施用给对象，而不引起任何不期望的生物学作用或者以有害的方式与含有其的药理学组合物中的任何其它成分相互作用。自然地将选择

载体以最小化活性组分的任何降解和最小化对象中的任何副作用，这也是本领域技术人员所熟知的。尽管通常优选局部鼻内给药或通过吸入剂给药，但是也可以口服，非肠道（例如，静脉内），通过肌内注射、通过腹膜内注射，经皮，体外（extracorporeally），局部地等等施用组合物。这里所用的“局部鼻内给药”意指通过一个或两个鼻孔向鼻子或鼻通道中递送组合物，并且可以包括通过喷雾机制或滴液机制，或者通过核酸或载体的气雾化进行递送。当同时处理大量动物时，后者是有效的。通过吸入剂施用组合物可以通过喷雾或滴液机制经过鼻子或口递送。也可以通过插管法直接向呼吸系统的任何区域（例如，肺）进行递送。根据对象的物种、年龄、重量和总的状况，治疗的过敏性疾病的严重性，使用的特定核酸或载体，其给药的方式等等，所需组合物的确切量随着对象而变化。因此，不可能为每种组合物规定确切的量。然而，利用这里的教导只要常规试验，本领域的普通技术人员就可以确定适当的量。

如果使用组合物的非肠道给药，一般地其特征在于注射。可以制备常规形式的注射剂，例如液体溶液或悬浮液，适于注射前在液体中制成悬浮溶液的固体形式或者乳液。最近修改的用于非肠道给药的方法包括使用缓释或持续释放系统，以维持恒定剂量。参见，例如美国专利号 3,610,795，这里将这篇专利并入为参考文献。

物质可以在溶液、悬浮液中（例如，掺入到微颗粒、脂质体或细胞中的）。这些物质可以通过抗体、受体或受体配体靶向特定的细胞类型。下面的参考文献是使用这种技术将特定蛋白质靶向肿瘤组织的例子(Senter等, Bioconjugate Chem., 2: 447-451, (1991); Bagshawe, K. D., Br. J. Cancer, 60: 275-281, (1989); Bagshawe 等, Br. J. Cancer, 58: 700-703, (1988); Senter 等, Bioconjugate Chem., 4: 3-9, (1993); Battelli 等, Cancer Immunol.Immunother., 35: 421-425, (1992); Pietersz 和 McKenzie, Immunolog. Reviews, 129: 57-80, (1992);和 Roffler 等, Biochem. Pharmacol, 42: 2062-2065,(1991)。载体诸如“隐形”和其它抗体缀合的脂质体（包括靶向结肠癌的脂质介导的药物），受体介导的通过细胞特异性

配体实现的 DNA 靶向, 淋巴细胞引导的肿瘤靶向, 靶向体内鼠神经胶质瘤细胞的高度特异性治疗性逆转录病毒。下面的参考文献是这项技术用于将特异性蛋白质靶向肿瘤组织的例子 (Hughes 等, Cancer Research, 49:6214-6220, (1989); Litzinger 和 Huang, Biochimica et Biophysica Acta, 1104:179-187, (1992))。一般地, 受体参与组成性或配体诱导的胞吞作用途径。这些受体在披有网格蛋白的小窝中簇集, 通过披有网格蛋白的小泡进入细胞, 穿过酸化的内体, 在内体中分拣受体, 然后受体或者重新回到细胞表面、或者细胞内贮藏, 或者在溶酶体中降解。内在化途径具有各种功能, 诸如营养物质摄取, 活化蛋白的移除, 大分子的清除, 病毒和毒素的机会进入, 配体的解离和降解, 及受体水平的调节。根据细胞类型、受体浓度、配体类型、配体价和配体浓度, 许多受体遵循一条以上的途径。已经综述了受体介导的胞吞作用的分子和细胞机制 (Brown 和 Greene, DNA and Cell Biology 10:6, 399-409 (1991))。

#### a) 药理学可接受载体

可以联合药理学可接受载体在治疗上使用组合物, 包括抗体。

药理学载体是本领域技术人员已知的。这些药理学载体大多数一般是向人施用药物时的标准载体, 包括溶液诸如无菌水、盐水和在生理 pH 的缓冲溶液。可以肌肉或皮下施用组合物。而其它的化合物可以根据本领域技术人员使用的标准方法施用。

药理学组合物除了包含选定的分子外, 还可以包含载体、增稠剂、稀释剂、缓冲液、防腐剂、表面活性剂等。药理学组合物也可以含有一种或多种活性组分诸如抗微生物剂, 抗炎剂、麻醉剂等等。

根据是否期望局部或系统的治疗及待要治疗的区域, 可以用多种方式给药药组合物。给药可以是局部 (包括眼睛、阴道、直肠、鼻内), 口服, 吸入, 或非肠道, 例如静脉滴注、皮下、腹膜内或肌肉注射。可以静脉内、腹膜内、肌肉、皮下、腔内或经皮给药公开的抗体。

非肠道给药的制剂包括无菌水性或非水性溶液、悬浮液和乳液。非水溶剂的例子是丙二醇、聚乙二醇、植物油诸如橄榄油和注射用有机酯诸如

油酸乙酯。含水载体包括水、醇/水溶液、乳液或悬浮液，包括盐水和缓冲介质。非肠道赋形药包括氯化钠溶液、林格葡萄糖、葡萄糖和氯化钠、乳酸林格氏液或固定油。静脉内赋形药包括液体和营养补充剂、电解质补充剂（诸如基于林格葡萄糖的）等等。也可以存在防腐剂和其它添加剂诸如，抗微生物剂、抗氧化剂、螯合剂和惰性气体等等。

用于局部给药的制剂可以包括软膏、洗液、乳膏、凝胶、滴剂、栓剂、喷雾剂、液体和粉末。常规的药理学载体、水性的，粉末的或油性的基质、增稠剂等可能是必需的或期望的。

口服给药的组合物包括粉末或颗粒，水或非水介质中的悬浮液或溶液，胶囊，小药囊或片剂。也可能期望有增稠剂、调味剂、稀释剂、乳化剂、分散助剂或粘合剂。

有些组合物可以以药理学可接受的酸或碱加成盐形式给药，所述盐通过与无机酸诸如盐酸、氢溴酸、高氯酸、硝酸、硫氰酸、硫酸和磷酸，及有机酸诸如甲酸、醋酸、丙酸、乙醇酸、乳酸、丙酮酸、草酸、丙二酸、琥珀酸、马来酸和延胡索酸反应形成，或者通过与无机碱诸如氢氧化钠、氢氧化铵、氢氧化钾和有机碱诸如单、二、三烷基和芳基胺和取代的乙醇胺反应形成。

#### b) 治疗应用

组合物给药的剂量范围是足可以产生影响症状疾病的期望效果的剂量。剂量不应该大到引起副作用，诸如不希望的交叉反应、过敏性反应等等。一般地，剂量将随着患者的年龄、状况、性别和疾病的程度而变化，并且本领域的技术人员可以确定该剂量。如果发生禁忌症，个体的医师可以调整该剂量。剂量可以改变，并且可以每日施用一次或多次剂量，施用1天或几天。

#### 8、计算机可读介质

应理解公开的核酸和蛋白质可以表示为由核苷酸或氨基酸组成的序列。有各种方法显示这些序列，例如可以用 G 或 g 表示鸟苷酸。同样地，可以用 Val 或 V 表示缬氨酸氨基酸。本领域技术人员理解如何用现有的各

种方法之任一显示和表示任何核酸或蛋白质序列，这里考虑公开了所述的每种方法。这里尤其考虑的是在计算机可读介质，诸如商购的软盘、磁带、芯片、硬驱、光盘和视盘上或其它计算机可读介质上显示这些序列。也公开了所公开序列的二进制码表示。本领域技术人员理解什么计算机可读介质记录、储存或保存核酸或蛋白质序列。

公开了计算机可读介质，该介质包含关于这里所述序列的序列和信息。也公开了计算机可读介质，该介质包含关于这里所述序列的序列和信息，其中该序列不包括 SEQ ID Nos: 37,38, 39,40, 41 和 42。

## 9、通过用公开的组合物筛选而鉴定的组合物

### a) 计算机辅助的药物设计

公开的组合物可以用作任何分子建模技术的靶以用于鉴定公开组合物的结构或用于鉴定以期望的方式与公开的组合物相互作用的潜在或实际的分子，诸如小分子。这里公开的核酸、肽和相关分子可以用做任何分子建模程序或方法中的靶。

可以理解当在建模技术中使用公开的组合物时，可以鉴定具有特定期望的特性诸如抑制或刺激靶分子的功能的性质的分子，诸如大分子。因此，也公开了通过使用公开的组合物诸如 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠)或 CR2- CVF 而鉴定和分离的分子。因此，利用涉及公开的组合物诸如 CR2, DAF, CD59, CR1, MCP, Crry, IgG1, IgM, IgG3, CVF, CR2-DAF, DAF-CR2, CR2-CD59, CD59-CR2, CR2-CR1, CR1-CR2, CR2-MCP, MCP-CR2, CR2-Crry, Crry-CR2, CR2-IgG1 Fc (人), CR2-IgM Fc, CR2-IgG3 Fc (鼠)或 CR2- CVF 的分子建模方法产生的产物也被认为在这里公开。

因此，分离与选定分子结合的分子的一种方法是合理的设计。通过结构信息和计算机建模获得该合理设计。计算机建模技术允许观察选定分子的三维原子结构和合理设计与该分子相互作用的新化合物。三维结构通常

取决于来自于选定分子的 x 射线结晶分析或 NMR 成像数据。分子动力学需要力场数据。计算机制图系统能够预测新的化合物将如何连接靶分子，从而允许实验操作化合物及靶分子的结构以完善结合特异性。当在分子和化合物之一或两者上实施小的改变后预测分子-化合物的相互作用需要分子力学软件和计算密集型计算机，该计算机通常偶联有在分子设计程序和用户之间的用户友好、菜单驱动的用户界面。

分子建模系统的例子是 CHARMM 和 QUANTA 程序，Polygen Corporation, Waltham, MA。CHARMM 执行能量最小化和分子动力学函数。QUANTA 执行分子结构的构建、图形建模和分析。QUANTA 允许交互式构建、修饰、观察和分析彼此相互作用的分子的行为。

大量的论文综述了与特定蛋白质相互作用的药物的计算机建模，诸如 Rotivinen 等, 1988 *Acta Pharmaceutica Fennica* 97, 159-166; Ripka, *New Scientist* 54-57 (1988 年 6 月 16 日); McKinaly 和 Rossmann, 1989 *Annu. Rev. Pharmacol. Toxicol.* 29, 111-122; Perry 和 Davies, QSAR: Quantitative Structure-Activity Relationships in Drug Design 189-193 页 (Alan R. Liss, Inc. 1989); Lewis 和 Dean, 1989 *Proc. R. Soc. Lond.* 236, 125-140 和 141-162; 以及, 关于用于核酸成分模型酶, Askew 等, 1989 *J. Am. Chem. Soc.* 111, 1082-1090。可以从公司诸如 BioDesign, Inc., Pasadena, CA., Allelix, Inc., Mississauga, Ontario, Canada 和 Hypercube, Inc., Cambridge, Ontario 得到用于筛选和图形描述化学制品的其它计算机程序。

尽管主要设计这些程序用于对特定蛋白质具有特异性的药物，但是一旦鉴定了 DNA 或 RNA 的特定区，就可以改变这些程序以设计与 DNA 或 RNA 特定区特异性相互作用的分子。

尽管上述提及到设计和产生可以改变结合的化合物，但是也可以筛选已知化合物文库，包括天然产物或合成化学制品，和生物活性物质，诸如蛋白质，以选择改变底物结合或酶促活性的化合物。

## 10、试剂盒

这里公开了试剂盒，该试剂盒涉及在实施这里所述方法时可以使用的试剂。试剂盒可以包含这里所述的或者可以理解在实施所公开的方法时需要或有益的任何试剂或试剂组合。例如，试剂盒可以包含在一些方法的实施方式中进行扩增反应的引物，及按预期使用引物所需的缓冲液和酶。例如，公开了用于评价对象患上癌症、哮喘、系统性红斑狼疮、类风湿性关节炎、反应性关节炎、脊椎关节炎、系统性血管炎、胰岛素依赖型糖尿病、多发性硬化、实验性变应性脑脊髓炎、舍格伦综合征、移植物抗宿主病、炎性肠病包括局限性回肠炎、溃疡性结肠炎、局部缺血再灌注损伤、心肌梗塞、阿尔茨海默氏病、移植排斥（同种异体和异种）、烧伤、任何免疫复合物诱导的炎症、肾小球肾炎、重症肌无力、多发性硬化、脑狼疮、吉-巴综合征、血管炎、系统性硬化、过敏症、导管反应（catheter reactions）、粉瘤（atheroma）、不育、甲状腺炎、ARDS、post-bypass 综合症、血液透析、幼年型类风湿、Behcets 综合症、溶血性贫血、天疱疮、大疱性类天疱疮、中风、动脉粥样硬化和硬皮病的风险的试剂盒。

#### 11、具有相似功能的组合物

应理解这里公开的组合物具有一些功能，诸如调节补体活性或结合 CR2,CR3 或 C3b。这里公开了执行该公开的功能的一些结构要求，并且可以理解存在与所公开的结构相关的可以执行相同功能的多种结构，并且这些结构最终将获得相同的结果，例如刺激或抑制补体活性。

#### E、制备组合物的方法

除了另外特别说明外，可以利用用于特定试剂或化合物的本领域技术人员已知的任何方法制备这里公开的组合物和实施所公开的方法所必需的组合物。

公开了制备包含构建体的组合物的方法，其中构建体包含 CR2 和补体调节剂。也公开了制备组合物的方法，其中组合物是本发明的组合物。

#### 1、肽合成

一种产生此处公开的蛋白质，诸如 SEQ ID NO:6 的方法是通过蛋白质化学技术将 2 个或多个肽或多肽连接在一起。例如，利用目前可用的实验

室设备, 利用 Fmoc (9-芴基甲氧羰基)或 Boc(叔丁氧羰基)化学可以化学合成肽或多肽 (Applied Biosystems, Inc., Foster City, CA)。本领域技术人员容易理解可以例如, 通过标准化学反应合成对应于所公开的蛋白质的肽或多肽。例如, 可以合成肽或多肽, 并且不从合成树脂上切下该肽或多肽, 而可以合成肽或蛋白质的其它片段, 并且随后从树脂上切下所述的片段, 由此暴露所述其它片段上被功能性封闭的端基。通过肽缩合反应, 这两个片段可以分别在它们的羧基和氨基末端通过肽键共价连接以形成抗体或其片段(Grant GA (1992) *Synthetic Peptides : A User Guide*. W. H. Freeman and Co., N. Y. (1992);Bodansky M 和 Trost B., 编辑(1993) *Principles of Peptide Synthesis*. Springer-Verlag Inc., NY (这里将其并入以至少作为涉及肽合成的材料的参考文献)。做为替代的方案, 如本文所述可以在体内独立合成肽或多肽。一旦分离这些独立的肽或多肽, 就可以通过相似的肽缩合反应, 连接它们以形成肽或其片段。

例如, 克隆或合成的肽片段的酶促连接允许相对短的肽片段被连接以产生更大的肽片段、多肽或整个蛋白质结构域(Abrahmsen L 等, *Biochemistry*, 30: 4151 (1991) )。做为可替代的方案, 合成肽的天然化学连接可以用于从较短的肽片段合成地构建大的肽或多肽。这种方法有 2 步化学反应(Dawson 等. *Synthesis of Proteins by Native Chemical Ligation*. *Science*, 266: 776-779 (1994) )。第一步是未保护的合成肽-硫酯与另一含有氨基末端 Cys 残基的未保护肽片段的化学选择性反应, 以产生硫酯连接的中间体做为最初的共价产物。不改变反应条件, 该中间体经受了自发的, 快速分子内反应以在连接位点形成天然的肽键 (Baggiolini M 等. (1992) *FEBS Lett.* 307: 97-101;Clark-Lewis I 等, *J. Biol. Chem.* , 269: 16075 (1994); Clark-Lewis I 等, *Biochemistry*, 30 :3128 (1991) );Rajaratnam K 等, *Biochemistry* 33: 6623-30 (1994))。

做为可替代的方案, 可以化学连接未保护的肽片段, 其中由于化学连接在肽片段之间形成的键是非天然(非肽)键(Schnolzer, M 等. *Science*, 256: 221 (1992) )。这种技术已经用于合成蛋白质结构域类似物及大量具有全

部生物活性的相对纯化的蛋白质(deLisle Milton RC 等, *Techniques in Protein Chemistry IV*. Academic Press, New York, 257-267 页(1992))。

## 2、制备组合物的方法

公开了制备组合物及制备产生组合物的中间体的方法。例如,公开了 SEQ ID NOs:5 中的核酸。有各种方法诸如合成的化学方法和标准分子生物学方法可以用于制备这些组合物。应理解具体地公开了制备这些和其它公开的组合物的方法。

公开了通过下述方法产生的核酸分子,该方法包括以可操作方式连接包含 SEQ ID NO:25 中所示序列的核酸和控制该核酸表达的序列。

也公开了通过下述方法产生的核酸分子,该方法包括以可操作方式连接包含与 SEQ ID NO: 25 中所示序列具有 80% 同一性的序列的核酸分子和控制该核酸表达的序列。

公开了通过用这里所公开的任何一种核酸分子转染动物体内细胞的方法而产生的动物。公开了通过用这里所公开的任何一种核酸分子转染动物体内细胞的方法而产生的动物,其中该动物是哺乳动物。也公开了通过用这里所公开的任何一种核酸分子转染动物体内细胞的方法而产生的动物,其中哺乳动物是小鼠、大鼠、兔子、母牛、绵羊、猪和灵长类。

也公开了通过向动物添加任何一种这里所公开的细胞而产生的动物。

在整个申请中,参考了各种出版物。这里将这些出版物公开的内容整体并入为本申请的参考文献以更充分地描述本发明涉及的现有技术的情况。这里也针对提及参考文献的句子中讨论的该参考文献中所含的材料,而将公开的参考文献单独和具体地并入为参考文献。

本领域技术人员显而易见,可以在本发明中进行各种修饰和改变,而不背离本发明的范围或精神。在考虑说明书和实践这里公开的本发明时,本领域技术人员显而易见本发明的其它实施方式。说明书和实施例仅仅被认为是示例性的,下面的权利要求表明了本发明的真正的范围和精神。

## F、实施例

给出了下面的实施例以提供给本领域普通技术人员关于如果制备和评

估这里所述的化合物、组合物、物品、装置和/或方法的完整公开和描述，这些实施例旨在仅仅示例本发明，而不意在限制 Tomlinson 博士所认为的其发明的范围。尽管努力确保数值的精确性（例如，量、温度等），但是应该考虑一些误差和偏差。除了另外指明外，份是按重量计算的份，温度是℃或者环境温度，并且压力是或大约是大气压。

## 1、实施例 1：补体受体 2 (CR2) 介导的补体抑制剂靶向补体激活位点

### a) 方法

#### (1) 细胞系和 DNA

在来源于 p118-mIgG1 (30) 的哺乳动物表达载体 PBM 中进行所有的 DNA 操作，其中通过缺失小鼠 IgG1 Fc 编码区产生该 PBM。中国仓鼠卵巢 (CHO) 细胞用于蛋白质表达，并且在补加有 10% FCS 的 Dulbecco 改进的 Eagle 培养基 (DMEM) (GIBCO Invitrogen Corp, Carlsbad, CA) 中维持该中国仓鼠卵巢 (CHO) 细胞。在 G418 存在的情况下培养稳定转染的 CHO 细胞克隆，并且在没有 FCS 的 CHO-S-SFM II (GIBCO) 中悬浮培养重组蛋白质表达细胞。在 RPMI (GIBCO), 10% FCS 中培养 U937 细胞。

#### (2) 抗体、试剂和血清

通过标准技术 (31) 制备抗 CHO 细胞膜、纯化的人 DAF 和 CD59 的兔抗血清。描述了小鼠抗 DAF mAb 1H4 (32), 大鼠抗 CD59 mAb YTH53.1 (33) 和小鼠抗人 CR2 mAb 171 (结合 SCR1-2)。抗绵羊红细胞 IgM 来源于 Research Diagnostic Inc. (Flanders, NJ)。从 Sigma (St. Louis, MO) 购买所有二级抗体。由 B. P. Morgan 博士 (University of Wales, Cardiff, UK) 惠赠纯化的重组 sCD59。从 Quidel (San Diego, CA) 购买 C6 耗竭的人血清，并且在实验室中从健康志愿者血液获得正常的人血清 (NHS)。

#### (3) 表达质粒的构建和蛋白质表达

图 1 中描述制备的重组融合蛋白和可溶补体抑制剂。通过将编码 4 个 N 末端 SCR 单元 (成熟蛋白质的残基 1-250, Swissprot 登录号 P20023) 的 CR2 序列和编码 DAF 或 CD59 细胞外区的序列连接而制备 cDNA 构建

体。所用的补体抑制剂序列编码成熟 DAF 蛋白质序列的残基 1-249 (Swissprot 登录号 P08174) 和成熟 CD59 蛋白质序列的残基 1-77 (Swissprot 登录号 P13987)。为了连接 CR2 和补体抑制剂序列, 编码 SS (GGGGS)<sub>3</sub> 和 (GGGS)<sub>2</sub> 的连接序列分别用于在 C 末端和 N 末端含有 CR2 的融合蛋白中。通过标准 PCR 方法学制备基因构建体 (35)。在 PBM 载体中进行所有的克隆步骤, 该 PBM 载体也用于蛋白质表达 (30)。为了表达, 利用脂质转染胺试剂, 根据制造商 (GIBCO) 的说明书, 将质粒转染至 CHO 细胞中。通过所述 (30) 的有限稀释来筛选稳定转染的克隆, 并且通过 ELISA 定量克隆的蛋白质表达。

#### (4) ELISA 和蛋白质测定法

利用标准 ELISA 技术 (31) 进行培养悬浮液中重组蛋白质的检测和相对蛋白质浓度的测定。根据要测定的重组蛋白质的类型, 捕获抗体是抗-DAF mAb1H4 或抗-CD59 mAb YTH53.1。一级检测抗体是抗-DAF 或抗 CD59 兔多克隆抗体。在一些 ELISAs, 抗-CR2 mAb A-3 也用做一级检测抗体, 并且尽管不太灵敏, 仍获得了相似的数据。通过 UV 吸光度或通过利用 BCA 蛋白质测定试剂盒 (Pierce Chemical Company, Rockford111) 测定重组蛋白质的蛋白质浓度。

#### (5) 蛋白质纯化

通过亲和层析, 从培养上清液纯化重组蛋白。如制造商所述, 通过将抗 DAF 1H4 mAb 或抗 CD59 YTH53.1 mAb 偶联到 HiTrap NHS 活化的亲和柱 (Pharmacia Biotech, New Jersey, USA) 上制备亲和柱。将含有重组蛋白的培养上清液调整到 pH8.0, 并且以 0.5ml/min 的流速加样到亲和柱上。用 6 到 8 倍柱体积的 PBS 冲洗柱子, 用 pH2.4 的 2 到 3 倍柱体积的 0.1 M 甘氨酸洗脱重组蛋白。将含有融合蛋白的级份收集到含有 1M Tris 缓冲液, pH8.0 的试管中, 并且用 PBS 透析。

#### (6) SDS-PAGE 和 Western 印迹

在非还原条件下, 在 SDS-PAGE 10% 丙烯酰胺凝胶 (Bio-Rad Life Science, Hercules, CA) 中分离纯化的重组蛋白。用考马斯亮蓝染色凝胶。

对于 Western 印迹, 标准方法如下 (31)。简而言之, 将分离的蛋白质转移到聚偏氟乙烯膜上, 并且通过抗-DAF mAb 1H4 或抗-CD59 mAb YTH53.1 的方法检测转移的蛋白质。用 ECL 检测试剂盒 (Amersham Biosciences, Piscataway, NJ) 显影膜。也通过 SDS-PAGE, 接着进行聚糖酶 (glycanase) 处理而分析了 CR2-CD59。在含有 0.1% SDS, 10mM 2-巯基乙醇和 5mM EDTA 的 15mM 磷酸钠缓冲液 (pH7.5) 中, 95℃ 加热 CR2-CD59 (2mg) 3 分钟。冷却后, 在 1% Nonidet P40 和 0.3mM PMSF 存在的情况下, 在 37℃ 温育 CR2-CD59 和 3U 脑膜脓毒性黄杆菌 (*Flavobacterium meningosepticum*) N-聚糖酶 (EC 3.5. 1.52, Sigma) 20 小时。

#### (7) 流式细胞术

通过流式细胞术测定重组融合蛋白与 C3 调理的细胞的结合。在 10% 抗 CHO 抗血清中温育 CHO 细胞 (30 分钟/4℃), 在 10% C6 耗竭的 NHS 中洗涤和温育 (45 分钟/37℃)。然后, 洗涤 C3 调理的细胞, 并且与 1μM 重组蛋白温育 (60 分钟/4℃)。洗后, 在适当的情况下, 细胞与 10μg/ml 抗 DAF mAb 1H4 或抗 CD59 mAb YTH53.1 温育 (30 分钟/4℃), 接着与 FITC 缀合的二级抗体温育 (1: 100, 30min/4℃)。然后洗细胞, 在 PBS 中用 2% 多聚甲醛固定, 然后用 FACSsan 流式细胞仪 (Becton Dickinson Immunocytometry Systems, San Jose, CA) 分析。在 DMEM 中进行所有的温育和洗涤。

#### (8) 分析 C3 配体与 CR2 融合蛋白的结合

在 BIAcore 3000 仪器上进行表面等离子体共振 (SPR) 测定以进行 CR2 融合蛋白与 C3dg-生物素相互作用的动力学分析。通过在 2μl/min 在芯片的一个流动室 (flow cell) 表面上以 50μg/ml 注射 C3dg-生物素, 从而将如所述 (36) 制备的人 C3dg-生物素结合在 BIAcore 链霉抗生物素蛋白 (SA) 传感器芯片的表面。流动缓冲液是 0.5X PBS + 0.05% Tween 20。来自于捕获的 C3dg 的 SPR 信号产生了 250 到 500 的 BIAcore 应答单位 (response unit)。在没有蛋白质的情况下运行对照链霉抗生物素蛋白包被的

流动室。在 25℃, 0.5X PBS, 0.05% Tween 20 中, 以流速 25 $\mu$ l/分钟评估在一系列 CR2 融合蛋白浓度 (15.6-500nM) 下的结合。利用 Kinject 指令, 以 50 $\mu$ l 等份注射 CR2 融合蛋白样品。监测融合蛋白与配体的缔合 120 秒, 之后在仅有缓冲液存在的情况下, 允许复合物解离另外 120 秒。以 50 $\mu$ l/min, 通过 10 秒 200mM 碳酸钠 (pH9.5) 脉冲, 在不同融合蛋白浓度分析之间再生结合表面。相对于与对照流动室的结合, 校正 CR2 融合蛋白片段与固定有 C3d 的流动室的结合。利用 BIAevaluation 版本 3.1 软件 (BIAcore), 结合数据拟合 1:1 Langmuir 结合模型, 并且通过低残值 (residual) 和  $\chi^2$  值, 评估最佳拟合。利用 BIAevaluation 版本 3.1 程序, 所获得的动力学解离曲线用于计算结合速率和解离速率 ( $k_a$  和  $k_d$ ) 和亲和常数 ( $k_D$ )。在实验之间, 以 50 $\mu$ l/min, 用 60 秒 50mM 氢氧化钠 (pH9.5) 脉冲再生链霉抗生物素蛋白表面, 并且如上所述再次应用 C3dg-生物素。

#### (9) 补体裂解试验

用 versene (GIBCO) 处理 60%-80% 铺满的 CHO 细胞, 洗细胞 2 次, 在 DMEM 中重悬浮到 10<sup>6</sup>/ml。通过向细胞添加 10% 兔抗 CHO 细胞膜抗血清, 使细胞对补体敏感 (30min/4℃)。然后移除抗血清, 并且在稀释在 DMEM 中的 NHS 中重悬浮细胞。最终的试验体积是 50 或 100 $\mu$ l。在 37℃ 放置 45 分钟后, 通过台盼蓝排阻测定法 (计数活和死细胞) 或 <sup>51</sup>Cr 释放试验测定细胞的生活力 (37)。两种测定产生了相似的结果。为了测定重组蛋白的补体抑制活性, 在 DMEM 中稀释蛋白质, 并且在向 CHO 细胞添加前, 向 NHS 添加蛋白质。使用 10% NHS 的终浓度, 其引起了约 90% 未保护的抗体敏化的 CHO 细胞裂解。利用抗体敏化的绵羊红细胞 (EA) (Advanced Research Technologies, San Diego, CA) 检测对补体介导的溶血作用的抑制。在 300 $\mu$ l 终体积中在明胶 veronal 缓冲液 (GVB<sup>++</sup>) (Advanced Research Technologies) 中进行溶血测定, 其中 300 $\mu$ l 终体积中含有 2.5x 10<sup>7</sup>EA, 1/300 终稀释的 NHS 和递增浓度的融合蛋白。在 37℃ 温育反应混合物 60min, 并且通过添加含有 10mM EDTA 的 300 $\mu$ l PBS 终止反应。通过离心移除细胞, 并且在 413nm, 通过上清液中血红蛋白的分光光度定量

分析细胞裂解。

#### (10) U937 细胞粘着红细胞

基本如所述(38)分析与 C3 调理的红细胞的依赖于 CR3 的粘着。简而言之,在 37°C,在 GVB(Advanced Research Technologies)中,用预定的亚凝集量的兔抗 SRBC IgM 敏化新鲜的绵羊红细胞(SRBC)30 分钟。洗 2 次后,通过在 GVB 中温育 IgM 敏化的 SRBC 和等体积的 1:2 稀释的 C6 耗竭的人血清(120 分钟/37°C)制备 C3b 调理的 SRBC。洗细胞 2 次,并且在 GVB 中重悬浮沉淀。该处理后,由于血清中 C3b 的短半衰期,结合红细胞的大多数 C3 是 iC3b 或 C3d 降解产物的形式(CR2 配体)。U937 细胞(200 $\mu$ l 中  $4 \times 10^5$  个细胞)添加到 50 $\mu$ l C3 调理的 SRBC ( $2 \times 10^6$  个细胞)中,并且离心混合物(4min/40 x g),在室温放置 90 分钟。然后,用相差显微技术检查细胞,并且确定粘着红细胞的 U937 细胞的数量。每个样品对至少 100 个红细胞进行评分,计算每个红细胞结合的平均 U937 数量。对每次试验进行一式三份的测定。在一些试验中,收获细胞前,在 50ng/ml 乙酸肉豆蔻佛波醇(PMA)存在的情况下培养 U937 细胞 3 天,这是引起 CR3 上调的处理(39, 40)。单独与 IgM 包被的 SRBC 温育的细胞,或者直接与 C6 耗竭的人血清温育的 SRBC 用作对照。

#### (11) 生物分布研究

接着进行用于检测注射的放射标记蛋白质的组织分布的标准方法(41, 42)。简而言之,将 1.7 $\mu$ g  $^{125}$ I 标记的 CR2-DAF ( $4.20 \times 10^6$  cpm/mg) 或 sDAF ( $4.84 \times 10^6$  cpm/mg) 注射到 34 周龄雌性 NZB/NZW F1 小鼠(Jackson Labs, Bar Harbor, ME)的尾静脉中。24 小时后,取血液样品,并且移出主要的器官、切碎并在含有 10mM EDTA 的 PBS 中洗涤,称重和计算。靶向特异性被评估为每克组织的百分比注射剂量。根据制造商的说明书,使用 iodogen 方法碘化蛋白质(Pierce Chemical Co.)。

#### (12) 免疫荧光显微术

向 24 周龄的 MRL/lpr 小鼠尾静脉注射 CR2-DAF 或 sDAF (270 $\mu$ g)。24 小时后,移出肾脏,并且快速冷冻。在丙酮中固定从冷冻肾脏制备的冰

冻切片 (5 $\mu$ m)，并且进行处理用于间接的免疫荧光显微术。等摩尔的小鼠抗人 DAF 1A10 和 1H6 mAb 混合物用做一级检测抗体 (终浓度, 10 $\mu$ g/ml)，而抗小鼠 IgG Fc 特异性 FITC 缀合的抗体作为二级抗体 (F4143, Sigma-Aldrich)。接着是标准方法 (49, 这里将其教导的抗体染色技术并入为参考文献)，除了为减少最可能由小鼠肾脏中沉积的免疫复合物引起的背景染色，以 1:800 稀释二级 FITC 标记的抗体 (10 倍于推荐的稀释度)。获得数字图像，并且利用相同设置，用 Adobe Photoshop 优化该数字图像。

## b) 结果

### (1) 构建体设计、表达和纯化

重组融合蛋白含有人 CR2 的 4 个 N 末端 SCR 单元，其连接到可溶形式人 CD59 或 DAF 的 N 或 C 末端 (图 1 中所示的构建体)。从稳定转染的 CHO 细胞克隆的培养上清液纯化重组蛋白，产率是 100-200 $\mu$ g/l。通过 SDS-PAGE 和 Western 印迹分析纯化的重组蛋白，显示了期望分子量范围内的蛋白质 (图 2)，并且除了 CR2-CD59，所有蛋白质都做为单条带迁移。观察到的 CR2-CD59 的两条带是由于糖基化差异引起，因为在聚糖酶处理后，CR2-CD59 做为单条带迁移。

### (2) 融合蛋白靶向补体调理的细胞

通过 CHO 细胞与补体激活抗体和 C6 耗尽的血清 (以阻止 MAC 形成和细胞裂解) 温育，在 CHO 细胞上沉积 CR2 的 C3 配体。所有含有 CR2 的融合蛋白，但不是 sCD59 或 sDAF，均结合 C3 包被的 CHO 细胞 (图 3)。

### (3) 融合蛋白和 C3dg 配体之间相互作用的动力学分析

通过表面等离子体共振测定法比较不同重组融合蛋白对 CR2 配体 C3dg 的亲合力。使可变浓度的融合蛋白通过含有捕获的 C3dg-生物素的 Biacore 链霉抗生物素蛋白芯片进行试验 (约 2000 个应答单位)。数据的动力学分析表明利用 Global 拟合参数，对 1:1 (Langmuir) 结合相互作用模型的最佳拟合 (图 4)。在 N 末端具有 CR2 的两种融合蛋白 (CR2-DAF 和 CR2-CD59) 显示了相似的结合曲线，具有快速的结合和快速的解离速

率。相反，在 C 末端具有 CR2 的融合蛋白 (DAF-CR2 和 CD59-CR2) 的结合显示了慢的结合和解离速率 (图 4, 表 1)。然而，N 末端 CR2 融合蛋白以最高亲和力结合 (表 1)。CD59 融合蛋白比 DAF 融合蛋白具有更高的结合亲和力。可溶的 DAF 和 sCD59 不结合固定的 C3dg。

#### (4) 融合蛋白的补体抑制活性

通过测定定向和未定向的补体抑制剂对补体介导的 CHO 细胞和红细胞裂解作用的影响，分析它们的补体抑制活性。在这些试验中，在人血清 (以引起 90-100% 未保护细胞裂解的浓度) 中温育抗体敏化细胞和重组蛋白。对于两种细胞类型，在抑制补体介导的裂解方面，定向的补体抑制剂比它们相应的未定向蛋白质更显著有效。定向的 DAF 蛋白质是比定向的 CD59 更有效的抑制剂 (图 5 和 6)。含有连接到 DAF 或 CD59 N 末端的 CR2 的融合蛋白比 C 末端 CR2 融合蛋白为更有效的抑制剂。补体裂解的最有效抑制剂是 CR2-DAF，需要 18nM 浓度以抑制 50% CHO 细胞裂解。相反，未定向的 sDAF 需要 375nM 以抑制 50% CHO 细胞裂解，有 20 倍差异 (图 5a)。sCD59 是特别差的补体抑制剂，并且在 500nM 只提供 25% 防止 CHO 细胞裂解的保护作用，该浓度是检测的最高浓度。然而，CR2-CD59 在 102nM 就提供了 50% CHO 细胞裂解抑制作用，其比未定向的 sDAF 更有效 (图 6a)。表 4 比较了不同重组补体抑制剂的抑制活性。N 末端 CR2 融合蛋白的更高补体抑制活性与这些蛋白质显示的对 C3dg 配体更高的亲和性相关 (表 3)。

在保护 CHO 细胞和红细胞免受补体介导的裂解时，补体抑制剂的相对效力有一些差异。对于 DAF 抑制剂，这特别正确；尽管定向的 DAF 更有效，但是与 sDAF 保护 CHO 细胞免受补体介导的裂解比较，sDAF 更显著有效地保护红细胞免受补体介导的裂解。当红细胞是补体裂解的靶细胞时，CR2-DAF 和 DAF-CR2 的抑制活性也有很小的差别。

#### (5) CR2 融合蛋白对细胞粘着的影响

补体受体 3 是白细胞受体，该受体参与内皮粘附、血细胞渗出和细胞的细胞溶解机制 (吞噬作用和脱粒) 的活化。因为 CR2 和 CR3 共有相同

的 iC3b 补体配体,测定是否 CR2 融合蛋白干扰 CR3 介导的细胞结合。对于这些试验使用 U937,一种充分表征的幼单核细胞系 (CR2<sup>-</sup>, CR3<sup>+</sup>),该细胞系以依赖 CR3 的机制结合 iC3b 包被的红细胞 (40)。所有的 CR2 融合蛋白,但不是 sDAF 或 sCD59,都显著抑制 U937 细胞与 C3 调理的绵羊红细胞的结合 ( $P < 0.01$ )。每种 CR2 融合蛋白在 500nM 浓度以相似的程度抑制 U937 结合 (图 7)。在利用 PMA 刺激 U937 细胞的试验中获得了相似的数据, PMA 刺激是引起 CR3 上调的处理 (39, 40)。对于红细胞的补体调理作用,由于红细胞上沉积的 IgG 将结合 U937 细胞上表达的 Fc $\gamma$  受体,所以 IgM 用于激活补体。U937 细胞也表达 CR4 (p150, 95, CD11c/CD18),其是共有 iC3b 配体的第三种受体。然而, U937 细胞与 C3 调理的红细胞的结合不依赖于 CR4,这可能是由于 CR4 与细胞骨架结合并且其在细胞膜上是固定的 (40)。

#### (6) CR2-DAF 靶向肾炎小鼠的肾脏

为了确定是否 CR2 融合蛋白将在体内靶向补体激活和疾病位点,进行雌性 NZB/W F1 小鼠中 CR2-DAF 和 sDAF 的生物分布研究。NZB/W F1 小鼠患有与人系统性红斑狼疮 (SLE) 非常相似的自发性自身免疫病,从 26 到 28 周龄,产生自身抗体和形成与补体沉积有关的严重免疫复合物介导的肾小球肾炎 (4, 52)。在注射后 24 小时和 48 小时检测 34 周龄 NZB/W F1 小鼠中 [<sup>125</sup>I]CR2-DAF 和 [<sup>125</sup>I]sDAF 的生物分布。尾静脉注射 [<sup>125</sup>I]CR2-DAF 后 24 小时,与检测的其它器官相比,显著更高比例的放射性定位到肾脏 (图 25a)。在 [<sup>125</sup>I]CR2-DAF 注射后 48 小时,肾脏中放射性水平与 24 小时时肾脏中放射性水平相似,但是肝脏和脾脏中放射性增加,血液放射性降低 (图 25b)。肝脏和脾脏是免疫复合物清除的位点,这可能解释在后面的时间点 [<sup>125</sup>I]CR2-DAF 向这些器官增加的靶向。 [<sup>125</sup>I]sDAF 在肾脏或任何其它器官中没有显示优选的结合 (图 25, a 和 b)。在 8 周龄肾炎前 NZB/W F1 小鼠中,没有 [<sup>125</sup>I]CR2-DAF 靶向肾脏的迹象 (图 25c)。更有意义的是, [<sup>125</sup>I]sDAF 从循环中的清除比 [<sup>125</sup>I]CR2-DAF 从循环中的清除要快得多,表明 CR2 部分有延长融合蛋白的循环半衰期的

作用。然而，在 24 小时，年轻小鼠血液中的 $[^{125}\text{I}]\text{CR2-DAF}$ 水平是年老小鼠中记录的水平约一半，并且 $[^{125}\text{I}]\text{CR2-DAF}$ 的长循环半衰期可能至少部分是其结合循环免疫复合物的结果。

通过肾脏切片的直接检查，在另一个 SLE 鼠模型中也检查了 CR2-DAF 靶向肾脏中沉积的补体。与雌性 NZB/W F1 小鼠相似，MRL/lpr 小鼠到 24 周龄出现严重增生性肾小球肾炎，伴有与肾小球免疫沉积物有关的补体沉积（53，这里将其关于该小鼠模型的教导并入为参考文献）。将 CR2-DAF 和 sDAF 注射到 24 周龄 MRL/lpr 小鼠尾静脉中，并且 24 小时后，通过荧光显微镜方法分析肾脏切片的人 DAF 免疫反应性。来源于用 CR2-DAF 注射的小鼠的肾脏切片显示了高水平的 DAF 染色，在肾小球中具有优先的定位，定位模式与观察到的免疫复合物的定位模式相同。在来源于用 sDAF 注射的小鼠的肾小球中，DAF 染色不明显（图 26）。

### c) 结论

本研究描述了含有可溶人 DAF 和 CD59 的蛋白质的产生和表征，其中该蛋白质靶向补体激活位点。该靶向蛋白质比它们未靶向的对应物显著更有效地抑制补体。通过连接抑制剂和结合补体 C3 激活产物的人 CR2 片段获得 CD59 和 DAF 的靶向。CR2 的 C3 配体相对长寿，并且常常大量共价结合在补体激活位点。因此，CR2 介导的定向补体抑制作用对大量与补体相关的疾病或疾病状态具有治疗益处。与这种假说一致，已证明 CR2-DAF 靶向肾炎 NZB/W F1 小鼠的肾脏。这些小鼠产生自身抗体，随后在肾脏中形成和沉积免疫复合物，引起补体激活和沉积（2，43）。人 CR2 以相似的亲和性结合人和小鼠 C3 配体（44），并且生物分布研究确立 CR2 融合蛋白在体内保留靶向功能。该研究确立了这种方法抑制人补体的可行性。此靶向方法对其它补体激活抑制剂诸如可溶性 CR1 也可能有效，所述可溶性 CR1 处在临床试验中，并且在体外是比 DAF 更有效的补体抑制剂（9）。

不同 CR2 融合蛋白对 C3dg 的相对亲和性使人想到 CR2 的 SCR1-2 和 CR2 的 SCR1-15 对 C3dg 的亲和性。CR2 SCR1-2 和 CR2 SCR1-15 与 C3dg 相互作用的 KD 值相似，但是 CR2 SCR1-2 的结合和解离快得多，表明另

外的 SCR 结构域对总体亲和性的贡献 (36)。对另一种含有 SCR 的蛋白质, 因子 H 的溶液结构分析表明 SCR 结构域本身回折, 并且 SCR 结构域之间的相互作用可以调节 C3 配体结合特征 (45)。预测 SCR 结构域之间的构象可变性是不同 (天然) 接头长度造成的结果, 接头越长提供越大的构象柔性。在本发明上下文中, 用相对长的 ser-gly 接头连接 CR2 和 DAF SCR 结构域, 这可以允许融合配偶体回折于彼此之上, 引起可以调节 CR2 结合亲和性的 SCR-SCR 相互作用。

利用抗体敏化的 CHO 细胞或绵羊红细胞做为靶进行补体介导的裂解试验。在保护不同的细胞免受补体介导的裂解方面, 有些补体抑制剂的相对活性具有显著的差异。与 sDAF, DAF-CR2 和 CD59-CR2 对 CHO 细胞的保护相比, 它们显著更有效地保护绵羊红细胞免受补体介导的裂解。不象红细胞, 补体介导的有核细胞裂解不完全由胶体渗透压下调引起, 其需要质膜中多个 MACs 的沉积 (46-48)。利用红细胞做为补体介导的裂解的靶细胞, 大多数以前研究已经进行了研究可溶 (未靶向的) 补体抑制剂的抑制活性。然而, 对于体外试验, CHO 细胞可能代表一种更生理相关的靶。

补体介导的损害的不同机制涉及不同的疾病状况, 并且不同疾病可以从作用在该途径中不同点的抑制策略中获益。例如, 如果可应用于疾病, 在途径中后面步骤阻断补体的一个特别益处是宿主防御功能和补体的免疫内稳态机制将保持完整。因此, 在发病机理主要涉及终末溶细胞途径的疾病中, 基于 CD59 的抑制剂将提供优于补体激活抑制剂的优点。由于可溶 CD59 体外非常差的活性, 其不可能具有治疗益处, 但是这里证明 CR2 介导的 CD59 靶向显著地增加了其补体抑制活性。实际上, CR2-CD59 比 sDAF 更有效地抑制补体介导的裂解, 而 sDAF 的体内治疗效力是已经得以证明的 (8)。CR2-CD59 的啮齿动物类似物也可以是用于仔细分析早期补体激活产物与 MAC 形成在疾病病理中的相对作用的有用工具。目前不太理解许多疾病状态中不同补体激活产物对组织损害的相对贡献, 并且对此存在争议。

CR2 融合蛋白抑制 U937 细胞与 C3 调理的红细胞结合。CR2 和 CR3

都结合 iC3b, 并且该数据表明 CR2 融合蛋白作为 CR3 拮抗剂, 因为 U937 与 C3 调理的红细胞的结合是依赖于 CR3 的 (40)。做为一种粘着分子, CR3 在炎性位点通过其与细胞间粘着分子-1 (ICAM-1) 的高亲和性相互作用而介导内皮粘着和血细胞渗出。做为补体受体, CR3 通过其与 iC3b 相互作用, 促进和增强了吞噬和脱粒作用。ICAM-1 和 iC3b 都结合 CR3 上重叠的表位 (参见 Ross 综述)。因此, CR3 可能是在炎性位点促进细胞介导的组织损害的重要决定因素, 并且在几种炎症病症中已证明了阻断 CR3 的抗体的效力 (参见 Ross 综述)。因此, CR2 对 CR3 结合的拮抗作用提示 CR2 补体抑制剂融合蛋白作用的第二抗炎机制, 这种抗炎机制与补体抑制协同地起作用。

将补体抑制剂靶向补体激活和疾病位点可以相当大地增强它们的效力。实际上, 对于将获益于基于 CD59 的治疗的疾病状态, CD59 靶向补体激活位点将是一个必要条件。CR2 介导的靶向与其它靶向方法, 诸如抗体介导的靶向相比的优点是 CR2 部分可以靶向任何可接近的补体激活位点, 并且具有广泛的治疗应用。CR2 融合蛋白也可以做为 CR3 拮抗剂, 并且这代表第二重要的治疗益处。人 CR2-补体抑制剂融合蛋白比含有抗体可变区的重组抑制剂更不可能是免疫原性的。预测定向的补体激活抑制剂能够提供有效的局部浓度同时具有低水平系统抑制, 这种能力也减少了危害宿主防御机制的可能性, 特别是由于长期全身性补体抑制而危害宿主防御机制的可能性 (对于基于 CD59 的抑制剂, 这是不太重要的考虑因素)。通过 CR2 定向的抑制剂也可以靶向激活补体的感染因子。

## 2、实施例 2: 靶向的补体抑制和激活

### a) 补体抑制剂 (炎性/生物不相容性)

补体抑制剂具有相当大的希望治疗许多自身免疫病和炎症疾病及与生物不相容性相关的疾病状态。目前尚未有安全有效的药物补体抑制剂。研究大多数关注于开发基于宿主细胞膜-结合的补体-调节蛋白质的可溶抑制剂。通过移除细胞膜连接区, 已经产生了重组形式的可溶 CR1, MCP, DAF 和 Crry, 并且在多种疾病模型中, 已经证明所有的蛋白质均有效地减少炎

症和补体介导的组织损害。阻断补体蛋白质 C5 的功能的抗体和可溶 CR1 处于临床试验中。然而，在系统施用的可溶性补体抑制剂的临床应用中存在严重的问题。补体在天然和适应性免疫性中起到至关重要的作用，并且 C3b 的产生对调理作用和白细胞介导的许多病原微生物的清除至关重要。此外，也证明液相补体激活产物 C5a 在控制感染方面是重要的，并且可能在从循环中清除病原物质方面也是重要的。因此，补体的全身性抑制可能会对宿主控制感染的能力造成严重的后果。补体对于免疫复合物的有效分解代谢也至关重要，因此在使用补体抑制剂治疗自身免疫病和免疫复合物疾病时，这是特别重要的考虑因素。

补体抑制剂靶向补体激活和疾病位点可以允许低得多的有效血清浓度，并且显著降低系统性补体抑制的水平。增加的效力是定向补体抑制剂的一个重要益处，并且靶向也可以解决可溶性重组补体抑制剂在循环中的短半衰期问题。

除了上述关于补体抑制剂靶向的考虑因素外，选择性阻断补体途径的不同部分也可以允许有益的补体激活产物的产生，而抑制参与疾病病理的补体激活产物的产生。例如，补体激活的抑制剂（诸如 CR1, DAF, Crry）抑制 C3b, C5a 和 C5b-9 产生。C5 抗体抑制 C5a 和 C5b-9 产生。另一方面，基于 CD59 的抑制剂不影响 C3b 和 C5a 产生，而是只阻断 C5b-9 形成（参见，图 8）。已证明终末补体途径和 C5b-9 产生在促进炎症方面是重要的，并且尤其是参与一些肾脏疾病（诸如免疫复合物肾小球肾炎）的进展。因此，对于一些疾病，基于 CD59 的抑制剂可以抑制疾病病理，而不干扰对于宿主防御和免疫复合物清除而言重要的早期补体激活产物的产生。然而，可溶 CD59 不是补体的有效抑制剂（不同于激活抑制剂 DAF, CR1, MCP 或 Crry），并且不可能具有任何临床应用。然而，靶向细胞的 CD59 可以是可行的治疗剂。有关抗体介导的 CD59 靶向的数据 [Zhang 等 1999, *J. Clin. Invest.* 103,55-61]，和这里有关 CR2 介导的 CD59 靶向所给出的数据表明，靶向细胞膜的 CD59 比可溶性未靶向的 CD59 显著更有效。

## b) 补体激活剂 (癌症)

还没有实现抗癌补体激活单克隆抗体做为癌症免疫治疗剂的最初希望。这种情况的一个原因是肿瘤细胞上补体抑制性蛋白的表达(补体抑制剂在肿瘤细胞上常常被上调)。因此,尽管已经证明一些抗体靶向人体中的肿瘤,并且激活肿瘤细胞表面上的补体,但是肿瘤细胞仍抵抗补体介导的破坏作用。有大量来自于体外研究的证据表明补体抑制剂在肿瘤对抗体治疗的抗性中起重要作用。此外,有报道[Caragine 等, 2002, *Cancer Res*, **62**,1110-15; Chenet 等, 2000, *Cancer Res.*, **60**,3013-18; Baranyi 等, 1994, *Immunology*,**82**, 522-8]已经确立肿瘤细胞表面上体内表达的补体抑制剂对补体沉积和肿瘤发生具有功能作用。增加肿瘤细胞上补体沉积允许更有效地通过免疫介导清除肿瘤细胞,并且提高利用补体激活抗肿瘤抗体进行成功免疫治疗的前景。增加的补体激活将压到肿瘤细胞表达的补体抑制性蛋白质。

## c) 结果-1

### (1) 定向的补体抑制剂融合蛋白

如前述,已经被表达、纯化和进行了靶向表征和被估测了体外补体抑制性功能的人融合蛋白的例子包括: CR2-DAF, CR2-CD59, DAF-CR2 和 CD59-CR2。成熟人融合蛋白的核苷酸序列和预测的氨基酸序列如图 8-11 中所示。

### (2) 表达和纯化

编码融合蛋白的 cDNA 质粒构建体被转染到 CHO 细胞中,并且分离稳定的表达克隆。选择表达最高水平融合蛋白的克隆。在生物反应器中培养选定的克隆,并且通过亲和层析从培养上清液分离融合蛋白。利用缀合抗 DAF 和抗 CD59 抗体的琼脂糖凝胶制备亲和性柱子。通过 SDS-PAGE 和 Western 印迹分析重组蛋白(图 2)。

### (3) 融合蛋白与 C3 配体的结合

#### (a) 流式细胞术

如前所述进行流式细胞术试验。如通过流式细胞术所分析的,所有含

有 CR2 的融合蛋白都结合 C3 包被的 CHO 细胞(图 12)。sDAF 和 sCD59 不结合 C3 包被的 CHO 细胞。

#### (b) ELISA

如前所述进行 ELISA 试验。在 ELISA 试验中, 向用纯化的 C3dg 包被的孔添加含有 CR2 的构建体。通过抗补体抑制剂抗体和酶缀合的二级抗体检测结合。所有含有 CR2 的构建体都结合 C3d, 但是 sCD59 和 sDAF 不结合 C3d。

#### (c) 表面等离子体共振

生物素酰化的 C3dg (CR2 配体) 结合链霉抗生物素蛋白包被的 BIAcore 芯片, 并且测定含有 CR2 的融合蛋白的结合动力学(图 13-16)。sDAF 和 sCD59 不结合捕获的 C3dg。在 N-末端具有 CR2 的融合蛋白以最高亲和力结合。含有 CD59 的融合蛋白比相应含有 DAF 的融合蛋白以更高的亲和力结合。

#### (4) 融合蛋白的补体抑制功能

通过测定蛋白质对补体介导的细胞裂解的影响, 分析融合蛋白和可溶未靶向的补体抑制剂的功能性活性。利用中国仓鼠卵巢 (CHO) 细胞(图 17 和 18) 和绵羊红细胞 (E) (图 19 和 20) 进行试验。

定向的补体抑制剂比可溶未定向的补体抑制剂提供了显著更多的避免补体介导的细胞裂解的保护作用。对于含有 DAF 和 CD59 的融合蛋白, 在 N 末端含有 CR2 的融合蛋白最有效。在 BIAcore 试验中, N 末端 CR2 融合蛋白也比 C 末端 CR2 融合蛋白以更高的亲和力结合 C3d(参见上述)。在这些试验中, CR2-DAF 比 CR2-CD59 提供了显著更有效的防止补体介导的裂解的保护作用。然而, 值得注意的是即使在高浓度时, 未靶向的 sCD59 也只具有非常弱的补体抑制活性(不象 sDAF), 并且 CD59 靶向细胞表面是 CD59 功能所必需的。

定向与未定向的补体抑制剂比较, 它们对 CHO 细胞和 E 的相对效力不同。然而, 红细胞通过“一次击中”裂解(即, 单个 MAC 的形成引起 E 裂解), 而有核细胞(诸如 CHO)具有其它抗性机制(诸如 MAC 的加

帽和脱落), 并且要求沉积多个 MAC 以进行裂解。这些差异可能解释裂解抑制数据的差异, 并且对于这些体外试验, CHO 细胞可能代表更生理学相关的靶。

#### D) 结果-2

##### (1) 靶向的补体激活融合/缀合蛋白质

已经表达和纯化了人 CR2-IgG1 Fc, 并且证明其适当地体外靶向 C3 调理的细胞。制备了含有(用于与 CVF 缀合的)人和小鼠 sCR2 及小鼠融合蛋白的编码序列的表达质粒。图 21 显示成熟人融合蛋白的核苷酸序列和预测的氨基酸序列。

##### (2) 表达和纯化

编码 CR2 最初 4 个 SCR 的 cDNA 连接编码人 IgG1 Fc 区的基因组序列。编码融合蛋白的质粒被转染到 CHO 细胞中, 并且分离稳定的表达克隆。选择表达最高水平融合蛋白的克隆。在生物反应器中培养选定的克隆, 并且通过蛋白质 A 亲和层析从培养上清液分离融合蛋白。通过 SDS-PAGE(图 22) 和 Western 印迹分析重组蛋白。在还原和非还原的条件下, 蛋白质以预期的分子量迁移(CR2-Fc 是二硫化物连接的二聚体)。构建了编码 CR2-小鼠 IgG3 的鼠质粒构建体。

##### (3) 融合蛋白与 C3 配体的结合

###### (a) 流式细胞术

如通过流式细胞术所进行的分析, CR2-Fc 结合 C3 包被的 CHO 细胞(图 23)。

###### (b) ELISA

在 ELISA 试验中, 向用纯化的 C3dg 包被的孔添加 CR2-Fc。通过抗人 Fc 抗体和酶缀合的二级抗体检测结合。CR2-Fc 结合 C3d。

###### (c) 表面等离子体共振

生物素酰化的 C3dg(CR2 配体)结合链霉抗生物素蛋白包被的 BIAcore 芯片, 并且证明了 CR2-Fc 的结合(图 24)。

### 3、实施例 3

### a) 在急性肾小管间质损伤的大鼠模型中抗体靶向的补体抑制剂

使用一组很好地被表征的小鼠抗大鼠肾脏单克隆抗体（49, 50, 这里将其中关于这些抗体和它们的序列的教导并入为参考文献）。通过标准 PCR 技术从总共 5 种抗体分离可变区 DNA（35, 这里将其中关于 PCR 的教导并入为参考文献）。成功地克隆了所有这些 DNA, 并且一些表达为单链抗体。所有单链抗体都体外识别大鼠肾脏上皮或内皮细胞系。一种 mAb, K9/9 结合上皮细胞表面鉴定的糖蛋白, 并且对体内肾小球毛细血管壁和近端小管具有特异性（49）。这种抗体被选作靶向载体, 用于在急性肾小管间质损伤的大鼠模型中研究定向的 Crry 和 CD59 介导的补体抑制。尽管在以前的研究（49）中表明 K9/9 mAb 诱导肾小球损伤, 但是只是当与弗氏佐剂一起给药时抗体具有致病性。实际上, K9/9 mAb（与佐剂）的致病性质未再现。

在蛋白尿和进行性肾损害之间具有联系, 并且有数据支持蛋白尿本身引起间质纤维化和炎症的假说。不知道蛋白尿引起肾中毒损害的机理, 但是有证据表明补体起重要的作用, 并且 MAC 是由于蛋白尿引起的肾小管间质损伤的主要介体。（最近已经综述了在肾小管间质损伤中补体和蛋白尿的作用（51, 52））。对 K9/9 mAb 的先前表征（参见上述（49））提示在由蛋白尿引起的大鼠肾小管间质损伤模型中, 该 mAb 适于定向研究靶向的补体抑制剂的治疗用途。可以特异性阻断 MAC 形成的抑制剂的可获得性允许在临床相关条件下估测 MAC 在肾小管间质损伤中的作用。

制备编码连接大鼠 Crry 或大鼠 CD59 的单链 K9/9 抗体的质粒构建体（图 27 中所示）。也制备表达可溶性大鼠 Crry（sCrry）和单链 K9/9（仅靶向载体）的构建体。通过在 15 升 New Brunswick 发酵罐中进行毕赤酵母（*pichia*）发酵, 在培养基中以 15mg/l 以上表达所有重组蛋白质。

如上所述, 利用大鼠上皮细胞系做为靶细胞, 表征重组蛋白的体外靶向及补体抑制活性。单链 K9/9, K9/9-Crry 和 K9/9-CD59 体外特异性结合大鼠上皮细胞。sCrry 和 K9/9-Crry 抑制补体沉积和裂解, 并且 K9/9-CD59 抑制补体介导的裂解。在大鼠嘌呤霉素氨基核苷（PAN）肾变病模型中表

征了两种定向的补体抑制剂，sCrry 和 K9/9 单链 Ab (53, 这里将其中关于 sCrry 和 K9/9 的教导并入为参考文献) (因为未靶向的 CD59 只有非常弱的补体抑制活性 (54), 所以没有评估 sCD59)。首先, 为了证实 K9/9 融合蛋白的肾脏靶向, 如上所述 (54, 41), 通过碘化蛋白质的生物分布确定体内结合特异性。单链 K9/9 和 K9/9 融合蛋白, 但不是 sCrry, 特异性靶向大鼠肾脏, 并且在给药后 48 小时可以检测到 (图 28)。给药后 24 小时的生物分布是相似的, 并且 24 小时时在血液中没有放射标记残留。

在治疗研究中, 在 0 天由 4 只大鼠组成的组接受了 PAN (150mg/kg), 在 4, 7 和 10 天, 它们接受了 PBS 或补体抑制剂 (40mg/kg)。收集尿 (代谢笼子) 和血液, 并且在 11 天处死动物。如通过肌酸酐清除所测定的, PAN 处理显著地损害了肾功能 (图 29, 从左起第 2 条柱)。在接受 sCrry 治疗的大鼠中, 有轻微但不显著的肾功能改善。然而, 在接受靶向的 Crry 或 CD59 治疗的 PAN 处理的大鼠中显著改进了肌酸酐的清除 ( $p < 0.01$ )。在对照 (非蛋白尿症) 大鼠和接受两种靶向抑制剂中任一种的 PAN 处理大鼠之间肌酸酐清除率没有显著差异 (图 29)。如预期的, 在所有不论用补体抑制剂治疗还是没有用补体抑制剂治疗的大鼠中, PAN 均诱导高的蛋白尿 (表 1)。如通过刷状缘损失所估测的, 从用 PAN 处理但没有接受治疗的大鼠制备的肾脏切片表明管腔扩大及小管和上皮细胞退化 (参见图 30b, 也在附录中)。用 sCrry 治疗观察到了最小的改进 (图 30d)。相比较, 在接受靶向 Crry 和 CD59 的 PAN 处理大鼠中显著地抑制了小管扩大和退化 (图 30c 表示 K9/9-Crry, 但是对于 K9/9-CD59, 组织学无差别)。这些数据证明了该模型中补体抑制的治疗效力, 证明了靶向比非靶向的补体抑制显著有益, 并且直接证明了在由蛋白尿诱导的肾小管间质损伤中 MAC 介导的损害的重要作用。sCD59 不是有效的抑制剂, 但该研究证明适当靶向的 CD59 允许体内特异性抑制 MAC。

在单独的试验中, 如所述测定了碘化重组蛋白的循环半衰期 (54, 41, 这里将教导该技术的一部分并入为参考文献)。蛋白质的半衰期 ( $t_{1/2}$ ) 如下: sCrry: 19 分钟, K9/9-Crry: 23 分钟, K9/9-CD59, 29 分钟, 单链 K9/9:

21分钟。为了测定重组蛋白对系统性补体抑制的影响,用蛋白质以40mg/kg注射大鼠,并且在对应于1, 3, 5和7x t<sub>1/2</sub>的时间收集血液。通过测定溶血活性(敏化的绵羊红细胞)检测血清中补体抑制活性。如所预期的,在血清中,K9/9-CD59具有最小的抑制活性(未靶向的试验系统)(图31)。到sCrry和K9/9-Crry注射后约3小时(7x t<sub>1/2</sub>),血清中残留了极小的补体抑制活性。靶向和未靶向的抑制剂的短t<sub>1/2</sub>,以及生物分布数据和sCrry不具有保护性的事实证明了结合肾脏的补体抑制剂有效地局部长期抑制补体。

这些数据确立了靶向的补体抑制剂的体内用途,并且证明在疾病模型中靶向的相对于未靶向的系统补体抑制的重要益处。尽管在这些研究中使用了不同的靶向载体(抗体片段),相同的原理适用于其它靶向载体,诸如CR2。

#### 4、实施例4

这里公开了根据这里所教导的方法制备的本发明构建体的例子。所用的术语具有下列含义:SCR = 短共有重复;LP = 前导序列。所有构建体都有CR2-接头-补体调节剂或补体调节剂-接头-CR2的基本式。括弧里的注释表明了组合物特定部分的详细情况。例如“(完整的)”意思是在该构建体中使用了完整成熟的蛋白质,而“(SCR2-4)”表示SCR1不是构建体的一部分。可以理解接头可以是化学接头、天然接头肽或氨基酸连接序列(例如(Gly<sub>4</sub>Ser)<sub>3</sub>)。可以理解此列表不是限制性的,只提供了本申请中公开的一些构建体的例子。

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 CD59

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - MCP

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CRI

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 小鼠 CD59

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgG1 Fc

CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgM Fc  
CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgG3 Fc  
CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgM Fc  
CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CVF  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 CD59  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - MCP  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CR1  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - Crry  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 小鼠 CD59  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgG1 Fc  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgM Fc  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgG3 Fc  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgM Fc  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CVF  
CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF(SCR 2-4)  
CR2 (完整的)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF(SCR 2-4)  
CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1 (LP--SCR1-4 - SCR8-11 - SCR15-18)  
CR2 (完整的)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry (5 个 N-末端 SCR)

CR2 (完整的)-VSVFPLE - DAF  
CR2 (完整的)-VSVFPLE - 人 CD59  
CR2 (完整的)-VSVFPLE - MCP  
CR2 (完整的)-VSVFPLE - CR1  
CR2 (完整的)-VSVFPLE - Crry  
CR2 (完整的)-VSVFPLE - 小鼠 CD59  
CR2 (完整的)-VSVFPLE - 人 IgG1 Fc  
CR2 (完整的)-VSVFPLE - 人 IgM Fc

CR2 (完整的)-VSVFPLE - 鼠 IgG3 Fc

CR2 (完整的)-VSVFPLE - 鼠 IgM Fc

CR2 (完整的)-VSVFPLE - CVF

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - DAF

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 CD59

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - MCP

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CR1

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - Crry

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 小鼠  
CD59

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgG1  
Fc

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgM  
Fc

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgG3  
Fc

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgM  
Fc

CR2 (完整的)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CVF

CR2 (完整的)-双马来酰亚氨基己烷 - DAF

CR2 (完整的)-双马来酰亚氨基己烷 - 人 CD59

CR2 (完整的)-双马来酰亚氨基己烷 - MCP

CR2 (完整的)-双马来酰亚氨基己烷 - CR1

CR2 (完整的)-双马来酰亚氨基己烷 - Crry

CR2 (完整的)-双马来酰亚氨基己烷 - 小鼠 CD59

CR2 (完整的)-双马来酰亚氨基己烷 - 人 IgG1 Fc

CR2 (完整的)-双马来酰亚氨基己烷 - 人 IgM Fc

CR2 (完整的)-双马来酰亚氨基己烷 - 鼠 IgG3 Fc

CR2 (完整的)-双马来酰亚氨基己烷 - 鼠 IgM Fc

CR2 (完整的)-双马来酰亚氨基己烷 - CVF

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 CD59

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - MCP

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 小鼠 CD59

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgG1 Fc

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgM Fc

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgG3 Fc

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgM Fc

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CVF

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 CD59

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - MCP

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CR1

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - Crry

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 小鼠 CD59

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgG1 Fc

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgM Fc

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgG3 Fc

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgM Fc

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CVF

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF (SCR 2-4)

CR2 (SCR1-2)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF (SCR 2-4)

CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1 (LP--SCR1-4 - SCR8-11 - SCR15-18)  
CR2 (SCR1-2)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry (5 个 N-末端 SCR)  
CR2 (SCR1-2)-VSVFPLE - DAF  
CR2 (SCR1-2)-VSVFPLE - 人 CD59  
CR2 (SCR1-2)-VSVFPLE - MCP  
CR2 (SCR1-2)-VSVFPLE - CR1  
CR2 (SCR1-2)-VSVFPLE - Crry  
CR2 (SCR1-2)-VSVFPLE - 小鼠 CD59  
CR2 (SCR1-2)-VSVFPLE - 人 IgG1 Fc  
CR2 (SCR1-2)-VSVFPLE - 人 IgM Fc  
CR2 (SCR1-2)-VSVFPLE - 鼠 IgG3 Fc  
CR2 (SCR1-2)-VSVFPLE - 鼠 IgM Fc  
CR2 (SCR1-2)-VSVFPLE - CVF  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - DAF  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 CD59  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - MCP  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CR1  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - Crry  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 小鼠  
CD59  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgG1  
Fc  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgM  
Fc  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgG3  
Fc  
CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgM  
Fc

CR2 (SCR1-2)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CVF  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - DAF  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - 人 CD59  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - MCP  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - CR1  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - Crry  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - 小鼠 CD59  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - 人 IgG1 Fc  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - 人 IgM Fc  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - 鼠 IgG3 Fc  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - 鼠 IgM Fc  
CR2 (SCR1-2)-双马来酰亚氨基己烷 - CVF  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 CD59  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - MCP  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 小鼠 CD59  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgG1 Fc  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgM Fc  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgG3 Fc  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgM Fc  
CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CVF  
CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF  
CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 CD59  
CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - MCP  
CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CR1  
CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - Crry

**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 小鼠 CD59**  
**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgG1 Fc**  
**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgM Fc**  
**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgG3 Fc**  
**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgM Fc**  
**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CVF**  
**CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF (SCR 2-4)**  
**CR2 (SCR1-3)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF (SCR 2-4)**  
**CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1 (LP--SCR1-4 - SCR8-11 - SCR15-18)**  
**CR2 (SCR1-3)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry (5 个 N-末端 SCR)**  
**CR2 (SCR1-3)-VSVFPLE - DAF**  
**CR2 (SCR1-3)-VSVFPLE - 人 CD59**  
**CR2 (SCR1-3)-VSVFPLE - MCP**  
**CR2 (SCR1-3)-VSVFPLE - CR1**  
**CR2 (SCR1-3)-VSVFPLE - Crry**  
**CR2 (SCR1-3)-VSVFPLE - 小鼠 CD59**  
**CR2 (SCR1-3)-VSVFPLE - 人 IgG1 Fc**  
**CR2 (SCR1-3)-VSVFPLE - 人 IgM Fc**  
**CR2 (SCR1-3)-VSVFPLE - 鼠 IgG3 Fc**  
**CR2 (SCR1-3)-VSVFPLE - 鼠 IgM Fc**  
**CR2 (SCR1-3)-VSVFPLE - CVF**  
**CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - DAF**  
**CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 CD59**  
**CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - MCP**  
**CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CR1**  
**CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - Crry**  
**CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 小鼠 CD59**

CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgG1 Fc

CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgM Fc

CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgG3 Fc

CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgM Fc

CR2 (SCR1-3)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CVF

CR2 (SCR1-3)-双马来酰亚氨基己烷 - DAF

CR2 (SCR1-3)-双马来酰亚氨基己烷 - 人 CD59

CR2 (SCR1-3)-双马来酰亚氨基己烷 - MCP

CR2 (SCR1-3)-双马来酰亚氨基己烷 - CR1

CR2 (SCR1-3)-双马来酰亚氨基己烷 - Crry

CR2 (SCR1-3)-双马来酰亚氨基己烷 - 小鼠 CD59

CR2 (SCR1-3)-双马来酰亚氨基己烷 - 人 IgG1 Fc

CR2 (SCR1-3)-双马来酰亚氨基己烷 - 人 IgM Fc

CR2 (SCR1-3)-双马来酰亚氨基己烷 - 鼠 IgG3 Fc

CR2 (SCR1-3)-双马来酰亚氨基己烷 - 鼠 IgM Fc

CR2 (SCR1-3)-双马来酰亚氨基己烷 - CVF

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 CD59

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - MCP

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 小鼠 CD59

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgG1 Fc

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 人 IgM Fc

CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgG3 Fc  
CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - 鼠 IgM Fc  
CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CVF  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 CD59  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - MCP  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CR1  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - Crry  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 小鼠 CD59  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgG1 Fc  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 人 IgM Fc  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgG3 Fc  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - 鼠 IgM Fc  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - CVF  
CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - DAF (SCR 2-4)  
CR2 (SCR1-4)-(Gly<sub>3</sub>Ser)<sub>4</sub> - DAF (SCR 2-4)  
CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - CR1 (LP--SCR1-4 - SCR8-11 - SCR15-18)  
CR2 (SCR1-4)-(Gly<sub>4</sub>Ser)<sub>3</sub> - Crry (5个 N-末端 SCR)  
CR2 (SCR1-4)-VSVFPLE - DAF  
CR2 (SCR1-4)-VSVFPLE - 人 CD59  
CR2 (SCR1-4)-VSVFPLE - MCP  
CR2 (SCR1-4)-VSVFPLE - CR1  
CR2 (SCR1-4)-VSVFPLE - Crry  
CR2 (SCR1-4)-VSVFPLE - 小鼠 CD59  
CR2 (SCR1-4)-VSVFPLE - 人 IgG1 Fc  
CR2 (SCR1-4)-VSVFPLE - 人 IgM Fc  
CR2 (SCR1-4)-VSVFPLE - 鼠 IgG3 Fc  
CR2 (SCR1-4)-VSVFPLE - 鼠 IgM Fc

**CR2 (SCR1-4)-VSVFPLE - CVF**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - DAF**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 CD59**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - MCP**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CR1**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - Crry**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 小鼠 CD59**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgG1 Fc**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 人 IgM Fc**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgG3 Fc**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - 鼠 IgM Fc**

**CR2 (SCR1-4)-间-马来酰亚氨基苯甲酰基-N-羧基琥珀酰亚胺酯 - CVF**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - DAF**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - 人 CD59**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - MCP**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - CR1**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - Crry**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - 小鼠 CD59**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - 人 IgG1 Fc**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - 人 IgM Fc**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - 鼠 IgG3 Fc**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - 鼠 IgM Fc**

**CR2 (SCR1-4)-双马来酰亚氨基己烷 - CVF**

## G. 参考文献

1. Wang, Y. , Rollins, S. A. , Madri, J. A. , 和 Matis, L. A. 1995. 抗-C5 单克隆抗体治疗阻止胶原诱导的关节炎和改善确立的疾病.*Proc Natl Acad Sci USA* 92: 8955-8959.
2. Wang, Y. , Hu, Q. , Madri, J. A. , Rollins, S. A. , Chodera, A. ,和 Matis, L. A. 1996. 利用对补体成分 C5 特异的阻断单克隆抗体治疗后 NZB/W F1 小鼠中狼疮样自身免疫病的改善.*Proc. Natl.Acad. Sci. USA* 93 : 8563-8568.
3. Kaplan, M. 2002. Eculizumab (Alexion).*Curr Opin Investig Drugs* 3: 1017- 1023.
4. Whiss, P. A. 2002. Pexelizumab Alexion. *Curr Opina Investig Drugs* 3: 870-877.
5. Weisman, H. F. , Bartow, T. , Leppo, M. K. , Marsh, H. C. , Carson, G. R. , Consino, M. F. , Boyle, M. P. , Roux, K. H. , Weisfeldt, M. L. , 和 Fearon, D. T. 1990. 可溶性 1 型人补体受体: 阻抑缺血后心肌炎症和坏死的体内补体抑制剂. *Science* 249: 146-151.
6. Rioux, P. 2001. TP-10 (AVANT 免疫治疗剂). *Curr Opin Investig Drugs* 2: 364-371.
7. Higgins, P. J. , Ko, J. -L., Lobell, R. , Sardonini, C. , Alessi, M. K. , 和 Yeh, C. G. 1997. 具有衰变加速因子和因子 I 辅因子活性的可溶嵌合补体激活抑制蛋白.*J. Immunol.* 158: 2872-2881.
8. Moran, P. , Beasly, H. , Gorrel, A. , Martin, E. , Gribbling, P. , Fuchs, H. , Gillet, N., Burton, L. E. ,和 Caras, I. W. 1992. 人重组可溶衰变加速因子体外和体内抑制补体活性. *J. Immunol.* 149: 1736-1743.
9. Christiansen, D., Milland, J. , Thorley, B. R. , McKenzie,I. F., 和 Loveland, B. E. 1996. 重组可溶 CD46 的体内功能分析和与体外重组可溶形式的 CD55 和 CD35 的比较.*ejl* 26: 578-585.
10. Salerno, C. T. , Kulick, D. M. , Yeh, C. G.,Guzman-Paz, M. , Higgins, P. J.,

- Benson, B. A., Park, S. J., Shumway, S. J., Bolman, R. M., 第三版,和 Dalmaso, A. P. 2002. 在猪到恒河猴的心脏移植中 C3 和 C5 转化酶的可溶嵌合抑制剂, 补体激活阻断剂 2 延长了移植存活 *Xenotransplantation* 9: 125-134.
11. Kroshus, T. J., Salerno, C. T., Yeh, C. G., Higgins, P. J., Bolman, R. M., 第三版,和 Dalmaso, A. P. 2000. 在猪到人心脏移植模型中, 由人 CD46 和 CD55 组成的重组可溶嵌合补体抑制剂减少了急性心脏组织的损伤. *Transplantation* 69: 2282-2289.
12. Rushmere, N. K., Van Den Berg, C. W., 和 Morgan, B. P. 2000. 可溶重组形式的小鼠 CD59 的产生和功能表征. *Immunology* 99: 326-332.
13. Quigg, R. J., He, C., Hack, B. K., Alexander, J. J., 和 Morgan, B. P. 2000. 在巴斯德毕赤酵母 (*Pichia pastoris*) 中, 大鼠 CD59 和嵌合 CD59-Crry 做为活性可溶蛋白质的产生和功能性分析. *Immunology* 99: 46-53.
14. Sugita, Y., Ito, K., Shiozuka, K., Gushima, H., Tomita, M., 和 Masuho, Y. 1994. 重组可溶 CD59 抑制补体的反应性溶血作用. *Immunol.* 82: 34-41.
15. Meri, S., Lehto, T., Sutton, C. W., Tynnela, J., 和 Baumann, M. 1996. 可溶 CD59 的结构组成和功能特征: 通过激光解吸质谱分析揭示出寡糖和糖基磷脂酰肌醇 (GPI) 锚的异质性. *Biochem J* 316 (Pt 3): 923-935.
16. Mulligan, M. S., Warner, R. L., Ritterhaus, C. W., Thomas, L. J., Ryan, U. S., Foreman, K. E., Crouch, L. D., Till, G. O., 和 Ward, P. A. 1999. 具有唾液酸 lewis X 部分的补体抑制剂的內皮靶向和增加的抗炎性作用. *J. Immunol.* 162: 4952-4959.
17. Ritterhaus, C. W., Thomas, L. J., Miller, D. P., Picard, M. D., Georhegan-Barek, K. M., Scesney, S. M., Henry, L. D., Sen, A. C., Bertino, A. M., Hannig, G., 等. 1999. 抑制补体活化和结合选择蛋白粘着分子的重组糖蛋白. *J. Biol. Chem.* 274: 11237-11244.
18. Zhang Hf, H., Lu, S., Morrison, S. L., 和 Tomlinson, S. 2001. 功能性抗

体-衰变加速因子 (DAF) 融合蛋白靶向细胞表面. *J Biol Chem* 14: 14.

19. Zhang, H.-F., Yu, J., Bajwa, E., Morrison, S. L., 和 Tomlinson, S. 1999. 功能性抗体-CD59 融合蛋白靶向细胞表面. *J. Clin. Invest.* 103: 55-66.

20. Linton, S. M., Williams, A. S., Dodd, I., Smith, R., Williams, B. D., 和 Morgan, B. P. 2000. 在大鼠的抗原诱导的关节炎中, 新的细胞膜靶向的补体调节剂的治疗效力. *Arthritis Rheum* 43: 2590-2597.

21. Smith, G. P., 和 Smith, R. A. 2001. 细胞膜靶向的补体抑制剂. *Mol Immunol* 38 : 249-255.

22. Carroll, M. C. 1998. 在免疫性诱导和调节中补体和补体受体的作用. *Annu. Rev. Immunol.* 16: 545-568.

23. Carroll, M. C. 2000. 在 B 细胞激活和耐受中补体的作用. *Adv Immunol* 74 : 61-88.

24. Holers, V. M. 1989. 补体受体. *Year Immunol* 4: 231-240.

25. Dierich, M. P., Schulz, T. F., Eigentler, A., Huemer, H., 和 Schwable, W. 1988. 补体系统的受体和调节剂之间的结构和功能关系. *Mol Immunol* 25 :1043-1051.

26. Lowell, C. A., Klickstein, L. B., Carter, R. H., Mitchell, J. A., Fearon, D. T., 和 Ahearn, J. M. 1989. 将 EB 病毒和 C3dg 结合位点作图定位至 2 型补体受体的共同结构域 *J Exp Med* 170: 1931-1946.

27. Szakonyi, G., Guthridge, J. M., Li, D., Young, K., Holers, V. M., 和 Chen, X. S. 2001. 补体受体 2 与其 C3 配体复合的结构. *Science* 292: 1725-1728.

28. Law, S. K., Fearon, D. T., 和 Levine, R. P. 1979. C3b 失活剂对结合细胞的 C3b 的作用. *J Immunol* 122 : 759-765.

29. Seya, T., 和 Nagasawa, S. 1985. 通过调节性酶 C3b 失活剂限制补体蛋白 C3b 的蛋白酶解: 生物活性片段 C3dg 的分离和表征. *J Biochem(Tokyo)* 97: 373-382.

30. Quigg, R. A., Kozono, Y., Berthiaume, D., Lim, A., Salant, J.,

- Weinfeld, A., Griffin, P., Kremmer, E., 和 Holers, V. M. 1998. 用可溶鼠补体抑制剂 Crry-Ig 封闭抗体诱导的肾小球肾炎. *J Immunol.* 160: 4553-4560.
31. Harlow, E., 和 Lane, D. 1988. *Antibodies. A laboratory manual.* New York: Cold Spring Harbor Laboratory.
32. Kim, S. 1993. 脂质体做为癌症化疗的载体, 目前的状况和未来的前景. *Drugs* 46 : 618-638.
33. Davies, A., Simmons, D. L., Hale, G., Harrison, R. A., Tighe, H., Lachmann, P. J., 和 Waldmann, H. 1989. 人淋巴细胞中表达的 Ly-6 蛋白质, CD59 调节同源细胞的补体膜攻击复合物的作用. *Journal of Experimental Medicine* 170:637-654.
34. Guthridge, J. M., Young, K., Gipson, M. G., Sarrias, M. R., Szakonyi, G., Chen, X. S., Malaspina, A., Donoghue, E., James, J. A., Lambris, J. D., 等. 2001. 利用 2 型补体受体 (CR2) /CD21 的 X 射线晶体结构进行表位作图: 直接识别 CR2-C3d 介面的高抑制性单克隆抗体的鉴定. *J Immunol* 167 : 5758- 5766.
35. 1995. *PCR Primer. A laboratory manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
36. Guthridge, J. M., Rakstang, J. K., Young, K. A., Hinshelwood, J., Aslam, M., Robertson, A., Gipson, M. G., Sarrias, M. R., Moore, W. T., Meagher, M., 等. 2001. 2 型补体受体 (CR2/CD21) 短共有/补体重复结构域的重组 N 末端对的溶液中结构研究和与其配体 C3dg 的相互作用. *Biochemistry* 40: 5931-5941.
37. Yu, J., Caragine, T., Chen, S., Morgan, B. P., Frey, A. F., 和 Tomlinson, S. 1999. 通过表达异源 CD59 保护人乳腺癌细胞免受补体介导的裂解. *Clin. Exp. Immunol.* 115 : 13-18.
38. Duits, A. J., Jainandunsing, S. M., van de Winkel, J. G., 和 Capel, P. J. 1991. 在人幼单核细胞 U937 细胞分化期间, 通过白细胞介素 6 选择性增加 Leu-CAM 表达. *Scand J Immunol* 33 : 151-159.

39. Rothlein, R.,和 Springer, T. A. 1986.由佛波酯刺激的同型白细胞粘着中需要淋巴细胞功能相关抗原 1. *J Exp Med* 163: 1132-1149.
40. Ross, G. D. , Reed, W. , Dalzell, J. G. , Becker, S. E. ,和 Hogg, N. 1992.巨噬细胞细胞骨架与 CR3 和 CR4 缔合可以调节受体运动性和 iC3b 调理的红细胞的吞噬作用. *J Leukoc Biol* 51: 109-117.
41. Sharkey, R. M. , Motta-Hennessy, C. , Pawlyk, D. , Siegel, J. A. ,和 Goldenberg, D. M. 1990.在具有人结肠肿瘤异种移植物的裸鼠中钇和碘标记的单克隆抗体 IgG 和片段的生物分布和剂量估算. *Cancer Res.* 50: 2330-2336.
42. Sharkey, R. M. , Natale, A. , Goldenberg, D. M. ,和 Mattes, M. J. 1991.在裸鼠中快速清除血液中免疫球蛋白 G2a 和免疫球蛋白 G2b. *Cancer Res* 51 : 3102-3107.
43. Andrews, B. S. , Eisenberg, R. A. , Theofilopoulos, A. N., Izui, S. , Wilson, C. B., McConahey, P. J. , Murphy, E. D. , Roths, J. B. ,和 Dixon, F. J. 1978. 自发鼠狼疮样综合症.在几种株系中的临床和免疫病理学表现. *J Exp Med* 148: 1198-1215.
44. Hebell, T.,Ahearn, J. M. ,和 Fearon, D. T. 1991.通过 B 淋巴细胞的可溶补体受体抑制免疫应答. *Science* 254: 102- 105.
45. Aslam, M.,和 Perkins, S. J. 2001.通过同步加速器 X 射线和中子散射、分析性超离心和限定的分子模建,人补体单体因子 H 的回折溶液结构. *J Mol Biol* 309 : 1117-1138.
46. Koski, C. L. , Ramm, L. E. , Hammer, C. H. , Mayer, M. M. ,和 Shin, M. L. 1983.通过补体对有核细胞的细胞溶解: 细胞死亡显示多次击中的特征. *Proc Natl Acad Sci USA* 80: 3816-3820.
47. Ramm,L. E. , Whitlow, M. B. , 和 Mayer, M. M. 1982.通过补体形成跨膜通道:单通道所需的 C5b6, C7, C8 和 C9 分子数量的功能分析. *Proc Natl Acad Sci U S A* 79:4751-4755.
48. Takeda, J. , Kozono, H.,Takata, Y. , Hong, K. , Kinoshita, T. , Sayama,

- K., Tanaka, E. ,和 Inoue, K. 1986.补体介导的溶血所必需的击中数量. *Microbiol Immunol* 30 : 461-468.
49. Mendrick, D. L. ,和 H. G.Rennke. 1988. 通过抗 SGP-115/107 的单克隆抗体对大鼠中蛋白尿的诱导. *Kidney Int* 33: 818-830.
50. Mendrick, D. L. , H. G. Rennke, R. S. Cotran, T. A. Springer,和 A. K. Abbas. 1983.抗大鼠肾小球抗原的单克隆抗体:产生和特异性.*Laboratory Investigation* 49: 107-117.
51. Hsu, S. I.,和 W. G. Couser. 2003.通过补体激活介导蛋白尿肾病中肾小管间质损伤的慢性进展: 补体抑制剂的治疗作用? *J Am Soc Nephrol* 14 : S186-191.
52. Sheerin, N. S. ,和 S. H. Sacks. 2002.肾脏中泄漏的蛋白质和间质损伤: 补体是丢失的连接? *Clin Exp Immunol* 130 : 1-3.
53. Hori, Y. , K. Yamada, N. Hanafusa, T. Okuda, N. Okada, T. Miyata, W. G. Couser, K. Kurokawa, T.Fujita,和 M. Nangaku. 1999. Crry, 一种补体调节蛋白, 调节由蛋白尿诱导的肾间质疾病. *Kidney Int* 56: 2096-2106.
54. Song, H. , C. He, C. Knaak, J. M. Guthridge, V. M. Holers,和 S. Tomlinson. 2003.补体受体 2 介导补体抑制剂靶向补体激活位点. *J Clin Invest* 111: 1875-1885.

## **H. 序列**

### **1、 DAF**

核苷酸序列对应于 SEQ ID NO:1

氨基酸序列对应于 SEQ ID NO:2

### **2、 CD59**

核苷酸序列对应于 SEQ ID NO:3

氨基酸序列对应于 SEQ ID NO:4

### **3、 CR2-DAF**

核苷酸序列对应于 SEQ ID NO:5

氨基酸序列对应于 SEQ ID NO:6

### **4、 CR2 - 人 CD59**

核苷酸序列对应于 SEQ ID NO:7

氨基酸序列对应于 SEQ ID NO:8

### **5、 DAF - CR2**

核苷酸序列对应于 SEQ ID NO:9

氨基酸序列对应于 SEQ ID NO:10

### **6、 人 CD59 - CR2**

核苷酸序列对应于 SEQ ID NO:11

氨基酸序列对应于 SEQ ID NO:12

### **7、 CR1**

核苷酸序列对应于 SEQ ID NO:13

氨基酸序列对应于 SEQ ID NO:14

### **8、 MCP**

核苷酸序列对应于 SEQ ID NO:15

氨基酸序列对应于 SEQ ID NO:16

### **9、 小鼠 Crry**

氨基酸序列对应于 SEQ ID NO:17

### **10、 人 IgG1 Fc**

氨基酸序列对应于 SEQ ID NO:18

**11、人 IgM Fc**

氨基酸序列对应于 SEQ ID NO:19

**12、CR2 - 人 IgG1 Fc**

核苷酸序列对应于 SEQ ID NO:20

氨基酸序列对应于 SEQ ID NO:21

**13、小鼠 IgG3 Fc**

氨基酸序列对应于 SEQ ID NO:22

**14、眼镜蛇毒因子**

核苷酸序列对应于 SEQ ID NO:23

氨基酸序列对应于 SEQ ID NO:24

**15、人 CR2**

核苷酸序列对应于 SEQ ID NO:25

氨基酸序列对应于 SEQ ID NO:26

**16、小鼠 CR2**

核苷酸序列对应于 SEQ ID NO:27

氨基酸序列对应于 SEQ ID NO:28

**17、人 CR2**

氨基酸序列对应于 SEQ ID NO:29

## 序列表

<110> MUSC研究发展基金会 (MUSC Foundation For Research Development)  
科罗拉多大学董事会 (Regents of University of Colorado)

<120> 通过补体受体2定向的补体调节剂

<130> 57771-20002.40

<140> PCT/US2003/036459

<141> 2003-11-13

<150> 60/426,676

<151> 2002-11-15

<160> 29

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1041

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 1

```

gactgtggcc ttccccaga tgtacctaat gccagccag ctttggagg ccgtacaagt 60
tttcccagg atactgtaat aacgtacaaa tgtgaagaaa gctttgtgaa aattcctggc 120
gagaaggact cagtgatctg ccttaagggc agtcaatggt cagatattga agagttctgc 180
aatcgtagct gcgaggtgcc aacaaggcta aattctgcat ccctcaaca gccttatac 240
actcagaatt attttccagt cggtagctgt gtggaatatg agtgccgtcc aggttacaga 300
agagaacctt ctctatcacc aaaactaact tgccttcaga atttaaaatg gtccacagca 360
gtcgaathtt gtaaaaagaa atcatgccct aatccgggag aaatacgaaa tggtcagatt 420
gatgtaccag gtggcatatt atttgggtgca accatctcct tctcatgtaa cacagggtac 480
aaattathtt gctcgacttc tagtttttgt cttatttcag gcagctctgt ccagtgagg 540
gaccggttgc cagagtgcag agaaatztat tgtccagcac caccacaaat tgacaatgga 600
ataattcaag gggaacgtga ccattatgga tatagacagt ctgtaacgta tgcattgta 660
aaaggattca ccatgattgg agagcactct atttattgta ctgtgaataa tgatgaagga 720
gagtggagtg gccaccacc tgaatgcaga ggaaaatctc taacttccaa ggtcccacca 780
acagttcaga aacctaccac agtaaagtgt ccaactacag aagtctcacc aacttctcag 840
aaaaccacca caaaaaccac cacacaaat gctcaagca cacggagtac acctgtttcc 900
aggacaacca agcattttca tgaaacaacc ccaaataaag gaagtggaac cacttcaggt 960
actaccgctc ttctatctgg gcacacgtgt ttcacgttga caggtttgct tgggacgcta 1020
gtaaccatgg gcttgctgac t                                     1041

```

<210> 2

<211> 380

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 2

Met Thr Val Ala Arg Pro Ser Val Pro Ala Ala Leu Pro Leu Leu Gly

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

<210> 3

<211> 306

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 3

cagtgtctaca actgtcctaa cccaactgct gactgcaaaa cagccgtcaa ttgttcatct 60  
gattttgatg cgtgtctcat taccaagct gggttacaag tgtataacaa gtgttgaag 120

```

tttgagcatt gcaatttcaa cgacgtcaca acccgcttga gggaaaatga gctaacgtac 180
tactgctgca agaaggacct gtgtaacttt aacgaacagc ttgaaaatgg tgggacatcc 240
ttatcagaga aacagttct tctgctggtg actccatttc tggcagcagc ctggagcctt 300
catccc                                           306

```

<210> 4  
 <211> 126  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

```

<400> 4
Met Gly Ile Gln Gly Gly Ser Val Leu Phe Gly Leu Leu Leu Val Leu
 1           5           10           15
Ala Val Phe Cys His Ser Gly His Gln Cys Tyr Asn Cys Pro Asn Pro
 20           25           30
Thr Ala Asp Cys Lys Thr Ala Val Asn Cys Ser Ser Asp Phe Asp Ala
 35           40           45
Cys Leu Ile Thr Lys Ala Gly Leu Gln Val Tyr Asn Lys Cys Trp Lys
 50           55           60
Phe Glu His Cys Asn Phe Asn Asp Val Thr Thr Arg Leu Arg Glu Asn
 65           70           75           80
Glu Leu Thr Tyr Tyr Cys Cys Lys Lys Asp Leu Cys Asn Phe Asn Glu
 85           90           95
Gln Leu Glu Asn Gly Gly Thr Ser Leu Ser Glu Lys Thr Val Leu Leu
 100          105          110
Leu Val Thr Pro Phe Leu Ala Ala Trp Ser Leu His Pro
 115          120          125

```

<210> 5  
 <211> 1485  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

```

<400> 5
atttcttgtg gctctcctcc gcctatccta aatggccgga ttagttatta ttctaccccc 60
attgctgttg gtaccgtgat aaggtacagt tgttcaggta cttccgcct cattggagaa 120
aaaagtctat tatgcataac taaagacaaa gtggatggaa cctgggataa acctgctcct 180
aaatgtgaat atttcaataa atattcttct tgccctgagc ccatagtacc aggaggatac 240
aaaattagag gctctacacc ctacagacat ggtgattctg tgacatttgc ctgtaaacc 300
aacttctcca tgaacggaaa caagtctggt tgggtgtcaag caataatat gtgggggccg 360
acacgactac caacctgtgt aagtgttttc cctctcgagt gtccagcact tcctatgatc 420
cacaatggac atcacacaag tgagaatggt ggctccattg ctccaggatt gtctgtgact 480
tacagctgtg aatctgggta cttgcttgtt ggagaaaaga tcattaactg tttgtcttcg 540
ggaaaatgga gtgctgtccc cccacatgt gaagaggcac gctgtaaact tctaggacga 600
tttcccaatg ggaaggtaaa ggagcctcca attctccggg ttggtgtaac tgcaaacttt 660
ttctgtgatg aagggtatcg actgcaaggc ccaccttcta gtcggtgtgt aattgctgga 720
cagggagttg cttggaccaa aatgccagta tgtggagggt ggtcgggtgg cggcggatcc 780
gactgtggcc tcccccaaga tgtacctaat gccagccag ctttgggaagg cgtacaagt 840
tttcccagag atactgtaat aacgtacaaa tgtgaagaaa gctttgtgaa aattcctggc 900
gagaaggact cagtgatctg ccttaagggc agtcaatggt cagatattga agagttctgc 960
aatcgtagct gcgaggtgcc aacaaggcta aattctgcat ccctcaaca gccttatac 1020

```

```

actcagaatt attttccagt cgg tactgtt gtggaatatg agtgccgtcc aggttacaga 1080
agagaacctt ctctatcacc aaaactaact tgccttcaga atttaaaatg gtccacagca 1140
gtcgaatttt gtaaaaagaa atcatgccct aatccgggag aaatacgaaa tggtcagatt 1200
gatgtaccag gtggcatatt atttggtgca accatctcct tctcatgtaa cacagggtac 1260
aaattatattg gctcgacttc tagtttttgt cttatttcag gcagctctgt ccagtgaggat 1320
gacccgttgc cagagtgcag agaaatttat tgtccagcac caccacaaat tgacaatgga 1380
ataattcaag gggaacgtga ccattatgga tatagacagt ctgtaacgta tgcgatgtaat 1440
aaaggattca ccatgattgg agagcactct atttattgta ctgtg 1485

```

<210> 6

<211> 495

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 6

```

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

```

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

<210> 7  
 <211> 1002  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 7  
 atttcttggtg gctctcctcc gcctatccta aatggcggga ttagttatta ttctaccccc 60  
 attgctggtg gtaccgtgat aaggtagcagt tgttcaggta ccttccgcct cattggagaa 120  
 aaaagtctat tatgcataac taaagacaaa gtggatggaa cctgggataa acctgctcct 180  
 aaatgtgaat atttcaataa atattcttct tgccctgagc ccatagtacc aggaggatac 240  
 aaaattagag gctctacacc ctacagacat ggtgattctg tgacatttgc ctgtaaaacc 300  
 aacttctcca tgaacggaaa caagtctggt tgggtgtcaag caaataatat gtggggggccg 360  
 acacgactac caacctgtgt aagtgttttc cctctcgagt gtccagcact tcctatgatc 420  
 cacaatggac atcacacaag tgagaatggt ggctccattg ctccaggatt gtctgtgact 480  
 tacagctgtg aatctgggta cttgcttggt ggagaaaaga tcattaactg tttgtcttcg 540  
 ggaaaatgga gtgctgtccc ccccatatgt gaagaggcac gctgtaaate tctaggacga 600  
 tttcccaatg ggaaggtaaa ggagcctcca attctccggg ttggtgtaac tgcaaaacttt 660  
 ttctgtgatg aagggtatcg actgcaaggc ccaccttcta gtcggtgtgt aattgctgga 720  
 caggaggttg cttggaccaaa aatgccagta tgttcaggag gaggagggttc cctgcagtgc 780  
 tacaactgtc ctaacccaac tgctgactgc aaaacagccg tcaattgttc atctgatttt 840  
 gatgcgtgtc tcattaccaa agctgggtta caagtgtata acaagtgttg gaagtttgag 900  
 cattgcaatt tcaacgacgt cacaaccgc ttgagggaaa atgagctaac gtactactgc 960  
 tgcaagaagg acctgtgtaa ctttaacgaa cagcttgaaa at 1002

<210> 8  
 <211> 334  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 8

```

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

```

<210> 9

<211> 1554

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 9

```

gactgtggcc ttccccaga tgtacctaat gccagccag ctttgaagg ccgtacaagt 60
tttcccagg atactgtaat aacgtacaaa tgtgaagaaa gctttgtgaa aattcctggc 120
gagaaggact cagtgatctg ccttaagggc agtcaatggc cagatattga agagttctgc 180
aatcgtagct gcgagggtgcc aacaaggcta aattctgcat ccctcaaca gccttatatc 240
actcagaatt attttccagt cggtactgtt gtggaatatg agtgccgtcc aggttacaga 300
agagaacctt ctctatcacc aaaactaact tgccttcaga atttaaaatg gtccacagca 360
gtcgaatttt gtaaaaagaa atcatgccct aatccgggag aaatacgaaa tggtcagatt 420

```

```

gatgtaccag gtggcatatt atttggtgca accatctcct tctcatgtaa cacagggtag 480
aaattatttg gctcgacttc tagtttttgt cttatttcag gcagctctgt ccagtggagt 540
gacccgttgc cagagtgcag agaaatttat tgtccagcac caccacaaat tgacaatgga 600
ataattcaag gggaacgtga ccattatgga tatagacagt ctgtaacgta tgcatgtaat 660
aaaggattca ccatgattgg agagcactct atttattgta ctgtgaataa tgatgaagga 720
gagtggagtg gccaccacc tgaatgcaga tctcttggtg gcggtggctc gggcggaggt 780
gggtcgggtg geggcgatc catttcttgt ggctctcctc cgcctatcct aaatggccgg 840
attagttatt attctacccc cattgctgtt ggtaccgtga taaggtacag ttgttcaggt 900
accttccgcc tcattggaga aaaaagtcta ttatgcataa ctaaagacaa agtggatgga 960
acctgggata aacctgctcc taaatgtgaa tatttcaata aatattcttc ttgccctgag 1020
cccatagtac caggaggata caaaattaga ggctctacac cctacagaca tgggtgattct 1080
gtgacatttg cctgtaaaac caacttctcc atgaacggaa acaagtctgt ttggtgtcaa 1140
gcaaataata tgtggggggcc gacacgacta ccaacctgtg taagtgtttt ccctctcgag 1200
tgtccagcac ttcctatgat ccacaatgga catcacacaa gtgagaatgt tggctccatt 1260
gctccaggat tgtctgtgac ttacagctgt gaatctgggt acttgcttgt tggagaaaag 1320
atcattaact gtttgtcttc gggaaaatgg agtgctgtcc cccccacatg tgaagaggca 1380
cgctgtaaat ctctaggacg atttcccaat gggaaggtaa aggagcctcc aattctccgg 1440
gttggtgtaa ctgcaaactt tttctgtgat gaagggtatc gactgcaagg cccaccttct 1500
agtcggtgtg taattgctgg acagggagtt gcttggacca aaatgccagt atgt 1554

```

<210> 10

<211> 518

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 10

```

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

```

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

<210> 11  
 <211> 990  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 11  
 ctgcagtgct acaactgtcc taaccctaact gctgactgca aaacagccgt caattgttca 60  
 tctgattttg atgcggtgtct cattaccaaa gctgggttac aagtgtataa caagtgttgg 120  
 aagtttgagc attgcaattt caacgacgtc acaaccgct tgagggaaaa tgagctaacg 180  
 tactactgct gcaagaagga cctgtgtaac tttaacgaac agcttgaaaa ttcctctggt 240  
 ggcggtggct ccggcggagg tgggtccggt ggcggcggt ccatttcttg tggctctcct 300  
 ccgcctatcc taaatggccg gattagttat tattctaccc ccattgctgt tggtagcgtg 360  
 ataaggtaaca gttgttcagg taccttccgc ctcatggag aaaaaagtct attatgcata 420  
 actaaagaca aagtggatgg aacctgggat aaacctgctc ctaaatgtga atatttcaat 480  
 aaatattctt cttgccctga gccatagta ccaggaggat acaaaattag aggctctaca 540  
 ccctacagac atggtgattc tgtgacattt gcctgtaaaa ccaacttctc catgaacgga 600  
 aacaagtctg tttgggtgtca agcaataat atgtgggggc cgacacgact accaacctgt 660  
 gtaagtgttt tccctctcga gtgtccagca cttcctatga tccacaatgg acatcacaca 720  
 agtgagaatg ttggctccat tgctccagga ttgtctgtga cttacagctg tgaatctggt 780  
 tacttgcttg ttggagaaaa gatcattaac tgtttgtctt cgggaaaatg gagtgtctgc 840

```

ccccccacat gtgaagaggc acgctgtaaa tctctaggac gatttcccaa tgggaaggta 900
aaggagcctc caattctccg ggttggtgta actgcaaact ttttctgtga tgaagggtat 960
cgactgcaag gccaccttc tagtcggtgt 990

```

<210> 12  
 <211> 330  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

```

<400> 12
Leu Gln Cys Tyr Asn Cys Pro Asn Pro Thr Ala Asp Cys Lys Thr Ala
 1                    5                    10                    15
Val Asn Cys Ser Ser Asp Phe Asp Ala Cys Leu Ile Thr Lys Ala Gly
                20                    25                    30
Leu Gln Val Tyr Asn Lys Cys Trp Lys Phe Glu His Cys Asn Phe Asn
                35                    40                    45
Asp Val Thr Thr Arg Leu Arg Glu Asn Glu Leu Thr Tyr Tyr Cys Cys
 50                    55                    60
Lys Lys Asp Leu Cys Asn Phe Asn Glu Gln Leu Glu Asn Ser Ser Gly
 65                    70                    75                    80
Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Ile Ser
                85                    90                    95
Cys Gly Ser Pro Pro Ile Leu Asn Gly Arg Ile Ser Tyr Tyr Ser
                100                    105                    110
Thr Pro Ile Ala Val Gly Thr Val Ile Arg Tyr Ser Cys Ser Gly Thr
                115                    120                    125
Phe Arg Leu Ile Gly Glu Lys Ser Leu Leu Cys Ile Thr Lys Asp Lys
 130                    135                    140
Val Asp Gly Thr Trp Asp Lys Pro Ala Pro Lys Cys Glu Tyr Phe Asn
 145                    150                    155                    160
Lys Tyr Ser Ser Cys Pro Glu Pro Ile Val Pro Gly Gly Tyr Lys Ile
                165                    170                    175
Arg Gly Ser Thr Pro Tyr Arg His Gly Asp Ser Val Thr Phe Ala Cys
                180                    185                    190
Lys Thr Asn Phe Ser Met Asn Gly Asn Lys Ser Val Trp Cys Gln Ala
                195                    200                    205
Asn Asn Met Trp Gly Pro Thr Arg Leu Pro Thr Cys Val Ser Val Phe
 210                    215                    220
Pro Leu Glu Cys Pro Ala Leu Pro Met Ile His Asn Gly His His Thr
 225                    230                    235                    240
Ser Glu Asn Val Gly Ser Ile Ala Pro Gly Leu Ser Val Thr Tyr Ser
                245                    250                    255
Cys Glu Ser Gly Tyr Leu Leu Val Gly Glu Lys Ile Ile Asn Cys Leu
                260                    265                    270
Ser Ser Gly Lys Trp Ser Ala Val Pro Pro Thr Cys Glu Glu Ala Arg
                275                    280                    285
Cys Lys Ser Leu Gly Arg Phe Pro Asn Gly Lys Val Lys Glu Pro Pro
 290                    295                    300
Ile Leu Arg Val Gly Val Thr Ala Asn Phe Phe Cys Asp Glu Gly Tyr
 305                    310                    315                    320
Arg Leu Gln Gly Pro Pro Ser Ser Arg Cys
                325                    330

```

<210> 13  
 <211> 5994  
 <212> DNA

## &lt;213&gt; 人工序列

&lt;220&gt;

<223> 人工序列的描述:/注释=  
合成构建体

&lt;400&gt; 13

```

caatgcaatg cccagaatg gcttcattt gccaggccta ccaacctaac tgatgagttt 60
gagtttccca ttgggacata tctgaactat gaatgccgcc ctggttattc cggaagaccg 120
ttttctatca tctgcctaaa aaactcagtc tggactgggtg ctaaggacag gtgcagacgt 180
aaatcatgtc gtaatcctcc agatcctgtg aatggcatgg tgcattgtgat caaaggcatc 240
cagttcggat cccaaattaa atattcttgt actaaaggat accgactcat tggttcctcg 300
tctgccacat gcatcatctc aggtgatact gtcatttggg ataatgaaac acctatttgt 360
gacagaattc cttgtgggct acccccacc atcaccaatg gagatttcat tagcaccaac 420
agagagaatt ttcactatgg atcagtgggtg acctaccgct gcaatcctgg aagcggaggg 480
agaaagggtg ttgagcttgt gggtaggcc tccatatact gcaccagcaa tgacgatcaa 540
gtgggcatct ggagcggccc cgcccctcag tgcattatac ctaacaaatg cacgcctcca 600
aatgtggaaa atggaatatt ggtatctgac aacagaagct tattttcctt aatgaagtt 660
gtggagttta ggtgtcagcc tggctttgtc atgaaaggac cccgccgtgt gaagtgccag 720
gccctgaaca aatgggagcc ggagctacca agctgtctca gggtatgtca gccacctcca 780
gatgtcctgc atgtgagcg tacccaaagg gacaaggaca acttttcacc tgggcaggaa 840
gtgttctaca gctgtgagcc cggctacgac ctcaagggg ctgctctat gcgctgcaca 900
ccccaggag actggagccc tgcagcccc acatgtgaag tgaatcctg tgatgacttc 960
atgggccaac ttcttaatgg ccgtgtgcta tttccagtaa atctccagct tggagcaaaa 1020
gtggattttg tttgtgatga aggatttcaa ttaaaggca gctctgctag ttactgtgtc 1080
ttggctggaa tggaaagcct ttggaatagc agtgttccag tgtgtgaaca aatcttttgt 1140
ccaagtcctc cagttattcc taatgggaga cacacaggaa aacctctgga agtctttccc 1200
tttgaaaag cagtaaatta cacatgcgac cccaccag acagaggag gagcttcgac 1260
ctcattggag agagcaccat ccgctgcaca agtgcacctc aagggaatgg ggtttggagc 1320
agccctgccc ctgctgtgg aattctgggt cactgtcaag cccagatca ttttctgttt 1380
gccaagttga aaacccaaac caatgcatct gactttccca ttgggacatc tttaaagtac 1440
gaatgccgtc ctgagtacta cgggaggcca ttctctatca catgtctaga taacctggtc 1500
tggtaagtc ccaaagatgt ctgtaaactg aaatcatgta aaactcctcc agatccagtg 1560
aatggcatgg tgcattgtgat cacagacatc caggttggat ccagaatcaa ctattcttgt 1620
actacagggc accgactcat tggtcactca tctgctgaat gtatcctctc gggcaatgct 1680
gccattgga gcacgaagcc gccaatttgt caacgaattc cttgtgggct acccccacc 1740
atcgccaatg gagatttcat tagcaccaac agagagaatt ttcactatgg atcagtgggtg 1800
acctaccgct gcaatcctgg aagcggaggg agaaagggtg ttgagcttgt gggtagagccc 1860
tccatatact gcaccagcaa tgacgatcaa gtgggcatct ggagcggccc ggcccctcag 1920
tgcattatac ctaacaaatg cacgcctcca aatgtggaaa atggaatatt ggtatctgac 1980
aacagaagct tattttcctt aatgaagtt gtggagttta ggtgtcagcc tggctttgtc 2040
atgaaaggac cccgccgtgt gaagtgccag gccctgaaca aatgggagcc ggagctacca 2100
agctgtctca gggtatgtca gccacctcca gatgtcctgc atgtgagcg tacccaaagg 2160
gacaaggaca acttttcacc cgggcaggaa gtgttctaca gctgtgagcc cggctatgac 2220
ctcagagggg ctgctctat gcgctgcaca cccagggag actggagccc tgcagcccc 2280
acatgtgaag tgaatcctg tgatgacttc atgggccaac ttcttaatgg ccgtgtgcta 2340
ttccagtaa atctccagct tggagcaaaa gtggattttg tttgtgatga aggatttcaa 2400
ttaaaggca gctctgctag ttattgtgtc ttggctggaa tggaaagcct ttggaatagc 2460
agtgttccag tgtgtgaaca aatcttttgt ccaagtcctc cagtattcc taatgggaga 2520
cacacaggaa aacctctgga agtctttccc tttggaaaag cagtaaatta cacatgcgac 2580
cccaccctc acagaggag gagcttcgac ctcatggag agagcaccat ccgctgcaca 2640
agtaccctc aagggaatgg ggtttggagc agccctgccc ctgctgtgg aattctgggt 2700
cactgtcaag cccagatca ttttctgttt gccaaagtga aaacccaac caatgcatct 2760
gactttccca ttgggacatc tttaaagtac gaatgccgtc ctgagtacta cgggaggcca 2820
ttctctatca catgtctaga taacctggtc tggtaagtc ccaaagatgt ctgtaaactg 2880
aaatcatgta aaactcctcc agatccagtg aatggcatgg tgcattgtgat cacagacatc 2940
caggttggat ccagaatcaa ctattcttgt actacagggc accgactcat tggtcactca 3000
tctgctgaat gtatcctctc aggcaatact gccattgga gcacgaagcc gccaatttgt 3060
caacgaattc cttgtgggct accccaacc atcgccaatg gagatttcat tagcaccaac 3120
agagagaatt ttcactatgg atcagtgggt acctaccgct gcaatcttgg aagcagaggg 3180
agaaagggtg ttgagcttgt gggtagagccc tccatatact gcaccagcaa tgacgatcaa 3240

```

```

gtgggcatct ggagcggccc cgcccctcag tgcattatac ctaacaaatg cacgcctcca 3300
aatgtggaaa atggaatatt ggtatctgac aacagaagct tattttcctt aaatgaagtt 3360
gtggagttta ggtgtcagcc tggctttgtc atgaaaggac cccgccgtgt gaagtgccag 3420
gccctgaaca aatgggagcc agagttacca agctgctcca ggggtgtgtca gccgcctcca 3480
gaaatcctgc atggtgagca taccccaagc catcaggaca acttttcacc tgggcaggaa 3540
gtgttctaca gctgtgagcc tggctatgac ctgagagggg ctgctgtctt gcaactgcaca 3600
ccccagggag actggagccc tgaagccccg agatgtgcag tgaatcctg tgatgacttc 3660
ttgggtcaac tccctcatgg ccgtgtgcta tttccactta atctccagct tggggcaaag 3720
gtgtcctttg tctgtgatga agggtttcgc ttaaagggca gttccgtagt tcattgtgtc 3780
ttggttggaa tgagaagcct ttggaataac agtgttctct tgtgtgaaca tatcttttgt 3840
ccaaatcctc cagctatcct taatgggaga cacacaggaa ctccctctgg agatattccc 3900
tatggaaaag aaatatctta cacatgtgac cccaccag acagagggat gaccttcaac 3960
ctcattgggg agagcacctt ccgctgcaca agtgaccctc atgggaatgg ggtttggagc 4020
agccctgccc ctgctgtgta actttctgtt cgtgctggtc actgtaaaac cccagagcag 4080
tttccatttg ccagtcctac gatcccaatt aatgactttg agtttccagt cgggacatct 4140
ttgaattatg aatgccgtcc tgggtathtt gggaaaatgt tctctatctc ctgcctagaa 4200
aacttgggtt ggtcaagtgt tgaagacaac tgtagacgaa aatcatgtgg acctccacca 4260
gaacccttca atggaatggt gcatataaac acagatacac agtttggatc aacagttaat 4320
tattcttgta atgaagggtt tgcactcatt ggttccccat ctactacttg tctcgtctca 4380
ggcaataatg tcacatggga taagaaggca cctatttgtg agatcatact ttgtgagcca 4440
cctccaacca tatccaatgg agacttctac agcaacaata gaacatcttt tcacaatgga 4500
acgggtggtaa cttaccagtg ccacactgga ccagatggag aacagctggt tgagcttgtg 4560
ggagaacggt caatatattg caccagcaaa gatgatcaag ttggtgtttg gagcagccct 4620
ccccctcggg gtatttctac taataaatgc acagctccag aagtggaaa tgcaattaga 4680
gtaccaggaa acaggagttt cttttccctc actgagatca tcagatttag atgtcagccc 4740
gggtttgtca tggtagggtc ccacactgtg cagtgccaga ccaatggcag atgggggccc 4800
aagctgccac actgctccag ggtgtgtcag ccgctccag aaatcctgca tgggtgagcat 4860
accctaagcc atcaggacaa cttttcacct gggcaggaag tgttctacag ctgtgagccc 4920
agctatgacc tcagaggggc tgcgtctctg cactgcacgc cccagggaga ctggagccct 4980
gaagccccta gatgtacagt gaaatcctgt gatgacttcc tgggccaact ccctcatggc 5040
cgtgtgctac ttccacttaa tctccagctt ggggcaaagg tgcctttgt ttgcgatgaa 5100
gggttccgat taaaaggcag gtctgctagt cattgtgtct tggctggaat gaaagccctt 5160
tggaatagca gtgttccagt gtgtgaacaa atcttttgtc caaatcctcc agctatcctt 5220
aatgggagac acacaggaac tccctttgga gatattccct atggaaaaga aatatcttac 5280
gcattgcgaca cccaccaga cagagggatg accttcaacc tcattgggga gagctccatc 5340
cgctgcacaa gtgaccctca agggaatggg gtttggagca gccctgcccc tgcgtgtgaa 5400
ctttctgttc ctgctgctg cccacatcca cccaagatcc aaaacgggca ttacattgga 5460
ggacacgtat ctctatatct tcttgggatg acaatcagct acacttgtga ccccggtctc 5520
ctgttagtgg gaaagggctt cattttctgt acagaccagg gaatctggag ccaattggat 5580
cattattgca aagaagtaaa ttgtagcttc ccactgttta tgaatggaat ctogaaggag 5640
ttagaaatga aaaaagtata tcaactatgga gattatgtga ctttgaagtg tgaagatggg 5700
tatactctgg aaggcagtc ctggagccag tgccaggcgg atgacagatg ggaccctcct 5760
ctggccaaat gtacctctcg tgcacatgat gctctcatag ttggcacttt atctggtacg 5820
atcttcttta ttttactcat cattttcctc tcttgataa ttctaagca cagaaaaggc 5880
aataatgcac atgaaaacc taaagaagtg gctatccatt tacattctca aggaggcagc 5940
agcgttcac cccgaactct gcaaacaaat gaagaaaata gcagggctct tcct 5994

```

<210> 14

<211> 2048

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 14

Met Cys Leu Gly Arg Met Gly Ala Ser Ser Pro Arg Ser Pro Glu Pro  
1 5 10 15

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

Leu Lys Thr Gln Thr Asn Ala Ser Asp Phe Pro Ile Gly Thr Ser Leu  
 515 520 525  
 Lys Tyr Glu Cys Arg Pro Glu Tyr Tyr Gly Arg Pro Phe Ser Ile Thr  
 530 535 540  
 Cys Leu Asp Asn Leu Val Trp Ser Ser Pro Lys Asp Val Cys Lys Arg  
 545 550 555 560  
 Lys Ser Cys Lys Thr Pro Pro Asp Pro Val Asn Gly Met Val His Val  
 565 570 575  
 Ile Thr Asp Ile Gln Val Gly Ser Arg Ile Asn Tyr Ser Cys Thr Thr  
 580 585 590  
 Gly His Arg Leu Ile Gly His Ser Ser Ala Glu Cys Ile Leu Ser Gly  
 595 600 605  
 Asn Ala Ala His Trp Ser Thr Lys Pro Pro Ile Cys Gln Arg Ile Pro  
 610 615 620  
 Cys Gly Leu Pro Pro Thr Ile Ala Asn Gly Asp Phe Ile Ser Thr Asn  
 625 630 635 640  
 Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro  
 645 650 655  
 Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile  
 660 665 670  
 Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala  
 675 680 685  
 Pro Gln Cys Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn  
 690 695 700  
 Gly Ile Leu Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val  
 705 710 715 720  
 Val Glu Phe Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro Arg Arg  
 725 730 735  
 Val Lys Cys Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys  
 740 745 750  
 Ser Arg Val Cys Gln Pro Pro Pro Asp Val Leu His Ala Glu Arg Thr  
 755 760 765  
 Gln Arg Asp Lys Asp Asn Phe Ser Pro Gly Gln Glu Val Phe Tyr Ser  
 770 775 780  
 Cys Glu Pro Gly Tyr Asp Leu Arg Gly Ala Ala Ser Met Arg Cys Thr  
 785 790 795 800  
 Pro Gln Gly Asp Trp Ser Pro Ala Ala Pro Thr Cys Glu Val Lys Ser  
 805 810 815  
 Cys Asp Asp Phe Met Gly Gln Leu Leu Asn Gly Arg Val Leu Phe Pro  
 820 825 830  
 Val Asn Leu Gln Leu Gly Ala Lys Val Asp Phe Val Cys Asp Glu Gly  
 835 840 845  
 Phe Gln Leu Lys Gly Ser Ser Ala Ser Tyr Cys Val Leu Ala Gly Met  
 850 855 860  
 Glu Ser Leu Trp Asn Ser Ser Val Pro Val Cys Glu Gln Ile Phe Cys  
 865 870 875 880  
 Pro Ser Pro Pro Val Ile Pro Asn Gly Arg His Thr Gly Lys Pro Leu  
 885 890 895  
 Glu Val Phe Pro Phe Gly Lys Ala Val Asn Tyr Thr Cys Asp Pro His  
 900 905 910  
 Pro Asp Arg Gly Thr Ser Phe Asp Leu Ile Gly Glu Ser Thr Ile Arg  
 915 920 925  
 Cys Thr Ser Asp Pro Gln Gly Asn Gly Val Trp Ser Ser Pro Ala Pro  
 930 935 940  
 Arg Cys Gly Ile Leu Gly His Cys Gln Ala Pro Asp His Phe Leu Phe  
 945 950 955 960  
 Ala Lys Leu Lys Thr Gln Thr Asn Ala Ser Asp Phe Pro Ile Gly Thr  
 965 970 975  
 Ser Leu Lys Tyr Glu Cys Arg Pro Glu Tyr Tyr Gly Arg Pro Phe Ser  
 980 985 990  
 Ile Thr Cys Leu Asp Asn Leu Val Trp Ser Ser Pro Lys Asp Val Cys  
 995 1000 1005

Lys Arg Lys Ser Cys Lys Thr Pro Pro Asp Pro Val Asn Gly Met Val  
 1010 1015 1020  
 His Val Ile Thr Asp Ile Gln Val Gly Ser Arg Ile Asn Tyr Ser Cys  
 1025 1030 1035 1040  
 Thr Thr Gly His Arg Leu Ile Gly His Ser Ser Ala Glu Cys Ile Leu  
 1045 1050 1055  
 Ser Gly Asn Thr Ala His Trp Ser Thr Lys Pro Pro Ile Cys Gln Arg  
 1060 1065 1070  
 Ile Pro Cys Gly Leu Pro Pro Thr Ile Ala Asn Gly Asp Phe Ile Ser  
 1075 1080 1085  
 Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys  
 1090 1095 1100  
 Asn Leu Gly Ser Arg Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro  
 1105 1110 1115 1120  
 Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly  
 1125 1130 1135  
 Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val  
 1140 1145 1150  
 Glu Asn Gly Ile Leu Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn  
 1155 1160 1165  
 Glu Val Val Glu Phe Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro  
 1170 1175 1180  
 Arg Arg Val Lys Cys Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro  
 1185 1190 1195 1200  
 Ser Cys Ser Arg Val Cys Gln Pro Pro Pro Glu Ile Leu His Gly Glu  
 1205 1210 1215  
 His Thr Pro Ser His Gln Asp Asn Phe Ser Pro Gly Gln Glu Val Phe  
 1220 1225 1230  
 Tyr Ser Cys Glu Pro Gly Tyr Asp Leu Arg Gly Ala Ala Ser Leu His  
 1235 1240 1245  
 Cys Thr Pro Gln Gly Asp Trp Ser Pro Glu Ala Pro Arg Cys Ala Val  
 1250 1255 1260  
 Lys Ser Cys Asp Asp Phe Leu Gly Gln Leu Pro His Gly Arg Val Leu  
 1265 1270 1275 1280  
 Phe Pro Leu Asn Leu Gln Leu Gly Ala Lys Val Ser Phe Val Cys Asp  
 1285 1290 1295  
 Glu Gly Phe Arg Leu Lys Gly Ser Ser Val Ser His Cys Val Leu Val  
 1300 1305 1310  
 Gly Met Arg Ser Leu Trp Asn Asn Ser Val Pro Val Cys Glu His Ile  
 1315 1320 1325  
 Phe Cys Pro Asn Pro Pro Ala Ile Leu Asn Gly Arg His Thr Gly Thr  
 1330 1335 1340  
 Pro Ser Gly Asp Ile Pro Tyr Gly Lys Glu Ile Ser Tyr Thr Cys Asp  
 1345 1350 1355 1360  
 Pro His Pro Asp Arg Gly Met Thr Phe Asn Leu Ile Gly Glu Ser Thr  
 1365 1370 1375  
 Ile Arg Cys Thr Ser Asp Pro His Gly Asn Gly Val Trp Ser Ser Pro  
 1380 1385 1390  
 Ala Pro Arg Cys Glu Leu Ser Val Arg Ala Gly His Cys Lys Thr Pro  
 1395 1400 1405  
 Glu Gln Phe Pro Phe Ala Ser Pro Thr Ile Pro Ile Asn Asp Phe Glu  
 1410 1415 1420  
 Phe Pro Val Gly Thr Ser Leu Asn Tyr Glu Cys Arg Pro Gly Tyr Phe  
 1425 1430 1435 1440  
 Gly Lys Met Phe Ser Ile Ser Cys Leu Glu Asn Leu Val Trp Ser Ser  
 1445 1450 1455  
 Val Glu Asp Asn Cys Arg Arg Lys Ser Cys Gly Pro Pro Pro Glu Pro  
 1460 1465 1470  
 Phe Asn Gly Met Val His Ile Asn Thr Asp Thr Gln Phe Gly Ser Thr  
 1475 1480 1485  
 Val Asn Tyr Ser Cys Asn Glu Gly Phe Arg Leu Ile Gly Ser Pro Ser  
 1490 1495 1500

Thr Thr Cys Leu Val Ser Gly Asn Asn Val Thr Trp Asp Lys Lys Ala  
 1505 1510 1515 1520  
 Pro Ile Cys Glu Ile Ile Ser Cys Glu Pro Pro Pro Thr Ile Ser Asn  
 1525 1530 1535  
 Gly Asp Phe Tyr Ser Asn Asn Arg Thr Ser Phe His Asn Gly Thr Val  
 1540 1545 1550  
 Val Thr Tyr Gln Cys His Thr Gly Pro Asp Gly Glu Gln Leu Phe Glu  
 1555 1560 1565  
 Leu Val Gly Glu Arg Ser Ile Tyr Cys Thr Ser Lys Asp Asp Gln Val  
 1570 1575 1580  
 Gly Val Trp Ser Ser Pro Pro Pro Arg Cys Ile Ser Thr Asn Lys Cys  
 1585 1590 1595 1600  
 Thr Ala Pro Glu Val Glu Asn Ala Ile Arg Val Pro Gly Asn Arg Ser  
 1605 1610 1615  
 Phe Phe Ser Leu Thr Glu Ile Ile Arg Phe Arg Cys Gln Pro Gly Phe  
 1620 1625 1630  
 Val Met Val Gly Ser His Thr Val Gln Cys Gln Thr Asn Gly Arg Trp  
 1635 1640 1645  
 Gly Pro Lys Leu Pro His Cys Ser Arg Val Cys Gln Pro Pro Pro Glu  
 1650 1655 1660  
 Ile Leu His Gly Glu His Thr Leu Ser His Gln Asp Asn Phe Ser Pro  
 1665 1670 1675 1680  
 Gly Gln Glu Val Phe Tyr Ser Cys Glu Pro Ser Tyr Asp Leu Arg Gly  
 1685 1690 1695  
 Ala Ala Ser Leu His Cys Thr Pro Gln Gly Asp Trp Ser Pro Glu Ala  
 1700 1705 1710  
 Pro Arg Cys Thr Val Lys Ser Cys Asp Asp Phe Leu Gly Gln Leu Pro  
 1715 1720 1725  
 His Gly Arg Val Leu Leu Pro Leu Asn Leu Gln Leu Gly Ala Lys Val  
 1730 1735 1740  
 Ser Phe Val Cys Asp Glu Gly Phe Arg Leu Lys Gly Arg Ser Ala Ser  
 1745 1750 1755 1760  
 His Cys Val Leu Ala Gly Met Lys Ala Leu Trp Asn Ser Ser Val Pro  
 1765 1770 1775  
 Val Cys Glu Gln Ile Phe Cys Pro Asn Pro Pro Ala Ile Leu Asn Gly  
 1780 1785 1790  
 Arg His Thr Gly Thr Pro Phe Gly Asp Ile Pro Tyr Gly Lys Glu Ile  
 1795 1800 1805  
 Ser Tyr Ala Cys Asp Thr His Pro Asp Arg Gly Met Thr Phe Asn Leu  
 1810 1815 1820  
 Ile Gly Glu Ser Ser Ile Arg Cys Thr Ser Asp Pro Gln Gly Asn Gly  
 1825 1830 1835 1840  
 Val Trp Ser Ser Pro Ala Pro Arg Cys Glu Leu Ser Val Pro Ala Ala  
 1845 1850 1855  
 Cys Pro His Pro Pro Lys Ile Gln Asn Gly His Tyr Ile Gly Gly His  
 1860 1865 1870  
 Val Ser Leu Tyr Leu Pro Gly Met Thr Ile Ser Tyr Thr Cys Asp Pro  
 1875 1880 1885  
 Gly Tyr Leu Leu Val Gly Lys Gly Phe Ile Phe Cys Thr Asp Gln Gly  
 1890 1895 1900  
 Ile Trp Ser Gln Leu Asp His Tyr Cys Lys Glu Val Asn Cys Ser Phe  
 1905 1910 1915 1920  
 Pro Leu Phe Met Asn Gly Ile Ser Lys Glu Leu Glu Met Lys Lys Val  
 1925 1930 1935  
 Tyr His Tyr Gly Asp Tyr Val Thr Leu Lys Cys Glu Asp Gly Tyr Thr  
 1940 1945 1950  
 Leu Glu Gly Ser Pro Trp Ser Gln Cys Gln Ala Asp Asp Arg Trp Asp  
 1955 1960 1965  
 Pro Pro Leu Ala Lys Cys Thr Ser Arg Ala His Asp Ala Leu Ile Val  
 1970 1975 1980  
 Gly Thr Leu Ser Gly Thr Ile Phe Phe Ile Leu Leu Ile Ile Phe Leu  
 1985 1990 1995 2000

Ser Trp Ile Ile Leu Lys His Arg Lys Gly Asn Asn Ala His Glu Asn  
 2005 2010 2015  
 Pro Lys Glu Val Ala Ile His Leu His Ser Gln Gly Gly Ser Ser Val  
 2020 2025 2030  
 His Pro Arg Thr Leu Gln Thr Asn Glu Glu Asn Ser Arg Val Leu Pro  
 2035 2040 2045

<210> 15  
 <211> 1029  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 15  
 tgtgaggagc caccaacatt tgaagctatg gagctcattg gtaaaccaaa accctactat 60  
 gagattggtg aacgagtaga ttataagtgt aaaaaaggat acttctatat acctcctctt 120  
 gccaccata ctatttgtga tcggaatcat acatggctac ctgtctcaga tgacgcctgt 180  
 tatagagaaa catgtccata tatacgggat cctttaaatg gccaaagcagt ccctgcaaat 240  
 gggacttacg agtttgggta tcagatgcac tttatttcta atgagggtta ttacttaatt 300  
 ggtgaagaaa ttctatatg tgaacttaaa ggatcagtag caatttggag cggtaagccc 360  
 ccaatatgtg aaaaggtttt gtgtacacca cctccaaaaa taaaaaatgg aaaacacacc 420  
 tttagtgaag tagaagtatt tgagtatctt gatgcagtaa cttatagttg tgatcctgca 480  
 cctggaccag atccattttc acttattgga gagagcacga tttattgttg tgacaattca 540  
 gtgtggagtc gtgctgctcc agagtgtaaa gtggtcaaat gtcgatttcc agtagtcgaa 600  
 aatgaaaac agatatcagg atttggaaaa aaattttact acaaagcaac agttatgttt 660  
 gaatgcgata agggttttta cctcgatggc agcgacacaa ttgtctgtga cagtaacagt 720  
 acttgggatc ccccagttcc aaagtgtctt aaagtgtcga cttcttccac tacaaaatct 780  
 ccagcgtcca gtgcctcagg tcctaggcct acttacaagc ctccagtctc aaattatcca 840  
 ggatatccta aacctgagga aggaatactt gacagtttgg atgtttgggt cattgctgtg 900  
 attgttattg ccatagttgt tggagttgca gtaatttgtg ttgtcccgta cagatatctt 960  
 caaaggagga agaagaaagg cacataccta actgatgaga cccacagaga agtaaaattt 1020  
 acttctctc 1029

<210> 16  
 <211> 378  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 16  
 Met Glu Pro Pro Gly Arg Arg Glu Cys Pro Phe Pro Ser Trp Arg Phe  
 1 5 10 15  
 Phe Pro Gly Leu Leu Ala Ala Met Val Leu Leu Leu Tyr Ser Phe  
 20 25 30  
 Ser Asp Ala Cys Glu Glu Pro Pro Thr Phe Glu Ala Met Glu Leu Ile  
 35 40 45  
 Gly Lys Pro Lys Pro Tyr Tyr Glu Ile Gly Glu Arg Val Asp Tyr Lys  
 50 55 60  
 Cys Lys Lys Gly Tyr Phe Tyr Ile Pro Pro Leu Ala Thr His Thr Ile  
 65 70 75 80  
 Cys Asp Arg Asn His Thr Trp Leu Pro Val Ser Asp Asp Ala Cys Tyr  
 85 90 95  
 Arg Glu Thr Cys Pro Tyr Ile Arg Asp Pro Leu Asn Gly Gln Ala Val

```

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

```

<210> 17

<211> 440

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 17

```

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

```

```

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

```

<210> 18

<211> 232

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 18

```

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1                      5                      10                      15
Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
      20                      25                      30
Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
      35                      40                      45
Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val

```

```

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

```

<210> 19  
 <211> 454  
 <212> PRT  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

```

<400> 19
Gly Ser Ala Ser Ala Pro Thr Leu Phe Pro Leu Val Ser Cys Glu Asn
 1              5              10              15
Ser Pro Ser Asp Thr Ser Ser Val Ala Val Gly Cys Leu Ala Gln Asp
      20              25              30
Phe Leu Pro Asp Ser Ile Thr Phe Ser Trp Lys Tyr Lys Asn Asn Ser
      35              40              45
Asp Ile Ser Ser Thr Arg Gly Phe Pro Ser Val Leu Arg Gly Gly Lys
 50              55              60
Tyr Ala Ala Thr Ser Gln Val Leu Leu Pro Ser Lys Asp Val Met Gln
65              70              75              80
Gly Thr Asp Glu His Val Val Cys Lys Val Gln His Pro Asn Gly Asn
      85              90              95
Lys Glu Lys Asn Val Pro Leu Pro Val Ile Ala Glu Leu Pro Pro Lys
      100              105              110
Val Ser Val Phe Val Pro Pro Arg Asp Gly Phe Phe Gly Asn Pro Arg
      115              120
Ser Lys Ser Lys Leu Ile Cys Gln Ala Thr Gly Phe Ser Pro Arg Gln
      130              135              140
Ile Gln Val Ser Trp Leu Arg Glu Gly Lys Gln Val Gly Ser Gly Val
145              150              155              160
Thr Thr Asp Gln Val Gln Ala Glu Ala Lys Glu Ser Gly Pro Thr Thr
      165              170              175
Tyr Lys Val Thr Ser Thr Leu Thr Ile Lys Glu Ser Asp Trp Leu Ser
      180              185              190
Gln Ser Met Phe Thr Cys Arg Val Asp His Arg Gly Leu Thr Phe Gln
      195              200              205

```

Gln Asn Ala Ser Ser Met Cys Val Pro Asp Gln Asp Thr Ala Ile Arg  
 210 215 220  
 Val Phe Ala Ile Pro Pro Ser Phe Ala Ser Ile Phe Leu Thr Lys Ser  
 225 230 235 240  
 Thr Lys Leu Thr Cys Leu Val Thr Asp Leu Thr Thr Tyr Asp Ser Val  
 245 250 255  
 Thr Ile Ser Trp Thr Arg Gln Asn Gly Glu Ala Val Lys Thr His Thr  
 260 265 270  
 Asn Ile Ser Glu Ser His Pro Asn Ala Thr Phe Ser Ala Val Gly Glu  
 275 280 285  
 Ala Ser Ile Cys Glu Asp Asp Trp Asn Ser Gly Glu Arg Phe Thr Cys  
 290 295 300  
 Thr Val Thr His Thr Asp Leu Pro Ser Pro Leu Lys Gln Thr Ile Ser  
 305 310 315 320  
 Arg Pro Lys Gly Val Ala Leu His Arg Pro Asp Val Tyr Leu Leu Pro  
 325 330 335  
 Pro Ala Arg Glu Gln Leu Asn Leu Arg Glu Ser Ala Thr Ile Thr Cys  
 340 345 350  
 Leu Val Thr Gly Phe Ser Pro Ala Asp Val Phe Val Gln Trp Met Gln  
 355 360 365  
 Arg Gly Gln Pro Leu Ser Pro Glu Lys Tyr Val Thr Ser Ala Pro Met  
 370 375 380  
 Pro Glu Pro Gln Ala Pro Gly Arg Tyr Phe Ala His Ser Ile Leu Thr  
 385 390 395 400  
 Val Ser Glu Glu Glu Trp Asn Thr Gly Glu Thr Tyr Thr Cys Val Val  
 405 410 415  
 Ala His Glu Ala Leu Pro Asn Arg Val Thr Glu Arg Thr Val Asp Lys  
 420 425 430  
 Ser Thr Gly Lys Pro Thr Leu Tyr Asn Val Ser Leu Val Met Ser Asp  
 435 440 445  
 Thr Ala Gly Thr Cys Tyr  
 450

<210> 20  
 <211> 1530  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 20  
 atgggagccg cgggctgct cggggtttct ttggctctcg tcgcaccggg ggtcctcggg 60  
 atttcttggtg gctctcctcc gcctatccta aatggccgga ttagttatta ttctaccccc 120  
 attgctggtg gtaccgtgat aaggtacagt tgttcaggta ccttccgcct cattggagaa 180  
 aaaagtctat tatgcataac taaagacaaa gtggatggaa cctgggataa acctgctcct 240  
 aaatgtgaat atttcaataa atattcttct tgccctgagc ccatagtacc aggaggatac 300  
 aaaattagag gctctacacc ctacagacat ggtgattctg tgacatttgc ctgtaaaacc 360  
 aacttctcca tgaacggaaa caagtctggt tgggtgcaag caaataatat gtgggggccg 420  
 acacgactac caacctgtgt aagtgttttc cctctcgagt gtccagcact tcctatgatc 480  
 cacaatggac atcacacaag tgagaatggt ggctccattg ctccaggatt gtctgtgact 540  
 tacagctgtg aatctgggta cttgcttggt ggagaaaaga tcattaactg tttgtcttcg 600  
 ggaaaatgga gtgctgtccc ccccacatgt gaagaggcac gctgtaaact tctaggacga 660  
 tttcccaatg ggaaggtaaa ggagcctcca attctccggg ttgggtgtaac tgcaaacttt 720  
 ttctgtgatg aagggtatcg actgcaaggc ccaccttcta gtcggtgtgt aattgctgga 780  
 cagggagttg cttggacca aatgccagta tgtgaagaaa ttttttgccc actgcggccg 840  
 cagtctagag acaaaactca cacatgccc cctgcccag cacctgaact cctgggggga 900  
 ccgtcagctc tcctcttccc cccaaaacc aaggacacc tcatgatctc ccggaccct 960  
 gaggtcacat gcgtgggtgt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 1020

```

tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgggagga gcagtacaac 1080
agcacgtacc gtgtggtcag cgtcctcacc gtcctgcacc aggactggct gaatggcaag 1140
gagtacaagt gcaaggctct caacaaagcc ctcccagtc ccacgcgagaa aaccatctcc 1200
aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggaggag 1260
atgaccaaga accaggtcag cctgacctgc ctgggtcaaag gcttctatcc cagcgacatc 1320
gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 1380
ctggactccg acggctcctt cttcctctat agcaagctca ccgtggacaa gagcagggtg 1440
cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg 1500
cagaagagcc tctccctgtc cccgggtaaa 1530

```

<210> 21

<211> 510

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 21

```

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

```

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

&lt;210&gt; 22

&lt;211&gt; 233

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

<223> 人工序列的描述:/注释=  
合成构建体

&lt;400&gt; 22

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

Phe Thr Cys Ser Val Val His Glu Ala Leu His Asn His His Thr Gln  
 210 215 220  
 Lys Asn Leu Ser Arg Ser Pro Gly Lys  
 225 230

<210> 23  
 <211> 4860  
 <212> DNA  
 <213> 人工序列

<220>  
 <223> 人工序列的描述:/注释=  
 合成构建体

<400> 23  
 gctctctaca ccctcatcac ccctgctggt ttgccaacag acacagaaga gcaaattttg 60  
 gtggaggccc atggagacag tactccaaaa cagcttgaca tctttgttca tgattttcca 120  
 cggaagcaga aaaccttggt ccaaaccaga gtagatatga atccagcagg aggcattgctt 180  
 gtcactccaa ctatagagat tccagcaaaa gaagttagta cggactccag gcaaaatcaa 240  
 tatgtggttg tgcaagtaac tggctcctcaa gtgagattgg aaaagggtgtt tctcctttct 300  
 taccagagta gctttctggt tatccagaca gataaaggca tctatacacc agggctctcca 360  
 gtactctatc gtgttttttc tatggatcac aacacaagca agatgaacaa aactgtgatt 420  
 gttgagtttc agactccaga aggcattctt gtcagttcta attcagttga cctaaacttc 480  
 ttctggcctt acaatttacc agaccttgtc agtttgggga cttggaggat tgtggccaaa 540  
 tatgaacatt cccagagaa ttatactgca tattttgatg tcaggaaata tgtgttgcca 600  
 agctttgaag tccgtctgca accatcagag aagttttttt acattgacgg caatgaaaat 660  
 ttccacgtgt ctactactgc aaggtaactt tatggagagg aagtggaaagg tgtggccttt 720  
 gtcctctttg gagtgaaaat agatgatgct aaaaagagta ttccagactc actcacgaga 780  
 attccgatta ttgatggaga tgggaaagca aactaaaaa gagatacatt ccgttctcga 840  
 tttccaaatc tcaatgagct tgttgggcat actctgtatg catctgtaac agtcatgaca 900  
 gaatcaggca gtgatatggt agtgactgag caaagcggca ttcataattgt ggcattctccc 960  
 tatcagatcc acttcacaaa aacccccaaa tatttcaagc caggaatgcc atatgaactg 1020  
 acgggtgatg ttaccaacc tgatggctca ccagctgcc atgtgccagt ggtatcagag 1080  
 gcctttcatt ctatgggaac cactttgagt gatgggactg ctaagctcat cctgaacata 1140  
 ccattgaatg ctcaaagcct accaatcact gttagaacta accatggaga cctccaaga 1200  
 gaacgccagg caacaaagtc catgacagcc atagcctacc aaaccaggg aggatctgga 1260  
 aactatcttc atgtagccat tacatctaca gagattaagc ccggagataa cttacctgtc 1320  
 aatttcaatg tgaagggcaa tgcaaatca ctgaagcaga tcaaatattt cacatacctc 1380  
 atattgaata aaggggaagat tttcaagggt ggcaggcaac ccaggagaga tgggcagaat 1440  
 ctggtgacca tgaatctgca tactactcca gatctcatcc cttccttccg gtttgtggct 1500  
 tactaccaag tgggaaacaa cgaaattgtg gctgattctg tctgggtgga tgtgaaggat 1560  
 acctgcatgg gaacgttggt tgtgaaagga gacaatctaa tacaatgcc agggagcaga 1620  
 atgaaaatca aattggaagg ggatccaggt gctcgggttg gtcttgtggc tctggacaaa 1680  
 gcagtatatg ttctcaatga taaatataag attagccaag ctaagatatg ggacacaata 1740  
 gaaaagagtg actttggctg tacagctggc agtggccaga ataatctggg tgtgtttgaa 1800  
 gatgctggac tggctctgac aaccagcact aatctcaaca ccaaacagag atcagctgca 1860  
 aagtgtcctc agcctgcaaa tccgaggcgt cgcagttctg ttttgctgct tgacagcaac 1920  
 gcaagcaaaag cggcagaatt tcaggatcaa gacctgcgta aatgctgtga agatgtcatg 1980  
 catgagaacc ccatggggta cacttgtgaa aagcgtgcaa aatacatcca ggaggagat 2040  
 gcttgtaagg ctgccttctt tgaatgctgt cgctacatca aggggtccg agatgaaaac 2100  
 caacgggaga gcgagtgtgt tctggcaaga gatgataatg aagatggttt catagcagat 2160  
 agtgatatca tctcaaggct tgatttcccc aagagttggt tgtggctaac aaaggacttg 2220  
 accgaggagc ctaacagtca agggatttca agcaagacaa tgtcttttta tctgagggat 2280  
 tccatcacia cctgggtggt gctggctgta agctttacac ccaccaaagg gatctgtgtg 2340  
 gctgaacctt atgaaataag agtcatgaaa gtcttcttca ttgatcttca aatgccatat 2400  
 tcagtagtga agaattgagca ggtggagatt cgagctattc tgcacaacta cgttaacgag 2460  
 gatatttatg tgcgagtgga actgttatac aaccagcct tctgcagtgc ttccacaaaa 2520  
 ggacaaagat accgacagca gttcccaatt aaagccctgt cctccagagc agtaccgttt 2580  
 gtgatagtcc cattagagca aggattgcat gatgttgaga ttaaagcaag tgtccaggaa 2640  
 gcgttgtggt cagacggtgt gaggaagaaa ctgaaagttg tacctgaagg ggtacagaaa 2700

```

tccattgtga ctattgttaa actggacca agggcaaaag gagttggtgg aacacagcta 2760
gaagtgatca aagcccgcaa attagatgac agagtgcctg acacagaaat tgaaccaag 2820
attatcatcc aaggtgaccc tgtggctcag attattgaaa actcaattga tggagtaaa 2880
ctcaaccatc tcattatcac tccttctggc tgtggggagc aaaatatgat cgcgatggcc 2940
gcaccagtta ttgccaccta ctacctggac accacagagc agtgggagac tctcggcata 3000
aatcgcagga ctgaagctgt caatcagatc gtgactggtt atgccagca gatggtgtac 3060
aagaaagcag atcattccta tgcagcattt acaaaccgtg catctagttc ttggctaaca 3120
gcatatgtcg taaaagtctt tgccatggct gccaaaatgg tagcaggcat tagtcatgaa 3180
atcattttgtg gaggtgtgag gtggctgatt ctgaacaggc aacaaccaga tggagcgttc 3240
aaagaaaatg cccctgtact ttctggaaca atgcagggag gaattcaagg tgctgaagaa 3300
gaagtatatt taacagcttt cattctgggt gcggtgttgg aatccaaaac aatctgcaat 3360
gactatgtca atagtctaga cagcagcatc aagaaggcca caaattatct actcaaaaag 3420
tatgagaaac tgcaaaggcc ttacactaca gccctcacag cctatgcttt ggctgtgca 3480
gaccaactca atgatgacag ggtactcatg gcagcatcaa caggaaggga tcattgggaa 3540
gaatacaatg ctcacacca caacattgaa ggcacttcct atgccttggt gccctgctg 3600
aaaatgaaga aatttgatca aactggtccc atagtcatg ggctgacaga tcagaatctt 3660
tatggggaaa catatggaca aaccaagca acagttatgg catttcaagc tcttgctgaa 3720
tatgagattc agatgcctac ccataaggac ttaaacttag atattactat tgaactgcca 3780
gatcggagaag tacctataag gtacagaatt aattatgaaa atgctctcct ggctcggaca 3840
gtagagacca aactcaacca agacatcact gtgacagcat caggtgatgg aaaagcaaca 3900
atgaccattt tgacattcta taacgcacag ttgcaggaga aggcaaatgt ttgcaataaa 3960
tttcatctta atgtttctgt tgaaaacatc cacttgaatg caatgggagc caagggagcc 4020
ctcatgctca agatctgcac aaggtatctg ggagaagttg attctacaat gacaataatt 4080
gatatttcta tgctgactgg ttttctccct gatgctgaag accttacaag gctttctaaa 4140
ggagtggaca gatacatctc cagatatgaa gttgacaata atatggctca gaaagtagct 4200
gttatcattt acttaaaaaa ggtctcccac tctgaagatg aatgcctgca cttaagatt 4260
ctcaagcatt ttgaagttgg cttcattcag ccaggatcag tcaaggtgta cagctactac 4320
aatctagatg aaaaatgtac caagttctac catccagata aaggaacagg ctttctcaat 4380
aagatatgta ttggtaacgt ttgccgatgt gcaggagaaa cctgttcctc gctcaaccat 4440
caggaaagga ttgatgttcc attacaaatt gaaaaagcct gcgagacgaa tgtggattat 4500
gtctacaaaa ccaagctgct tcgaatagaa gaacaagatg gtaatgatat ctatgtcatg 4560
gatgttttag aagttattaa acaaggtact gacgaaaatc cacgagcaaa gaccaccag 4620
tacataagtc aaaggaaatg ccaggaggct ctgaatctga aggtgaatga tgattatctg 4680
atctgggggt ccaggagtga cctggtgccc acgaaagata aaatttccta catcattaca 4740
aagaacacat ggattgagag atggccacat gaagacgaat gtcaggaaga agaattccaa 4800
aagttgtgtg atgactttgc tcagtttagc tacacattga ctgagtttgg ctgcctact 4860

```

<210> 24

<211> 1620

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 24

```

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

```

			100					105				110				
Gly	Ile	Tyr	Thr	Pro	Gly	Ser	Pro	Val	Leu	Tyr	Arg	Val	Phe	Ser	Met	
		115					120					125				
Asp	His	Asn	Thr	Ser	Lys	Met	Asn	Lys	Thr	Val	Ile	Val	Glu	Phe	Gln	
		130				135					140					
Thr	Pro	Glu	Gly	Ile	Leu	Val	Ser	Ser	Asn	Ser	Val	Asp	Leu	Asn	Phe	
145					150					155					160	
Phe	Trp	Pro	Tyr	Asn	Leu	Pro	Asp	Leu	Val	Ser	Leu	Gly	Thr	Trp	Arg	
				165					170						175	
Ile	Val	Ala	Lys	Tyr	Glu	His	Ser	Pro	Glu	Asn	Tyr	Thr	Ala	Tyr	Phe	
			180					185							190	
Asp	Val	Arg	Lys	Tyr	Val	Leu	Pro	Ser	Phe	Glu	Val	Arg	Leu	Gln	Pro	
		195					200					205				
Ser	Glu	Lys	Phe	Phe	Tyr	Ile	Asp	Gly	Asn	Glu	Asn	Phe	His	Val	Ser	
		210				215					220					
Ile	Thr	Ala	Arg	Tyr	Leu	Tyr	Gly	Glu	Glu	Val	Glu	Gly	Val	Ala	Phe	
225					230					235					240	
Val	Leu	Phe	Gly	Val	Lys	Ile	Asp	Asp	Ala	Lys	Lys	Ser	Ile	Pro	Asp	
				245					250						255	
Ser	Leu	Thr	Arg	Ile	Pro	Ile	Ile	Asp	Gly	Asp	Gly	Lys	Ala	Thr	Leu	
			260					265							270	
Lys	Arg	Asp	Thr	Phe	Arg	Ser	Arg	Phe	Pro	Asn	Leu	Asn	Glu	Leu	Val	
		275					280								285	
Gly	His	Thr	Leu	Tyr	Ala	Ser	Val	Thr	Val	Met	Thr	Glu	Ser	Gly	Ser	
		290				295					300					
Asp	Met	Val	Val	Thr	Glu	Gln	Ser	Gly	Ile	His	Ile	Val	Ala	Ser	Pro	
305					310					315					320	
Tyr	Gln	Ile	His	Phe	Thr	Lys	Thr	Pro	Lys	Tyr	Phe	Lys	Pro	Gly	Met	
				325						330					335	
Pro	Tyr	Glu	Leu	Thr	Val	Tyr	Val	Thr	Asn	Pro	Asp	Gly	Ser	Pro	Ala	
				340				345							350	
Ala	His	Val	Pro	Val	Val	Ser	Glu	Ala	Phe	His	Ser	Met	Gly	Thr	Thr	
		355					360					365				
Leu	Ser	Asp	Gly	Thr	Ala	Lys	Leu	Ile	Leu	Asn	Ile	Pro	Leu	Asn	Ala	
		370				375					380					
Gln	Ser	Leu	Pro	Ile	Thr	Val	Arg	Thr	Asn	His	Gly	Asp	Leu	Pro	Arg	
385					390					395					400	
Glu	Arg	Gln	Ala	Thr	Lys	Ser	Met	Thr	Ala	Ile	Ala	Tyr	Gln	Thr	Gln	
				405					410						415	
Gly	Gly	Ser	Gly	Asn	Tyr	Leu	His	Val	Ala	Ile	Thr	Ser	Thr	Glu	Ile	
			420					425							430	
Lys	Pro	Gly	Asp	Asn	Leu	Pro	Val	Asn	Phe	Asn	Val	Lys	Gly	Asn	Ala	
		435					440					445				
Asn	Ser	Leu	Lys	Gln	Ile	Lys	Tyr	Phe	Thr	Tyr	Leu	Ile	Leu	Asn	Lys	
		450				455					460					
Gly	Lys	Ile	Phe	Lys	Val	Gly	Arg	Gln	Pro	Arg	Arg	Asp	Gly	Gln	Asn	
465					470					475					480	
Leu	Val	Thr	Met	Asn	Leu	His	Ile	Thr	Pro	Asp	Leu	Ile	Pro	Ser	Phe	
				485					490						495	
Arg	Phe	Val	Ala	Tyr	Tyr	Gln	Val	Gly	Asn	Asn	Glu	Ile	Val	Ala	Asp	
			500					505							510	
Ser	Val	Trp	Val	Asp	Val	Lys	Asp	Thr	Cys	Met	Gly	Thr	Leu	Val	Val	
		515					520					525				
Lys	Gly	Asp	Asn	Leu	Ile	Gln	Met	Pro	Gly	Ala	Ala	Met	Lys	Ile	Lys	
		530				535					540					
Leu	Glu	Gly	Asp	Pro	Gly	Ala	Arg	Val	Gly	Leu	Val	Ala	Val	Asp	Lys	
545					550					555					560	
Ala	Val	Tyr	Val	Leu	Asn	Asp	Lys	Tyr	Lys	Ile	Ser	Gln	Ala	Lys	Ile	
				565					570						575	
Trp	Asp	Thr	Ile	Glu	Lys	Ser	Asp	Phe	Gly	Cys	Thr	Ala	Gly	Ser	Gly	
			580					585							590	
Gln	Asn	Asn	Leu	Gly	Val	Phe	Glu	Asp	Ala	Gly	Leu	Ala	Leu	Thr	Thr	

	595						600					605				
Ser	Thr	Asn	Leu	Asn	Thr	Lys	Gln	Arg	Ser	Ala	Ala	Lys	Cys	Pro	Gln	
	610					615				620						
Pro	Ala	Asn	Arg	Arg	Arg	Arg	Ser	Ser	Val	Leu	Leu	Leu	Asp	Ser	Asn	
625					630					635					640	
Ala	Ser	Lys	Ala	Ala	Glu	Phe	Gln	Asp	Gln	Asp	Leu	Arg	Lys	Cys	Cys	
			645						650					655		
Glu	Asp	Val	Met	His	Glu	Asn	Pro	Met	Gly	Tyr	Thr	Cys	Glu	Lys	Arg	
			660					665					670			
Ala	Lys	Tyr	Ile	Gln	Glu	Gly	Asp	Ala	Cys	Lys	Ala	Ala	Phe	Leu	Glu	
		675					680					685				
Cys	Cys	Arg	Tyr	Ile	Lys	Gly	Val	Arg	Asp	Glu	Asn	Gln	Arg	Glu	Ser	
	690					695					700					
Glu	Leu	Phe	Leu	Ala	Arg	Asp	Asp	Asn	Glu	Asp	Gly	Phe	Ile	Ala	Asp	
705					710					715					720	
Ser	Asp	Ile	Ile	Ser	Arg	Ser	Asp	Phe	Pro	Lys	Ser	Trp	Leu	Trp	Leu	
				725					730					735		
Thr	Lys	Asp	Leu	Thr	Glu	Glu	Pro	Asn	Ser	Gln	Gly	Ile	Ser	Ser	Lys	
			740					745					750			
Thr	Met	Ser	Phe	Tyr	Leu	Arg	Asp	Ser	Ile	Thr	Thr	Trp	Val	Val	Leu	
		755					760					765				
Ala	Val	Ser	Phe	Thr	Pro	Thr	Lys	Gly	Ile	Cys	Val	Ala	Glu	Pro	Tyr	
	770					775					780					
Glu	Ile	Arg	Val	Met	Lys	Val	Phe	Phe	Ile	Asp	Leu	Gln	Met	Pro	Tyr	
785					790					795					800	
Ser	Val	Val	Lys	Asn	Glu	Gln	Val	Glu	Ile	Arg	Ala	Ile	Leu	His	Asn	
				805						810				815		
Tyr	Val	Asn	Glu	Asp	Ile	Tyr	Val	Arg	Val	Glu	Leu	Leu	Tyr	Asn	Pro	
			820					825					830			
Ala	Phe	Cys	Ser	Ala	Ser	Thr	Lys	Gly	Gln	Arg	Tyr	Arg	Gln	Gln	Phe	
		835					840					845				
Pro	Ile	Lys	Ala	Leu	Ser	Ser	Arg	Ala	Val	Pro	Phe	Val	Ile	Val	Pro	
	850					855					860					
Leu	Glu	Gln	Gly	Leu	His	Asp	Val	Glu	Ile	Lys	Ala	Ser	Val	Gln	Glu	
865					870					875					880	
Ala	Leu	Trp	Ser	Asp	Gly	Val	Arg	Lys	Lys	Leu	Lys	Val	Val	Pro	Glu	
				885					890					895		
Gly	Val	Gln	Lys	Ser	Ile	Val	Thr	Ile	Val	Lys	Leu	Asp	Pro	Arg	Ala	
			900					905					910			
Lys	Gly	Val	Gly	Gly	Thr	Gln	Leu	Glu	Val	Ile	Lys	Ala	Arg	Lys	Leu	
		915					920					925				
Asp	Asp	Arg	Val	Pro	Asp	Thr	Glu	Ile	Glu	Thr	Lys	Ile	Ile	Ile	Gln	
	930					935					940					
Gly	Asp	Pro	Val	Ala	Gln	Ile	Ile	Glu	Asn	Ser	Ile	Asp	Gly	Ser	Lys	
945					950					955					960	
Leu	Asn	His	Leu	Ile	Ile	Thr	Pro	Ser	Gly	Cys	Gly	Glu	Gln	Asn	Met	
				965					970					975		
Ile	Arg	Met	Ala	Ala	Pro	Val	Ile	Ala	Thr	Tyr	Tyr	Leu	Asp	Thr	Thr	
			980					985					990			
Glu	Gln	Trp	Glu	Thr	Leu	Gly	Ile	Asn	Arg	Arg	Thr	Glu	Ala	Val	Asn	
		995					1000					1005				
Gln	Ile	Val	Thr	Gly	Tyr	Ala	Gln	Gln	Met	Val	Tyr	Lys	Lys	Ala	Asp	
	1010					1015					1020					
His	Ser	Tyr	Ala	Ala	Phe	Thr	Asn	Arg	Ala	Ser	Ser	Ser	Trp	Leu	Thr	
1025					1030					1035					1040	
Ala	Tyr	Val	Val	Lys	Val	Phe	Ala	Met	Ala	Ala	Lys	Met	Val	Ala	Gly	
				1045					1050					1055		
Ile	Ser	His	Glu	Ile	Ile	Cys	Gly	Gly	Val	Arg	Trp	Leu	Ile	Leu	Asn	
			1060					1065					1070			
Arg	Gln	Gln	Pro	Asp	Gly	Ala	Phe	Lys	Glu	Asn	Ala	Pro	Val	Leu	Ser	
		1075					1080					1085				
Gly	Thr	Met	Gln	Gly	Gly	Ile	Gln	Gly	Ala	Glu	Glu	Glu	Val	Tyr	Leu	

1090	1095	1100
Thr Ala Phe Ile Leu Val Ala Leu Leu Glu Ser Lys Thr Ile Cys Asn		
1105	1110	1115
Asp Tyr Val Asn Ser Leu Asp Ser Ser Ile Lys Lys Ala Thr Asn Tyr		1120
	1125	1130
Leu Leu Lys Lys Tyr Glu Lys Leu Gln Arg Pro Tyr Thr Thr Ala Leu		1135
	1140	1145
Thr Ala Tyr Ala Leu Ala Ala Ala Asp Gln Leu Asn Asp Asp Arg Val		1150
	1155	1160
Leu Met Ala Ala Ser Thr Gly Arg Asp His Trp Glu Glu Tyr Asn Ala		1165
	1170	1175
His Thr His Asn Ile Glu Gly Thr Ser Tyr Ala Leu Leu Ala Leu Leu		1180
1185	1190	1195
Lys Met Lys Lys Phe Asp Gln Thr Gly Pro Ile Val Arg Trp Leu Thr		1200
	1205	1210
Asp Gln Asn Phe Tyr Gly Glu Thr Tyr Gly Gln Thr Gln Ala Thr Val		1215
	1220	1225
Met Ala Phe Gln Ala Leu Ala Glu Tyr Glu Ile Gln Met Pro Thr His		1230
	1235	1240
Lys Asp Leu Asn Leu Asp Ile Thr Ile Glu Leu Pro Asp Arg Glu Val		1245
	1250	1255
Pro Ile Arg Tyr Arg Ile Asn Tyr Glu Asn Ala Leu Leu Ala Arg Thr		1260
1265	1270	1275
Val Glu Thr Lys Leu Asn Gln Asp Ile Thr Val Thr Ala Ser Gly Asp		1280
	1285	1290
Gly Lys Ala Thr Met Thr Ile Leu Thr Phe Tyr Asn Ala Gln Leu Gln		1295
	1300	1305
Glu Lys Ala Asn Val Cys Asn Lys Phe His Leu Asn Val Ser Val Glu		1310
	1315	1320
Asn Ile His Leu Asn Ala Met Gly Ala Lys Gly Ala Leu Met Leu Lys		1325
	1330	1335
Ile Cys Thr Arg Tyr Leu Gly Glu Val Asp Ser Thr Met Thr Ile Ile		1340
1345	1350	1355
Asp Ile Ser Met Leu Thr Gly Phe Leu Pro Asp Ala Glu Asp Leu Thr		1360
	1365	1370
Arg Leu Ser Lys Gly Val Asp Arg Tyr Ile Ser Arg Tyr Glu Val Asp		1375
	1380	1385
Asn Asn Met Ala Gln Lys Val Ala Val Ile Ile Tyr Leu Asn Lys Val		1390
	1395	1400
Ser His Ser Glu Asp Glu Cys Leu His Phe Lys Ile Leu Lys His Phe		1405
	1410	1415
Glu Val Gly Phe Ile Gln Pro Gly Ser Val Lys Val Tyr Ser Tyr Tyr		1420
1425	1430	1435
Asn Leu Asp Glu Lys Cys Thr Lys Phe Tyr His Pro Asp Lys Gly Thr		1440
	1445	1450
Gly Leu Leu Asn Lys Ile Cys Ile Gly Asn Val Cys Arg Cys Ala Gly		1455
	1460	1465
Glu Thr Cys Ser Ser Leu Asn His Gln Glu Arg Ile Asp Val Pro Leu		1470
	1475	1480
Gln Ile Glu Lys Ala Cys Glu Thr Asn Val Asp Tyr Val Tyr Lys Thr		1485
	1490	1495
Lys Leu Leu Arg Ile Glu Glu Gln Asp Gly Asn Asp Ile Tyr Val Met		1500
1505	1510	1515
Asp Val Leu Glu Val Ile Lys Gln Gly Thr Asp Glu Asn Pro Arg Ala		1520
	1525	1530
Lys Thr His Gln Tyr Ile Ser Gln Arg Lys Cys Gln Glu Ala Leu Asn		1535
	1540	1545
Leu Lys Val Asn Asp Asp Tyr Leu Ile Trp Gly Ser Arg Ser Asp Leu		1550
	1555	1560
Leu Pro Thr Lys Asp Lys Ile Ser Tyr Ile Ile Thr Lys Asn Thr Trp		1565
	1570	1575
Ile Glu Arg Trp Pro His Glu Asp Glu Cys Gln Glu Glu Glu Phe Gln		1580



```

agccaacctg cccctcattg taaagaggta aactgtagct caccagcaga tatggatgga 2700
atccagaaag ggctggaacc aaggaaaatg tatcagtatg gagctgttgt aactctggag 2760
tgtgaagatg ggtatatgct ggaaggcagt cccagagacc agtgccaatc ggatcaccaa 2820
tggaaccctc ccctggcggt ttgcagatcc cgttcacttg ctctgtcctt ttgtggtatt 2880
gctgcagggt tgatacttct taccttcttg attgtcatta ccttatacgt gatatcaaaa 2940
cacagagaac gcaattatta tacagataca agccagaaag aagcttttca tttagaagca 3000
cgagaagtat attctgttga tccatacaac ccagccagc 3039

```

<210> 26

<211> 1033

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 26

```

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

```

			340					345				350			
His	Pro	Gln	Ile	Leu	Arg	Gly	Arg	Met	Val	Ser	Gly	Gln	Lys	Asp	Arg
		355					360					365			
Tyr	Thr	Tyr	Asn	Asp	Thr	Val	Ile	Phe	Ala	Cys	Met	Phe	Gly	Phe	Thr
	370					375					380				
Leu	Lys	Gly	Ser	Lys	Gln	Ile	Arg	Cys	Asn	Ala	Gln	Gly	Thr	Trp	Glu
385					390					395					400
Pro	Ser	Ala	Pro	Val	Cys	Glu	Lys	Glu	Cys	Gln	Ala	Pro	Pro	Asn	Ile
				405					410					415	
Leu	Asn	Gly	Gln	Lys	Glu	Asp	Arg	His	Met	Val	Arg	Phe	Asp	Pro	Gly
			420					425					430		
Thr	Ser	Ile	Lys	Tyr	Ser	Cys	Asn	Pro	Gly	Tyr	Val	Leu	Val	Gly	Glu
	435						440					445			
Glu	Ser	Ile	Gln	Cys	Thr	Ser	Glu	Gly	Val	Trp	Thr	Pro	Pro	Val	Pro
	450					455					460				
Gln	Cys	Lys	Val	Ala	Ala	Cys	Glu	Ala	Thr	Gly	Arg	Gln	Leu	Leu	Thr
465					470					475					480
Lys	Pro	Gln	His	Gln	Phe	Val	Arg	Pro	Asp	Val	Asn	Ser	Ser	Cys	Gly
				485					490					495	
Glu	Gly	Tyr	Lys	Leu	Ser	Gly	Ser	Val	Tyr	Gln	Glu	Cys	Gln	Gly	Thr
			500					505					510		
Ile	Pro	Trp	Phe	Met	Glu	Ile	Arg	Leu	Cys	Lys	Glu	Ile	Thr	Cys	Pro
	515						520					525			
Pro	Pro	Pro	Val	Ile	Tyr	Asn	Gly	Ala	His	Thr	Gly	Ser	Ser	Leu	Glu
	530					535					540				
Asp	Phe	Pro	Tyr	Gly	Thr	Thr	Val	Thr	Tyr	Thr	Cys	Asn	Pro	Gly	Pro
545					550					555					560
Glu	Arg	Gly	Val	Glu	Phe	Ser	Leu	Ile	Gly	Glu	Ser	Thr	Ile	Arg	Cys
				565					570					575	
Thr	Ser	Asn	Asp	Gln	Glu	Arg	Gly	Thr	Trp	Ser	Gly	Pro	Ala	Pro	Leu
			580					585					590		
Cys	Lys	Leu	Ser	Leu	Leu	Ala	Val	Gln	Cys	Ser	His	Val	His	Ile	Ala
	595						600					605			
Asn	Gly	Tyr	Lys	Ile	Ser	Gly	Lys	Glu	Ala	Pro	Tyr	Phe	Tyr	Asn	Asp
	610					615						620			
Thr	Val	Thr	Phe	Lys	Cys	Tyr	Ser	Gly	Phe	Thr	Leu	Lys	Gly	Ser	Ser
625					630					635					640
Gln	Ile	Arg	Cys	Lys	Ala	Asp	Asn	Thr	Trp	Asp	Pro	Glu	Ile	Pro	Val
				645					650					655	
Cys	Glu	Lys	Glu	Thr	Cys	Gln	His	Val	Arg	Gln	Ser	Leu	Gln	Glu	Leu
			660					665					670		
Pro	Ala	Gly	Ser	Arg	Val	Glu	Leu	Val	Asn	Thr	Ser	Cys	Gln	Asp	Gly
		675						680					685		
Tyr	Gln	Leu	Thr	Gly	His	Ala	Tyr	Gln	Met	Cys	Gln	Asp	Ala	Glu	Asn
	690					695					700				
Gly	Ile	Trp	Phe	Lys	Lys	Ile	Pro	Leu	Cys	Lys	Val	Ile	His	Cys	His
705					710					715					720
Pro	Pro	Pro	Val	Ile	Val	Asn	Gly	Lys	His	Thr	Gly	Met	Met	Ala	Glu
				725					730					735	
Asn	Phe	Leu	Tyr	Gly	Asn	Glu	Val	Ser	Tyr	Glu	Cys	Asp	Gln	Gly	Phe
			740					745					750		
Tyr	Leu	Leu	Gly	Glu	Lys	Lys	Leu	Gln	Cys	Arg	Ser	Asp	Ser	Lys	Gly
	755						760					765			
His	Gly	Ser	Trp	Ser	Gly	Pro	Ser	Pro	Gln	Cys	Leu	Arg	Ser	Pro	Pro
	770					775					780				
Val	Thr	Arg	Cys	Pro	Asn	Pro	Glu	Val	Lys	His	Gly	Tyr	Lys	Leu	Asn
785					790					795					800
Lys	Thr	His	Ser	Ala	Tyr	Ser	His	Asn	Asp	Ile	Val	Tyr	Val	Asp	Cys
				805					810					815	
Asn	Pro	Gly	Phe	Ile	Met	Asn	Gly	Ser	Arg	Val	Ile	Arg	Cys	His	Thr
			820				825						830		
Asp	Asn	Thr	Trp	Val	Pro	Gly	Val	Pro	Thr	Cys	Met	Lys	Lys	Ala	Phe

	835						840					845				
Ile	Gly	Cys	Pro	Pro	Pro	Pro	Lys	Thr	Pro	Asn	Gly	Asn	His	Thr	Gly	
	850						855				860					
Gly	Asn	Ile	Ala	Arg	Phe	Ser	Pro	Gly	Met	Ser	Ile	Leu	Tyr	Ser	Cys	
865					870					875					880	
Asp	Gln	Gly	Tyr	Leu	Leu	Val	Gly	Glu	Ala	Leu	Leu	Leu	Cys	Thr	His	
				885					890					895		
Glu	Gly	Thr	Trp	Ser	Gln	Pro	Ala	Pro	His	Cys	Lys	Glu	Val	Asn	Cys	
			900					905					910			
Ser	Ser	Pro	Ala	Asp	Met	Asp	Gly	Ile	Gln	Lys	Gly	Leu	Glu	Pro	Arg	
		915					920					925				
Lys	Met	Tyr	Gln	Tyr	Gly	Ala	Val	Val	Thr	Leu	Glu	Cys	Glu	Asp	Gly	
	930					935					940					
Tyr	Met	Leu	Glu	Gly	Ser	Pro	Gln	Ser	Gln	Cys	Gln	Ser	Asp	His	Gln	
945					950					955					960	
Trp	Asn	Pro	Pro	Leu	Ala	Val	Cys	Arg	Ser	Arg	Ser	Leu	Ala	Pro	Val	
				965					970					975		
Leu	Cys	Gly	Ile	Ala	Ala	Gly	Leu	Ile	Leu	Leu	Thr	Phe	Leu	Ile	Val	
			980					985					990			
Ile	Thr	Leu	Tyr	Val	Ile	Ser	Lys	His	Arg	Glu	Arg	Asn	Tyr	Tyr	Thr	
		995					1000					1005				
Asp	Thr	Ser	Gln	Lys	Glu	Ala	Phe	His	Leu	Glu	Ala	Arg	Glu	Val	Tyr	
	1010					1015					1020					
Ser	Val	Asp	Pro	Tyr	Asn	Pro	Ala	Ser								
1025					1030											

&lt;210&gt; 27

&lt;211&gt; 3042

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

<223> 人工序列的描述:/注释=  
合成构建体

&lt;400&gt; 27

```

atctctgtg accctcctcc tgaagtcaaa aatgctcgga aaccctatta ttctcttccc 60
atagttcctg gaactgttct gaggtacact tgttcaccta gctaccgctt cattggagaa 120
aaggctatct tttgtataag tgaaaatcaa gtgcatgcca cctgggataa agctcctcct 180
atatgtgaat ctgtgaataa aaccatttct tgctcagatc ccatagtacc agggggattc 240
atgaataaag gatctaaggc accattcaga catggtgatt ctgtgacatt tacctgtaaa 300
gccaacttca ccatgaaagg aagcaaaact gtctggtgcc aggcaaatga aatgtgggga 360
ccaacagctc tgccagtctg tgagagtgat ttccctctgg agtgcccac acttccaacg 420
attcataatg gacaccacac aggacagcat gttgaccagt ttgttgctgg gttgtctgtg 480
acatacagtt gtgaacctgg ctatttgctc actggaaaa agacaattaa gtgcttatct 540
tcaggagact gggatggtgt catcccgaca tgcaaagagg cccagtgtga acatccagga 600
aagtttccca atgggcaggt aaaggaacct ctgagccttc aggttggcac aactgtgtac 660
ttctcctgta atgaagggtt ccaattaca ggacaaccct ctagtcaagt tgtaattggt 720
gaacagaaag ccaatctggc taagaagcca gtatgtaaag aaattctctg cccaccacct 780
ccacctgttc gtaatggaag tcatacaggc agcttttcag aaaatgtacc atatggaagc 840
acagttacct acacctgtga cccaagccca gagaaaggcg tgagcttcac tcttattgga 900
gagaagacta tcaattgtac tactggtagt cagaagactg ggatctggag tggccctgct 960
ccatattgtg tactttcaac ttctgcagtt ctgtgtttac aaccgaagat caaaagaggg 1020
caaatattat ctattttgaa agatagttat tcatataatg acactgtggc attttcttgt 1080
gaacctggct tcacctgaa gggcaacagg agcattcgat gcaatgctca tggcacatgg 1140
gagccaccgg taccagtgtg tgaaaaagga tgtcaggctc ctccctaaaat tatcaatggg 1200
caaaaagaag atagttactt gctcaacttt gaccctggta catccataag atatagctgt 1260
gacctggct atttactggt gggagaggac actatacatt gcaccctga ggggaagtgg 1320
acaccatta ctcccagtg cacagttgca gagtgtaagc cagtaggacc acatctcttt 1380
aagaggcctc agaatcagtt tattaggaca gctgttaatt cttcttgtga tgaagggttc 1440

```

```

cagttaagtg agagtgctta tcaactgtgt caaggtacaa ttccttggtt tatagaaatc 1500
cgtctttgta aagaaatcac ctgcccacca cctcctgtta tacacaacgg gacacataca 1560
tggagttcct cagaagatgt cccatatgga actgtggtca catacatgtg ctatcctggg 1620
ccagaggaag gcgtaaaatt caaactcatc ggggagcaaa ccatccactg tacaagtgac 1680
agcagaggaa gaggctcctg gagtagcctt gctcctctct gtaaactttc cctcccagct 1740
gtccagtgca cagacgttca tgttgaaaat ggagtcaagc tcaactgacaa taaagcccca 1800
tattttctaca atgatagtgt gatgttcaag tgtgatgatg gatataatctt gagggtgagc 1860
agtcagatcc ggtgtaaagc caataatacc tgggatcctg aaaaaccact ttgtaaaaaa 1920
gaaggatgtg agcctatgag agtacatggc cttccagatg attcacatat aaaactagtg 1980
aaaagaacct gtcaaaatgg gtaccagttg actggatata cttatgagaa gtgtcaaaat 2040
gctgagaatg ggacttggtt taaaaagatt gaagtttgta cagttattct ctgtcaacct 2100
ccaccaaaaa ttgcaaatgg tggtcacaca ggcattgatg caaagcactt cctatatgga 2160
aatgaagtgt cttatgaatg tgatgaaggg ttctatcttt tgggagagaa aagtttgag 2220
tgcgtaaatg attctaagg tcatggctct tggagtggac ctccaccaca atgcttacia 2280
tcttctcctc taactcattg ccccgatcca gaagtcaaac atggttacia actcaataaa 2340
actcattctg ctttttctca taatgacata gtacattttg tctgcaatca aggcttcatc 2400
atgaacggca gccacttgat aagggtgcat actaataaca catgggtacc aggtgtacca 2460
acttgatca gaaaggcttc tttagggtgt cagtctccat ccacaatccc caatgggaat 2520
catactgggtg ggagtatagc tcgatttccc cctggaatgt cagtcatgta cagttgctac 2580
caaggcttcc ttatggctgg agaggcacgt cttatctgta ctcatgaggg tacctggagt 2640
caacctcccc ctttttgcaa agaggtaaac tgtagcttcc ctgaagatac aaatggaatc 2700
cagaagggat ttcaacctgg gaaaacctat cgatttgggg ctactgtgac tctggaatgt 2760
gaggatgggt ataccttggg ggggaagtccc cagagccagt gccaggatga cagccaatgg 2820
aacctcctcct tggctctttg caaataccgt aggtggtcaa ctattcctct tatttgggt 2880
atctctgtgg gctcagcact tatcattttg atgagtgtcg gcttctgtat gatattaaaa 2940
cacagagaaa gcaattatta tacaagaca agacccaaag aaggagctct tcatttagaa 3000
acacgagaag tatattctat tgatccatat aaccagcaa gc 3042

```

<210> 28

<211> 1014

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 28

```

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

```

			180					185				190			
Glu	Ala	Gln	Cys	Glu	His	Pro	Gly	Lys	Phe	Pro	Asn	Gly	Gln	Val	Lys
			195					200				205			
Glu	Pro	Leu	Ser	Leu	Gln	Val	Gly	Thr	Thr	Val	Tyr	Phe	Ser	Cys	Asn
			210					215				220			
Glu	Gly	Tyr	Gln	Leu	Gln	Gly	Gln	Pro	Ser	Ser	Gln	Cys	Val	Ile	Val
225						230					235				240
Glu	Gln	Lys	Ala	Ile	Trp	Thr	Lys	Lys	Pro	Val	Cys	Lys	Glu	Ile	Leu
				245						250					255
Cys	Pro	Pro	Pro	Pro	Pro	Val	Arg	Asn	Gly	Ser	His	Thr	Gly	Ser	Phe
				260				265							270
Ser	Glu	Asn	Val	Pro	Tyr	Gly	Ser	Thr	Val	Thr	Tyr	Thr	Cys	Asp	Pro
				275				280							285
Ser	Pro	Glu	Lys	Gly	Val	Ser	Phe	Thr	Leu	Ile	Gly	Glu	Lys	Thr	Ile
				290				295				300			
Asn	Cys	Thr	Thr	Gly	Ser	Gln	Lys	Thr	Gly	Ile	Trp	Ser	Gly	Pro	Ala
305						310					315				320
Pro	Tyr	Cys	Val	Leu	Ser	Thr	Ser	Ala	Val	Leu	Cys	Leu	Gln	Pro	Lys
				325											335
Ile	Lys	Arg	Gly	Gln	Ile	Leu	Ser	Ile	Leu	Lys	Asp	Ser	Tyr	Ser	Tyr
				340											350
Asn	Asp	Thr	Val	Ala	Phe	Ser	Cys	Glu	Pro	Gly	Phe	Thr	Leu	Lys	Gly
				355				360							365
Asn	Arg	Ser	Ile	Arg	Cys	Asn	Ala	His	Gly	Thr	Trp	Glu	Pro	Pro	Val
				370				375				380			
Pro	Val	Cys	Glu	Lys	Gly	Cys	Gln	Ala	Pro	Pro	Lys	Ile	Ile	Asn	Gly
385						390					395				400
Gln	Lys	Glu	Asp	Ser	Tyr	Leu	Leu	Asn	Phe	Asp	Pro	Gly	Thr	Ser	Ile
				405						410					415
Arg	Tyr	Ser	Cys	Asp	Pro	Gly	Tyr	Leu	Leu	Val	Gly	Glu	Asp	Thr	Ile
				420				425							430
His	Cys	Thr	Pro	Glu	Gly	Lys	Trp	Thr	Pro	Ile	Thr	Pro	Gln	Cys	Thr
				435				440							445
Val	Ala	Glu	Cys	Lys	Pro	Val	Gly	Pro	His	Leu	Phe	Lys	Arg	Pro	Gln
				450				455				460			
Asn	Gln	Phe	Ile	Arg	Thr	Ala	Val	Asn	Ser	Ser	Cys	Asp	Glu	Gly	Phe
465						470					475				480
Gln	Leu	Ser	Glu	Ser	Ala	Tyr	Gln	Leu	Cys	Gln	Gly	Thr	Ile	Pro	Trp
				485							490				495
Phe	Ile	Glu	Ile	Arg	Leu	Cys	Lys	Glu	Ile	Thr	Cys	Pro	Pro	Pro	Pro
				500											510
Val	Ile	His	Asn	Gly	Thr	His	Thr	Trp	Ser	Ser	Ser	Glu	Asp	Val	Pro
				515				520							525
Tyr	Gly	Thr	Val	Val	Thr	Tyr	Met	Cys	Tyr	Pro	Gly	Pro	Glu	Glu	Gly
				530				535				540			
Val	Lys	Phe	Lys	Leu	Ile	Gly	Glu	Gln	Thr	Ile	His	Cys	Thr	Ser	Asp
545						550					555				560
Ser	Arg	Gly	Arg	Gly	Ser	Trp	Ser	Ser	Pro	Ala	Pro	Leu	Cys	Lys	Leu
				565						570					575
Ser	Leu	Pro	Ala	Val	Gln	Cys	Thr	Asp	Val	His	Val	Glu	Asn	Gly	Val
				580				585							590
Lys	Leu	Thr	Asp	Asn	Lys	Ala	Pro	Tyr	Phe	Tyr	Asn	Asp	Ser	Val	Met
				595				600							605
Phe	Lys	Cys	Asp	Asp	Gly	Tyr	Ile	Leu	Ser	Gly	Ser	Ser	Gln	Ile	Arg
				610				615				620			
Cys	Lys	Ala	Asn	Asn	Thr	Trp	Asp	Pro	Glu	Lys	Pro	Leu	Cys	Lys	Lys
625						630					635				640
Glu	Gly	Cys	Glu	Pro	Met	Arg	Val	His	Gly	Leu	Pro	Asp	Asp	Ser	His
				645											655
Ile	Lys	Leu	Val	Lys	Arg	Thr	Cys	Gln	Asn	Gly	Tyr	Gln	Leu	Thr	Gly
				660											670
Tyr	Thr	Tyr	Glu	Lys	Cys	Gln	Asn	Ala	Glu	Asn	Gly	Thr	Trp	Phe	Lys

```

        675                680                685
Lys Ile Glu Val Cys Thr Val Ile Leu Cys Gln Pro Pro Pro Lys Ile
      690                695                700
Ala Asn Gly Gly His Thr Gly Met Met Ala Lys His Phe Leu Tyr Gly
705                710                715                720
Asn Glu Val Ser Tyr Glu Cys Asp Glu Gly Phe Tyr Leu Leu Gly Glu
      725                730                735
Lys Ser Leu Gln Cys Val Asn Asp Ser Lys Gly His Gly Ser Trp Ser
      740                745                750
Gly Pro Pro Pro Gln Cys Leu Gln Ser Ser Pro Leu Thr His Cys Pro
      755                760                765
Asp Pro Glu Val Lys His Gly Tyr Lys Leu Asn Lys Thr His Ser Ala
770                775                780
Phe Ser His Asn Asp Ile Val His Phe Val Cys Asn Gln Gly Phe Ile
785                790                795                800
Met Asn Gly Ser His Leu Ile Arg Cys His Thr Asn Asn Thr Trp Leu
      805                810                815
Pro Gly Val Pro Thr Cys Ile Arg Lys Ala Ser Leu Gly Cys Gln Ser
      820                825                830
Pro Ser Thr Ile Pro Asn Gly Asn His Thr Gly Gly Ser Ile Ala Arg
      835                840                845
Phe Pro Pro Gly Met Ser Val Met Tyr Ser Cys Tyr Gln Gly Phe Leu
850                855                860
Met Ala Gly Glu Ala Arg Leu Ile Cys Thr His Glu Gly Thr Trp Ser
865                870                875                880
Gln Pro Pro Pro Phe Cys Lys Glu Val Asn Cys Ser Phe Pro Glu Asp
      885                890                895
Thr Asn Gly Ile Gln Lys Gly Phe Gln Pro Gly Lys Thr Tyr Arg Phe
900                905                910
Gly Ala Thr Val Thr Leu Glu Cys Glu Asp Gly Tyr Thr Leu Glu Gly
915                920                925
Ser Pro Gln Ser Gln Cys Gln Asp Asp Ser Gln Trp Asn Pro Pro Leu
930                935                940
Ala Leu Cys Lys Tyr Arg Arg Trp Ser Thr Ile Pro Leu Ile Cys Gly
945                950                955                960
Ile Ser Val Gly Ser Ala Leu Ile Ile Leu Met Ser Val Gly Phe Cys
      965                970                975
Met Ile Leu Lys His Arg Glu Ser Asn Tyr Tyr Thr Lys Thr Arg Pro
980                985                990
Lys Glu Gly Ala Leu His Leu Glu Thr Arg Glu Val Tyr Ser Ile Asp
995                1000                1005
Pro Tyr Asn Pro Ala Ser
1010

```

<210> 29

<211> 1033

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:/注释=  
合成构建体

<400> 29

```

Met Gly Ala Ala Gly Leu Leu Gly Val Phe Leu Ala Leu Val Ala Pro
 1                5                10                15
Gly Val Leu Gly Ile Ser Cys Gly Ser Pro Pro Pro Val Leu Asn Gly
      20                25                30
Arg Ile Ser Tyr Tyr Ser Thr Pro Ile Ala Val Gly Thr Val Ile Arg
      35                40                45

```

Tyr Ser Cys Ser Gly Thr Phe Arg Leu Ile Gly Glu Lys Ser Leu Leu  
 50 55 60  
 Cys Ile Thr Lys Asp Lys Val Asp Gly Thr Trp Asp Lys Pro Ala Pro  
 65 70 75 80  
 Lys Cys Glu Tyr Phe Asn Lys Tyr Ser Ser Cys Pro Glu Pro Ile Val  
 85 90 95  
 Pro Gly Gly Tyr Lys Ile Arg Gly Ser Thr Pro Tyr Arg His Gly Asp  
 100 105 110  
 Ser Val Thr Phe Ala Cys Lys Thr Asn Phe Ser Met Asn Gly Asn Lys  
 115 120 125  
 Ser Val Trp Cys Gln Ala Asn Asn Met Trp Gly Pro Thr Arg Leu Pro  
 130 135 140  
 Thr Cys Val Ser Val Phe Pro Leu Glu Cys Pro Ala Leu Pro Met Ile  
 145 150 155 160  
 His Asn Gly His His Thr Ser Glu Asn Val Gly Ser Ile Ala Pro Gly  
 165 170 175  
 Leu Ser Val Thr Tyr Ser Cys Glu Ser Gly Tyr Leu Leu Val Gly Glu  
 180 185 190  
 Lys Ile Ile Asn Cys Leu Ser Ser Gly Lys Trp Ser Ala Val Pro Pro  
 195 200 205  
 Thr Cys Glu Glu Ala Arg Cys Lys Ser Leu Gly Arg Phe Pro Asn Gly  
 210 215 220  
 Lys Val Lys Glu Pro Pro Ile Leu Arg Val Gly Val Thr Ala Asn Phe  
 225 230 235 240  
 Phe Cys Asp Glu Gly Tyr Arg Leu Gln Gly Pro Pro Ser Ser Arg Cys  
 245 250 255  
 Val Ile Ala Gly Gln Gly Val Ala Trp Thr Lys Met Pro Val Cys Glu  
 260 265 270  
 Glu Ile Phe Cys Pro Ser Pro Pro Pro Ile Leu Asn Gly Arg His Ile  
 275 280 285  
 Gly Asn Ser Leu Ala Asn Val Ser Tyr Gly Ser Ile Val Thr Tyr Thr  
 290 295 300  
 Cys Asp Pro Asp Pro Glu Glu Gly Val Asn Phe Ile Leu Ile Gly Glu  
 305 310 315 320  
 Ser Thr Leu Arg Cys Thr Val Asp Ser Gln Lys Thr Gly Thr Trp Ser  
 325 330 335  
 Gly Pro Ala Pro Arg Cys Glu Leu Ser Thr Ser Ala Val Gln Cys Pro  
 340 345 350  
 His Pro Gln Ile Leu Arg Gly Arg Met Val Ser Gly Gln Lys Asp Arg  
 355 360 365  
 Tyr Thr Tyr Asn Asp Thr Val Ile Phe Ala Cys Met Phe Gly Phe Thr  
 370 375 380  
 Leu Lys Gly Ser Lys Gln Ile Arg Cys Asn Ala Gln Gly Thr Trp Glu  
 385 390 395 400  
 Pro Ser Ala Pro Val Cys Glu Lys Glu Cys Gln Ala Pro Pro Asn Ile  
 405 410 415  
 Leu Asn Gly Gln Lys Glu Asp Arg His Met Val Arg Phe Asp Pro Gly  
 420 425 430  
 Thr Ser Ile Lys Tyr Ser Cys Asn Pro Gly Tyr Val Leu Val Gly Glu  
 435 440 445  
 Glu Ser Ile Gln Cys Thr Ser Glu Gly Val Trp Thr Pro Pro Val Pro  
 450 455 460  
 Gln Cys Lys Val Ala Ala Cys Glu Ala Thr Gly Arg Gln Leu Leu Thr  
 465 470 475 480  
 Lys Pro Gln His Gln Phe Val Arg Pro Asp Val Asn Ser Ser Cys Gly  
 485 490 495  
 Glu Gly Tyr Lys Leu Ser Gly Ser Val Tyr Gln Glu Cys Gln Gly Thr  
 500 505 510  
 Ile Pro Trp Phe Met Glu Ile Arg Leu Cys Lys Glu Ile Thr Cys Pro  
 515 520 525  
 Pro Pro Pro Val Ile Tyr Asn Gly Ala His Thr Gly Ser Ser Leu Glu  
 530 535 540

Asp Phe Pro Tyr Gly Thr Thr Val Thr Tyr Thr Cys Asn Pro Gly Pro  
 545 550 555 560  
 Glu Arg Gly Val Glu Phe Ser Leu Ile Gly Glu Ser Thr Ile Arg Cys  
 565 570 575  
 Thr Ser Asn Asp Gln Glu Arg Gly Thr Trp Ser Gly Pro Ala Pro Leu  
 580 585 590  
 Cys Lys Leu Ser Leu Leu Ala Val Gln Cys Ser His Val His Ile Ala  
 595 600 605  
 Asn Gly Tyr Lys Ile Ser Gly Lys Glu Ala Pro Tyr Phe Tyr Asn Asp  
 610 615 620  
 Thr Val Thr Phe Lys Cys Tyr Ser Gly Phe Thr Leu Lys Gly Ser Ser  
 625 630 635 640  
 Gln Ile Arg Cys Lys Ala Asp Asn Thr Trp Asp Pro Glu Ile Pro Val  
 645 650 655  
 Cys Glu Lys Glu Thr Cys Gln His Val Arg Gln Ser Leu Gln Glu Leu  
 660 665 670  
 Pro Ala Gly Ser Arg Val Glu Leu Val Asn Thr Ser Cys Gln Asp Gly  
 675 680 685  
 Tyr Gln Leu Thr Gly His Ala Tyr Gln Met Cys Gln Asp Ala Glu Asn  
 690 695 700  
 Gly Ile Trp Phe Lys Lys Ile Pro Leu Cys Lys Val Ile His Cys His  
 705 710 715 720  
 Pro Pro Pro Val Ile Val Asn Gly Lys His Thr Gly Met Met Ala Glu  
 725 730 735  
 Asn Phe Leu Tyr Gly Asn Glu Val Ser Tyr Glu Cys Asp Gln Gly Phe  
 740 745 750  
 Tyr Leu Leu Gly Glu Lys Lys Leu Gln Cys Arg Ser Asp Ser Lys Gly  
 755 760 765  
 His Gly Ser Trp Ser Gly Pro Ser Pro Gln Cys Leu Arg Ser Pro Pro  
 770 775 780  
 Val Thr Arg Cys Pro Asn Pro Glu Val Lys His Gly Tyr Lys Leu Asn  
 785 790 795 800  
 Lys Thr His Ser Ala Tyr Ser His Asn Asp Ile Val Tyr Val Asp Cys  
 805 810 815  
 Asn Pro Gly Phe Ile Met Asn Gly Ser Arg Val Ile Arg Cys His Thr  
 820 825 830  
 Asp Asn Thr Trp Val Pro Gly Val Pro Thr Cys Met Lys Lys Ala Phe  
 835 840 845  
 Ile Gly Cys Pro Pro Pro Pro Lys Thr Pro Asn Gly Asn His Thr Gly  
 850 855 860  
 Gly Asn Ile Ala Arg Phe Ser Pro Gly Met Ser Ile Leu Tyr Ser Cys  
 865 870 875 880  
 Asp Gln Gly Tyr Leu Leu Val Gly Glu Ala Leu Leu Leu Cys Thr His  
 885 890 895  
 Glu Gly Thr Trp Ser Gln Pro Ala Pro His Cys Lys Glu Val Asn Cys  
 900 905 910  
 Ser Ser Pro Ala Asp Met Asp Gly Ile Gln Lys Gly Leu Glu Pro Arg  
 915 920 925  
 Lys Met Tyr Gln Tyr Gly Ala Val Val Thr Leu Glu Cys Glu Asp Gly  
 930 935 940  
 Tyr Met Leu Glu Gly Ser Pro Gln Ser Gln Cys Gln Ser Asp His Gln  
 945 950 955 960  
 Trp Asn Pro Pro Leu Ala Val Cys Arg Ser Arg Ser Leu Ala Pro Val  
 965 970 975  
 Leu Cys Gly Ile Ala Ala Gly Leu Ile Leu Leu Thr Phe Leu Ile Val  
 980 985 990  
 Ile Thr Leu Tyr Val Ile Ser Lys His Arg Glu Arg Asn Tyr Tyr Thr  
 995 1000 1005  
 Asp Thr Ser Gln Lys Glu Ala Phe His Leu Glu Ala Arg Glu Val Tyr  
 1010 1015 1020  
 Ser Val Asp Pro Tyr Asn Pro Ala Ser  
 1025 1030

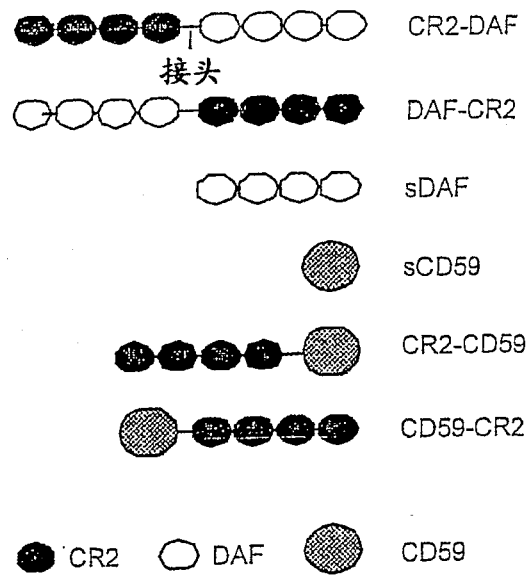


图 1

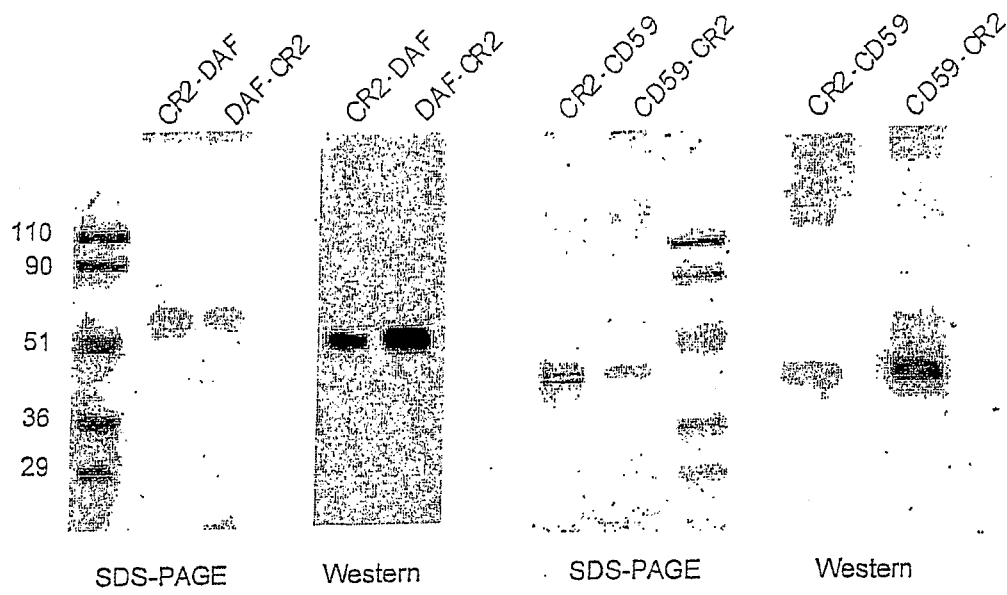


图 2

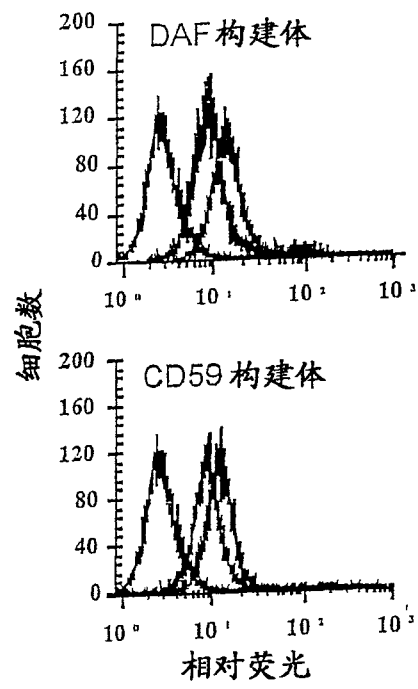


图 3

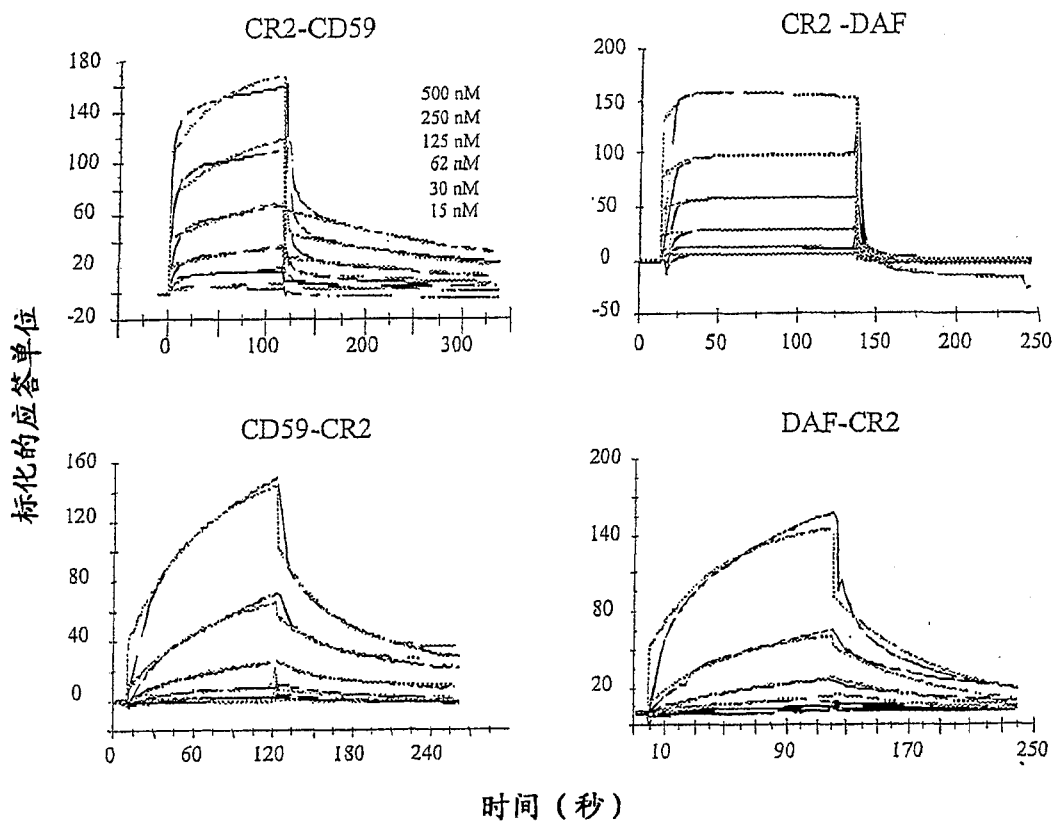


图 4

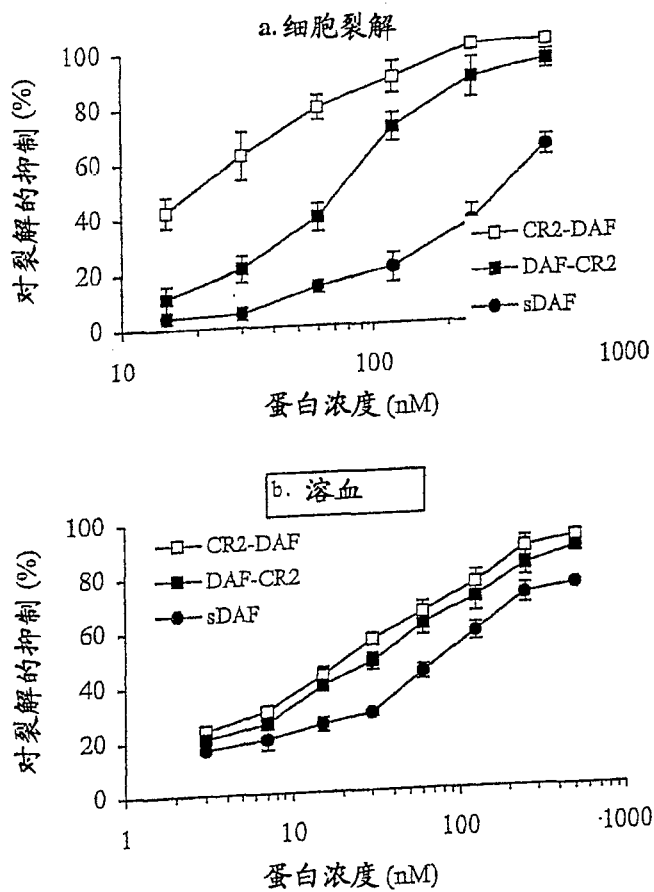


图 5

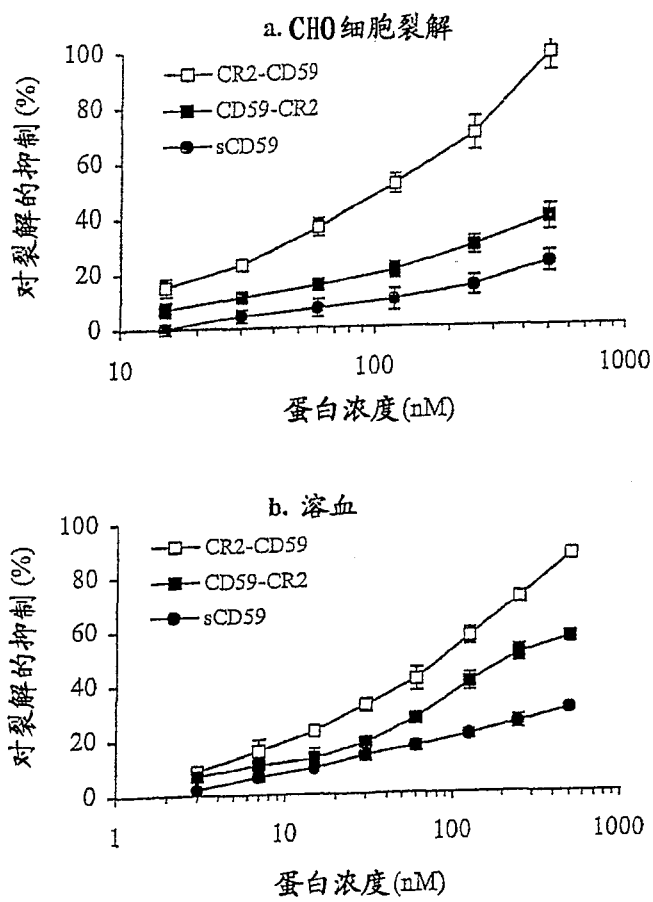


图 6

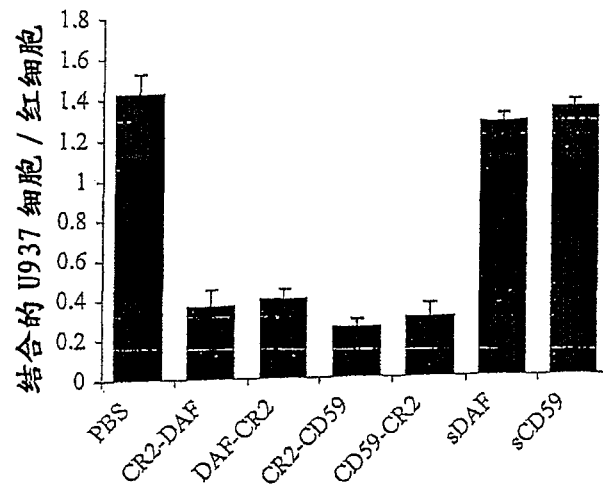


图 7

序列 :CR2DAF 长度 : 1530

```

ATT TCT TGT GGC TCT CCT CCG CCT ATC CTA AAT GGC CCG ATT AGT 45
I S G G S P P P I L N G R I S
TAT TAT TCT ACC CCC ATT GCT GTT GGT ACC GTG ATA AGG TAC AGT 90
Y V S T P I A V G T V I R Y S
TGT TCA GGT ACC TTC CGC CTC ATT GGA GAA AAA AGT CTA TTA TGC 135
C S G T F R L I G E K S L L C
ATA ACT AAA GAC AAA GTG GAT GGA ACC TGG GAT AAA CCT GCT CCT 180
I T K D K V D G T W D K P A P
AAA TGT GAA TAT TTC AAT AAA TAT TCT TCT TGC CCT GAG CCC ATA 225
K C E Y F N K Y S S C P E P I
GTA CCA GGA GGA TAC AAA ATT AGA GGC TCT ACA CCC TAC AGA CAT 270
V P G G Y K I R G S T P Y R H
GGT GAT TCT GTG ACA TTT GCC TGT AAA ACC AAC TTC TCC ATG AAC 315
G D S V T F A C K T N F S M N
GGA AAC AAG TCT GTT TGG TGT CAA GCA AAT AAT ATG TGG GGG CCG 360
G N K S V W C Q A N N M W G P
ACA CGA CTA CCA ACC TGT GTA AGT GTT TTC CCT CTC GAG TGT CCA 405
T R L P T C V S V F P L E C P
GCA CTT CCT ATG ATC CAC AAT GGA CAT CAC ACA AGT GAG AAT GTT 450
A L P M I H N G H H T S E N V
GGC TCC ATT GCT CCA GGA TTG TCT GTG ACT TAC AGC TGT GAA TCT 495
G S I A P G L S V T Y S C E S
GGT TAC TTG CTT GTT GGA GAA AAG ATC ATT AAC TGT TTG TCT TCG 540
G Y L L V G E K I J N C L S S
GGA AAA TGG AGT GCT GTC CCC CCC ACA TGT GAA GAG GCA CGC TGT 585
G K W S A V P P T C E E A R C
AAA TCT CTA GGA CGA TTT CCC AAT GGG AAG GTA AAG GAG CCT CCA 630
K S L G R F P N G K V K E P P
ATT CTC CGG GTT GGT GTA ACT GCA AAC TTT TTC TGT GAT GAA GGG 675
I L R V G V T A N F F C D E G
TAT CGA CTG CAA GGC CCA CCT TCT AGT CGG TGT GTA ATT GCT AGA 720
Y R L Q G P P S S R C V I A G
CAG GGA GTT GCT TGG ACC AAA ATG CCA GTA TGT GGA GGT GGG TCG 765
Q G V A W T K M P V C G G G S
GGT GGC GGC GGA TCC GAC TGT GGC CTT CCC CCA GAT GTA CCT AAT 810
G G G S D C G L P P D V P N
GCC CAG CCA GCT TTG GAA GGC CGT ACA AGT TTT CCC GAG GAT ACT 855
A Q P A L E G R T S F P E D T
GTA ATA ACG TAC AAA TGT GAA GAA AGC TTT GTG AAA ATT CCT GGC 900
V I T Y K C E E S F V K I P G
GAG AAG GAC TCA GTG ATC TGC CTT AAG GGC AGT CAA TGG TCA GAT 945
E K D S V I C L K G S Q W S D
ATT GAA GAG TTC TGC AAT CGT AGC TGC GAG GTG CCA ACA AGG CTA 990
I E E F C N R S C E V P T R L
AAT TCT GCA TCC CTC AAA CAG CCT TAT ATC ACT CAG AAT TAT TTT 1035
N S A S L K Q P Y I T Q N Y F
CCA GTC GGT ACT GTT GTG GAA TAT GAG TGC CGT CCA GGT TAC AGA 1080
P V G T V V E Y E G R P G Y R
AGA GAA CCT TCT CTA TCA CCA AAA CTA ACT TGC CTT CAG AAT TTA 1125
R E P S L S P K L T C L Q N L
AAA TGG TCC ACA GCA GTC GAA TTT TGT AAA AAG AAA TCA TGC CCT 1170
K W S T A V E F C K K K S C P
AAT CCG GGA GAA ATA CGA AAT GGT CAG ATT GAT GTA CCA GGT GGC 1215
N P G E I R N G Q I D V P G G
ATA TTA TTT GGT GCA ACC ATC TCC TTC TCA TGT AAC ACA GGG TAC 1260
I L F G A T I S F S C N T G Y
AAA TTA TTT GGC TCG ACT TCT AGT TTT TGT CTT ATT TCA GGC AGC 1305
K L F G S T S S F C L I S G S
TCT GTC CAG TGG AGT GAC CCG TTG CCA GAG TGC AGA GAA ATT TAT 1350
S V Q W S D P L P E C R E I Y
TGT CCA GCA CCA CCA CAA ATT GAC AAT GGA ATA ATT CAA GGG GAA 1395
C P A P P Q I D N G I I Q G E
CGT GAC CAT TAT GGA TAT AGA CAG TCT GTA ACG TAT GCA TGT AAT 1440
R D H Y G Y R Q S V T Y A C N
AAA GGA TTC ACC ATG ATT GGA GAG CAC TCT ATT TAT TGT ACT GTG 1485
K G F T M I G E H S I Y C T V
AAT AAT GAT GAA GGA GAG TGG AGT GGC CCA CCA CCT GAA TGC AGA 1530
N N D E G E W S G P P P E C R

```

图 8

序列: CR2CD59                      长度: 1002

ATT TCT TGT GGC TCT CCT CCG CCT ATC CTA AAT GGC CGG ATT AGT 45  
 I S C G S P P P I L N G R I S  
 TAT TAT TCT ACC CCC ATT GCT GTT GGT ACC GTG ATA AGG TAC AGT 90  
 Y Y S T P I A V G T V I R Y S  
 TGT TCA GGT ACC TTC CGC CTC ATT GGA GAA AAA AGT CTA TTA TGC 135  
 C S G T F R L I G E K S L L C  
 ATA ACT AAA GAC AAA GTG GAT GGA ACC TGG GAT AAA CCT GCT CCT 180  
 I T K D K V D G T W D K P A P  
 AAA TGT GAA TAT TTC AAT AAA TAT TCT TCT TGC CCT GAG CCC ATA 225  
 K C E Y F N K Y S S C P E P I  
 GTA CCA GGA GGA TAC AAA ATT AGA GGC TCT ACA CCC TAC AGA CAT 270  
 V P G G Y K I R G S T P Y R H  
 GGT GAT TCT GTG ACA TTT GCC TGT AAA ACC AAC TTC TCC ATG AAC 315  
 G D S V T F A C K T N F S M N  
 GGA AAC AAG TCT GTT TGG TGT CAA GCA AAT AAT ATG TGG GGG CCG 360  
 G N K S V W C Q A N N M W G P  
 ACA GGA CTA CCA ACC TGT GTA AGT GTT TTC CCT CTC GAG TGT CCA 405  
 T R L P T C V S V F P L E C P  
 GCA CTT CCT ATG ATC CAC AAT GGA CAT CAC ACA AGT GAG AAT GTT 450  
 A L P M J H N G H H T S E N V  
 GGC TCC ATT GCT CCA GGA TTG TCT GTG ACT TAC AGC TGT GAA TCT 495  
 G S I A P G L S V T Y S C E S  
 GGT TAC TTG CTT GTT GGA GAA AAG ATC ATT AAC TGT TTG TCT TCG 540  
 G Y L L V G E K J I N C L S S  
 GGA AAA TGG AGT GCT GTC CCC CCC ACA TGT GAA GAG GCA CGC TGT 585  
 G K W S A V P P T C E E A R C  
 AAA TCT CTA GGA CCA TTT CCC AAT GGG AAG GTA AAG GAG CCT CCA 630  
 K S L G R F P N G K V K E P P  
 ATT CTC CGG GTC GGT GTA ACT GCA AAC TTT TTC TGT GAT GAA GGG 675  
 I L R V G V T A N F F C D E G  
 TAT CGA CTG CAA GGC CCA CCT TGT AGT CGG TGT GTA ATT GCT GGA 720  
 Y R L Q G P P S S R C V I A G  
 CAG GGA GTT GCT TGG ACC AAA ATG CCA GTA TGT TCA GGA GGA GGA 765  
 Q G V A W T K M P V C S G G G  
 GGT TCC CTG CAG TGC TAC AAC TGT CCT AAG CCA ACT GCT GAC TGC 810  
 G S L Q C Y N C P N P T A D C  
 AAA ACA GCC GTC AAT TGT TCA TCT GAT TTT GAT GCG TGT CTC ATT 855  
 K T A V N C S S D F D A C L I  
 ACC AAA GCT GGG TTA CAA GTG TAT AAC AAG TGT TGG AAG TTT GAG 900  
 T K A G L Q V Y N K C W K F E  
 CAT TGC AAT TTC AAC GAC GTC ACA ACC CGC TTG AGG GAA AAT GAG 945  
 H C N F N D V T T R L R E N E  
 CTA ACG TAC TAC TGC TGC AAG AAG GAC CTG TGT AAC TTT AAC GAA 990  
 L T Y Y C C K K D L C N F N E  
 CAG CTT GAA AAT 1035  
 Q L E N

图 9

序列:DAFCR2 长度:1554

```

GAC TGT GGC CTT CCC CCA GAT GTA CCT AAT GCC CAG CCA GCT TTG 45
D C G L P P D V P N A Q P A T L AAA 90
GAA G GGT ACA AGT TTT CCC GAG GAT ACT GTA ATA T ACG TAC AAA 90
E R T S F P E D T V I T V Y K
TGT G GAA S T AGC TTT GTG AAA ATT CCT GGC GAG AAG GAC S TCA V GTG 135
C E E S F V K I P G E E K K D S V
ATC TGC CTT AAG GGC S AGT CAA TGG TCA GAT ATT GAA GAG TTC TGC 180
I L K G S Q W S D I E T GCA F C
AAT CGT AGC TGC GAG GTG CCA ACA AGG CTA AAT TCT GCA TCC CTC 225
N R S C T E ATC ACT Q CAG AAT TAT TTT CCA GTC GGT ACT GTT 270
AAA CAG P CCT TAT I T T CCA GGT TAC AGA AGA GAA CCT TCT CTA 315
K Q CCA AAA CTA ACT TGC CTT CAG AAT TTA AAA TGG TCC ACA GCA 360
S P K L T C AAG AAA TCA TGC CCT AAT CCG GGA GAA ATA 405
GTC GAA F TTT TGT AAA K K G A S C P N P T GGA GAA ATA 405
V CGA AAT GGT CAG ATT D V G A CCA GGT GGC ATA TTA TTT GGT GCA 450
R N G Q T TCA TGT AAC ACA GGG TAC AAA L TTA TTT GGC TCG 495
ACC ATC TCC TTC TCA CT N T TCA GGC AGC TCT GTC CAG TGG AGT 540
T TCT S AGT TTT TGT CTT ATT S CAA G S S V CCA GCA CCA CCA 585
D CCG TTG CCA E GAG C TGA AGA GAA ATT Y C P A P P P
CAA ATT GAC AAT GGA ATA ATT CAA GGG GAA R GGT GAC CAT TAT GGA 630
Q I D N G I ACG TAT GCA TGT AAT AAA GGA TTC ACC ATG 675
Y R Q S V T Y A C N K G F T M
ATT GGA GAG CAC TCT ATT TAT TGT ACT GTG AAT AAT GAT GAA GGA 720
G E H S I C T V N N N D E G
GAG TG AGT GGC CCA CCA CCT GAA TGC AGA TCC TCT GGT GGC GGT 765
S G P P E C R S S G G G
GGC TCG GGC GGA GGT GGC TCG GGT GGC GGC GGA TCC ATT TCT TGT 810
G S G G G G G S I S C
GGC TCT CCT CCG CCT ATC CTA AAT GGC CCG ATT AGT TAT TAT TCT 855
P P P P N G R S Y Y Y Y Y
ACC CCC ATT GCT GTT GGT ACC GTG ATA AGG TAC AGT TGT TCA GGT 900
T S A T V I R Y S C S K
ACC TTC CCG CTC ATT GGA GAA AAA AGT CTA TTA TGC ATA ACT AAA 945
T F R L I G E K S L L C J T K
GAC AAA GTG GAT GGA ACC TGG GAT AAA P A P K C C TGT GAA 990
D K V D G T W D K P A P K C C
TAT TTC AAT AAA TAT TCT TCT TGC CCT GAG CCC ATA GTA CCA GGA 1035
Y F N K Y S S T C E Y R H G
GGA TAC AAA ATT AGA GGC TCT ACA CCC Y TAC AGA CAT GGT GAT TCT 1080
GTG ACA TTT GCC TGT AAA ACC AAC TTC TCC ATG AAC GGA AAC AAG 1125
V T F A C K T N F S M N K
TCT GTT TGG TGT CAA GCA AAT AAT ATG TGG GGC CCG ACA CGA CTA 1170
S V W C Q A N N M W G P T T R L
CCA ACC TGT GTA AGT GTT TTC CCT CTC GAG TGT CCA GCA CTT CCT 1215
P T C V S V F P L E C P A L P
ATG ATC CAC AAT GGA CAT CAC ACA AGT GAG AAT GTT GGC TCC ATT 1260
I H N G H H T S E N V G S I
GCT CCA GGA TTG TCT GTG ACT TAC AGC TGT GAA TCT GGT TAC TTG 1305
A P G I S V T I S C E S G Y L
CTT GTT GGA GAA AAG ATC ATT AAC TGT TTG TCT TCG GGA AAA TGG 1350
L V G E I L N C L S S G K W
AGT GCT GTC CCC CCC ACA TGT GAA GAG CCA CGC TGT AAA TCT CTA 1395
S A V P P C T GAG CCT CCA ATT CTC CGG 1440
GGA CGA TTT CCC AAT GGG AAG GTA AAG GAG CCT CCA ATT CTC CGG 1440
R F P N N G K V K E P P I L R
GTT GGT GTA ACT GCA AAC TTT TTC TGT GAT GAA GGG TAT CGA CTG 1485
V G V T A N F C D E G Y R R G
CAA GGC CCA CCT TCT AGT CGG TGT GTA ATT GCT GGA CAG GGA GTT 1530
P P S S R C V I A G Q G
GCT TGG ACC AAA ATG CCA GTA TGT 1554
A W T K M P V C

```

图 10

序列: CD59CR2                      长度: 1035

```

CTG CAG TGC TAC AAC TGT CCT AAC CCA ACT GCT GAC TGC AAA ACA 45
L Q C Y N C P N P T A D C K T
GCC GTC AAT TGT TCA TCT GAT TTT GAT GCG TGT CTC ATT ACC AAA 90
A V N C S S D F D A C L I T K.
GCT GGG TTA CAA GTG TAT AAC AAG TGT TGG AAG TTT GAG CAT TGC 135
A G L Q V Y N K C W K F E H C
AAT TTC AAC GAC GTC ACA ACC CGC TTG AGG GAA AAT GAG CTA ACG 180
N F N D V T T R L R E N E L T
TAC TAC TGC TGC AAG AAG GAC CTG TGT AAC TTT AAG GAA CAG CTT 225
Y Y C C K K D L C N F N E Q L
GAA AAT TCC TCT GGT GGC GGT GGC TCC GGC GGA GGT GGG TCC GGT 270
E N S S G G G G S G G G G S G
GGC GGC GGA TCC ATT TCT TGT GGC TCT CCT CCG CCT ATC CTA AAT 315
G G G S I S C G S P P P I L N
GCC CGG ATT AGT TAT TAT TCT ACC CCC AIT GCT GTT GGT ACC GTG 360
G R I S Y Y S T P I A V G T V
ATA AGG TAC AGT TGT TCA GGT ACC TTC CGC CTC ATT GGA GAA AAA 405
I R Y S C S G T F R L I G E K
AGT CTA TTA TGC ATA ACT AAA GAC AAA GTG GAT GGA ACC TGG GAT 450
S L L C I T K D K V D G T W D
AAA CCT GCT CCT AAA TGT GAA TAT TTC AAT AAA TAT TCT TCT TGC 495
K P A P K C E Y F N K Y S S C
CCT GAG CCC ATA GTA CCA GGA GGA TAC AAA ATT AGA GGC TCT ACA 540
P E P I V P G G Y K I R G S T
CCC TAC AGA CAT GGT GAT TCT GTG ACA TTT GCC TGT AAA ACC AAC 585
P Y R H G D S V T F A C K T N
TTC TCC ATG AAC GGA AAC AAG TCT GTT TGG TGT CAA GCA AAT AAT 630
F S M N G N K S V W C Q A N N
ATG TGG GGG CCG ACA CGA CTA CCA ACC TGT GTA AGT GTT TTC CCT 675
M W G P T R L P T C V S V F P
CTC GAG TGT CCA GCA CIT CCT ATG ATC CAC AAT GGA CAT CAC ACA 720
L E C P A L P M I H N G H H T
AGT GAG AAT GTT GGC TCC ATT GCT CCA GGA TTG TGT GTG ACT TAC 765
S E N V G S I A P G L S V T Y
AGC TGT GAA TCT GGT TAC TTG CTT GTT GGA GAA AAG ATC ATT AAC 810
S C E S G Y L L V G E K I I N
TGT TTG TCT TCG GGA AAA TGG AGT GCT GTC CCC CCC ACA TGT GAA 855
C L S S G K W S A V P P T C E
GAG GCA CGC TGT AAA TCT CTA GGA CGA TTT CCC AAT GGG AAG GTA 900
E A R C K S L G R F P N G K V
AAG GAG CCT CCA ATT CTC CGG GTT GGT GTA ACT GCA AAC TTT TTC 945
K E P P J L R V G V T A N F F
TGT GAT GAA GGG TAT CGA CTG CAA GGC CCA CCT TCT AGT CGG TGT 990
C D E G Y R L Q G P P S S R C.
1035

```

图 11

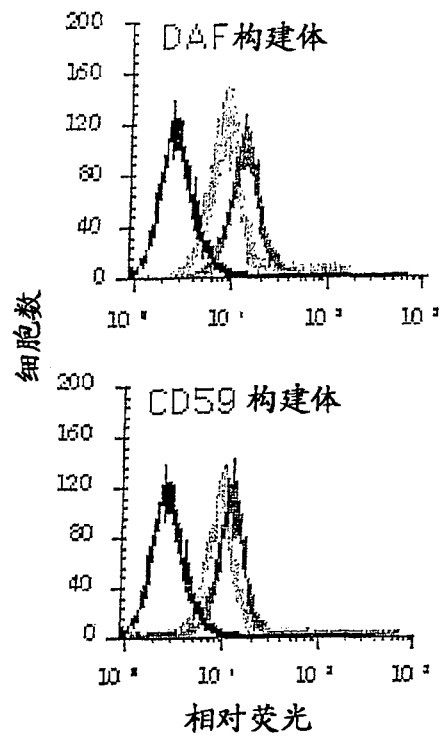


图 12

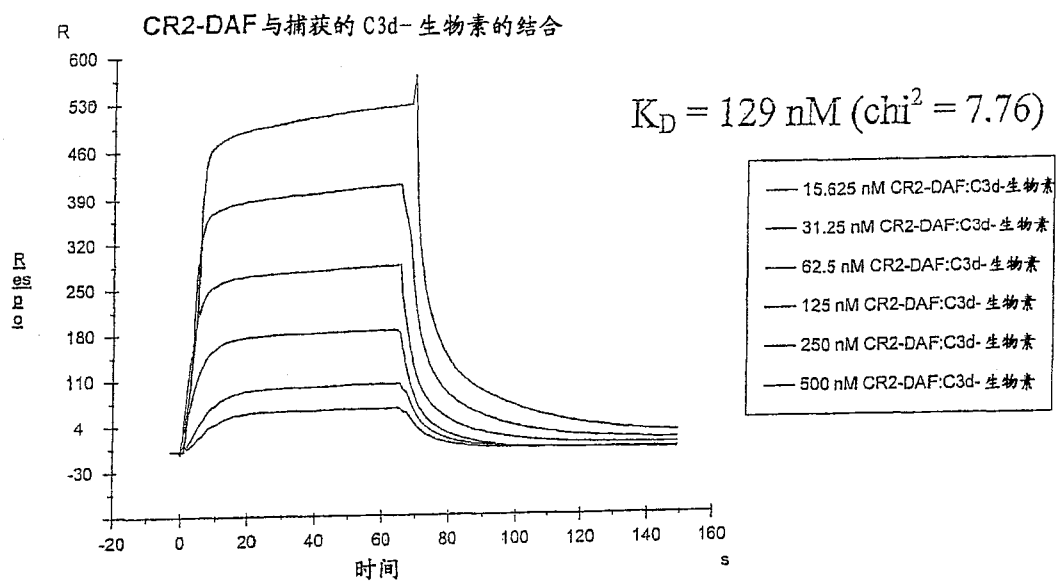


图 13

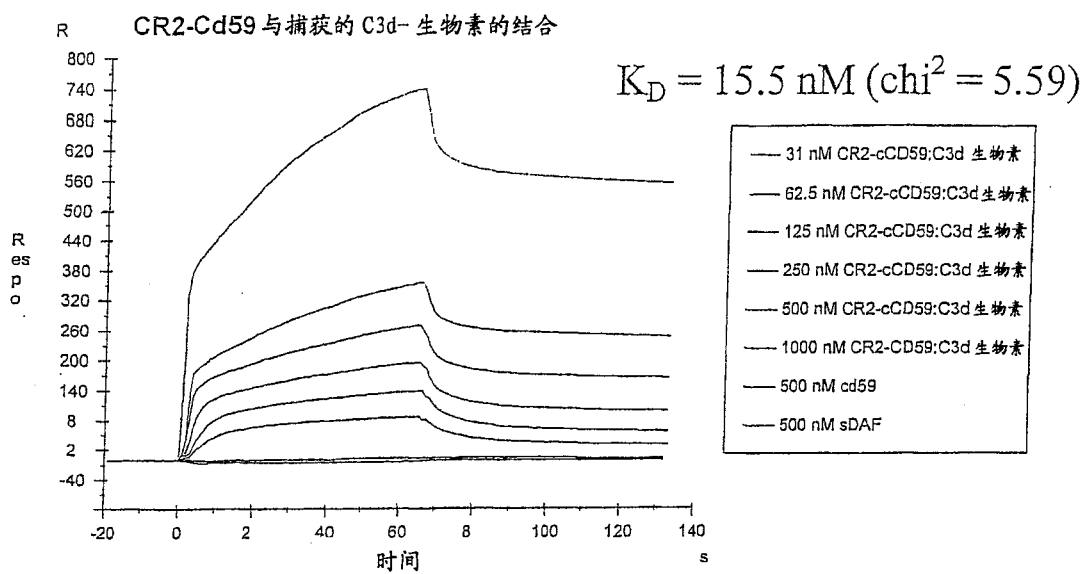


图 14

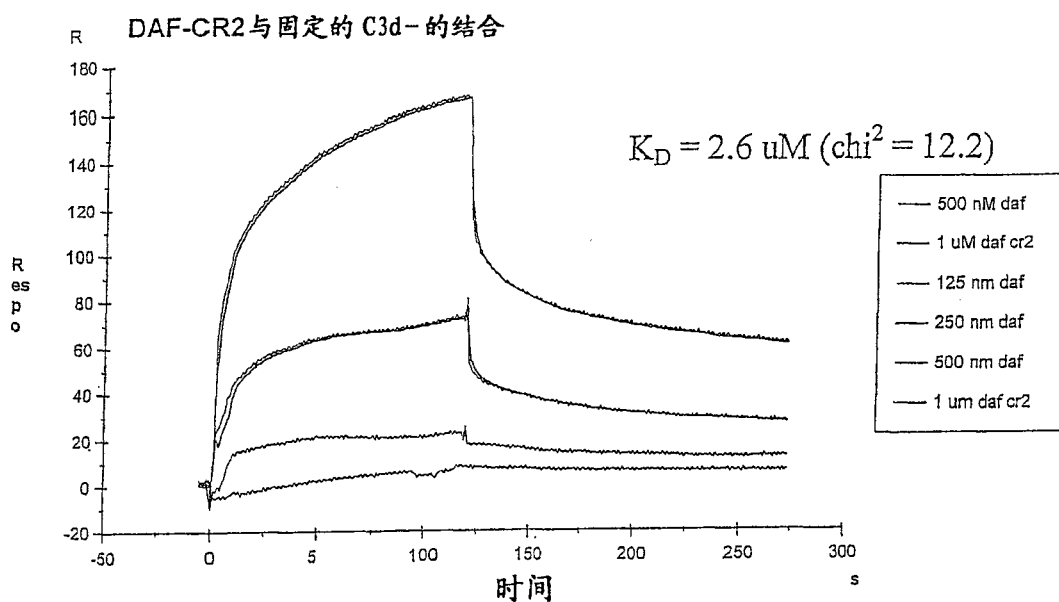


图 15

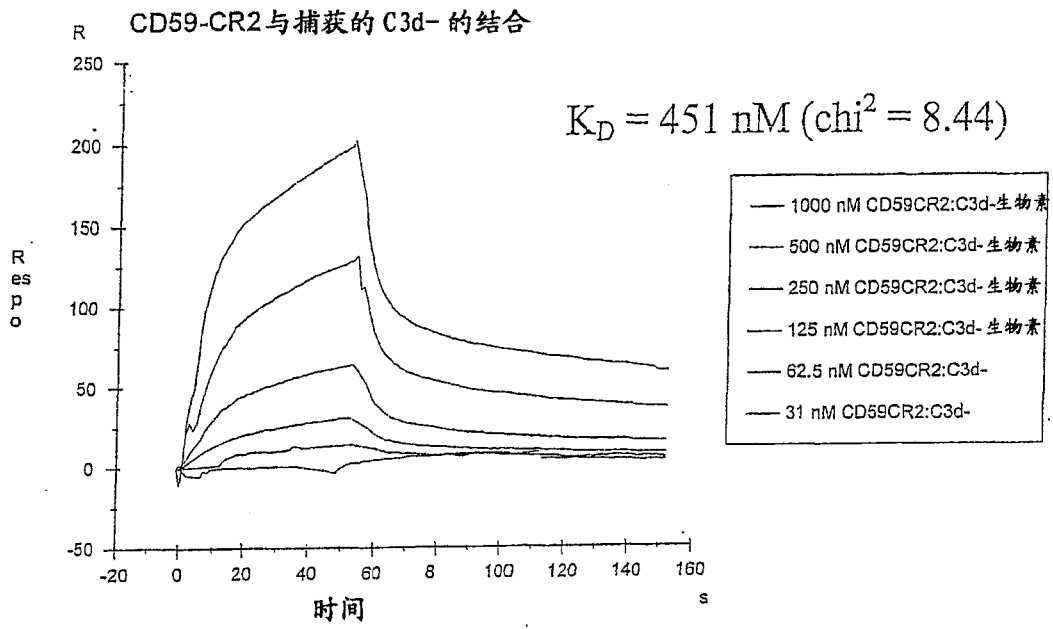


图 16

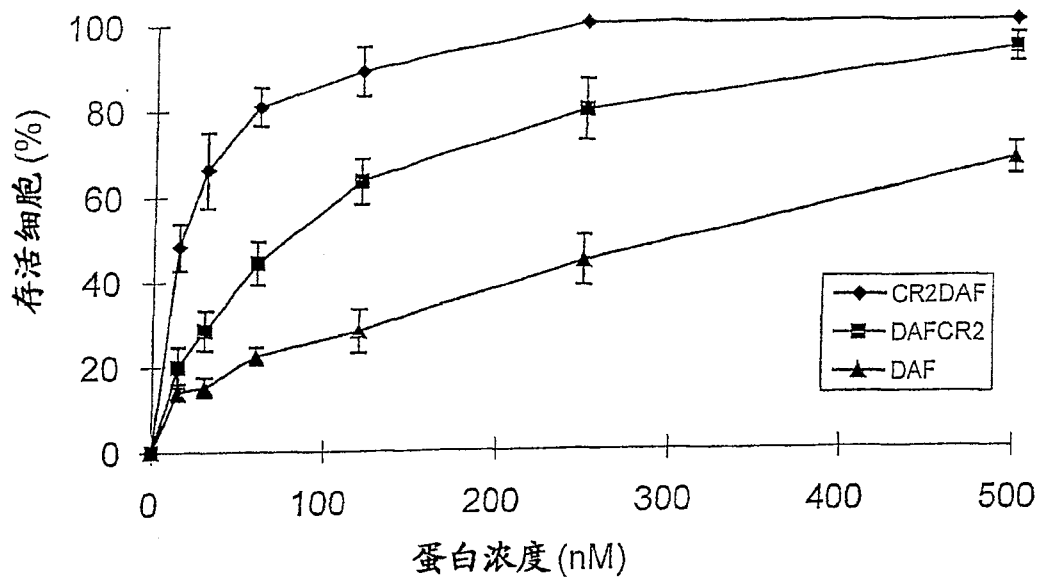


图 17

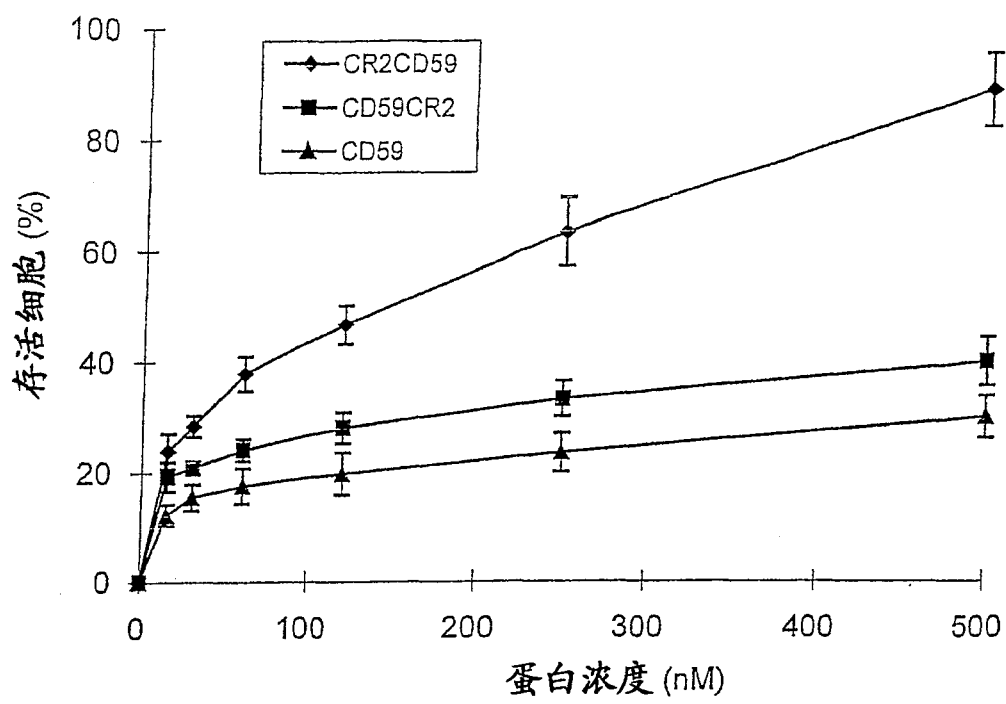


图 18

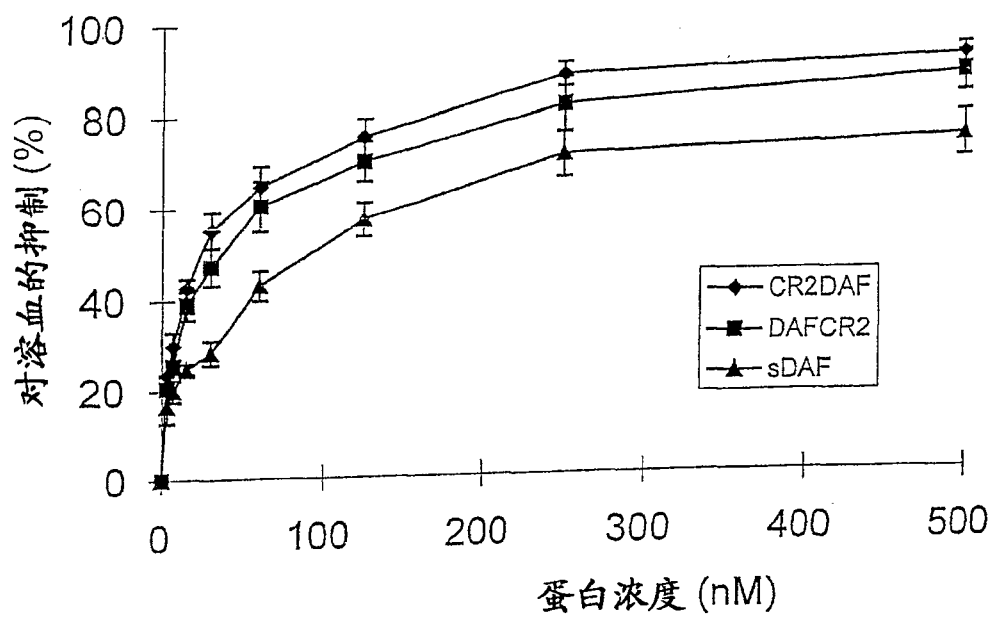


图 19

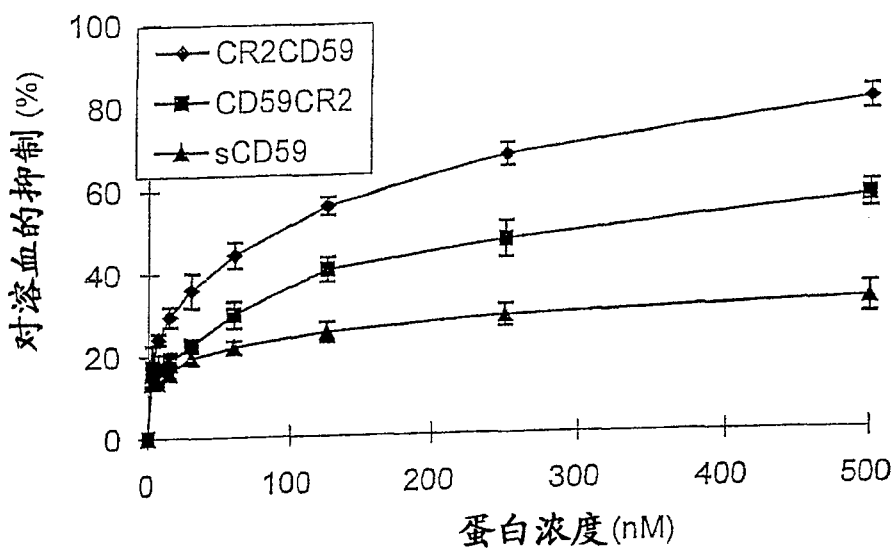


图 20

序列：CR2Fc 长度：1530

```

ATG GGC GCC GCG GGC CTG CTC GGG GTT TTC TTG GCT CTC GTC GCA 45
M G A A G L L G V F L A L V A
CCG GGG GTC CTC GGG ATT TCT TGT GGC TCT CCT CCG CCT ATC CTA 90
P G V L G I S C G S P P P I L
AAT GGC CGG ATT AGT TAT TAT TCT ACC CCC ATT GCT GTT GGT ACC 135
N G R I S Y Y S T P I A V G T
GTG ATA AGG TAC AGT TGT TCA GGT ACC TTC CGC CTC ATT GGA GAA 180
V I R Y S C S G T F R L I G E
AAA AGT CTA TTA TGC ATA ACT AAA GAC AAA GTG GAT GGA ACC TGG 225
K S L L C I T K D K V D G T W
GAT AAA CCT GCT CCT AAA TGT GAA TAT TTC AAT AAA TAT TCT TCT 270
D K P A P K C E Y F N K Y S S
TGC CCT GAG CCC ATA GTA CCA GGA GGA TAC AAA ATT AGA GGC TCT 315
C P E P I V P G G Y K I R G S
ACA CCC TAC AGA CAT GGT GAT TCT GTG ACA TTT GCC TGT AAA ACC 360
T P Y R H G D S V T F A C K T
AAC TTC TCC ATG AAC GGA AAC AAG TCT GTT TGG TGT CAA GCA AAT 405
N F S M N G N K S V W C Q A N
AAT ATG TGG GGC CCG ACA CGA CTA CCA ACC TGT GTA AGT GTT TTC 450
N M W G P T R L P T C V S V F
CCT CTC GAG TGT CCA GCA CTT CCT ATG ATC CAC AAT GGA CAT CAC 495
P L E C P A L P M I H N G H H
ACA AGT GAG AAT GTT GGC TCC ATT GCT CCA GGA TTG TCT GTG ACT 540
T S E N V G S I A P G L S V T
TAC AGC TGT GAA TCT GGT TAC TTG CTT GTT GGA GAA AAG ATC ATT 585
Y S G E S G Y L L V G E K I I
AAC TGT TTG TCT TCG GGA AAA TGG AGT GCT GTC CCC CCA ACA TTT 630
N C L S S G K W S A V P P T C
GAA GAG GCA CGC TGT AAA TCT CTA GGA CGA TTT CCC AAT GGG AAG 675
E E A R C K S L G R F P N G K
GTA AAG GAG CCT CCA ATT CTC CGG GTT GGT GTA ACT GCA AAC TTT 720
V K E P P I L R V G V T A N F
TTC TGT GAT GAA GGG TAT CGA CTG CAA GGC CCA CCT TCT AGT CGG 765
F C D E G Y R L Q G P P S S R
TGT GTA ATT GCT GGA CAG GGA GTT GCT TGG ACC AAA ATG CCA GTA 810
C V I A G Q G V A W T K M P V
TGT GAA GAA ATT TTT TGC CCA CTG CGG CCG CAG TCT AGA GAC AAA 855
C E E I F C P L R P O S R D K
ACT CAC ACA TGC CCA CCG TGC CCA GCA CCT GAA CTC CTG GGG GGA 900
T H T C P P C P A P E L L G G
CCG TCA GTC TTC CTC TTC CCC CCA AAA CCC AAG GAC ACC CTC ATG 945
P S V F L F P P K P K D T L M
ATC TCC CGG ACC CCT GAG GTC ACA TGC GTG GTG GTG GAC GTG AGC 990
I S R T P E V T C V V V D V S
CAC GAA GAC CCT GAG GTG AAG TTC AAC TGG TAC GTG GAC GGC GTG 1035
H E D P E V K F N W Y V D G V
GAG GTG CAT AAT GCC AAG ACA AAG CCG CGG GAG GAG CAG TAC AAC 1080
E V H N A K T K P R E E Q Y N
AGC AGC TAC CGT GTG GTC AGC GTC CTC ACC GTC CTG CAC CAG GAC 1125
S T Y R V V S V L T V L H Q D
TGG CTG AAT GGC AAG GAG TAC AAG TGC AAG GTC TCC AAC AAA GCG 1170
W L N G K E Y K C K V S N K A
CTC CCA GTC CCC ATC GAG AAA ACC ATC TCC AAA GCC AAA GGG CAG 1215
L P V P I E K T I S K A K G Q
CCC CGA GAA CCA CAG GTG TAC ACC CTG CCC CCA TCC CCG GAG GAG 1260
P R E E P Q V Y T L P P S R E E
ATG ACC AAG AAC CAG GTG AGC CTG ACC TGC CTG GTC AAA GGC TTC 1305
M T K N Q V S L T C L V K G F
TAT CCC AGC GAC ATC GCC GTG GAG TGG GAG AGC AAT GGG CAG CCG 1350
Y P S D I A V E W E S N G Q P
GAG AAC AAC TAC AAG ACC ACG CCT CCC GTG CTG GAC TCC GAC GGC 1395
E N N Y K T T P P V L D S D G
TCC TTC TTC CTC TAT AGC AAG CTC ACC GTG GAC AAG AGC AGG TGG 1440
S F F L Y S K L T V D K S R W
CAG CAG GGG AAC GTC TTC TCA TGC TCC GTG ATG CAT GAG GCT CTG 1485
Q Q G N V F S C S V M H E A L
CAC AAC CAC TAC ACG CAG AAG AGC CTC TCC CTG TCC CCG GGT AAA 1530
H N H Y T Q K S L S L S P G K

```

图 21

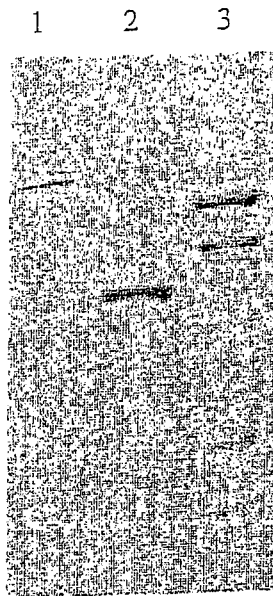


图 22

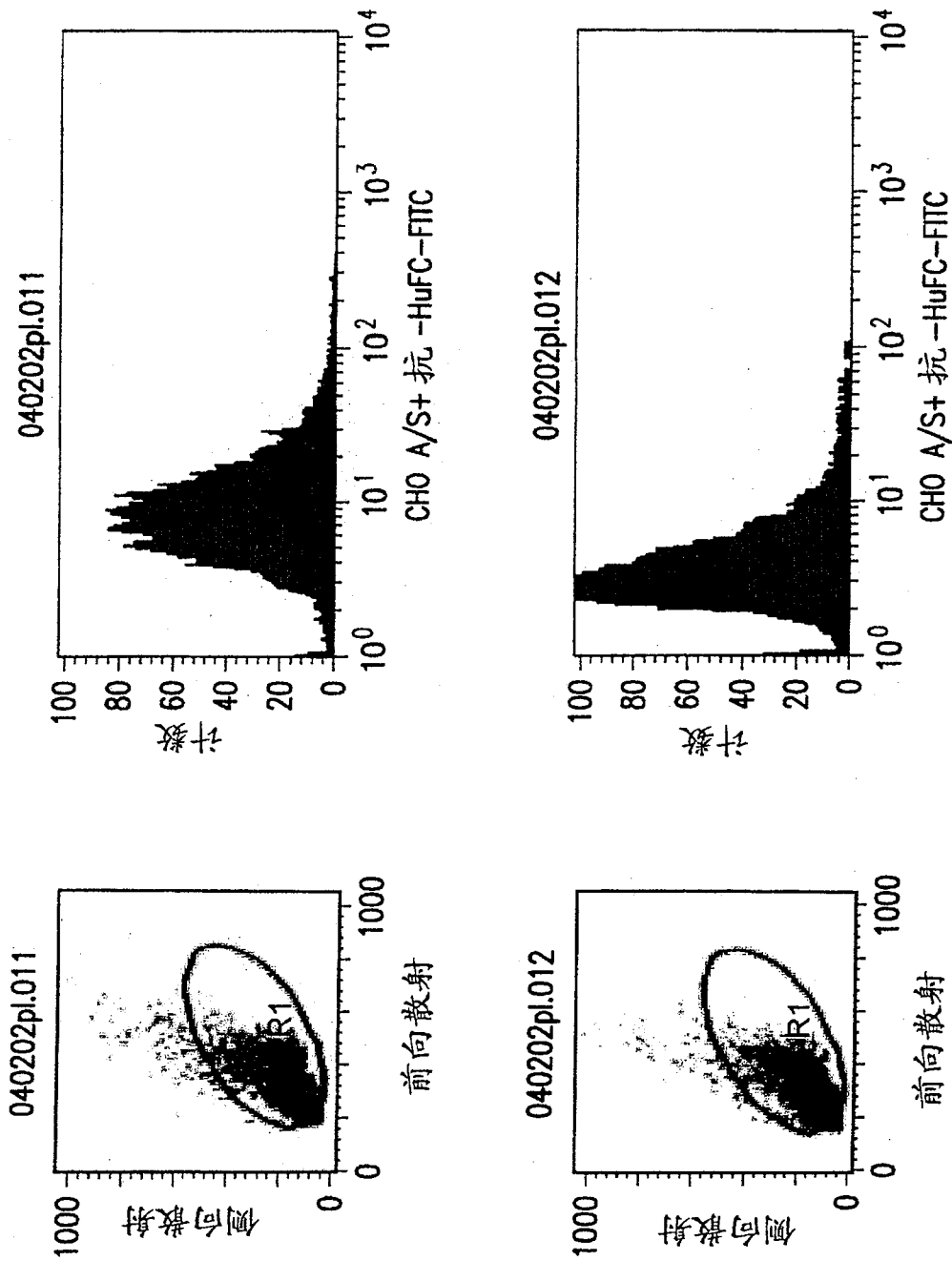


图 23

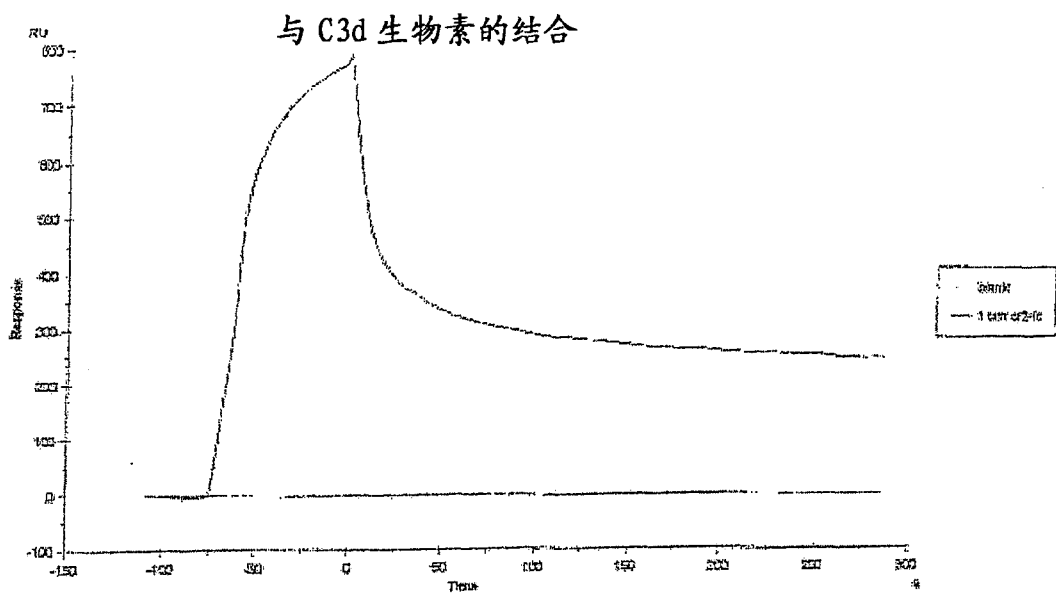


图 24

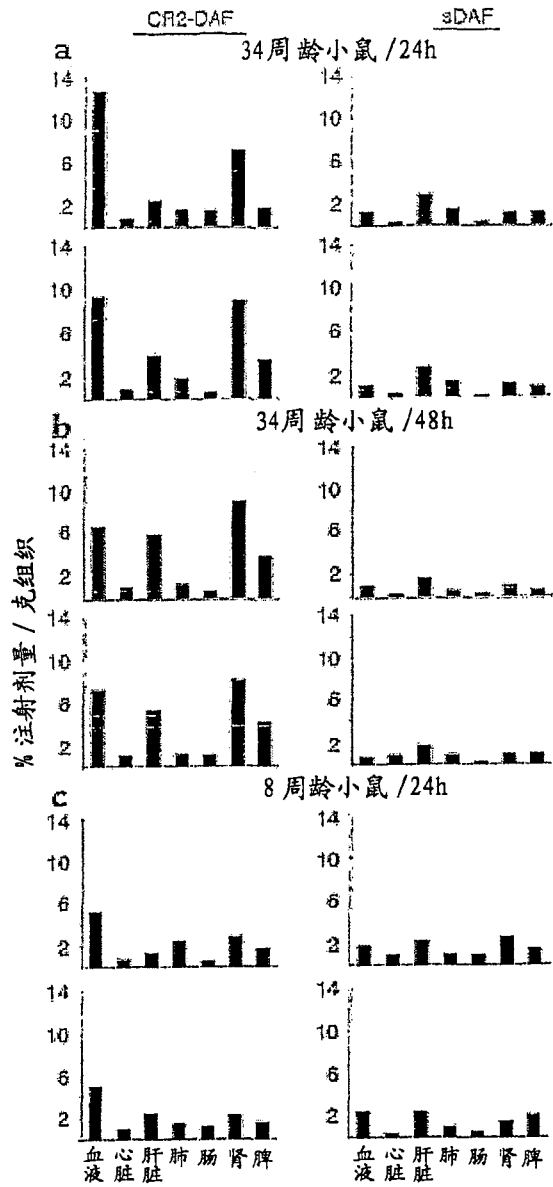


图 25

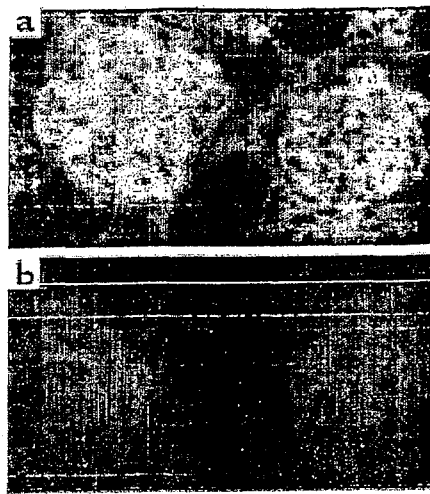


图 26

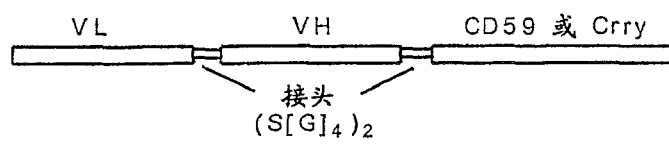


图 27

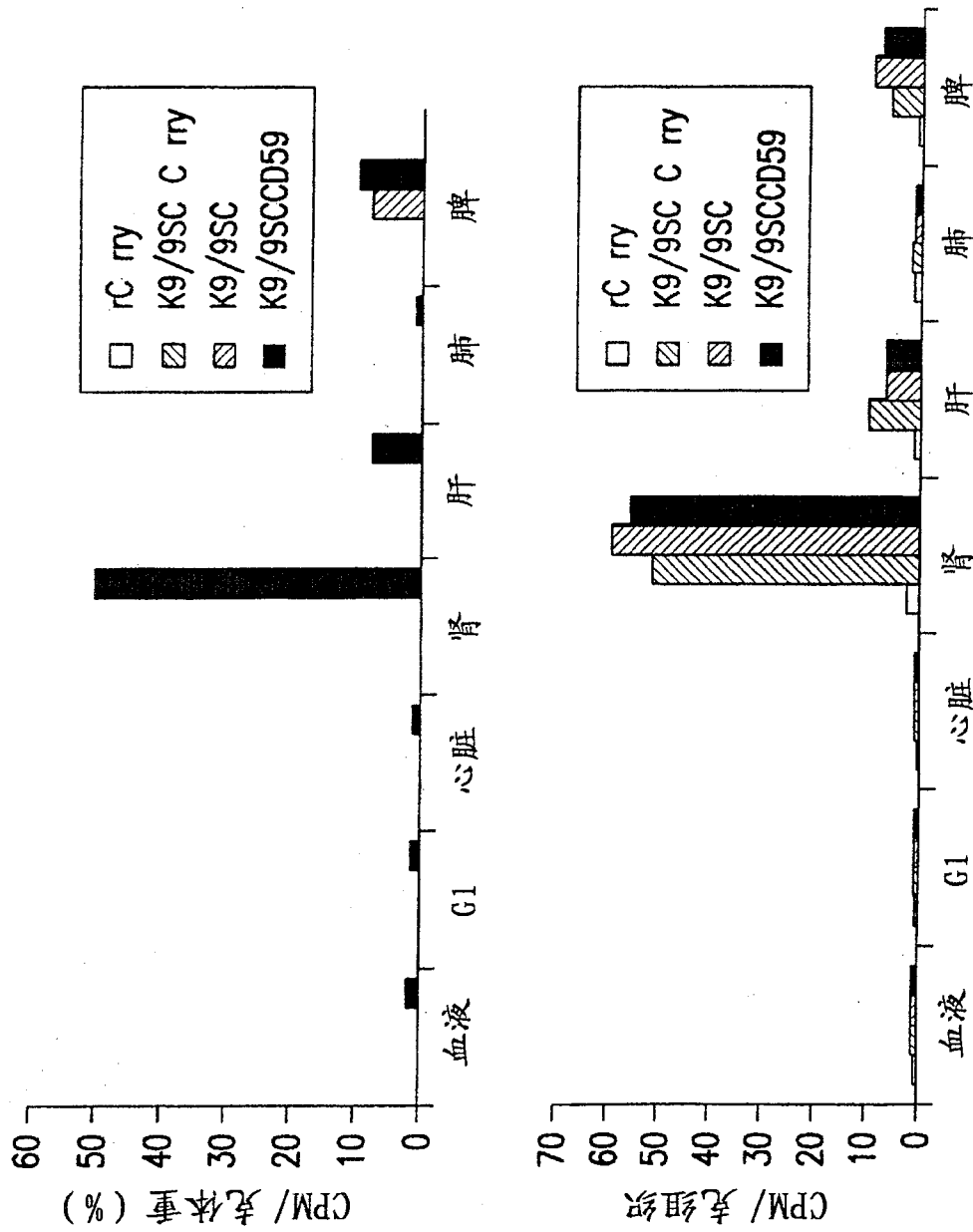


图 28

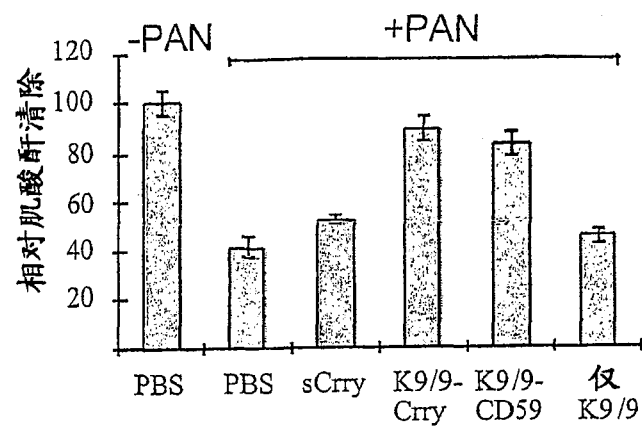


图 29

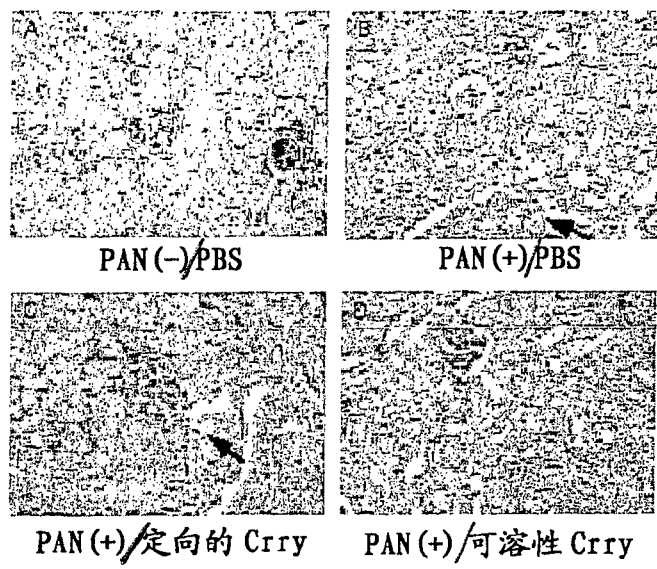


图 30

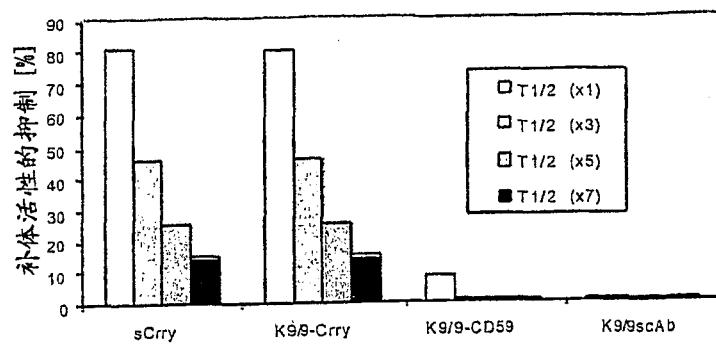


图 31