

[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 02819350.4

[51] Int. Cl.
A61K 39/21 (2006.01)
A61P 31/18 (2006.01)

[45] 授权公告日 2009年2月25日

[11] 授权公告号 CN 100463694C

[22] 申请日 2002.8.21 [21] 申请号 02819350.4

[30] 优先权

[32] 2001.8.21 [33] US [31] 09/934,060

[86] 国际申请 PCT/US2002/026543 2002.8.21

[87] 国际公布 WO2003/016333 英 2003.2.27

[85] 进入国家阶段日期 2004.3.31

[73] 专利权人 马里兰大学生物技术研究所

地址 美国马里兰州

[72] 发明人 安托尼·L·德维科

蒂莫西·R·福茨

罗伯特·G·图什坎

[56] 参考文献

EXPRESSION AND CHARACTERIZATION OF ASINGLE - CHAIN POLYPEPTIDE ANALOGUE OF THEHUMAN IMMUNODEFICIENCY VIRUS TYPE 1 GP120 - CD4RECEPTOR COMPLEX. TIMOTHY R. FOUTS et al. JOURNAL OF VIROLOGY, Vol. 74 No. 2. 2000

审查员 苏林

[74] 专利代理机构 中原信达知识产权代理有限责任公司

代理人 杨青 樊卫民

权利要求书4页 说明书99页 附图21页

[54] 发明名称

病毒外壳蛋白/受体嵌合体及使用方法

[57] 摘要

本发明涉及含有病毒外壳序列和受体序列的嵌合分子，这些序列彼此之间能够相互作用，形成一种能够结合辅助受体的复合物。这样的嵌合分子因此表现出受体-外壳蛋白复合物特征的功能性质，可以用做试剂，其由于占据了细胞上存在的辅助受体而抑制病毒对细胞的感染。具体来说，嵌合的多肽包括免疫缺陷病毒的外膜多肽，例如 HN、SN、FIV、FeL V、FPV 和疱疹病毒的外膜多肽。适合于在嵌合多肽中使用的受体序列包括例如 CD4 DID2 和 CD4M9 序列。

1. 一种嵌合多肽，由以下序列或区组成：

病毒外壳多肽序列；

病毒受体多肽序列，其对病毒外壳多肽序列具有结合亲和性；以及

氨基酸间隔区，该间隔区的长度足够使病毒外壳多肽序列和病毒受体多肽序列彼此结合，

其中，病毒受体多肽由 CD4 多肽序列组成，其中病毒外壳多肽是由 SEQ ID NO:30 组成的 HIV gp120，其中的病毒外壳多肽序列和受体多肽序列通过上述氨基酸间隔区相连。

2. 一种嵌合多肽，由 HIV 病毒外壳多肽序列以及对病毒外壳多肽具有结合亲和性的病毒受体多肽序列组成，其中该嵌合多肽由选自 SEQ ID NO:4 和 SEQ ID NO:6 的氨基酸序列组成。

3. 权利要求 2 的嵌合多肽，其中受体多肽序列选自 SEQ ID NO:26 和 SEQ ID NO:20。

4. 权利要求 1 的嵌合多肽，其中受体多肽序列是 CD4 多肽序列或 CD4 模拟物。

5. 权利要求 1 的嵌合多肽，其中 CD4 多肽序列含有 D1 和 D2 结构域。

6. 权利要求 1 的嵌合多肽，其中间隔区具有从 5 个到 200 个氨基酸。

7. 权利要求 1 的嵌合多肽，其中间隔区含有模拟肽的序列。

-
8. 一种嵌合多肽，由以下序列、区或结构域组成：
病毒外壳多肽序列；
病毒受体多肽序列，其对病毒外壳多肽序列具有结合亲和性；
氨基酸间隔区，该间隔区的长度足够使病毒外壳多肽序列和病毒受体多肽序列彼此结合；以及
由 SEQ ID NO:14 或 SEQ ID NO:32 组成的异源结构域，
其中，病毒受体多肽由 CD4 多肽序列组成，其中病毒外壳多肽是由 SEQ ID NO:30 组成的 HIV gp120，其中的病毒外壳多肽序列和受体多肽序列通过上述氨基酸间隔区相连。
9. 权利要求 1 的嵌合多肽，在可药用的载体中。
10. 权利要求 4 的嵌合多肽，在可药用的载体中。
11. 一种多核苷酸序列，由编码权利要求 2 的嵌合多肽的核酸序列组成。
12. 权利要求 11 的多核苷酸序列，其中核酸序列选自 SEQ ID NO:3 和 SEQ ID NO:5。
13. 权利要求 11 的多核苷酸序列，其中病毒外壳多肽的核酸序列为 SEQ ID NO:29。
14. 权利要求 13 的多核苷酸序列，其中受体多肽的核酸序列选自 SEQ ID NO:25 和 SEQ ID NO:19。
15. 一种宿主细胞，含有权利要求 12 的多核苷酸序列。
16. 一种宿主细胞，含有权利要求 14 的多核苷酸序列。

17. 一种生产与权利要求 1 的嵌合多肽结合的抗体的方法，包括将权利要求 1 的嵌合多肽、或者编码权利要求 1 的嵌合多肽的多核苷酸，以足够使受试对象产生针对权利要求 1 的嵌合多肽的抗体的量施用给受试对象。

18. 一种生产与权利要求 2 的嵌合多肽结合的抗体的方法，包括将权利要求 2 的嵌合多肽、或者编码权利要求 2 的嵌合多肽的多核苷酸，以足够使受试对象产生针对权利要求 2 的嵌合多肽的抗体的量施用给受试对象。

19. 权利要求 1-14 任一项的嵌合多肽或编码所述嵌合多肽的多核苷酸在生产治疗或抑制受试对象感染性疾病的药物中的应用，所述疾病选自人免疫缺陷病毒或 AIDS。

20. 权利要求 19 的应用，其中病毒是免疫缺陷病毒。

21. 权利要求 19 的应用，其中受试对象是人。

22. 权利要求 19 的应用，其中所述药物产生免疫反应。

23. 权利要求 22 的应用，其中免疫反应包含抗体。

24. 权利要求 22 的应用，其中抗体结合至由病毒外壳多肽序列和 HIV 病毒受体多肽序列结合所产生的表位。

25. 权利要求 22 的应用，其中抗体在体外中和病毒。

26. 一种鉴定抑制 HIV 病毒和位于 T 细胞上的 HIV 病毒辅助受体之间相互作用的试剂的方法，包括下列步骤：（a）在存在和不存在测试试剂的情况下，将权利要求 1-6 任一项的嵌合多肽与一种病毒辅助受

体在允许嵌合多肽和辅助受体结合的条件下相接触；以及（b）检测在存在和不存在测试试剂的情况下的结合，其中在存在测试试剂的情况下结合的减少可以鉴定抑制病毒和病毒辅助受体之间结合的试剂。

27. 权利要求 26 的方法，其中病毒是免疫缺陷病毒。

28. 权利要求 26 的方法，其中免疫缺陷病毒辅助受体是 CCR5 或 CXCR4 多肽序列。

29. 权利要求 26 的方法，其中病毒辅助受体存在于完整细胞的表面上。

30. 一种鉴定降低嵌合多肽内病毒外壳多肽序列和病毒受体多肽序列之间分子内结合的试剂的方法，包括下列步骤：（a）将权利要求 2 的嵌合多肽与测试试剂相接触；以及（b）检测病毒外壳多肽序列和病毒受体多肽序列之间结合的量，其中在存在测试试剂的情况下，病毒外壳多肽和病毒受体多肽之间结合量的减少可以鉴定抑制病毒和病毒受体之间结合的试剂，该试剂通过结合到病毒或病毒受体从而阻止病毒外壳多肽和病毒受体多肽之间的直接结合。

31. 权利要求 30 的方法，其中测试试剂选自肽、有机分子、抗体、抗病毒剂、免疫缺陷病毒受体或其功能片段。

32. 权利要求 31 的方法，其中免疫缺陷病毒受体多肽是 CD4 多肽序列。

33. 一种鉴定抑制 HIV 病毒对细胞感染的嵌合多肽序列的方法，包括下列步骤：

（a）在存在和不存在权利要求 1 的嵌合多肽的情况下，将易受 HIV 病毒感染的 T 细胞与感染性病毒粒子相接触；以及

（b）确定嵌合多肽是否抑制 HIV 病毒对细胞的感染，从而鉴定出抑制病毒感染的嵌合多肽序列。

病毒外壳蛋白/受体嵌合体及使用方法

发明背景

发明领域

本发明主要涉及受体配体相互作用，更具体来说，涉及含有病毒外壳多肽和细胞受体多肽序列的嵌合多肽，这些多肽序列彼此结合，模拟病毒蛋白与受体在体内相互作用时自然发生的结构、功能和免疫原性性质。

相关技术描述

初次感染 HIV-1 后产生的体液免疫不能够防止 AIDS 的发展(R. A. Koup 等, *Nature*, 370:416 (1994); R. A. Koup 等, *J. Virol.* 68:4650-5 (1994))。但是，如果一个人在暴露到病毒之前具有高滴度的中和抗体，体液免疫就有可能能够防止感染。这个概念得到了被动免疫研究的大力支持，在该研究中在对黑猩猩使用无细胞的病毒进行攻击的时间附近，给它输入中和的抗 V3 单克隆抗体或合并的高滴度的中和抗血清(E. A. Emini 等, *Nature*: 355:728-30 (1992); R. Shibata 等, *Nat. Med.*, 5:204-10 (1999))。在两组研究中都获得了保护，表明如果在进行攻击时或其后很短的时间内存在足够滴度的正确的抗体，体液免疫可以具有保护性。

其它的研究建议体液免疫可以针对 HIV-1 起保护作用。例如，使用 SCID-hu 小鼠系统进行的被动免疫已经表明，特异性针对 gp120 的 CD4 结合结构域的人单克隆抗体能够防止感染(M. C. Gauduin 等, *Nat. Med.*, 3:1389-93 (1997); P. W. Parren 等, *AIDS*, 9:F1-6 (1995))。此外已经显示二价 CD4-Ig “免疫粘附素”（一种在 CD4 和人 IgG2 重链之间形成的嵌合体）的被动转移，在 HIV-1 黑猩猩攻击系统中能够起保护作用(J. W. Eichberg 等, *AIDS Res. Hum. Retroviruses*, 8:1515-19

(1992); R. H. Ward 等, *Nature*, 352:434-6 (1991))。此外, 中和抗体与针对 SIV 的保护性免疫高度相关(J. L. Heeney 等, *Proc. Natl. Acad. Sci. U.S.A.*, 95:10803-8 (1998))。另外, 在恒河猴中进行的被动转移研究表明, 特异性针对 HIV-1_{DH12} 分离株的高滴度的黑猩猩抗体, 如果使用了足够浓度的抗体, 可以在恒河猴中提供针对 SHIV_{DH12} 的消除性免疫 (sterilizing immunity) (R. Shibata 等, *Nat. Med.*, 5:204-10 (1999))。同样, 在恒河猴中使用 HIVIg、2G12 和 2F5 的被动转移实验证明, 针对 SHIV-89.6P 的攻击, 接受组与非接受的对照组相比保护提高了 50% (Mascola 等, *J. Virol.*, 73:4009-18 (1999))。因此, 这些研究支持了一种想法, 即引发了持久的、高滴度 (或高效) 的、具有广泛特异性的中和抗体反应的免疫策略可能是有保护作用的。达到这一目的的一个成功的策略已经被阐述。重组的单体或寡聚体 HIV 外膜的亚基制剂已被测试能够引发针对较窄范围的分离株的中和反应 (J. P. Moore 等, *AIDS*, 9:S117-136 (1995); Q. J. Sattentau, *Curr. Opin. Immunol.*, 8:540-5 (1996); R. Wyatt 等, *Science*, 280:1884-8 (1998))。

发明概述

本发明涉及含有一种病毒外壳多肽序列和一种病毒受体多肽序列的嵌合多肽, 其中外壳多肽序列和受体多肽序列通过一种间隔区相连。嵌合多肽中的外壳多肽和病毒受体多肽序列能够彼此结合。本发明的嵌合多肽可用于诱导免疫反应和产生抗体。此外, 通过抵抗病毒感染的被动保护或通过给受试对象施用嵌合多肽产生免疫反应 (即抗体或 CTL 反应), 该嵌合多肽还可用于防止、抑制或改善病毒感染的。

在多种不同的实施方案中, 嵌合多肽的病毒外壳多肽序列是一种外膜多肽序列 (例如全长的 gp120 或其片段), 来自于一种结合辅助受体多肽的病毒、一种免疫缺陷病毒、包括 HIV (例如 HIV-1 或 HIV-2)、SIV、FIV、FeLV、FPV 和疱疹病毒。在其它各种实施方案中, 病毒受体多肽序列是 CD4 多肽序列, 全长的或其片段, 例如 D1、D2 结构域或其突变体。引入来自利用可选择的辅助受体的病毒的外

膜基因能够进一步扩展这些单链分子的潜力，为表达不同辅助受体的不同细胞类型提供保护以防止病毒感染。

也提供了具有异源结构域的嵌合多肽。这样的异源结构域给予了独特的官能团，包括标志、粘附素和免疫增强剂。例如，异源结构域可以具有一种氨基酸序列，例如 c-myc 多肽序列或免疫球蛋白多肽序列（例如重链多肽序列）。

在本发明中提供了具有编码嵌合多肽的核酸序列的多核苷酸序列。多核苷酸可以包含在表达载体中并用于表达嵌合多肽。

在本发明中还提供了与本发明的嵌合多肽结合的抗体及其功能片段。抗体可用于治疗方法和诊断方法中。这样的抗体可以在体外或体内中和免疫缺陷病毒，也可以例如通过被动保护用于抑制免疫缺陷病毒的感染。这样的抗体可以结合到由病毒外壳多肽序列和病毒受体多肽序列的结合所产生的表位上。例如，这样的表位可以出现在外膜多肽序列上。

本发明的嵌合多肽、多核苷酸和抗体可用于治疗病毒感染，或用于诱导免疫反应。因此，在本发明中提供了在可药用载体中的嵌合多肽、多核苷酸和抗体。

生产抗体的方法包括将足够量的本发明的嵌合多肽给药到受试对象以产生针对嵌合多肽的抗体。这样的方法也可用于例如抑制或改善受试对象中的病毒感染，或者当抗体施用给接受的受试对象时用于被动保护。

在受试对象中抑制病毒感染的方法包括施用有效量的本发明的嵌合多肽或编码其的多核苷酸以抑制病毒对细胞的感染。施用的嵌合多肽能够通过受试对象细胞上的病毒辅助受体结合或产生保护性的免

疫反应来防止病毒的感染。嵌合多肽可以被施用足够的量以改善受试对象中病毒的感染。

产生免疫反应的方法能够产生抗体反应或 CTL 反应。产生的抗体能够在体外中和免疫缺陷病毒。抗体也可以结合到由嵌合多肽的两个多肽序列结合而暴露出的表位上。

还提供了鉴定能够调节病毒和病毒辅助受体、以及病毒和病毒受体之间的结合或相互作用的试剂的方法。在一个实施方案中，该方法包括将一种含有结合到受体上的病毒外壳蛋白的嵌合多肽与一种辅助受体多肽（例如 CCR5 或 CXCR4 多肽序列），在允许嵌合多肽与辅助受体多肽结合的条件下，在存在和不存在测试试剂的情况下相接触，检测在存在和不存在测试试剂的情况下的结合。在存在测试试剂的情况下结合减少说明该试剂抑制了病毒和病毒辅助受体多肽之间的结合。

在另一个实施方案中，该方法包括将一种嵌合多肽在允许嵌合多肽分子内结合的条件下，在存在和不存在测试试剂的情况下相接触，检测嵌合多肽内的分子内结合或相互作用。在存在测试试剂的情况下结合减少说明该试剂抑制了嵌合体内病毒和病毒受体多肽之间的分子内结合或相互作用。试剂可以在嵌合多肽与病毒辅助受体多肽接触之前或之后加入。病毒辅助受体或受体多肽可以出现在完整细胞的表面，可以出现在动物中，例如一个非人类的灵长动物。该方法可以使用一种免疫缺陷病毒，例如 HIV、SIV 等来进行。测试试剂包括了一种试剂文库，例如肽、有机分子、抗体及其片段、抗病毒剂、病毒辅助受体、功能片段或其模拟肽。

还提供了鉴定能够调节（抑制或刺激）病毒对细胞感染的嵌合多肽序列的方法。在一个实施方案中，该方法包括将一种对病毒感染敏感的细胞在存在和不存在本发明的嵌合多肽序列的情况下与一种感染

性病毒粒子相接触，然后确定嵌合多肽是否调节（抑制或刺激）了病毒对细胞的感染（体外或体内）。

附图简述

图 1 是一个编码了示例嵌合多肽的多核苷酸结构的简图。全长的单链（FLSC）嵌合多肽包含了 HIVgp120（BaL 株）、20 个氨基酸的间隔多肽、含有 D1 和 D2 结构域的 CD4 多肽序列、以及 myc 肽“标志”。一种截短了的单链（TsSC）嵌合体在 C1（恒定区 1）、V1（可变区 1）、V2 和 V5 含有缺失。在 TcSC 中显示的缺失按照 BaL gp120 的序列编号。FLSC R/T 嵌合体在弗林蛋白酶切割位点带有一个单一突变，位于 gp120 的 C-端的 R 被改变为 T。FLSC R/T CD4M9 嵌合体在弗林蛋白酶切割位点带有一个单一突变、21 个氨基酸的间隔多肽和 CD4M9 肽序列。

图 2 是含有由 293-SC 细胞表达的 FLSC 和 TcSC 可溶性嵌合多肽的细胞培养上清液的 Western 印迹分析。使用 gp120（第 1 到 4 道）和 CD4（第 5 到 8 道）进行免疫印迹，箭头所指按照凝胶迁移率的减小依次为 gp120-CD4 单链（单链）、切割的 gp120（gp120 片段）和切割的 CD4（CD4 片段）。

图 3 是对 293-SC 细胞表达的 gp120-CD4 的分析；第 1 道是未交联的 gp120-CD4，第 2 道是交联的 gp120-CD4。

图 4 是 FLSC 在交联后的免疫印迹分析。每种不同的 FLSC 浓度（1-0.03 μ M）的总蛋白的相对百分数（%）被显示在条形图中：（A）45% 172KD；（B）25% 302KD；以及（C）10% 高级寡聚体。

图 5A-5C 是 gp120-CD4 嵌合体的结合分析。（A）与抗-gp120 抗体保温的全长单链（FLSC）（17b、48d、A32 和 C1 1）与交联的 gp120/rsCD4 和未复合的 gp120.17b、48d 和 A32 相比，对于复合的 gp120（gp120）有偏好的亲和性。标准误差以横号表示。（B）FLSC 和 TcSC 中人抗-gp120 单克隆抗体的半数最大结合浓度的倒数（ELISA）。（C）与 gp120 CD4 结合结构域反应的单克隆抗体 IgG1b12、F91 和 205-469 的半数最大结合的倒数。

图 6 是对 gp120-CD4 嵌合体 (FLSC、TcSC) 与表达 CCR5 (R5) 或 CXCR4 (X4) 辅助受体的 L1.2 细胞的结合的分析。不表达 CCR5 或 CXCR4 的对照细胞被命名为 L1.2。结合的复合物使用 5 μ g/ml 抗-CD4 Mab45 通过流式细胞计检测。显示的值是实验三次的代表性数据。

图 7 是在存在 gp120 结合抗体 (17b、48d、A32、C11 和 2G12) 和 gp41 抗体 (F240) 的情况下, 对 gp120-CD4 与辅助受体的结合的分析。L1.2 细胞如所示表达辅助受体 CCR5 (R5)、CXCR4 (X4) 或不表达辅助受体 (L1.2)。无抗体的对照显示为“+”。用未处理的细胞获得的背景测量被显示为“-”。结合的复合物使用 5 μ g/ml Mab45 通过流式细胞计检测。结果表示为相对于在相应的对照分析中获得的平均荧光强度的结合百分数。显示值为三次单独实验的平均值。标准误差用横号表示。

图 8 是对 HIV-1₂₀₄₄ (一种 X4 特异性分离株) 和 HIV-Bal (一种 R5 特异性分离株) 病毒被 FLSC、TcSC、BaLgp120 和 BaLgp120-rsCD4 复合物中和的分析。U373 细胞在 HIV-1LTR 启动子的调控下表达 R5 或 X4 型的 CD4 以及 P-半乳糖苷酶。在最高的试验浓度下 FLSC 和 TcSC 针对 HIV-1₂₀₄₄ 的 ID₉₀ 也没有达到, 因此被表示为 >10 μ g/ml。

图 9 是显示编码结构域的嵌合的 gp120-CD4-IgG1 基因的简图。基本上是将原始的 gp120-CD4 亚克隆到含有 IgG1 重链铰链区 CH2 和 CH3 的质粒中, 从而使嵌合的 gp120-CD4-IgG1 多肽得以表达。

图 10 是对 293 细胞中表达的 gp120-CD4-IgG1 嵌合多肽的免疫印迹分析。显示了从培养上清液分离的嵌合的 gp120-CD4-IgG1 (第 1 道), 并与纯化的 HIV Bal 株的 gp120 多肽 (第 2 道) 进行比较。切割的 gp120 由箭头所示, 与纯化的 gp120 共同迁移。

图 11 是 gp120-CD4-IgG1 嵌合多肽与表达辅助受体的 L1.2 细胞结合的倒数稀释分析。表达 CCR5 和 CXCR4 的 L1.2 细胞如图所示。

图 12 是对 FLSC-IgG1 上的 MAbs (17b) 与表达 CCR5 细胞的结合的阻断的分析, 显示了 FLSC-IgG1 与 R5 辅助受体通过 gp120 上的 R5-结合结构域相互作用。

图 13 显示了在对弗林蛋白酶切割位点突变 (R-T) 后 gp120-CD4

(FLSC) 分子的稳定性增加。

图 14 显示了对 FLSC R/T CD4M9 与 BaLgp120、FLSC 和 FLSC R/T 的免疫印迹比较。FLSC R/T CD4M9 是通过将 FLSC R/T 中的 CD4 D1D2 序列用 CD4M9 基因序列转换后构建成的。

图 15 是对 FLSC R/T CD4M9 与 CCR5 (R5) 结合的分析。分析结果显示为平均荧光强度。该图显示 FLSC R/T CD4M9 以与 FLSC R/T 相等的效率结合到 R5 表达细胞上。

图 16 显示了当 gp120 与 CD4 相互作用时暴露逐渐增加的表位的结合, 以及暴露的 17b 表位是 FLSC R/T CD4M9, 并与 FLSC R/T 相当。

图 17 显示了原始的 R5 HIV-1 (92BR020) 被来自 FLSC 接种的小鼠的血清所中和。

图 18 显示了 BaLgp120/sCD4 复合物的共价交联封闭了暴露在 FLSC 上的表位。

图 19 显示了纯化的 R/T FLSC-IgG1 在还原和非还原条件下免疫印迹的比较。

图 20 显示了 R/T FLSC-IgG1 与人和恒河猴 CCR5 的结合。

图 21 显示了 RANTES 竞争性抑制 R/T FLSC-IgG1 与 CCR5 的结合。

发明详述

本发明是基于这样的发现, 即含有 HIV 外膜多肽和 CD4 受体的嵌合多肽可以形成一种能够与辅助受体结合的相互作用的复合物。在本发明的嵌合多肽中, 与 CD4 结合的 HIV gp120 模拟了当 HIV 结合了细胞上存在的 CD4 时出现的外膜蛋白-CD4 过渡态; gp120 展示出当复合物形成时暴露出的保守的表位, 它们与辅助受体 CCR5 直接相互作用。外膜-CD4 过渡态的形成和其后与细胞辅助受体的结合是 HIV 感染细胞中的一个关键步骤。因此, 能够阻止或抑制 gp120-CD4 与辅助受体结合的抗体或其它试剂, 例如通过与 gp120-CD4 复合物形成时暴露的表位结合, 能够抑制病毒与辅助受体的相互作用, 因而介导了

防止 HIV 感染。

因此，本发明的嵌合多肽或编码嵌合多肽的核酸，通过例如诱导针对病毒外壳蛋白结合到受体多肽上时形成的过渡态复合物的免疫反应，可以在治疗学上用于治疗、抑制、防止或改善病毒感染。这样的嵌合多肽在此也称为“单链”分子，可以用于筛选能够抑制、阻止或破坏外壳多肽序列与嵌合序列内的多肽受体序列结合、或嵌合体与辅助受体多肽序列结合的试剂，从而鉴定出潜在的用于治疗相应的病毒感染的治疗方法。例如，一种能够抑制、阻止或破坏免疫缺陷病毒外膜多肽 CD4 复合物与 CCR5 结合的试剂可能是一种治疗剂，用于治疗患有 HIV 或有患该病的危险的受试对象。

嵌合多肽也可用于生产特异性针对相互作用的外壳蛋白-受体复合物的抗体。这样的特异性抗体能够用于针对病毒感染或增殖的被动保护、用于诊断目的、以及用于鉴定和表征在复合物形成时暴露的表位（例如一种隐藏的表位）。即使在不存在病毒外壳蛋白和受体之间的分子内结合时，嵌合多肽在引发免疫反应方面也可能比单独的病毒外壳多肽序列更有效。因此，这样的非相互作用的嵌合多肽也是有价值的，并也包括在本发明内。

含有与受体和辅助受体结合的病毒外壳多肽的嵌合多肽还具有其它的优点，能够通过体内抑制病毒进入细胞辅助受体而被动防止病毒感染。此外，通过鉴定能够抑制、阻止或干扰嵌合多肽与辅助受体结合的试剂，这样的嵌合多肽可用于筛选治疗方法。例如，一种能够抑制、阻止或干扰免疫缺陷病毒外膜多肽 CD4 复合物与 CCR5 结合的试剂可能是一种治疗剂，用于治疗患有 HIV 或有患该病的危险的受试对象。因为病毒与细胞受体的结合是病毒感染任何细胞所必需的，因此含有任何病毒外壳蛋白和相应的受体的多肽序列的嵌合多肽都包括在本发明的组合物和方法中。

在本发明中提供了嵌合多肽，其中含有被间隔区连接起来的病毒外壳多肽序列和病毒受体多肽序列。嵌合多肽的外壳多肽序列和受体多肽序列被一种具有足够长度氨基酸的间隔区连接起来，以便嵌合多肽的两个多肽序列更好地结合或相互作用。在一个实施方案中，外壳多肽序列是一种免疫缺陷病毒的外膜多肽序列。在另一个实施方案中，外壳多肽序列来自于一种与辅助受体多肽结合的病毒。在其它各种不同的实施方案中，外壳多肽序列和受体多肽序列是相应的全长天然序列的活性片段。

在此所用的术语“外壳”是指来自病毒的能够与细胞结合的多肽序列。一般来说，病毒外壳蛋白出现在病毒粒子的外表面附近，使得病毒可以结合并随后穿进细胞膜。但是，外壳多肽序列包括任何能够与受体多肽结合或相互作用的病毒蛋白。此处定义的外壳多肽序列可以与其它分子实体，例如糖、脂肪酸、脂类等非共价地或共价地结合。外壳多肽序列可以含有多个病毒多肽序列。例如，在一个嵌合多肽中 gag 多肽序列也可以与外膜多肽序列包括在一起，以将外膜多肽序列维持在一种能够结合受体多肽序列的构象。

在本发明中有用的病毒外壳多肽序列可以是任何来源的，包括例如细菌、植物和动物病毒，只要相应的细胞受体是已知的或可以被鉴定的。具体的病毒的例子包括：反转录病毒科（例如人免疫缺陷病毒，例如 HIV）；小 RNA 病毒科（例如脊髓灰质炎病毒、甲肝病毒、呼肠病毒、人库克萨基病毒、鼻病毒、埃可病毒）； Calciviridae（例如引起胃肠炎的病毒株）； Togaviridae（例如马脑脊髓炎病毒、风疹病毒）；黄病毒科（例如登革热病毒、脑炎病毒、黄热病病毒）；冠状病毒科（例如冠状病毒）； Rhabdoviridae（例如疱疹性口炎病毒、狂犬病病毒）；线病毒科（例如伊波拉病毒）；副粘病毒科（例如副流感病毒、流行性腮腺炎病毒、麻疹病毒、呼吸道合胞体病毒）；正粘液病毒科（例如流感病毒）；布尼亚病毒科（例如汉坦病毒、bunga 病毒、白蛉病毒和 Nairo 病毒）；沙粒状病毒科（出血热病毒）；呼

肠孤病毒科（例如呼肠孤病毒、环状病毒和轮状病毒）；双 RNA 病毒科（Birnaviridae）；嗜肝 DNA 病毒科（Hepadnaviridae）（乙肝病毒）；细小病毒科（细小病毒）；乳多空病毒科（乳头瘤病毒、多瘤病毒）；腺病毒科（大多数腺病毒）；疱疹病毒科（单纯性疱疹病毒（HSV）1 和 2、带状疱疹病毒、细胞肥大病毒（CMV）、疱疹病毒）；痘病毒科（天花病毒、牛痘病毒、痘病毒）；以及彩虹病毒科（Iridoviridae）（例如非洲猪瘟病毒）；以及未分类的病毒（例如海绵状脑病的病原体、丁型肝炎的病原体（被认为是乙肝病毒的缺陷的附随体）、非甲非乙型肝炎的病原体（1 类为内部传播；2 类为肠胃外传播（即丙型肝炎）、诺沃克病毒及相关病毒、以及星状病毒）。（也参见表 1）。

在此所用的术语“受体”是指任何由细胞表达的可以被病毒结合的多肽。一般来说，这样的受体天然存在于细胞的表面，但是也可以是基因工程产生的。受体多肽可以与其它分子实体，例如糖、脂肪酸、脂类等非共价地或共价地结合。受体多肽可以含有一个或多个相邻的被非共价地或共价地结合的多肽区段。这样的分子实体或其它多肽序列对于受体的构象，例如与外壳多肽序列的结合可能是重要的。因此，其它的元件包括对于受体构象重要的分子，可以包括在本发明的嵌合多肽中。受体多肽序列可以是原核或真核来源的。

如果是真核的，植物核动物受体都可以考虑。优选的动物受体是哺乳动物，包括人和灵长动物，例如黑猩猩、猿、恒河猴、长臂猿、猩猩等，以及其它的动物物种，包括驯养动物和家畜。人受体的一个例子是 CD4。其它的受体的例子包括粘多糖和 CD2、CR1。其它的受体是已知的，并适用于本发明的组合物和方法中（参见例如表 1，也参见 Eckard Wimmer 编辑的“动物病毒的细胞受体”，冷泉港出版社（1994））。

表 1

受体 (结合亚基)	病毒 (科)	参考文献
免疫球蛋白类分子		
VCAM-1	EMC-D (小 RNA 病毒科)	Huber (1994)
[CAM-1] (第一个结构域)	大部分的 HRV, CAV 13、18 和 21 (小 RNA 病毒科)	Colonna 等 (1986) Greve 等 (1989) Staunton 等 (1989) Tomassini 等 (1989)
PVR (第一个结构域)	脊髓灰质炎病毒 (小 RNA 病毒科) (1989)	Koike 等 (1990) Mendelsotm 等 (1989)
CD4 (第一个结构域)	HIV-1、2; SIV (Lentiviridae) 人疱疹病毒 7	Daigleish 等 (1984) Klatzmain 等 (1984) Lusso 等 (1994)
CEA, 几个成员 (第一个结构域)	小鼠肝炎病毒 (冠状病毒科)	Williams 等 (1978)
MHC I	赛姆利基森林病毒 (Semliki forest virus) (披盖病毒科) Factate dehydrogenase 病毒 小鼠细胞肥大病毒 (疱疹病毒科) SV-40	Hetenius 等 (1978) Otdstone 等 (1980) Inada 和 Mims (1984) Wykes 等 (1993) Breau 等 (1992)
MHC II	髓鞘脱落病毒 (Lentiviridae)	Dalziel 等 (1991)
整合蛋白		
VLA-2 (α -链)	埃可病毒 1.8 (小 RNA 病毒科)	Bergelson 等 (1992, 1993)
(RGD-结合蛋白)	FMDV (小 RNA 病毒科)	Fox 等 (1989) Mason 等 (1994)
$\alpha\beta 3$ (vibronectin)	CAV9、埃可病毒 1.8 (小 RNA 病毒科)	Roivainen 等 (1994)
转运蛋白		
磷酸转运蛋白 Analogen	长臂猿白血病病毒 (反转录病毒科) Amphotropic murine (反转录病毒科)	Johann 等 (1992) Miller 等 (1994)
氨基酸阳离子转运蛋白	嗜亲性的鼠白血病病毒 (反转录病毒科)	Albritton 等 (1989)
信号受体		
LDL 受体蛋白家族	小部分的 HRV (小 RNA 病毒科) 鸟白血病性增生肉瘤病毒 A 亚类 (科未知)	Hofer 等 (1994) Bates 等 (1993) Connolly 等 (1994)
乙酰胆碱受体 (α -1)	狂犬病毒 (rhabdoriridae) 牛痘病毒 (痘病毒科)	Leniz 等 (1990) Marsh 和 Eppatein (1987)

EGF 受体		
白细胞分化抗原 [CD9]	猫免疫缺陷病毒 (Lentiviridae)	Willett 等 (1994)
其它		
氨肽酶 N	人冠状病毒 229E (冠状病毒科) TGEV (冠状病毒科)	Yeager 等 (1992) Delmas 等 (1992)
补体受体 CR2	EBV (疱疹病毒科)	McClure (1992)
高亲和性层粘蛋白受体	辛德比斯病毒 (披盖病毒科)	Wang 等 (1992)
腐败加速因子 [CD55]	埃可病毒 7 (6、11、12、20、21)	Bergelson 等 (1994)
膜辅助因子蛋白	麻疹病毒 (麻疹病毒科)	Dorig 等 (1993)
Moesin	麻疹病毒 (麻疹病毒科)	Dunster 等 (1994)
血型糖蛋白 A	EMCV (小 RNA 病毒科) 呼肠病毒 (呼肠病毒科)	Allaway 和 Barness (1986) Paul 和 Lee (1987)
半乳糖苷神经酰胺	HIV-1 (Lentiviridae)	Bhat 等 (1991)
红细胞 P 抗原	细小病毒 B19 (细小病毒科)	Brown 等 (1993)
BLV 受体 1	牛白血病病毒 (反转录病毒科)	Ban 等 (1993)
唾液酸糖蛋白	仙台病毒 (副流感病毒科)	Suzuki 等 (1985)
唾液酸	流感病毒 (Orthomyxoviridae) 呼肠病毒 (呼肠病毒科) 猪轮状病毒 A 类 (轮状病毒科) 人冠状病毒 OC43、牛冠状病毒 (冠状病毒科)	Herrler 等 (1985) Femandes 等 (1994) Roisma 等 (1994) Vlasak 等 (1988)
硫酸乙酰肝素 (heparan sulfate)	人细胞肥大病毒 (疱疹病毒科) HSV	Compton 等 (1993) WuDunn 和 Spear (1989)

在此所用的术语“辅助受体”是指任何在病毒与受体结合之后或一同结合的受体。因此，辅助受体包括了出现在细胞上的任何多肽或分子实体，其通过与病毒多肽-受体复合物结合而直接或间接地促进病毒进入。除了能够促进病毒进入细胞的辅助受体外，还包括那些介导细胞附着或向性而不直接或间接促进病毒进入的辅助受体。具体的辅助受体的例子是含有 7 个跨膜结构域 (7-TM) 的趋化因子受体，例如 CCR5 和 CXCR4，其能够结合免疫缺陷病毒。其它的辅助受体包括 CCR-2b、CCR3、CCR8、V28/CXCR1、US28、STRL 33/BOB/TYMSTR、GPR15/Bonzo 和 GPR1。

在此所用的术语“多肽”、“蛋白”和“肽”可以互换使用，用以指示至少由两个氨基酸通过酰胺键共价连接成的序列聚合物，而不论其长度或翻译后修饰（例如糖基化、磷酸化、脂化、肉豆蔻化、遍在蛋白化等）。D-和 L-氨基酸以及 D-和 L-氨基酸的混合物也包括在内。

嵌合多肽是指一个含有两个或多个部分的氨基酸序列，这些部分一般来说在天然的氨基酸序列中不在一起被发现。

正如在此处公开的，一种含有 CD4 多肽序列和与 CD4 结合的 HIV 外膜 gp120 多肽序列的嵌合多肽，当被一种氨基酸间隔区分开时，在嵌合体内可以互相结合。gp120-CD4 嵌合体能够结合辅助受体，例如 CCR5。因此，在另一个实施方案中，嵌合多肽含有能够结合辅助受体的病毒的外壳多肽序列。

CD4 似乎是多种与免疫缺陷有关的病毒进入的靶。例如，免疫系统的细胞，例如淋巴细胞和巨噬细胞表达 CD4，容易受到 HIV、SIV、疱疹病毒 7 和许多其它病毒的感染。在此所用的术语“免疫缺陷”，当用于病毒时，是指能够感染免疫来源的细胞或能够感染参与免疫反应的细胞的病毒，一般来说这样的感染能够损害被感染的宿主的免疫功能。因此，本发明适用于任何能够结合 CD4 的病毒或病毒株的任何病毒外壳多肽。

按照本发明，提供了具有免疫缺陷病毒外膜多肽序列的嵌合多肽。在不同的情况下，外膜多肽序列是 HIV、HTLV、SIV、FeLV、FPV 和疱疹病毒的多肽序列。在其他情况下，病毒是巨噬细胞向性 HIV 或淋巴细胞向性 HIV。在另一种情况下，HIV 是 HIV-1 或 HIV-2。在其它多种情况下，外膜多肽序列是 gp120、gp160 或 gp41。

本发明嵌合多肽的受体和病毒外壳多肽序列在它们之间需要一个间隔区域，以便例如在两个多肽之间形成一个相互作用的复合物。尽管不希望被理论所束缚，据信间隔区允许受体和病毒外壳多肽序列之间的运动或可变性以形成相互作用的复合物。

在此所用的术语“间隔区”是指一种物理的或化学的部分、或任何大小或性质的共价的或非共价的键，它将病毒外壳多肽序列和受体多肽序列连接起来，同时提供了必要的可变性或运动以形成相互作用的复合物。在本发明中，间隔区优选以“头对头”的方向连接两个多肽序列。“头对头”是指外壳多肽的氨基或羧基末端氨基酸被连接到受体多肽序列的氨基或羧基末端氨基酸。因此，一个间隔区能够将外壳多肽序列的羧基末端氨基酸连接到受体多肽序列的氨基末端氨基酸上，例如象在此处举例的 HIV gp120 和 CD4。此外，间隔区还能够将外壳多肽的氨基末端氨基酸连接到受体多肽的羧基末端氨基酸，或多肽序列的羧基末端氨基酸或多肽序列的两个氨基末端氨基酸上。

具体的间隔区的例子包括一个或多个氨基酸或肽模拟物。氨基酸间隔区事实上可以是任何长度，例如少至 5 个或多至 200 个或更多的氨基酸。因此，氨基酸间隔区可以含有从大约 10 到大约 100 个氨基酸，或含有从大约 15 到大约 50 个氨基酸。在优选情况下，间隔区具有从大约 20 到大约 40 个氨基酸。其它的间隔区的例子包括多肽序列的末端之间的二硫键。糖间隔区也可以考虑在内。本领域的专业技术人员将知道或能够容易地确定其它在功能上能够允许在病毒外壳多肽序列和受体多肽序列之间形成相互作用复合物的部分。

受体和外壳多肽序列可以是任何氨基酸长度。优选情况下，它们的长度允许多肽序列在嵌合多肽内相互结合。因此，受体和外壳多肽序列包括天然的全长受体和全长的外壳蛋白序列以及多肽序列的部分。例如，受体和外壳多肽序列的氨基酸截短、内部缺失或亚基也包括在内。在优选情况下，这样的修饰形式能够彼此间相互作用。例如，

优选的是截短的或缺失的外壳多肽序列能够与受体多肽序列相互作用。截短的受体多肽序列的一个例子是 CD4 的 D1 和 D2 结构域，它们能够与 HIV 外膜多肽序列相互作用（图 9）。截短的外壳多肽序列的一个例子是截短的 HIV gp120，它失去了氨基末端的 60 个氨基酸和羧基末端的 20 个氨基酸（例如在 TcSC 中）。

因此，根据本发明，提供了嵌合多肽，包括截短的或内部缺失的序列。在一个实施方案中，病毒外壳多肽序列或受体多肽序列与它们的相应的全长多肽序列相比含有一个或多个被去除的氨基酸。在一种情况下，截短的病毒外壳多肽序列是 HIV 外膜多肽序列，而在另一种情况下，截短的受体多肽序列是 CD4 序列。如同在此举例的，截短的 HIV 外膜多肽序列是失去了氨基末端的 60 个氨基酸或羧基末端的 20 个氨基酸的 gp120，以及是含有 D1 和 D2 结构域的截短的 CD4 多肽。在其它不同的情况下，嵌合多肽含有一个内部缺失的病毒外壳多肽序列或一个内部缺失的 CD4 多肽序列。

除了截短的、内部缺失的和亚基多肽序列，还包括了其它的多肽序列修饰。这样的修饰包括对含有嵌合多肽的一个或两个多肽序列的氨基酸序列进行少量的取代、变异或衍生，只要修饰的嵌合多肽与未修饰的嵌合多肽具有基本相同的活性或功能。例如，病毒外壳或受体多肽序列可以含有糖、脂肪酸（软脂酸、豆蔻酸）、脂类，可被磷酸化或可具有其它典型的与多肽序列相关的翻译后修饰。

另一个修饰的例子是添加异源的结构域，它赋予了与两种多肽或嵌合多肽都不同的功能。异源的结构域可以是任何小的有机或无机分子或大分子，只要它赋予了其它的功能。异源的结构域可以影响、也可以不影响病毒外壳多肽和受体多肽之间的相互作用或亲和性。赋予了不同功能的异源结构域的具体例子包括了能够赋予导向（例如受体配体、抗体等）、免疫增强功能（例如免疫球蛋白、佐剂）、能够纯化、分离或检测（例如 myc、T7 标志、多聚组氨酸、亲和素、生物

素、凝集素等)的氨基酸序列。

具体的异源结构域可以包括一个 c-myc 多肽序列和/或 IgG1 重链多肽序列。异源结构域可以有多种功能。例如, IgG1 能够在体内用作免疫增强剂,也可以用作附着分子从而被纯化、分离或检测(例如通过与具有酶活性,例如辣根过氧化物酶或碱性磷酸酶的二级抗体反应)。专业技术人员将知道其它的异源结构域,并根据所需的应用和功能选择适当的异源结构域。

因此,根据本发明,提供了具有一个或多个异源结构域的嵌合多肽。在一个实施方案中,异源结构域是一个 c-myc 多肽序列 glu-gln-lys-leu-ile-ser-glu-glu-asp-leu (SEQ ID NO:14)。在另一个实施方案中,异源结构域是一个含有重链的免疫球蛋白多肽序列 (SEQ ID NO:32)。

受体和外壳多肽序列可以是任何的氨基酸长度。优选情况下,它们的长度允许多肽序列在嵌合分子内时可以彼此结合。因此,受体和外壳多肽序列包括了天然的全长受体和全长外壳多肽序列以及多肽序列的部分。

在一种情况下,本发明包括了一种全长的单链 (FLSC) 嵌合多肽,其中含有 HIV gp120 (BaL 株)、氨基酸间隔多肽、含有 D1D2 结构域的 CD4 多肽序列和 myc 肽“标志”(SEQ ID NO:2),或与编码嵌合多肽的 SEQ ID NO:2 具有至少 95%的序列同一性。

在另一种情况下,本发明包括一种在其弗林蛋白酶切割位点具有单一突变的 FLSC 多肽,其中在 gp120 的 C 末端的 R 被改变成 T (FLSC-R/T),或者与编码嵌合多肽的 SEQ ID NO:2 具有至少 95%的序列同一性。具体来说,FLSC-R/T 在 506 位的氨基酸从精氨酸突变为苏氨酸 (SEQ ID NO:4)。

如同此处所举例的，多肽序列包括了对含有嵌合多肽的一个或两个多肽序列的氨基酸序列进行取代、突变或衍生，只要修饰的嵌合多肽与未修饰的嵌合多肽具有基本相同的活性或功能。例如，一个病毒外壳或受体多肽序列可以含有糖、脂肪酸（软脂酸、豆蔻酸）、脂类，可被磷酸化或可具有其它典型的与多肽序列相关的翻译后修饰。

在另一种情况下，病毒外壳多肽序列或受体多肽序列与它们相应的未修饰的多肽序列相比含有一个或多个氨基酸取代。例如，提供了一种核苷酸序列（SEQ ID NO:5），它编码的多肽包含了一个模拟 CD4 的受体，该受体显示了基本上相同的活性或增强的免疫反应。具体来说，编码嵌合多肽 FLSC-R/T（SEQ ID NO:4）的 CD4 D1D2 结构域中的氨基酸序列 KKVVLGKKGDTVELTCTASQKKSIQFHW 的基因序列被一个编码氨基酸序列 CNLARCQLRCKSLGLLGKCAGSFCACGP（氨基酸 528-556（SEQ ID NO:20））的核苷酸序列（SEQ ID NO:19）所取代，后者在此后称为 FLSC-R/T CD4M9（SEQ ID NO:6）。

在此所用的术语“基本上相同的活性或功能”，当用来指如此修饰的嵌合多肽时，正如此处描述的或本领域所熟知的，是指多肽保留了大部分、所有或多数的与未修饰的多肽有关的活性。同样地，不影响嵌合多肽与辅助受体相互作用的能力的修饰被包括在内。同样，不影响嵌合多肽诱导比单独施用病毒外壳蛋白更有效的免疫反应的能力的修饰也包括在内。

此处包括的“活性的”或“功能的”修饰的嵌合多肽可以通过常规的功能分析方法来鉴定。例如，通过使用抗体结合分析、辅助受体结合分析或者测定在过渡态复合物中暴露的表位的诱导，这些表位当两个多肽序列不结合时一般是隐藏的，人们可以容易地确定修饰的嵌合多肽是否具有活性。

能够诱导更有效的免疫反应的嵌合多肽，可以通过例如在将嵌合

体施用到受试对象后测量抗体的滴度来鉴定。破坏了病毒外壳多肽序列和受体多肽序列之间的相互作用或者破坏了含有不相互作用的病毒外壳多肽序列和受体多肽序列的嵌合多肽诱导更有效的免疫反应的能力的修饰，不具有与相应的未修饰的嵌合多肽基本相同的活性或功能，因此不包括在内。

在此所用的术语“同源性”或“同源的”，当用于多肽时，是指两个多肽之间的氨基酸序列的相似性。当两个多肽中的一个氨基酸位置被同样的氨基酸占据时，它们在该位置是同源的。因此，“基本上同源的”是指一种大部分但不是完全同源的氨基酸序列，它保留了与它同源的序列的大部分或全部的活性。

因为修饰的嵌合多肽将保留与未修饰的嵌合多肽相关的活性或功能，因此修饰的嵌合多肽一般具有与未修饰的多肽的氨基酸序列“基本相同的”或“基本同源的”氨基酸序列。正如此处所用，术语“基本相同的”或“基本同源的”，当用于多肽序列时，是指一个多肽的序列与参比的序列具有至少 50% 的同一性。修饰的多肽和基本相同的多肽典型地与参比的序列具有至少 70%、或 85%、更可能是 90%、最可能是 95% 的同源性。对于多肽，为了在序列之间获得上述的同源性百分数而比较的长度一般至少为 25 个氨基酸、或者至少 50 个氨基酸、更可能是至少 100 个氨基酸、而最可能是 200 个氨基酸或以上。

如本发明所述，基本相同的或同源的多肽包括在不破坏嵌合多肽的功能的氨基酸序列的位置上（通过例如此处描述的功能分析方法确定）进行添加、截短、内部缺失或插入、保守或非保守取代、或其它的修饰。取代的一个具体的例子是其中一个或多个氨基酸被另一个化学上或生物学上相似的残基代替。此处所用的术语“保守取代”是指一个残基被一个化学或生物学上相似的残基取代。保守取代的例子包括用一个疏水残基例如异亮氨酸、缬氨酸、亮氨酸或甲硫氨酸代替另一个疏水残基，用一个极性残基代替另一个极性残基、例如用精氨酸

代替赖氨酸、谷氨酸代替天冬氨酸、或谷氨酰胺代替天冬酰胺等。本领域的专业技术人员能够认识到大量的氨基酸能够被修饰或被其它化学相似的残基代替而基本上不改变活性。

基本相同的或同源的多肽还包括那些具有能够改进或赋予其它的功能或活性的修饰的多肽。例如，FLSC R/T 具有一个突变的弗林蛋白酶位点，它增加了修饰的 FLSC 的稳定性（参见例如图 13）。

修饰的多肽还含有“化学衍生物”，其中的一个或多个氨基酸的侧链被化学改变或衍生了。这样的衍生多肽包括例如那些其中游离的氨基基团形成了胺盐酸盐、对甲苯磺酰基团、carobenzoxy 基团，游离的羧基基团形成了盐、甲酯和乙酯，而游离的羟基基团形成 O-芳基或 O-烷基衍生物的氨基酸，以及天然存在的氨基酸衍生物，例如脯氨酸的衍生物 4-羟基脯氨酸、赖氨酸的衍生物 5-羟基赖氨酸、丝氨酸的衍生物高丝氨酸、赖氨酸的衍生物鸟氨酸等。还包括 D-氨基酸以及能够改变共价键的氨基酸衍生物，例如在两个半胱氨酸残基之间形成的二硫键产生了一个环化的多肽。

在此所用的术语“分离的”或“基本上纯的”，当用来修饰本发明的嵌合多肽、其序列片段和多核苷酸时，是指它们由人类的干预而产生并且与它们的体内天然的细胞环境分离开来。一般来说，这样分离的多肽和多核苷酸基本上不含天然情况下与它们相连的其它蛋白、核酸、脂类、糖或其它物质。

一般来说，当一种多肽与在天然情况下和其结合的蛋白及其它分子分开，至少占 60%重量时，可以被认为是基本上纯的。制备物的纯度按重量计有可能至少 75%、更可能至少 90%、而最可能至少 95%。基本纯的嵌合多肽可以通过例如在细胞中表达编码多肽的多核苷酸并分离产生的多肽来获得。例如，如同实施例中描述的，在哺乳动物细胞中表达编码 gp120-CD4 多肽的重组多核苷酸，允许从培养基中使用

免疫亲和柱分离嵌合多肽。此外，嵌合多肽也可以化学合成。纯度可以通过任何适当的方法测定，例如通过聚丙烯酰胺凝胶电泳，然后对胶进行染色（例如银染）或通过 HPLC 分析。

本发明的嵌合多肽及其修饰可以通过本领域的各种已知方法来制备。例如多肽的修饰可以通过定点突变（例如基于 PCR）或随机突变（例如 EMS）、外切核酸酶删除、化学修饰、或者与编码异源结构域的多核苷酸序列融合来导入。嵌合多肽的获得可以通过在一个宿主细胞，例如细菌、酵母或哺乳动物细胞中表达编码多肽的多核苷酸，然后通过使用典型的生化方法（例如免疫亲和纯化、凝胶纯化、表达筛选等）来纯化表达的嵌合多肽。其它众所周知的方法在 Deutscher 等（蛋白纯化指南，酶学方法，Vol. 182, Academic Press(1990), 在此引为参考）中有描述。

本发明还提供了编码嵌合多肽、其片段的多核苷酸序列和互补序列。在一个实施方案中，核酸编码了此处举例的嵌合的 gp120-CD4 多肽。例如，SEQ ID NO:1 定义了编码上述的 FLSC 的序列，其中含有编码 gp120（SEQ ID 23）和 CD4 D1D2（SEQ ID NO:25）的核苷酸序列。SEQ ID NO:3 定义了编码 FLSC R/T 的序列，其中在 gp120 的 C-末端、gp120 中推测的弗林蛋白酶切割位点上的一个精氨酸被苏氨酸取代，从而使 FLSC-R/T 的稳定性高于 FLSC。FLSC-RT 的核苷酸序列含有被 SEQ ID NO:29 编码的修饰的 gp120 和 CD4 D1D2（SEQ ID NO:25）。此外，本发明还提供了多核苷酸序列 SEQ ID NO:5，它编码含有一个被取代了的弗林蛋白酶切割位点的嵌合多肽 FLSC R/T CD4M9，并进一步提供了用编码模拟了 CD4 受体的氨基酸序列的序列取代了编码 CD4 D1D2 区域的基因序列，从而提供了相对于 FLSC 或 FLSC-R/T 而言更高的免疫反应和附加的稳定性。编码 FLSC R/T CD4M9 的核苷酸序列含有编码修饰的 gp120 的 SEQ ID NO:29 和编码 CD4M9 的 SEQ ID NO:19。FLSC R/T CD4M9 嵌合多肽还可以含有 SEQ ID NOs:23 和 19。

在另一个实施方案中，TsSC (SEQ ID NO:12) 编码了一个 gp120-CD4 多肽 (SEQ ID NO:13)，其中 gp120 的氨基酸序列从氨基和羧基端被截短了。TsSC 的核苷酸序列含有一个编码截短了的 gp120 的序列 (SEQ ID NO:27) 和编码 CD4D1D2 的序列 (SEQ ID NO:25)。在另一个实施方案中，嵌合多肽 gp120-CD4-IgD1 被带有一个附加的标志 (SEQ ID NO:31) 的核苷酸 SEQ ID NO:1 编码。

在此所用的术语“核酸”、“多核苷酸”、“寡核苷酸”和“引物”可以交替使用，用于指称脱氧核糖核酸(DNA)或核糖核酸(RNA)，不论是双链还是单链、线性还是环状的。RNA 可以是未拼接的或拼接的 mRNA、rRNA、tRNA 或反义 RNAi。DNA 可以是互补 DNA(cDNA)、基因组 DNA 或反义 DNA。具体包括核苷酸类似物和衍生物，例如那些能抵抗核酸酶降解、能够编码本发明的嵌合多肽的核苷酸类似物和衍生物。具有核酸酶抗性的寡核苷酸和多核苷酸对于本文描述的核酸疫苗是特别有用的。

“分离的”或“基本上纯的”多核苷酸是指核酸并不与在它所来自的生物体的天然存在的基因组中与其紧密相连的 5'或 3'端的编码序列继续紧密相连。因此该术语包括例如重组 DNA (例如通过 PCR 或在克隆过程中通过限制性内切酶处理产生的 cDNA 或基因组 DNA 片段)，以及整合在载体、自主复制的质粒或病毒中的重组 DNA，或者原核或真核生物的基因组 DNA。它也包括例如嵌合体或融合体的重组 DNA 部分。因此该术语不包括那些存在于基因组或 cDNA 文库的数百万个序列中但性质不明了的核酸，以及凝胶上分级的限制性酶消化的文库中的核酸。

本发明的多核苷酸还包括遗传密码简并的核酸。有 20 种天然的氨基酸，它们中大多数由一个以上的密码子确定。所有编码本发明的嵌合多肽的简并多核苷酸序列都包括在内。

本发明的多核苷酸序列可以使用本领域熟知的标准技术（例如分子克隆、化学合成）来获得，纯度可以通过聚丙烯酰胺或琼脂糖凝胶电泳、测序分析等来确定。也可以使用本技术领域众所周知的基于杂交或计算机的技术来分离多核苷酸。这样的技术包括但不限于：（1）将基因组 DNA 或 cDNA 文库与探针杂交以检测同源的核苷酸序列；（2）使用抗体筛选由 DNA 序列表达的多肽（例如使用表达文库）；（3）使用能够与目的核酸序列退火的引物在基因组 DNA 或 cDNA 上进行聚合酶链反应（PCR）；（4）使用计算机在序列数据库中搜索相关的序列；以及（5）对扣减的核酸文库进行差异筛选。因此，为了获得其它的编码受体的核苷酸，例如那些编码 CD4 的多核苷酸，可以从文库中筛选同源序列的存在。

本发明还包括了基本上同源的多核苷酸。此处所用的术语“同源的”，当用于核酸分子时，是指两个核苷酸序列之间的相似性。当在两个分子中的一个核苷酸位置被同样的核苷酸占据时，它们在该位置是同源的。“基本上同源的”核酸序列至少 50%同源、更可能至少 75%同源、而最可能 90%或以上同源。如同基本上同源的本发明的嵌合多肽一样，与编码嵌合多肽的本发明多核苷酸基本上同源的多核苷酸编码的多肽保留了大部分或全部的与它同源的序列相关的活性或功能。对于多核苷酸来说，序列间比较的长度一般至少 30 个核苷酸、或者至少 50 个核苷酸、更可能至少 75 个核苷酸、而最可能 110 个核苷酸或以上。根据多核苷酸序列的间隔和错配的寡核苷酸来识别同源序列的算法在本领域内是众所周知的，例如 BLAST（参见例如 Altschul 等，*J. Mol. Biol.* 15:403-10(1990)）。

此外，多核苷酸可用作杂交探针以鉴定编码嵌合多肽的多核苷酸，例如 mRNA 的存在或量（Sambrook 等，*分子克隆：实验室手册*，冷泉港实验室，N.Y. (1989)）。一般来说，这样的探针被设计成对所需的序列是特异性的，以便降低与无关的序列杂交的可能性。这样的

探针可以使用放射性核素、发光部分等修饰以便能够被检测。杂交的条件也可以被修饰以便获得所需的特异性。例如，中度严紧的杂交条件可以包括：2X SSC/0.1% SDS，大约 37°C 或 42°C（杂交条件）；0.5X SSC/0.1% SDS，在室温附近（低严紧性清洗）；0.5X SSC/0.1% SDS，在大约 42°C（中度严紧性清洗）。中度到高度严紧的杂交条件的例子如下：0.1X SSC/0.1% SDS，在大约 52°C（中度到高度严紧性清洗）。高度严紧的杂交条件的例子如下：0.1X SSC/0.1% SDS，在大约 65°C（高度严紧性清洗）。

如果需要，本发明的多核苷酸可以是裸露的或存在于适合于通过细胞膜的载体中（例如多核苷酸-脂质体复合物或胶体分散系统）、包含在载体中（例如逆转录病毒载体、腺病毒载体等）、连接在惰性的珠或其它异源的结构域上（例如抗体、配体、生物素、链亲和素、凝集素等）或其它本发明公开的或在本领域熟知的适当的成分。因此，递送多核苷酸的病毒和非病毒手段都能够成功并值得考虑。本发明的多核苷酸也可以含有其它与其连接的核酸序列，这些核酸序列编码具有不同的功能性，例如本文提出的多种异源结构域的多肽。

本发明的多核苷酸还能够被修饰成例如对核酸酶具有抗性，以提高它们在药物制剂中的稳定性。所述的多核苷酸可用于编码本发明的嵌合多肽，特别是当这样的多核苷酸整合在本文公开的或本领域熟知的表达系统中时。因此，含有表达载体的多核苷酸也包括在内。

为了在细胞内繁殖或表达，此处描述的多核苷酸可以插入到载体中。术语“载体”是指质粒、病毒或其它本领域熟知的媒介物，可以通过将核酸插入或整合来进行操作。这样的载体可用于遗传操作（即“克隆载体”）或用于转录或翻译插入的多核苷酸（即“表达载体”）。载体一般含有至少一个用于在细胞中繁殖的复制起点和一个启动子。还包含控制元件，包括表达载体内存在的启动子以促进适当的转录和翻译（例如内含子的拼接信号、维持基因正确的阅读框架以允许 mRNA

按阅读框架翻译、以及终止密码子)。此处描述的多核苷酸的体内或体外表达可以由一个与核酸可操作连接的启动子介导。“启动子”是指足以指导与启动子可操作连接的核酸的转录的最小核酸序列(参见例如 Bitter 等, *Methods in Enzymology*, 153:516-544 (1987))。启动子可以组成型地指导转录、可以是组织特异性的、或者可以提供可诱导的或可阻遏的转录;这样的元件一般位于所调控的基因的 5'或 3'区域。

在本发明中,对于结合辅助受体的病毒,在易受病毒感染的细胞(例如表达辅助受体的细胞)中导入和表达编码嵌合多肽的多核苷酸是有利的。采用这种方法,转化的易感细胞将在辅助受体附近分泌表达的嵌合多肽,从而抑制或阻止了病毒与辅助受体的接近,这又进一步抑制或阻止了病毒对细胞的感染。为此,可以将一个组织特异性启动子与多核苷酸序列可操作连接,以使嵌合多肽在适当的靶细胞中表达。

在此所用的词组“组织特异性启动子”是指一种启动子在特定的细胞或组织内有活性,能够在特定的细胞,例如肝细胞、造血细胞或动物体内特定组织的细胞中,表达与其可操作连接的多核苷酸。该术语也涵盖了所谓的“渗漏的”启动子,它主要在一个组织中调节选定的 DNA 的表达,但是也可以在一个或多个其它组织中表达。

可诱导的启动子也可用于调节在细胞中的表达。“可诱导的启动子”是指一种启动子,它对一个外部信号或试剂处理作出反应而提高活性水平(例如金属硫蛋白 IIA 启动子、热休克启动子)。“可阻遏的启动子”或“条件启动子”是指一种启动子,它对一个阻遏蛋白或等效的化合物作出反应而降低活性水平。当阻遏蛋白不再存在时,转录受到激活或去阻遏。这样的启动子可以组合使用,也可以包含其它为转录和表达所需的 DNA 序列,例如内含子和增强子序列。

在此所用的术语“可操作连接”是指一种选定的多核苷酸(例如

编码一种嵌合多肽)和调控序列,当适当的分子(例如转录激活蛋白)结合到调控序列上时以允许转录的方式被连接在一起。一般来说,启动子位于多核苷酸的5'端,可以紧邻转录起始位点以允许启动子调节多核苷酸的表达。但是,当第一个载体上的启动子控制一种蛋白的表达,而该蛋白又调控一个控制第二个载体上的多核苷酸的表达的启动子时,间接的可操作连接也是可能的。

当在细菌系统中克隆时,组成型启动子,例如 T7 等,以及诱导型启动子,例如 γ 噬菌体的 pL、plac、ptrp、ptac 等都可以使用。当在哺乳动物细胞系统中克隆时,组成型启动子,例如 SV40、RSV 等,或者来自于哺乳动物细胞基因组的诱导型启动子(例如金属硫蛋白启动子)或来自于哺乳动物病毒的诱导型启动子(例如小鼠乳腺癌病毒的长末端重复序列、腺病毒晚期启动子),都可以使用。通过重组 DNA 或合成技术产生的启动子也可以用来提供本发明的核酸序列的转录。

利用重组病毒或病毒元件指导表达的哺乳动物表达系统可以被改造。例如,当使用腺病毒表达载体时,核酸序列可以与一个腺病毒转录/翻译控制复合物,例如晚期启动子和三重前导序列连接起来。此外,痘苗病毒 7.5K 启动子也可以使用(参见例如 Mackett 等, Proc. Natl. Acad. Sci. USA, 79:7415-7419 (1982); Mackett 等, J. Virol., 49:857-864 (1984); Panicali 等, Proc. Natl. Acad. Sci. USA, 79:4927-4931 (1982))。

哺乳动物表达系统还包括专门为“基因治疗”方法设计的载体,包括腺病毒载体(美国专利 Nos: 5700470 和 5731172)、腺病毒相关载体(美国专利 No: 5604090)、单纯疱疹病毒载体(美国专利 No: 5501979)以及逆转录病毒载体(美国专利 Nos: 5624820、5693508 和 5674703, 以及 WIPO 公告 WO92/05266 和 WO92/14829)。编码嵌合多肽的基因可以被导入疫苗投送载体,例如减毒的牛痘(M. Girard 等, C R Acad Sci III., 322:959-66 (1999); B. Moss 等, AIDS, 2 Suppl 1:S103-5 (1988))、Semiliki 森林病毒(M. Girard 等, C R Acad Sci III.,

322:959-66 (1999); S. P. Mossman 等, *J. Virol.*, 70:1953-60 (1996)) 或沙门氏菌 (R. Powell 等, 传染病控制的分子方法, 183-187 页, F. Bran, E. Norrby, D. Burton, 和 J. Meckalanos 编辑, 冷泉港出版社, 冷泉港, NY (1996); M. T. Shata 等, *Mol Med Today*, 6:66-71 (2000)), 以提供一种有效和可靠的手段, 来表达适当地连接和折叠的病毒外壳蛋白和受体序列, 例如 gp120 和 CD4。基于牛乳头瘤病毒 (BPV) 的载体具有染色体外元件的复制能力 (Sarver 等, *Mol. Cell. Biol.*, 1:486(1981))。在染色体外载体进入小鼠细胞后不久, 载体复制到每个细胞大约 100 到 200 个拷贝。由于插入的 cDNA 的转录不需要将质粒整合到宿主染色体中, 可以进行高水平的表达。这样的载体也已经在基因治疗中使用 (美国专利 No: 5719054)。基于 CMV 的载体也包括在内 (美国专利 No: 5561063)。

对于酵母的表达, 有多种含有组成型或诱导型启动子的载体可以使用 (例如参见 *Current Protocols in Molecular Biology*, Vol. 2, Ch. 13, Ausubel 等编, Greene Publish. Assoc. & Wiley Interscience (1988); Grant 等, “酵母的表达和分泌载体”, *Methods in Enzymology*, Vol. 153, pp. 516-544, Wu & Grossman 编辑, 3 1987, Acad. Press, N.Y. (1987); Glover, DNA 克隆, Vol. II, Ch. 3, IRL Press, Wash., D.C. (1986); Bitter, “酵母中的异源基因表达”, *Methods in Enzymology*, Vol. 152, pp. 673-684, Berger & Kimmel 编辑, Acad. Press, N.Y. (1987); 以及“糖酵母的分子生物学”, Strathem 等编辑, 冷泉港出版社, Vols. I 和 II (1982))。组成型酵母启动子, 例如 ADH 或 LEU2, 或者诱导型启动子例如 GAL 可以使用 (“酵母中的克隆”, R. Rothstein, 在“DNA 克隆实验方法”中, Vol. 11, Ch. 3, D. M. Glover 编辑, IRL 出版社, Wash., D.C. (1986))。此外, 能够通过例如同源重组来促进外源核酸序列整合到酵母染色体中的载体, 在本技术领域内是熟知的, 也可以使用。一般来说, 当被插入的多核苷酸对于更常用的酵母表达载体来说太大时 (例如大于大约 12kb), 可以使用酵母人工染色体 (YAC)。多核苷酸可以被插入到表达载体中用于体外表达 (例如使用可购买到的体外转录/翻译试剂

盒)，或者可以插入到一种含有启动子序列的表达载体中，该启动子序列在体内通过将一种适当的核酸转入适当的细胞、器官、组织或生物体，能够促进在原核或真核生物中的表达。

当用于此处时，“转基因”是任何一个通过手段插入到宿主细胞中、并变成从该细胞发育而成的生物体的一部分的多核苷酸。转基因可以包括一个或多个启动子，以及其它任何为所选的 DNA 的表达所必需的 DNA，例如内含子，它们与所选的 DNA 都是可操作连接的，还可以包括增强子序列。转基因可以包括对被转基因的生物体而言是部分或完全异源（即外源）的多核苷酸，或者可以代表一种与生物体的内源基因同源的基因。转基因可以整合到宿主细胞的基因组中，或者维持为一个自我复制的质粒。

当用于此处时，“宿主细胞”是引入了多核苷酸的细胞，该多核苷酸可以在细胞内增殖、转录或表达出编码的多肽。该术语也包括这些宿主细胞的任何后代。应该理解所有的后代并不一定与亲本细胞一样，因为在复制时可能发生突变。宿主细胞包括但不限于细菌、酵母、昆虫和哺乳动物细胞。例如，用重组噬菌体多核苷酸、质粒核酸、或粘粒核酸表达载体转化的细菌；用重组酵母表达载体转化的酵母；用重组病毒表达载体（例如花椰菜花叶病毒 CaMV、烟草花叶病毒 TMV）感染、或用重组质粒表达载体（例如 Ti 质粒）转化的植物细胞系统，用重组病毒表达载体（例如杆状病毒）感染的昆虫细胞系统，或用重组病毒表达载体（例如逆转录病毒、腺病毒、痘苗病毒）感染的动物细胞系统，或者被改建后以稳定表达的转化的动物细胞系统。

为了长期表达本发明的多肽，优选稳定的表达。因此，使用含有病毒复制起点的表达载体，可以将被适当的控制元件（例如启动子/增强子序列、转录终止子、多聚腺苷化位点等）控制的核酸转化到细胞中。尽管不希望被任何具体的理论所束缚或限制，但据信在哺乳动物细胞中稳定地维持表达载体，需要将载体整合到宿主细胞的染色体

上。此外，表达载体也可以包含一种核酸，该核酸编码了一个赋予了选择压力有抗性的选择标记或指示细胞导入了基因的报告基因，因此允许带有载体的细胞被鉴定、生长和扩展。当用于此处时，“报告基因”是指一种其表达可以被分析的基因；这样的基因包括但不限于 lacZ、氨基酸生物合成基因例如酵母的 LEI2 基因、荧光素酶、或哺乳动物的氯霉素转乙酰酶（CAT）基因。报告基因可以被整合在染色体中，或者也可以携带在自主复制的质粒上（例如酵母 2 μ m 质粒）。此外，选择标记可以在第二个载体上，与含有本发明的多核苷酸的第一个载体共转染到宿主细胞中。

有许多选择系统可以使用，包括但不限于对氨基糖苷 G418 赋予抗性的新霉素基因（Colberre Garapin 等，*J. Mol. Biol.*, 150:1 (1981)）和对潮霉素赋予抗性的潮霉素基因（Santerre 等，*Gene*, 30: 147 (1984)）。最近，描述了其它的选择标记，即 trpB，它允许细胞利用吲哚代替色氨酸；hisD，它允许细胞利用 histinol 代替组氨酸（Hartman 等，*Proc. Natl. Acad. Sci. USA*, 85:8047 (1988)）；以及 ODC（鸟氨酸脱羧酶），它对鸟氨酸脱羧酶抑制剂 2-(二氟甲基)-DL-鸟氨酸、DFMO 赋予抗性（McConlogue, 分子生物学现代通讯，冷泉港实验室编辑，(1987)）。

在此所用的术语“转化”是指将一种对细胞来说是外源的多核苷酸（例如一种转基因）导入细胞后产生的遗传变化。因此，“转化的细胞”是指一种细胞，在其中或其后代之中已经通过重组技术导入了多核苷酸。转化的细胞不包括整个人。宿主细胞的转化可以通过本领域的专业人员所熟知的常规技术来进行。当宿主细胞是真核细胞时，DNA 转化的方法包括例如磷酸钙、微注射、电穿孔、脂质体和病毒载体。如同此处描述的或本领域所熟知的，真核细胞也可以使用本发明的多核苷酸序列或其片段和编码选择标记的第二个 DNA 分子共转化。另一种方法是使用一种真核病毒载体，例如猿病毒 40（SV40）或牛乳头瘤病毒来短暂地感染或转化真核细胞并表达蛋白（参见例如《真

核病毒载体》，冷泉港实验室，Gluzman 编辑（1982））。当宿主是原核细胞（例如大肠杆菌）时，能够吸入 DNA 的感受态细胞，可以使用本领域熟知的步骤，将指数生长期收获的细胞用氯化钙方法处理后制备。原核细胞的转化也可以通过宿主细胞的原生质体融合来进行。

本发明的嵌合多肽、多核苷酸以及含有它们的载体可以使用标准的技术包被在脂质体中，然后导入细胞或完整生物体中。对于多核苷酸的投送优选使用阳离子型的脂质体。使用脂质体在体内或体外导入各种成分，包括蛋白和多核苷酸，对于本领域的专业技术人员来说是熟知的（参见例如美国专利 Nos: 4844904、5000959、4863740 和 4975282）。

脂质体可以通过在脂质体制备液中加入一种其相应的细胞受体已被鉴定的配体，例如多肽，来靶向到目的细胞类型或组织中。例如，在感染 CD4+细胞的病毒的情况下，CD4+细胞是适当的靶，HIV gp120 是一个适当的配体，可以用于在细胞内导入含有此处描述的嵌合多肽或多核苷酸的脂质体。单克隆抗体也可用于寻靶，对于本领域的专业人员来说，许多这样的针对广范围的细胞表面蛋白的抗体是已知的并且可以获得。选定的配体被共价结合到预先制成的脂质体内的脂质锚上，或在脂质体制备过程中被掺入（参见 Lee & Low, *J Biol. Chem.*, 269:3198 (1994); Lee & Low *Biochem. Biophys. Acta*, 1233:134 (1995)）。

本发明的嵌合多肽和编码其的多核苷酸可以被导入完整的生物体中。具体来说，对于含有与辅助受体结合的病毒外壳蛋白的嵌合多肽，表达本发明的嵌合多肽的转基因动物可用于研究嵌合表达的长期效应，以及确定表达的嵌合多肽是否能够保护或抑制相应病毒的感染。

因此，在另一个实施方案中，本发明提供了表达嵌合多肽的非人

类的转基因动物。优选的动物易受到其相应的受体多肽序列是已知的病毒的攻击。优选的动物是那些易受到免疫缺陷病毒感染的动物，包括哺乳动物，例如非人类的灵长动物（例如恒河猴、黑猩猩、猿、长臂猿、猩猩等）、驯养动物以及家畜，如同本文描述的。

术语“转基因动物”是指任何其体细胞或生殖细胞带有通过在亚细胞水平上精心的遗传操作，例如通过微注射或用重组病毒感染而直接或间接地接受的遗传信息的动物。如本文所述，术语“转基因的”还包括从遗传操作的转基因动物中获得的细胞或组织（即“转基因细胞”、“转基因组织”）。在本发明上下文中，“转基因动物”不包括通过经典的杂交育种或体外受精产生的动物，但是包括其中一个或多个细胞接受了重组 DNA 分子的动物。转基因动物相对于转基因可以是杂合的也可以是纯合的。产生转基因动物的方法在本领域内是众所周知的（参见例如美国专利 Nos: 5721367、5695977、5650298 和 5614396）。

此处描述的嵌合多肽可以被用来生产其它的试剂，例如抗体。本发明的抗体可用于本文提出的各种治疗方法。例如，在一个免疫接种的受试对象体内产生的抗体可以保护受试对象抵抗病毒的攻击，或者可以被转移给一个接受的受试对象，从而被动地保护第二个受试对象抵抗感染。能够与病毒外壳多肽序列和受体多肽序列之间形成复合物时暴露的表位结合的抗体也可以被产生。此外，本发明的抗体可用于诊断方法、纯化方法以及筛选方法中（例如，鉴定隐藏的表位、辅助受体等），如同本文所公开的。

因此，在本发明中提供了与嵌合多肽结合的抗体，包括特异性针对上述的复合物形成时暴露出的隐藏的表位的抗体。在一个实施方案中，抗体在体外中和多种病毒分离株和来自不同的地理进化枝的病毒（术语称为“广谱中和”）。在另一个实施方案中，抗体在体内或体外抑制、阻止或阻断了病毒的攻击。在这些实施方案的不同情况下，

被中和的病毒是免疫缺陷病毒，包括本文提出的 HIV-1 和 HIV-2 免疫缺陷病毒。还提供了包括多克隆抗体、具有不同的表位特异性的合并的单克隆抗体、以及独特的单克隆抗体制剂在内的抗体。

针对嵌合多肽的抗体通过将嵌合多肽施用给动物而产生。可以使用本领域熟知的方法生产、分离和纯化抗体。因此，在另一个实施方案中，本发明提供了生产针对嵌合多肽的抗体的方法。本发明的一个方法包括将嵌合多肽施用给受试对象，然后分离与嵌合多肽结合的抗体。在一个实施方案中，产生的抗体与病毒外壳多肽序列和受体多肽序列之间结合时暴露出的隐藏的表位结合。

在优选情况下，抗体与病毒外壳多肽序列（例如外膜多肽序列）和受体多肽序列彼此结合时暴露出的隐藏的表位结合。例如，HIV 外膜多肽序列 gp120 当与 CD4 受体多肽序列结合时暴露出一个隐藏的表位，针对暴露的表位的抗体可以导致 HIV 的广谱中和。这样的表位可以在不同的病毒分离株和地理进化枝之间共有，因此解释了针对这些表位的抗体的广谱中和活性。

尽管不希望被理论所束缚，但似乎在未与 CD4 结合时，隐藏的表位不被暴露或不具有抗原性。此处所用的术语“表位”是指一种抗原上的抗原决定簇，抗体中的抗原互补位结合在其上。表位决定簇经常由化学活性的表面分子部分，例如氨基酸或糖侧链组成，并且经常具有特定的三维结构特性以及特定的电荷性质。此处所用的术语“隐藏的”是指一种性质或特点，它需要结构或构象的变化才能使该性质或特点变得明显；在不发生变化时，该特点或性质是“隐藏的”。隐藏的表位既可以存在于病毒外壳蛋白上也可以存在于受体多肽序列上。

术语“抗体”包括完整的分子及其片段，例如 Fab、F(ab')₂、和 Fv，它们能够结合到存在于此处描述的嵌合多肽的表位上。还包括了

其它的抗体片段，只要这些片段保留了与它的抗原选择性结合的能力。本发明的抗体片段（例如 Fab、F(ab')₂、和 Fv）可以通过抗体的蛋白切割作用来制备，例如通过用胃蛋白酶消化完整的抗体。与公开的嵌合多肽结合的抗体可以使用完整的嵌合多肽或其片段作为免疫抗原来制备。在嵌合多肽片段的情况下，优选的是病毒外壳多肽序列和受体多肽序列维持彼此结合的能力以便任何存在的隐藏的表位会被暴露。用于免疫动物的嵌合多肽来自被翻译的多核苷酸或者是化学合成的，并且如果需要，可以与载体连接。这样的常用的与免疫肽化学连接的载体包括例如匙孔血蓝蛋白（KLH）、甲状腺球蛋白、牛血清白蛋白（BSA）和破伤风类毒素。

单克隆抗体使用本领域的专业技术人员所熟知的方法来制备（Kohler 等，*Nature*, 256:495 (1975); 和 Harlow 等，*抗体：实验室手册*，726 页，冷泉港出版社（1988），在此引为参考）。简单地说，单克隆抗体的获得可以通过下面的步骤：用含有抗原的组合物注射小鼠，通过分析血清样品以证实抗体生产的存在，取脾获得 B 淋巴细胞，将 B 淋巴细胞与骨髓瘤细胞融合以产生杂交瘤细胞，克隆杂交瘤细胞，筛选产生针对抗原的抗体的阳性克隆，以及从杂交瘤细胞培养物中分离抗体。单克隆抗体可以通过多种成熟的技术从杂交瘤细胞培养物中分离和纯化，包括例如用蛋白 A Sepharose 亲和层析、孔径排阻层析和离子交换层析（参见例如 Coligan 等，“在兔、大鼠、小鼠和仓鼠中生产多克隆抗血清”，《现代免疫学方法》第 2.7.1-2.7.12 节和第 2.9.1-2.9.3 节；以及 Barnes 等，“免疫球蛋白 G (IgG) 的纯化”，《分子生物学方法》第 10 卷，79-104 页，Humana Press (1992)）。制备多克隆抗体的方法对于本领域的专业技术人员来说是众所周知的（参见例如 Green 等，“多克隆抗血清的生产”，《免疫化学方法》第 1-5 页，Manson 编辑，Humana Press (1992)；Harlow 等（1988），同上；以及 Coligan 等（1992），同上第 2.4.1 节，在此引为参考）。

为了治疗的目的，在一个物种中产生的针对嵌合多肽的抗体可以

被人源化，以便当抗体施用到宿主不会诱导免疫反应，例如进行被动免疫时。一般来说，人源化抗体是通过将非人类的恒定区用人类的恒定区取代而生产的。这样的抗体人源化方法在本领域内是众所周知的，在本发明的方法中特别有用(Morrison等, Proc. Natl. Acad. Sci. USA, 81:6851 (1984); Takeda等, Nature, 314:452 (1985); Singer等, J. Immunol., 150:2844 (1993))。

与嵌合多肽结合的抗体，特别是与隐藏的表位结合的抗体，可以在体外或体内（即受试对象体内）中和病毒。这样的抗体因此可以在体外或体内阻止或抑制病毒的感染，并且可以改善与感染相关的一些或全部症状。这样的抗体可以在一个受试对象体内生产然后导入另一个受试对象中，即被动免疫治疗。此外，与嵌合多肽结合的抗体，当在受试对象体内产生时，可以保护该受试对象抵抗感染或改善与感染相关的一些或全部症状。

因此，在本发明中提供了在受试对象中抑制、阻止和改善病毒感染的方法。在一个实施方案中，本发明的一种方法包括给受试对象施用有效量的与嵌合多肽结合的抗体，从而阻止或抑制了病毒在受试对象中的感染。在另一个实施方案中，本发明的一种方法包括给受试对象施用有效量嵌合多肽，从而在受试对象中产生了足够阻止或抑制病毒感染的免疫反应。在另一个实施方案中，本发明的一种方法包括给受试对象施用有效量的编码本发明的嵌合多肽的多核苷酸。在不同的情况下，嵌合多肽含有一种免疫缺陷病毒的外膜多肽，如本文所公开。

在施用一种嵌合多肽或编码嵌合多肽的多核苷酸以便在受试对象中抑制、阻止和改善病毒的感染的方法中，也可以产生免疫反应。免疫反应在本质上可能是体液的，尽管施用编码嵌合多肽的多核苷酸可以诱导 CTL 反应。还应该明白，如果合适，本发明的方法也可以与其它病毒疗法结合使用。

“有效量”对于抑制、阻止或改善受试对象中的病毒感染来说应该是足够的，或者对于在受试对象中产生免疫反应来说是足够的。因此，有效量的嵌合多肽可以是能够引发针对多肽或外壳蛋白所属的病毒的免疫反应所需的量。对于已经被病毒感染的受试对象施用的有效量也可以是减少病毒载量或增加 CD4+细胞数目所需的量。有效量可以是抑制病毒从一个受感染的受试对象传播到另一个受试对象（未被感染的或已被感染的）所需的量。

在本发明的给受试对象施用编码嵌合多肽的多核苷酸的方法中，可以产生针对嵌合多肽的 CTL 反应以抵抗含有相应外壳多肽序列的病毒。

因为本发明的嵌合多肽、多核苷酸和抗体将被施用给受试对象，包括人类，因此本发明也提供了含有公开的嵌合多肽、多核苷酸和抗体的药物制剂。被施用给受试对象的成分因此将包含在“可药用的”或“生理上可接受的”制剂中。

在此所用的术语“可药用的”和“生理上可接受的”是指能够施用到受试对象而最好不具有过份的不利的副作用（例如恶心、头痛等）的载体、稀释剂、赋形剂等。这样的给药制剂包括无菌水或非水性溶液、悬浮液和乳浊液。非水性溶剂的例子是丙二醇、聚乙二醇、植物油例如橄榄油、以及可注射的有机酯例如油酸乙酯。水性载体包括水、醇/水溶液、悬浮液或乳浊液，包括生理盐水和缓冲的介质。媒介物包括氯化钠溶液、任氏葡萄糖、葡萄糖和氯化钠、乳酸化的任氏溶液或硬化油。静脉内的媒介物包括流体和营养补充剂、电解质补充剂（例如基于任氏葡萄糖的那些补充剂）等。防腐剂和其它的添加剂也可以存在，例如抗微生物剂、抗氧化剂、螯合试剂和惰性气体等。多种本领域所熟知的适合于施用给受试对象的药物制剂可用于本发明的方法中（例如《Remington 药物科学》，第 18 版，Mack Publishing Co., Easton, PA (1990)；《The Merck Index》第 12 版，Merck Publishing Group,

Whitehouse, NJ (1996))。

控制作用的持续时间或者一种施用的成分的可控投送，可以通过将成分掺入到微粒或聚合的物质中而实现，例如聚酯、聚氨基酸、水凝胶、聚乙烯吡咯烷酮、乙烯-乙烯基乙酸酯、甲基纤维素、羧甲基纤维素、硫酸鱼精蛋白或乳酸交酯/乙醇酸交酯共聚物、聚丙烯交酯/乙醇酸交酯共聚物、或乙烯-乙烯基乙酸酯共聚物。成分释放的速度可以通过改变这些大分子的浓度或成分来控制。胶体分散系统包括大分子复合物、纳米胶囊、微球体、珠子和基于脂类的系统、包括水包油乳浊液、胶体微粒、混合胶束和脂质体。

通过本发明的方法施用的成分可以通过注射、随时间逐渐灌注、通过大丸剂给药（例如在针对由针头损伤引起的 HIV 感染的被动保护的情况下）或者通过一个微结构的可植入装置进行非肠道给药。成分也可以通过吸入、静脉内、腹膜内、肌内、皮下、腔内（例如阴道或直肠内）、经皮、局部的或血管内给药。成分可以分多剂量给药。为治疗、抑制或阻止病毒感染或传播、或者为诱导免疫反应所需的剂量或“有效量”，在优选情况下将足够改善感染的部分或全部症状，尽管对于许多病毒包括 HIV 的感染来说，阻止感染的发展或恶化也是令人满意的结果。有效的量可以被本领域的专业人员容易地确定（参见例如 Ansel 等, *Pharmaceutical Drug Delivery Systems*, 第 5 版, (Lea 和 Febiger (1990), Gennaro 编辑))。

本发明的嵌合多肽、多核苷酸和抗体也可用于诊断目的。例如，一种嵌合多肽，其中含有来自能够利用辅助受体来感染的病毒的病毒外壳多肽序列，可以被用来鉴定其表达的辅助受体与嵌合多肽的结合亲和性减少的受试对象。具有减少的结合亲和性的受试对象被该病毒感染危险可能降低了。此外，其表达的辅助受体与嵌合多肽的结合亲和性增加的受试对象，被该病毒感染危险可能增加了。采用这种方法，可以鉴定被病毒感染的危险减少或增加的受试对象。例如，受

试对象表达的 CCR5 或 CXCR4 辅助受体与含有 HIV gp120-CD4 的嵌合多肽的亲合性增加或减少时，其被 HIV 感染的危险性分别增加或降低了。因此，这样的方法对于预后的评估也是有用的；表达高结合亲和性的辅助受体的受试对象可能预后较差。

在此处公开的嵌合多肽具有能够利用辅助受体的病毒的病毒外壳多肽序列的情况下，这样的嵌合多肽可用于鉴定调节病毒与辅助受体的结合的试剂。这样的嵌合多肽也可用于鉴定在嵌合多肽内调节病毒外壳多肽序列和受体序列的分子内相互作用/结合的试剂。因此，所述的含有不能够利用辅助受体的病毒的外壳多肽的嵌合多肽，可以用于鉴定在嵌合多肽内调节外壳序列和受体序列的结合的试剂。

因此，本发明提供了鉴定能够调节病毒和病毒辅助受体之间结合的试剂的方法，以及鉴定能够调节病毒和病毒受体之间结合的试剂的方法。

在一个实施方案中，本发明的一个方法包括在存在和不存在测试试剂的情况下，将一个嵌合多肽与一个辅助受体多肽在允许嵌合多肽与辅助受体多肽结合的条件下相接触，然后检测在存在和不存在测试试剂的情况下的结合。在另一个实施方案中，本发明的一个方法包括将一种能够形成分子内复合物的嵌合多肽与一种测试试剂相接触，然后检测嵌合体内病毒外壳多肽序列和受体多肽序列之间的结合。因此在存在测试试剂的情况下结合量减少表明该试剂抑制了病毒和病毒辅助受体或受体之间的相互作用/结合。在存在测试试剂的情况下结合量增加表明该试剂刺激了病毒和病毒辅助受体或受体之间的相互作用/结合。

接触可以发生在溶液中、固相中、完整细胞上、或一个生物体例如非人类的灵长动物中。在各种实施方案中，病毒是一种免疫缺陷病毒例如 HIV，辅助受体是一种趋化因子例如 CCR5 或 CXCR4。能够利

用辅助受体穿过细胞的病毒的结合，对于后面的感染、病毒增殖、以及其最终导致的病理症状来说是一个关键的步骤。因此，在另一个实施方案中，提供了鉴定能够抑制病毒穿过细胞、感染、增殖的试剂的方法，以及鉴定能够改善与病毒感染相关的症状的试剂的方法。在本发明的一个鉴定这种试剂的方法中，测试试剂可以在嵌合多肽与辅助受体多肽接触之后加入，也可以在嵌合多肽与辅助受体多肽接触之前加入。

候选的试剂包括抗体、抗病毒剂、辅助受体多肽序列（例如来自CCR5或CXCR4）、肽模拟物或其活性片段。候选的试剂还包括大量的化学门类，包括有机分子例如分子量大于50但小于大约2500道尔顿的小有机化合物。候选的试剂包括与蛋白的结构相互作用、特别是氢键所必需的官能团，典型的包括至少一个氨基、羰基、羟基或羧基基团，在优选情况下至少两个化学官能团。候选的试剂通常包括用一个或多个上述的官能团取代的环状碳或杂环结构，和/或芳香或多芳香结构。候选的试剂在生物分子中也有发现，包括但不限于肽、糖、脂肪酸甾体、嘌呤、嘧啶、衍生物、结构类似物或其组合。

候选的试剂可以从广泛的来源获得，包括合成的或天然的化合物的库。例如，有大量的方法可用于随机地或定向地合成各种各样的有机化合物和生物分子，包括随机的寡核苷酸和寡聚肽的表达。此外，以细菌、真菌、植物和动物的抽提物形式存在的天然的化合物库也可以获得或容易地产生。另外，天然的或合成的库和化合物可以容易地通过常规的化学、物理和生物化学的方法进行修饰，并可用于产生组合文库。已知的药剂可得以进行随机的或定向的化学修饰，例如酰化、烷基化、酯化、酰胺化等，以产生结构类似物。在检测结合的方法中，一个或多个分子可以结合到一个标记物上，其中该标记物可以直接地或间接地提供一个可检测的信号。各种标记物包括放射性同位素、荧光剂、化学发光剂、酶、特异性结合分子、微粒例如磁性微粒等。特异性结合分子包括成对的试剂，例如生物素和链亲和素、地高辛和抗

地高辛等。对于特异性结合的成员来说，互补的成员一般将用一个根据现有的步骤提供检测的分子来标记。

多种其它的试剂可以被包括在分析方法中。它们包括象盐、中性蛋白例如白蛋白、去污剂等试剂，可用于促进最适的蛋白-蛋白结合和/或减少非特异性的或背景相互作用。改善分析效率的试剂，例如蛋白酶抑制剂、核酸酶抑制剂、抗微生物剂等，也可以使用。组分的混合物可以以任何能够提供所需的结合的次序加入。可以在任何适当的温度下保温，一般在 4°C 和 40°C 之间。保温的时间根据最适的活性加以选择，但也可以被最优化以促进快速的高通量筛选。一般来说，在 0.1 小时和 1 小时之间将是足够的。

在不同的实施方案中，如同此处描述的，病毒是一种免疫缺陷病毒，例如 HIV、HTLV、SIV、FeLV、FPV 或疱疹病毒。在其它的实施方案中，辅助受体是 CCR5、CXCR4、CCR-2b、CCR3、CCR8、V28/CX3CR1、US28（疱疹病毒编码的趋化因子类受体）、STRL33/BOB/TYMSTR、GPR15/Bonzo 或 GPR1 多肽序列。

通过此处描述的本发明的方法鉴定的试剂还可以对其在体外或体内抑制病毒结合或感染细胞的能力进一步加以测试。因此，本发明提供了鉴定能够抑制病毒对细胞感染的试剂的方法。本发明的一种方法包括将易受病毒感染的细胞与感染性的病毒粒子在存在和不存在测试试剂的情况下相接触，然后确定测试试剂是否抑制了病毒对细胞的结合或感染，因而鉴定出能够抑制病毒感染的试剂。在各种实施方案中，测试试剂在将细胞与感染性病毒粒子接触之前或之后加入。本方法也可以在任何适当的动物中进行，例如非人类的灵长动物。

此处描述的嵌合多肽也可用于鉴定新的辅助受体或表征蛋白作为辅助受体。使用这种方法，任何病毒的病毒感染和随后的病理可以被更好地理解，因此能够改善对感染的治疗。例如，一种鉴定新的辅助

受体或表征辅助受体功能的方法是双杂交系统，它可以通过其表达被相互作用的多肽诱导的报告蛋白的活化来检测蛋白-蛋白相互作用。因此，一种适当的嵌合多肽可以在酵母或哺乳动物双杂交系统中被用做饵序列，从而筛选文库以鉴定相互作用的蛋白，包括新的辅助受体。成熟的检测蛋白-蛋白相互作用的生化方法（例如柱层析、梯度离心、免疫共沉淀分析等）也适用于鉴定辅助受体或表征蛋白为具有潜在的辅助受体功能。

结合辅助受体的嵌合多肽也可用于鉴定辅助受体的结合位点。例如，通过产生辅助受体多肽片段，然后将该片段与适当的嵌合多肽相接触。接触可以在溶液中（例如共沉淀）、固相中（例如亲和柱）、或在完整细胞上（例如将辅助受体与细胞表面接触，然后检测辅助受体片段是否抑制了嵌合多肽与细胞的结合）进行。辅助受体结合位点一旦被鉴定，就可以例如用做一种抗病毒剂来治疗感染。

除非另有定义，此处所用的所有技术和科学术语都与本发明所属领域的专业技术人员所通常理解的意义相同。尽管与本文描述的方法和材料相似或等价的方法和材料都可以用于本发明的实践或试验中，在下面仍然描述了适合的材料和方法。本文提到的所有出版物、专利申请、专利和其它参考文献以其全文引为参考。在有冲突的情况下，以本文的说明包括定义为准。此外，材料、方法和实施例仅仅是为了说明而不具有任何限制。本发明的其它特点和优点在下面的详细描述中和权利要求中将变得明显。在下面的实施例中对本发明进行了进一步的描述，它对权利要求中所述的本发明的范围没有限制。

实施例 1

本实施例描述了编码单链 gp120-CD4 嵌合多肽 FLSC、TsSC、FLSC-R/T 和 RLSC-R/T CD4M9 的多核苷酸的构建。构建单链复合物的策略是基于在 gp120 的 C 端和 CD4 的 N 端之间放置一种 20 到 30 个氨基酸的连接序列。使用 Swiss PDB Viewer 对与可溶性 CD4 和 17b

Fab 结合的修饰的 gp120 的晶体结构 (Dwong, P.D.等, Nature, 393:648-59 (1998)) 进行分析, 结果建议嵌合分子应该能够发生分子内相互作用导致形成 gp120-CD4 复合物。通过以下列次序安排相应的编码序列, 构建了编码 gp120-CD4 嵌合多肽 (SEQ ID NO:1) 的单链核酸: (1) 在 5'端是合成的密码子, 编码巨噬细胞向性的 HIV BaL 的 gp120; (2) 一种编码 20 个氨基酸接头的序列, 其由甘氨酸、丙氨酸和丝氨酸组成; (3) 可溶性 CD4 结构域 1 和 2 (D1D2) 的序列; 以及 (4) 对于 FLSC 来说, 在 3'末端是编码一种来自 c-myc 癌基因的短多肽的序列。FLSC-R/T 核苷酸序列 (SEQ ID NO:3) 编码一种在 gp120 的 C 末端发生了一个突变的蛋白, 其中精氨酸被苏氨酸代替 (SEQ ID NO:4)。FLSC-R/T CD4M9 (SEQ ID NO:5) 还包括了本发明的嵌合多肽 (SEQ ID NO:6) 的核苷酸序列中的变化, 其中 CD4 D1D2 区被一个编码 CD4M9 的序列代替, 该序列编码了一种模拟 CD4 D1D2 区域的功能活性的肽。使用了密码子被最适化的 gp120 序列, 因为它能够以回转-独立的方式进行高水平的表达 (Haas, J.等, Curr. Biol. 6:3 15-24 (1996))。所用的人 CD4 序列来自 T4-pMV7 (Maddon, P. J.,等, Cell, 47:333-48 (1986); 国立卫生研究院爱滋病试剂库, Bethesda, MD)。myc 多肽序列使得嵌合多肽能够进行方便的分析、纯化及其它操作。

含有这些不同序列的完整的多核苷酸通过 PCR 产生并插入到 pEF6 中 (Invitrogen), 使用强的延伸因子启动子 (EF 1) 驱动表达。在这个构建体 (命名为 pEF6-SCBal) 中导入了限制性内切酶位点, 使得能够方便地与其它免疫缺陷病毒的其它外膜基因进行交换。

简单地说, 使用质粒 pMR1W1-9 和 T4-pMV7 为模板通过 PCR 构建了 FLSC 分子。gp120 的正向引物是 GGG-GGT-ACC-ATG-CCC-ATG-GGG-TCT-CTG-CAA-CCG-CTG-GCC (SEQ ID NO:7), 而反向引物是 GGG-TCC-GGA-GCC-CGA-GCC-ACC-GCC-ACC-AGA-GGA-TCC-ACG-CTT-CTC-GCG-CTG-CAC-CAC-GCG-GCG-CTT (SEQ ID

NO:8)。CD4 的正向引物是 GGG-TCC-GGA-GGA-GGT-GGG-TCG-GGT-GGC-GGC-GCG-GCC-GCT-AAG-AAA-GTG-GTG-CTG-GGC-AAA-AAA-GGG-GAT (SEQ ID NO:9)，而反向引物是 GGG-GTT-TAA-ACT-TAT-TAC-AGA-TCC-TCT-TCT-GAG-ATG-AGT-TTT-GTT-CAG-CTA-GCA-CCA-CGA-TGT-CTA-TTT-TGA-ACT-C (SEQ ID NO:10)。使用 Kpn1 和 Pme1 限制性内切酶位点将 PCR 产物亚克隆到 pEF6 中 (Invitrogen, Carlsbad, CA)。

为了构建 pEF6-TcSC 质粒，pEF6-FLSC 中的全长 gp120 表达的序列被一个截短的 gp120 序列 (DC1DC5DV1V2) 代替。截短的 gp120 使用 GGG-GGT-ACC-ATG-CCC-ATG-GGG-TCT-CTG-CAA-CCG-CTG-GCC-ACC-TTG-TAC-CTG-CTG-GGG-ATG-CTG-GTC-GCT-TCC-TGC-CTC-GGA-AAG-AAC-GTG-ACC-GAG-AAC-TTC-AAC-ATG-TGG (SEQ ID NO:15) 作为正向引物，以及 GGG-GGA-TCC-GAT-CTT-CAC-CAC-CTT-GAT-CTT-GTA-CAG-CTC (SEQ ID NO:16) 作为反向引物而产生。使用 CTG-TGC-GTG-ACC-CTG-GGC-GCG-GCC-GAG-ATG-AAG-AAC-TGC-AGC-TTC-AAC-ATC-GGC-GCG-GCC-CGC-CTG-ATC-AGC-TGC (SEQ ID NO:17) 作为正向引物，以及 GCA-GCT-GAT-CAG-GCG-GCC-CGC-GCC-GAT-GTT-GAA-GCT-GCA-GTT-CTT-CAT-CTC-GCC-CGC-GCC-CAG-GGT-CAC-GCA-CAG (SEQ ID NO:18) 作为反向引物来缺失 V1 和 V2 区域。

用于克隆到 FLSC R/T CD4M9 中的 CD4M9 序列 (SEQ ID NO:19) 通过使用 5'-3' 的引物 GCG-GCC-GCT-TGC-AAC-CTG-GCC-CGC-TGC-CAG-CTG-CGC-TGC-AAG-AGC-CTG-GGC-CTG-CTG-GGC-AAG-TGC-GCC-GGC-AGC-TTC-TGC-GCC-TGC-GGC-CCC-TAA-GAA-TTC (SEQ ID NO:21) 作为正向引物，以及 GAA-TTC-TTA-GGG-GCC-GCA-GGC-GCA-GAA-GCT-GCC-GGC-GCA-CTT-GCC-CAG-CAG-GCC-CAG-GCT-CTT-GCA-GCG-CAG-CTG-GCA-GCG-GGC-CAG-GTT-GCA-AGC-GGC-CGC (SEQ ID NO:22) 作为反向引

物并一起退火而产生。片段用 Not1 和 BamH1 酶切，然后亚克隆到用 Not1 和 BamH1 酶切制备并经凝胶纯化以从 FLSC R/T 序列中除去释放出的 hD1D2 的 pEF6-FLSC R/T 中。克隆通过测序确认。

重组的结构在图 1 中显示。含有 BaL gp120 (SEQ ID NO:24) 序列与间隔区域 (SEQ ID NO:11) 和 CD4D1D2 区域 (SEQ ID NO:26) 的嵌合重组子被命名为全长单链 (FLSC)。第二个结构被设计以产生更类似于用来解析 gp120 晶体结构的分子的复合物的。该结构被命名为截短的单链 (TcSC)，其构建除了用编码 $\Delta C1\Delta C5\Delta V1V2$ gp120 的序列代替全长的编码序列 (SEQ ID NO:28) 以外与 FLSC 的构建相同。同时还显示了被命名为 FLSC-R/T 的结构，其中的 BaL gp120 在 506 位的氨基酸被突变 (SEQ ID NO:30)，FLSC-R/T CD4M9 含有序列 SEQ ID NO:30 和 20。在本实施例中显示的间隔区域的氨基酸序列为 GSSGGGSGSGGGGSGGGAAA (SEQ ID NO:11)。

实施例 2

本实施例描述了用编码 gp120-CD4 嵌合多肽的多核苷酸转染细胞以及表达的可溶性多肽的性质。使用 Fugene，按照供应商的方案 (Boehringer-Manheim)，将重组的 pEF6-FLSC 和 pEF6-TcSC 转染到 293 细胞中。使用 5 μ g/ml 灭瘟素筛选获得稳定的转染子。稳定的细胞系 (293-SC) 在不同条件下培养，然后通过免疫印迹分析评估嵌合多肽的生产，使用抗 gp120 单克隆抗体 (Y. H. Abacioglu 等, AIDS Res. Hum. Retroviruses, 10:371-81 (1994)) 或抗人 CD4 多克隆血清 (T4-4) (K. C. Deen 等, Nature, 331:82-4 (1998); R. L. Willey 等, J. Viral., 66:226-34 (1992); NIH AIDS Reagent Repository) 的混合物。

简而言之，含有嵌合多肽的细胞培养上清液被收集，并在 SDS-PAGE 载样缓冲液 (75 mM Tris, 2% SDS, 10%甘油, 0.001%溴酚蓝, pH8.3) 中煮沸。然后将样品在 4-20%的 SDS-聚丙烯酰胺梯度凝胶中电泳。接着将凝胶分级的蛋白转移到硝酸纤维素膜上。在膜上的非特

异性结合位点用 2%脱脂奶粉的 Tris 缓冲盐, pH7 封闭 30 分钟。然后用抗 CD4 多克隆兔血清(T4-4; NIH AIDS Reagent Repository, Bethesda, MD) 或鼠抗 HIV gp120 的单克隆抗体的混合物来探测膜。如图 2 所示, 转染的细胞表达了预期大小(150kD)的可溶性蛋白。该多肽与抗 gp120 和抗 CD4 抗体都能够反应, 因此代表了完整的嵌合多肽。

在其他研究中, 检测了与抗 myc 抗体的反应性, 从而进一步验证了该 150kD 物种作为嵌合多肽的身份。除了该多肽, 与 gp120 和 CD4 D1D2/myc 标志的预计大小相匹配的条带也被观察到, 表明一部分的嵌合多肽已经在间隔区被切割。加入生物相容的蛋白酶抑制剂(Pefabloc; Boehringer-Mannheim) 获得了基本上未被切割的嵌合多肽分子。这表明 gp120-CD4 的切割是被一个丝氨酸蛋白酶催化的。测定了 293-SC 细胞系产生的 gp120-CD4 嵌合多肽的量, 使用抗 gp120 捕获 ELISA 和绵羊抗 gp120 抗体 D7324 (International Enzymes)、绵羊抗 gp120 C5 区中的一个高度保守的表位的多克隆 IgG (J. P. Moore 等, AIDS, 4:297-305(1990); J. P. Moore 等, J. Virol., 67:863-75 (1992); J. P. Moore 等, AIDS, 4:307-15 (1990)) 和一个 gp120 标准曲线。

简而言之, 2 μ g/ml D7324 的磷酸盐缓冲液被吸附在一个塑料板上。非特异性结合位点用 2%脱脂奶粉的缓冲盐封闭。然后将来自 293-SC 细胞系的饱和浓度的细胞培养上清液加到板上。使用来自感染了 HIV 的病人的失活的人血清和与辣根过氧化物酶结合的抗人 IgG 来检测被捕获的嵌合多肽。据估算 293-SC 细胞系分泌大约 3 μ g/ml 的 gp120-CD4 嵌合多肽。293-SC 细胞系已经适应了无血清条件中培养。因为免疫印迹研究表明 gp120-CD4 嵌合多肽有部分被切割了, 因此将纯化的单链样品进行了交联, 然后分析交联的样品以确定 gp120 和 CD4 分子是否仍保持相连。简单地说, 使用免疫亲和柱纯化了 293-SC 细胞系产生的来自上清液的单链 gp120-CD4。该柱通过将抗 gp120 人单克隆抗体 A32 连接到 CDBr 活化的 sepharose 4B (Amersham-Pharmacia Biotech, Piscataway, NJ) 上而形成。A32 对 gp120 上的高度

不连续的表位具有特异性，并倾向于识别与 CD4 结合的外膜。结合的 gp120-CD4 用 pH2.5 的 0.1M 乙酸洗脱、冷冻干燥、对 PBS 透析。蛋白的浓度通过 BCA 分析方法 (Bio-Rad, Hercules, CA) 按照生产商的步骤测定。然后将 20 μ l 等份试样的纯化的 gp120-CD4 与 1mM 同型双功能交联剂 BS3 溶液交联，与未交联的 gp120-CD4 一起在 4-20%聚丙烯酰胺凝胶上电泳。分级的蛋白被转移到硝酸纤维素上，用抗 gp120 单克隆抗体的混合物进行免疫印迹，接着用碱性磷酸酶标记的抗小鼠 IgG 反应，最后用商业化的 BCIP/NBT 混合物 (KPL) 显色。

图 3 显示了这些研究的结果；未交联的 gp12-CD4 在第 1 道，交联的 gp120-CD4 在第 2 道。第 1 道显示了免疫亲和柱纯化了切割的和未切割的单链 gp120-CD4。交联，如第 2 道中所示，产生了两个宽的带，位于 150kDa 和 300kDa，这种形式表明单链 gp120-CD4 在溶液中以结合的 150kDa 分子的形式存在。gp120 和 CD4 亚基即使在切割发生后仍保持连接。300kDa 的带表明一部分的 gp120-CD4 在溶液中是二聚体，可能代表了通过在不同分子的外膜和 CD4 结构域之间的分子间相互作用而连接起来的单链分子。在某些条件下 (图 2) 单链分子明显切割成 gp120 和 CD4 部分，对于 DNA 疫苗来说可能需要关注，因为这样的加工在体内有可能发生。但是，这些研究表明，尽管有切割，单链分子仍保持连接作为 gp120-CD4 复合物 (图 3)。为了更详细地研究天然 FLSC 的结构性质，不同浓度 (1 μ M-0.03 μ M) 的同样上述研究的蛋白制备液在 PBS 中共价交联以便固定溶液中存在的任何多聚体结构。然后将交联的材料通过免疫印迹分析方法使用抗 CD4 抗体进行分析。如图 4 所示，一条 172kD 的主要蛋白带 (插页；带 A) 始终可见，另外有两条具有较高分子量的弱带。一条较弱的条带 (插页；带 B) 的表观大小大约为 302kD，而另一条 (插页；带 C) 在胶中迁移得不够远，因而无法通过 SDS-PAGE 准确地估计大小。不同蛋白带的出现和比例不依赖于交联前 FLSC 的浓度。因此，密度计分析表明带 A、B 和 C 总是分别代表了大约 65%、25%和 10%的总蛋白。

与 FLSC 相比，交联的 TcSC 的色谱图更加复杂。在非变性条件下洗脱的 TcSC 是一个广泛系列的峰群，范围从 166kD 到 353kD。这样的图形表明较短的 TcSC 多肽在表达和/或纯化中形成了多种较高级的结构。这种行为表明，TcSC 主要以可变大小的多肽链的形式存在，它们通过不同分子上的 gp120 序列和 CD4 序列之间的相互作用连接在一起。因为 TcSC 是通过从 gp120 缺失了 20 个 C 端氨基酸而产生的，CD4 核心结构与 gp120 的 CD4 结合结构域 (CD4bd) 之间的距离缩短了，这可能阻碍了 TcSC 完成分子内 gp120-CD4 相互作用的能力，因此倾向于形成链间的复合物。然而，TcSC 也表现出 gp120-CD4 复合物的抗原和功能特性。可能因为分子间的相互作用涉及了多个 TcSC 分子，较少比例的总蛋白表达了能够与表面辅助受体相互作用的辅助受体结合位点。此外，在 TcSC 中缺失 V1/V2 结构域可以减少 BaL 外膜对 CCR5 的相对亲和性。对 TcSC 进行进一步的修饰延长了 gp120 和 CD4 部分之间的接头，可以形成较高比例的链内复合物。是否由于 TcSC 的多聚体性质使得该分子不如 FLSC 具有优势，仍然是一个未解决的问题，因为用其它多聚体分子进行的研究表明它们与其单体对应物相比是更具潜力的免疫原(A. L. DeVico 等, AIDS Rev., 1:4-14 (1999); S. A. Jeffs 等, J. Gen. Virol, 77:1403-1410 (1996); R. A. LaCasse 等, Science, 283:357-62 (1999))。

实施例 3

本实施例描述了证实 gp120-CD4 嵌合多肽与几种能够与 gp120 和 CD4 反应的不同抗体结合的数据。gp120 与 CD4 的结合引起分子中的构象变化，导致辅助受体结合结构域的暴露。因此，直接针对这个结构域中的表位的抗体将强烈地与适当折叠的单链分子反应。为了确定嵌合分子中暴露的表位，比较了 FLSC 和 TcSC 分子的抗原性质。纯化的 FLSC 和 TcSC 通过抗原捕获 ELISA 进行免疫化学分析。简单地说，使用一种纯化的针对来自 gp120 的 C 端 15 个氨基酸的肽 D7324 (J. P. Moore 等, AIDS Res. Hum. Retroviruses, 4:369-79 (198X)) 的多克隆羊抗体 (International Enzymes, Fallbrook, CA) 将 BaLgp120、

gp120-rsCD4 复合物或单链嵌合分子捕获，吸附到基体上。D7324 在 PBS 中稀释到 $2\mu\text{g/ml}$ ，通过在室温保温过夜吸附到 96 孔板上(Maxisorb plates, VWR Scientific, St. Louis, MO)。用 BLOTTO (5%脱脂奶粉的 Tris 盐缓冲液) 处理板以防止孔中的非特异性结合。用 TBS 洗板后，样品稀释在 BLOTTO 中， $200\mu\text{l}$ 的等份样品加到两个 D7324 包被的孔中，在室温保温 1 小时。用在 BLOTTO 中稀释 1000 倍的灭活的 HIV-I 与血清的混合物，然后再用辣根过氧化物酶标记的山羊抗人 IgG 来检测结合的抗原。

检测也可以使用以前显示的在衔接 CD4 后倾向于与 gp120 结合的单克隆抗体(MAbs A32、17b 和 48d)(M. Thali 等, *J. Virol.*, 67:3978-86 (1993))，然后使用适当标记的第二抗体来完成。两个抗体 17b 和 48d 结合在由 CD4 结合诱导的辅助受体结合位点中(N. Sullivan 等, *J. Virol.*, 72:4694-703 (1998); A. Trkola 等, *Nature*, 384:184-6 (1996); L. Wu 等, *Nature*, 384:179-183 (1996))。识别游离的 gp120 C1-C5 区域当中的保守表位的抗体 C11 也被测试。抗体在 BLOTTO 中稀释，在室温保温 1 小时。在每次保温步骤之间板用 TBS 洗 3 次。样品中存在的 gp120 序列的量根据用商品化的重组 HIV IIIB gp120 (Bartels, Issaquah, WA) 制作的标准曲线确定。在包含 BaLgp120-rsCD4 复合物的对比研究中，用饱和浓度的 gp120 处理 D7324 包被的板。在洗孔后，在孔中加入浓度过量的 rsCD4 ($1\mu\text{g/ml}$)，保温 1 小时以形成复合物。为了对缺少 D7324 表位的 TcSC 抗原进行评估，开发了另一种使用抗 CD4 MAb 45 (Bartels, Issaquah, WA) 捕获的 ELISA 形式。 $1\mu\text{g/ml}$ 的抗体被吸附到塑料孔中，并用 BLOTTO 封闭孔。然后如前所述用指示的人血清或人单克隆抗体进行分析。

如图 5A 所示，所有的抗体都能够与 FLSC 强烈地反应。但是，抗体 17b、48d 和 A32 与 FLSC 的半数最大结合浓度总是比与单独的 gp120 的半数最大结合浓度高，并与用可溶性非共价的 BaLgp120-rsCD4 复合物观察到的半数最大结合浓度相当。FLSC 较高的免疫反应性

特异于针对抗 CD4 诱导的表位的抗体，因为抗体 C11 与 FLSC 和抗体 C11-5 自由 gp120 的半数最大结合浓度没有显著差异。

如图 5B 所示，17b 和 48d 与 TcSC 反应性的水平与在平行分析实验中与 FLSC 反应所观察到的水平相当。正如预计的那样，抗体 C11 和 A32 不与 TcSC 反应，因为在 TcSC 的结构中它们相应的表位部分被缺失了。

在单链分子中 gp120 和 CD4 序列的结合也将阻断 CD4 结合位点中的表位在 gp120 上的暴露。为了证实这样的结合已经发生，以致 gp120 中的 CD4 结合位点不再能够用于结合，使用 Mab45 捕获方式以及一系列针对 gp120 上 CD4 结构域 (CD4bd) 的单克隆抗体 (IgG1b12、F91 和 205-469) 对 FLSC 和 TcSC 进行了评估。

如图 5C 所示，这些抗体中没有一个是能与 FLSC 或 TcSC 反应，尽管在用收集的 HIV+血清的平行试验中观察到了阳性反应性。这个数据表明在 FLSC 和 TcSC 分子内存在 CD4 序列和 gp120 CD4 结合结构域之间的相互作用。

总而言之，这些结果证明，gp120-CD4 嵌合多肽的反应性可以与使用由可溶性的 gp120 和 CD4 (未交联的) 结合形成的复合物所观察到的可比，并比单独使用 gp120 的高。这些数据表明，单链 gp120-CD4 分子形成的相互作用复合物与过渡态 HIV 外膜-CD4 复合物相似。被捕获的 gp120-CD4 在其它 ELISA 研究中也与抗 CD4 抗血清和抗 myc 抗体反应，与 western 印迹分析的结果一致。合在一起，这些数据表明大部分的单链 gp120-CD4 分子代表了适当折叠的 gp120-CD4 复合物。

实施例 4

本实施例描述了能够证明含有一种 CCR5 特异性的 HIV 外膜序

列的 gp120-CD4 嵌合分子与 CCR5 表达细胞结合的数据。

gp120-CD4 复合物的形成通常暴露出与适当的辅助受体相互作用的外膜结构域 (M. Thali 等, *J. Virol.*, 67:3978-86 (1993); M. A. Vodicka 等, *Virol.*, 233:193-8 (1997))。因此, 对 gp120-CD4 复合物的适当折叠及其抑制病毒对细胞感染的能力的另一种度量是与 CCR5 辅助受体结合的能力。

为了评估单链复合物与辅助受体结合的能力, 纯化的单链 gp120-CD4 分子被允许与表达 CCR5 或 CXCR4 的细胞相互作用。简单地说, 通过用 pEF6-SC 短暂转染 293 细胞产生了含有 gp120-CD4 单链的上清液。然后将上清液加入到 A32 免疫亲和柱上, 用 0.2M pH2.5 的乙酸洗脱纯化的单链, 如前所述用 D7324 捕获 ELISA 和免疫印迹进行分析。含有单链的级分被收集, 平衡到 pH7, 然后浓缩。

为了进行结合, 将纯化的单链制备物与表达 CCR5 的 L1.2 细胞相互作用 (L. Wu 等, *Nature*, 384:179-183 (1996); L. Wu 等, *J. Exp. Med.*, 186:1373-81 (1997))。不表达辅助受体、CXCR4 或 CCR5 的鼠 B 细胞系 L1.2、L1.2/X4 和 L1.2/R5 细胞与浓度逐渐减少的纯化的单链蛋白混合。在 37°C 保温 1 小时后, 洗涤细胞。用 1 μ g/ml MAb C11 (J. E. Robinson 等, *J. Cell. Biochem. Suppl.*, 16E:71 (1992); M. Thali 等, *J. Virol.*, 67:3978-86 (1993)) (一种抗 gp120 的 MAb), 然后用一种荧光分子藻红蛋白标记的抗人 IgG 来检测结合的单链分子。C11 识别由 C1-C4 区域形成的构象决定簇。结合的荧光素的水平通过荧光素激活的细胞分类 (FACS) 分析方法使用 FACS Calibur 仪器 (Becton Dickinson) 进行测定。使用 Cell Quest 3.1.3 程序 (Becton Dickinson) 计算每个样品的平均荧光强度。

如图 6 所示, 两种单链 gp120-CD4 复合物 (FLSC 和 TcSC) 都与表达 CCR5 的 L1.2 细胞结合, 但不与表达 CXCR4 的 L1.2 细胞结合。

用 FLSC 观察到的最大结合位于浓度为 $10\mu\text{g/ml}$ 时，这与使用可溶性 BaL gp120-rsCD4 复合物进行对照试验时观察到的相当。与此相比，为了达到饱和的结合，需要高大约 10 倍浓度的 TcSC。因此，gp120-CD4 嵌合多肽表现出对于 CCR5 的功能性辅助受体结合位点，如同对于含有巨噬细胞向性的 gp120 的分子所预计的那样。

考虑到在 gp120-CD4 嵌合体中 HIV 外膜多肽对 CD4 的明显的特异性，在这些研究中不与 CXCR4 结合并不是完全没有预料到的。因此，如同本文描述的，通过构建能够结合 CXCR4 或其它辅助受体的多肽嵌合体、或者通过对病毒外壳多肽进行修饰，以获得能够与另一辅助受体结合的嵌合多肽，则可以获得能够与其它辅助受体结合的其它病毒外壳多肽-受体多肽嵌合体。

为了证实单链 gp120-CD4 与 CCR5 的结合是通过它的辅助受体结合位点进行的，使用 17b 和 48d 抗体进行了竞争结合研究，这两种抗体已被显示能够与 gp120 的辅助受体结合位点相互作用并阻止 gp120/sCD4 复合物与辅助受体表达细胞相互作用。作为对照，使用了另一种 gp120 抗体 C11 和一种 gp41 抗体 F240。所有这些抗体都来自 HIV-1 感染的病人。每种抗体的使用浓度为 $10\mu\text{g/ml}$ ，与 $3\mu\text{g/ml}$ 纯化的单链分子一起加入到表达 CCR5 或 CXCR4 的 L1.2 细胞中。结合的 gp120-CD4 先用 C11、再用 PE 标记的抗人 IgG 检测。gp120-CD4 的量通过 FACS 测定，被表示为在没有竞争抗体的对照孔中总结合量的百分数。

如图 7 所示，17b 和 48d 强烈地抑制了两种单链复合物与细胞的结合。在存在这些抗体的情况下，在 CCR5 表达细胞上的结合信号与用 L1.2/CXCR4 和 L1.2 亲本细胞时看见的背景结合相同。有趣的是，2G12（一种潜在的中和抗体）也减少了所有复合形式与 CCR5 的相互作用。与此相比，能够识别辅助受体结合结构域之外的表位的抗 gp120 抗体 C11、A32 和抗 gp41 抗体 F240，均不能减少 FLSC 或 TcSC 与表

达 CCR5 的 L1.2 细胞的结合。

这些结果表明 gp120 辅助受体结合位点对于与辅助受体的结合是重要的。这些结果还表明，可以使用这样的分析方法来鉴定能够抑制 gp120-CD4 和辅助受体之间的结合/相互作用的试剂。这样的试剂具有作为治疗剂的潜在的价值。

总而言之，这些数据证实了复制病毒外壳-受体复合物的过渡态构象的可溶的嵌合多肽得到了成功的表达。有了这个成功，现在有可能利用嵌合多肽或编码多肽的多核苷酸来免疫接种受试对象，以产生针对病毒或具有同样的外壳多肽表位的病毒的免疫反应。产生的免疫反应可以是抗体（体液的）或 CTL 反应。此外，既然嵌合多肽可以与活细胞表面上适当的辅助受体结合，那么多肽就可以施用给急性暴露于免疫缺陷病毒的受试对象，以被动地保护表达辅助受体的细胞不受病毒感染。

实施例 5

本实施例描述的数据可以证实 gp120-CD4 嵌合分子可以中和由使用同样的辅助受体的 HIV 株引起的感染。单链分子被进一步检测它们中和 R5 和 X4 病毒的能力。总数为 10^4 的表达 CCR5 或 CXCR4 的 U373/CD4/MAGI 细胞 (M.A. Vodicka 等, *Virology*, 233:193-8 (1997)) 在平底组织培养孔中附着过夜。然后取出培养基并用 100 μ l 新鲜的含有不同浓度嵌合蛋白的培养基代替。然后向培养液中加入 100 μ l 含有 50 TCID₅₀ 病毒的培养基。然后将完整的混合物在 37°C 下保温直到可以观察到合胞体，一般在 3-5 天内。然后用一种 P-半乳糖苷酶化学发光试剂 Galatostar (Tropix, Bedford, MA)，根据制造商的流程处理培养孔。病毒感染作为化学发光的函数被测定，使用 Victor² 荧光读板器 (EG&G Wallac, Gaithersburg, MD) 进行定量。在不含病毒的情况下进行分析以确定背景信号。从测试分析获得的信号通过减去背景值来校正。用每个实验孔的校正的相对光单位除以只含有细胞和病毒的对

照孔的校正的光单位计算出感染的百分数。根据用测试蛋白的浓度对抑制感染的百分数所作的图确定 90%抑制剂量 (ID_{90}) 值。所有的试验条件被重复进行 3 次。

如图 8 所示, FLSC 和 TcSC 都潜在地并选择性地中和 R5 HIV-1 BaL 分离株, 而对 2044 分离株仅有轻微的抑制 ($ID_{90}>10\mu\text{g/ml}$)。与此相比, 未复合的 BaLgp120 由于直接与 CD4 相互作用, 正如预期的那样抑制了 HIV-1BaL 和 X4 (HIV-12044) 两种病毒的进入。因此, 这些数据证实了病毒外壳多肽-受体嵌合分子可以结合到细胞辅助受体上, 从而阻断了利用该辅助受体结合或感染的病毒对细胞的结合或感染。

实施例 6

本实施例描述了一种含有免疫球蛋白多肽序列的修饰的 gp120-CD4 嵌合多肽 gp120-CD4-IgG1 的构建和表达。这个示例性的异源结构域为 gp120-CD4 嵌合多肽添加了功能, 包括粘附和免疫刺激功能、延长的稳定性、增加的循环半衰期和透过胎盘屏障的能力。这个实施例也显示了 gp120-CD4-IgG1 嵌合体结合到在完整细胞表面表达的受体上并中和 HIV 病毒。Gp120 是 HIV-I 外膜蛋白的一个亚基, 结合到 CD4 上, 经历了一种允许复合物与辅助受体例如 CCR5 相互作用的构象的变化。这种相互作用允许 HIV-1 感染进入 CD4+细胞。能够干扰 HIV-1 与辅助受体的相互作用的抗体或其它试剂可以防止感染。

为了鉴定这样的试剂, 对单链 gp120-CD4 进行了修饰, 将其融合到形成 IgG1 的重链、铰链 CH2 和 CH3 的恒定区 (图 9)。Gp120-CD4-IgG1 可以被用于鉴定能够阻断、抑制或破坏 HIV-1 与辅助受体相互作用的试剂, 因而能够鉴定可以抑制 HIV 感染的试剂。含有 SEQ ID NOs:24、11、26 和 32 的 gp120-CD4-IgG1 多肽也可以用作被动免疫治疗剂, 以在急性暴露例如针头伤害后防止 HIV 感染。

用含有包括至少 SEQ ID NOs:23、25 和 31 的 gp120-CD4-IgG1 的质粒短暂地转染 293 细胞，通过对培养上清液的免疫印迹表征了表达的蛋白。简单地说，收集的上清液样品在 4-20%的梯度 PAGE 凝胶上电泳。分级的蛋白被转移到硝酸纤维素膜上，并用抗 gp120 单克隆抗体的混合物检测。如图 10 所示，短暂转染的细胞表达了 gp120-CD4-IgG1（第 1 道）。表达来自 HIV-1 BaL 的纯 gp120 的细胞的上清液被电泳以便进行相对大小的比较（第 2 道）。gp120-CD4-IgG1 多核苷酸编码的蛋白具有 gp120-CD4-IgG1 重链嵌合体的预期大小。象原来的 gp120-CD4 一样，gp120-CD4-IgG1 的一部分被切割产生了一个 120kDa 的蛋白片段，它最可能是 gp120（“切割的 gp120”）。这个片段的大小表明 gp120-CD4-IgG1 是在间隔区内被切割的。为了证实 gp120-CD4-IgG1 被折叠成允许结合辅助受体的构象，将稀释的上清液加入到表达 CCR5 或 CXCR4 辅助受体的 L1.2 细胞中。结合的 gp120-CD4-IgG1 用被一种荧光试剂标记的抗人 IgG 来检测。荧光的量与结合的物质直接相关。

如图 11 所示，gp120-CD4-IgG1 特异性地与表达 CCR5 的 L1.2 细胞结合。又一次，使用这种分析方法几乎检测不到与 CXCR4 的结合，这同 gp120-CD4 的结果一致。这些研究表明能够提供附加的或增强的功能性的异源结构域可以被加到嵌合分子中而不影响它们形成与细胞辅助受体结合的复合物的能力。为了证实嵌合 gp120-CD4-IgG1 重链与 CCR5 表达细胞的结合是由 gp120 的辅助受体结合位点介导的，在存在封闭抗体 17b 的情况下对结合进行了研究。简单地说，对于 MAb/FLSC-IgG1 竞争研究来说，丁酸钠活化的表达辅助受体的 L1.2 细胞以 10^5 个/孔的浓度加到 V 形底的板中。10 μ g/ml FLSC-IgG1 和 1 μ g/ml MAbs 被加入到细胞中。细胞和蛋白在 37 $^{\circ}$ C 下一起保温 1 小时。离心收集细胞并用 TBS 洗 3 次。结合的物质用藻红蛋白标记的抗人 IgG 以 5 μ g/ml 的浓度在 4 $^{\circ}$ C 反应 1 小时来检测。细胞用 TBS 洗 3 次，然后通过荧光激活的细胞分类（FACS）来分析。

如图 12 所示, 17b (一种识别 gp120 上的 CCR5 结合结构域的抗体) 阻断了 FLSC-IgG1 与 L1.2R.5 细胞的相互作用, 而对照抗体 F240 则不能。这些数据证实了 FLSC-IgG1 与 R5 辅助受体的相互作用是通过 gp120 上的 R5 结合结构域。为了证实嵌合的 gp120-CD4-IgG1 重链能够阻断病毒进入细胞, 然后进行了中和分析。简单地说, 表达 CCR5 或 CXCR4 的 U373/CD4/MAGI 细胞以 10^4 个/孔的浓度在平底组织培养板中附着过夜。取出培养基, 并向细胞中加入 $100\mu\text{l}$ 含有不同浓度的 MAbs 和免疫粘附素。然后加入病毒 (每个孔加入 $100\mu\text{l}$ 含有 50 TCID₅₀ 病毒的培养基), 将混合物在 37°C 下保温直到可以观察到合胞体, 一般需要 3-5 天。如前所述使用一种 P-半乳糖苷酶化学发光试剂 Galatostar, 根据制造商的操作流程读板, 产生的化学发光用 Victor² 荧光读板器进行定量。用每个实验孔减去无病毒的背景孔的相对光单位除以有病毒无蛋白的孔减去背景孔的相对光单位, 计算出病毒生长的百分数。ID₅₀ 和 ID₉₀ 根据作图来确定。

表 2

X4、R5 和 X4/R5 HIV 被 FLSC-IgG1 的中和

U373/CD4/CCR5					
	FLSC-IgG1	2G12	2F5	1 IgG1b12	对照 IgG
ID ₉₀ ($\mu\text{g/mL}$)					
BaL	3.1	>10	>10	1.57	>10
ADA	4.58	>10	>10	>10	>10
89.6	3.56	8.07	>10	3.39	>10

U373/CD4/CXCR4					
	SC1g	2G12	2F5	IgG1b12	对照 IgG
ID ₉₀ ($\mu\text{g/mL}$)					
2044	>10	>10	>10	1.57	>10
2005	>10	>10	>10	>10	>10
89.6	>10	>10	>10	5.34	>10

表 2 中的数据表明 FLSC-IgG 阻断了病毒利用 R5 进入细胞。

FLSC-IgG 在中和病毒方面与当前在被动免疫治疗试验中正加以评估的抗体 2G12、2F5 和 IgG1b12 同样有效。因此这些数据进一步证实了 gp120-CD4 嵌合体在抑制特别是 HIV 感染中是有用的，以及在普遍意义上病毒外壳蛋白-受体嵌合体可用作其它利用辅助受体结合或穿透细胞的病毒的抑制剂。

实施例 7

本实施例描述了证实弗林蛋白酶切割位点的突变能够提高 FLSC 复合物的稳定性的数据。分开 FLSC 片段的切割位点的位置可能位于仅存在于 FLSC 中的 gp120 序列的 C 端内，因为较短的 TcSC 不表现出降解。值得注意的是，这些序列含有通常被弗林蛋白酶切割的 gp120 gp41 连接 (M. Girard 等, C R Acad Sci III., 322:959-66 (1999))。FLSC 在天然弗林蛋白酶位点的切割与 FLSC 片段的行为一致，因为这对 gp120 和 CD4 部分的结构和它们相互作用的能力影响最小。

为了确定这种假设的弗林蛋白酶位点是否为切割负责，通过一种特异于 gp120 C 端的抗体 (抗体的结合不受 R/T 突变的影响)，将 BaLgp120、与由 CD4 的前两个结构域 (V1V2) 组成的 sCD4 分子复合的 BaLgp120、FLSC 和 FLSC R/T 捕获在塑料板上。四个结构域 V1-V4 的 sCD4 以起始浓度 30 μ g/ml 被滴加在被捕获的复合物上。四个结构域的 sCD4 比两个结构域 V1V2 的 sCD4 与 gp120 的亲合性高，因此将从复合物中竞争替换掉较小的单位。结合的四个结构域的 CD4 用仅与四个结构域的 CD4 结合的 OKT4 抗体来检测。图 13 的结果显示了弗林蛋白酶切割位点的突变阻止了 FLSC R/T 上的 V1V2 象被切割的 FLSC 那样容易地解离，因此提高了 FLSC R/T 复合物的稳定性。在 BaLgp120 的 C 端引入 RT 突变消除了 FLSC 中观察到的弗林蛋白酶介导的切割。减少这种切割通过增加 gp120 和 CD4 部分的局部浓度增加了连接序列的连续性和 FLSC 结构的稳定性 (参见图 13)。这种增加的实验结果是可溶性四个结构域的 CD4 与在 FLSC R/T 上发现的两

个结构域的 CD4 竞争能力的降低。

实施例 8

本实施例描述了用编码 gp120-CD4 修饰嵌合多肽的多核苷酸转染细胞，以及表达的可溶性多肽的性质。使用 Fugene，按照制造商的说明（Boehringer-Manheim）将重组的 pEF6-FLSC、pEF6-RLSC-R/T、pEF6-FLSC-R/T CD4M9 和 pEF6-BaLgp120 转染 293 细胞。通过用 5 μ g/ml 的灭瘟素筛选获得稳定的转染子。简单地说，收集含有嵌合多肽的细胞培养上清液，在 SDS-PAGE 载样缓冲液（75mM Tris, 2% SDS, 10%甘油, 0.001%溴酚蓝, pH8.3）中煮沸。然后将样品在 4-20% SDS-PAGE 梯度凝胶中电泳。然后将凝胶分级的蛋白转移到硝酸纤维素膜上。膜上的非特异性结合位点用 2%脱脂奶粉的 Tris 盐缓冲液, pH7 封闭 30 分钟。然后用抗 HIV gp120 的鼠单克隆抗体混合物探测膜，结合的抗体用碱性磷酸酶标记的山羊抗小鼠 IgG 检测。

如图 14 所示，BaLgp120（第 1 道）和 FLSC-R/T CD4M9（第 4 道）迁移到分子量大约为 120kDa 的位置。FLSC R/T CD4M9 的预计分子量为大约 130kDa, 10kDa 的差别在这种印迹上是难以看出的。FLSC（第 2 道）是一种 150kDa 的蛋白，在蛋白 C 端的弗林蛋白酶位点被切割。切割将 FLSC 的 gp120 和 CD4 部分分开。较小的 120kDa 的带是这种切割的结果。释放出的 CD4 部分在这个印迹中不能看见，因为用来检测蛋白的抗体是特异性针对 gp120 的。在某些条件下单链分子明显被切割成 gp120 和 CD4 部分，这在用于 DNA 疫苗时可能需要考虑，因为这样的加工在体内也可能发生。

本实施例描述了能够证实弗林蛋白酶切割位点的突变提高了 FLSC 复合物的稳定性的数据。分开 FLSC 片段的切割位点的位置可能位于仅存在于 FLSC 中的 gp120 序列的 C 端内。值得注意的是，这些序列含有通常被弗林蛋白酶切割的 gp120/gp41 连接（M. Girard 等, C R Acad Sci III., 322:959-66 (1999)）。FLSC 在天然弗林蛋白酶位点

的切割与 FLSC 片段的行为一致，因为这对 gp120 和 CD4 部分的结构和它们相互作用的能力影响最小。结果显示了弗林蛋白酶切割位点的突变阻止了 FLSC R/T 上发现的 V1V2 象被切割的 FLSC 那样容易地解离，因此提高了 FLSC R/T 复合物的稳定性。结果，用于产生 FLSC R/T 的 R/T 突变最小化了这种切割并稳定了蛋白。

实施例 9

本实施例描述了证实 gp120-CD4 嵌合多肽与能够同 gp120 和 CD4 反应的抗体结合的数据。Gp120 与 CD4 的结合引起分子构象的变化，导致辅助受体结合结构域的暴露。因此，直接针对该结构域中的表位的抗体将强烈地与适当折叠的单链分子反应。为了确定在嵌合分子中的暴露的表位，对 BaLgp120、FLSC、FLSC-R/T 和 FLSC-R/T CDM9 分子的抗原性质进行了比较。用以前显示的在与 CD4 接合后倾向于结合 gp120 的单克隆抗体 17b (M. Thali 等, *J. Virol.*, 67:3978-86 (1993))，然后用适当标记的第二抗体来完成检测。抗体 17b 是一种人单克隆抗体，它识别的表位当 gp120 与 CD4 相互作用并结合在辅助受体附着位点 (CCR5) 中时其暴露程度增加。(N. Sullivan 等, *J. Virol.*, 72:4694-703 (1998); A. Trkola 等, *Nature*, 384: 184-6 (1996); L. Wu 等, *Nature*, 384:179-183 (1996))。抗体在 BLOTTO 中稀释，在室温保温 1 小时。在每次保温步骤之间用 TBS 洗板 3 次。样品中 gp120 序列的量根据用商品化的重组 HIV IIIB gp120 (Bartels, Issaquah, WA) 制作的标准曲线来确定。1 μ g/ml 的抗体被吸附在塑料孔中，并用 BLOTTO 封闭孔。然后如前所述用指示的人单克隆抗体进行分析。

如图 16 所示，通过将 17b 与都含有 gp120 和 CD4 的 FLSC-R/T 或 FLSC 嵌合蛋白结合，17b 与 BaLgp120、FLSC、FLSC-R/T 和 FLSC-R/T CDM9 分子的结合曲线都被增强了。17b 与 FLSC-R/T CD4M9 的结合同与 FLSC-R/T 的结合相当，表明 17b 的表位在 FLSC-R/T CD4M9 蛋白中是暴露的。合在一起，这些数据表明单链的 gp120-CD4 分子 FLSC、FLSC-R/T 和 FLSC-R/T CDM9 代表了适当折叠的 gp120-CD4

复合物。

实施例 10

本实施例描述了证实含有一种 CCR5 特异性 HIV 外膜序列的 gp120-CD 嵌合分子与 CCR5 表达细胞结合的数据。gp120-CD4 复合物的形成通常暴露出能够与适当的辅助受体相互作用的外膜结构域 (M. Thali 等, *J. Virol.*, 67:3978-86 (1993); M. A. Vodicka 等, *Virol.*, 233:193-8 (1997))。因此, 对适当折叠的 gp120-CD4 复合物及其抑制病毒对细胞感染的能力的另一种度量是它与 CCR5 辅助受体结合的能力。

为了评估单链复合物与辅助受体结合的能力, 将纯化的单链 gp120-CD4 分子与表达 CCR5 或没有辅助受体的狗胸腺细胞 Cf2Th 相互作用。简单地说, 通过用 pEF6 短暂转染 293 细胞产生含有 gp120-CD4 单链嵌合多肽 FLSC-R/T 和 FLSC-R/T CDM9 分子的上清液。

为了结合, 将纯化的单链制备物与表达 CCR5 或没有辅助受体的狗胸腺细胞相互作用。使用抗 gp120 的 MAbs、A32, 然后用一种荧光素分子藻红蛋白标记的山羊抗人 IgG 来检测结合的单链分子。通过荧光激活的细胞分类 (FACS) 分析, 使用 FACS Calibur 仪器 (Becton Dickinson) 来确定结合荧光的水平。荧光的量直接与结合的物质相关。使用 Cell Quest 3.1.3 程序 (Becton Dickinson) 计算每个样品的平均荧光强度。图 15 所示的结果显示了 FLSC-R/T CD4M9 与 CCR5 表达细胞结合, 但是不与没有辅助受体的细胞结合, 结合的效率与 FLSC-R/T 的相当。

实施例 11

本实施例描述了原初的 R5 HIV-1 (92BR020) 被来自接种了 FLSC 的小鼠的血清中和。每只 C587B1/6 小鼠用与 10 μ g 霍乱毒素 (CT) 混合的 25 μ g FLSC 接种 4 次。接种以两个星期的间隔进行。在最后一次接种后 14 天, 收集每只小鼠的血清, 并分析针对原初的 R5 HIV-1

分离株 92BR020 的中和活性。从 1:2 开始的一系列稀释的血清与每个孔 50 TCID₅₀ 的病毒感染剂量和每个孔 10⁴ 个 U373/CD4/R5/MaGI 细胞混合。24 小时后，血清、病毒和培养基用 200μl 新鲜的培养基替换。分析试验被保温 5 天直到可以观察到合胞体。HIV-1 的生长由细胞裂解物中的 b-半乳糖苷酶的产生来指示，使用化学发光试剂 Galactostar (Tropix) 按照生产商的说明进行测量。病毒感染作为化学发光的函数被测定，使用 Victor² 荧光读板器 (EG&G Wallac, Gaithersburg, MD) 进行定量。在不含病毒和血清的情况下进行分析以确定背景信号。从测试分析获得的信号通过减去背景值来校正。用每个实验孔的校正的相对光单位除以只含有细胞和病毒的对照孔的校正的光单位计算出本发明的百分数。来自 FLSC 接种的小鼠的血清被标记为 #0、#1、#2、#3 和 #4，来自幼稚小鼠的被标记为 C。

如图 17 所示，当稀释因子增加时病毒的感染也增加。此外，从对照小鼠分离的血清对病毒的感染不显示影响，而高浓度的来自 #2 小鼠的血清显示了最少量的病毒感染。

合在一起，这些数据证实了，复制病毒外壳-受体复合物的过渡态构象的可溶性嵌合多肽得到了成功的表达。有了这样的结果，现在就可能利用嵌合多肽或编码多肽的多核苷酸来免疫接种受试对象，以产生针对病毒或具有同样的外壳多肽表位的病毒的免疫反应。产生的免疫反应可以是一种抗体（体液）或 CTL 反应。此外，既然事实上嵌合多肽可以与活细胞表面上适当的辅助受体结合，那么多肽就可以被施用给急性暴露于免疫缺陷病毒的受试对象，以被动地保护表达辅助受体的细胞不受病毒感染。

实施例 12

FLSC 和 BaLgp120 与 sCD4 的复合物被捕获到 D7324 包被的 ELISA 板上。D7324 是一种绵羊的多克隆 IgG，与 gp120 的 C 端区域具有反应性，是一种抗体，通常用于通过捕获 ELISA 来检测 HIV-1 外

膜蛋白的抗原性。BaLgp120/sCD4 复合物然后用 0.5mM 双（硫代琥珀酰亚胺基）辛二酸（Pierce）交联 30 分钟，并用 10mM Tris-HCl 处理以终止反应。然后用 TBS 洗涤 BaLgp120/sCD4 和 FLSC 平板。针对 BaLgp120 的 V3 环（39F）、C1-C5（C11）、C1-C4（A32）、辅助受体结合结构域（17b）和 C3-V4（2G12）区域的单克隆抗体被滴加到捕获的抗原上。使用标记有辣根过氧化物酶的山羊抗人 IgG 检测结合的抗体。

图 18 显示了交联反应改变了 BaLgp120/sCD4 复合物的结构并减少了 39F、C11、A32 和 17b 表位的抗原性。相反，这些表位在 FLSC 上没有被封闭。抗原性的变化会影响这些表位的功能。例如，被 17b 识别的表位与 R5 辅助受体相互作用。这个表位被交联剂的封闭将减少交联的复合物与辅助受体相互作用的能力。这个观察也提示交联的复合物不能用于筛选有可能通过其辅助受体阻断 HIV-1 的试剂。

实施例 13

纯化的 R/T FLSC-IgG1 用 0.5mM 双（硫代琥珀酰亚胺基）辛二酸（Pierce）交联 30 分钟，然后用 10mM Tris-HCl 处理以终止反应。然后将交联的物质与未交联的物质在还原的和未还原的 SDS-PAGE 条件中运行进行比较。如图 19 所示，在还原的凝胶上未交联的物质（中间的道）运行在 180kDa 的位置，这是 BaLgp120-CD4-IgG1 嵌合体的预期大小。较小的带是适当大小的 CD4-IgG，表明嵌合体在分子的 BaLgp120 和 CD4-IgG 部分之间被切割。这个观察提示，尽管 R/T 突变消除了由弗林蛋白酶引起的切割，但另一蛋白酶也能够作用于 gp120 的 C 端。在未还原的凝胶上未交联的物质（右边的道）运行在 360kDa 的位置，这是完全装配的免疫粘附素的预期大小。这个观察说明，当物质的一部分被切割时（参见中间的道），免疫粘附素保持结合。物质的交联稳定了装配结构，证实了这个观察（左边的道）。在这里物质运行在预期的 360kDa 的位置。还可以看见一个分子量较高的形式，说明有一部分纯化的制备物聚集了。

实施例 14

在使用前 24 小时用 pcDNA-人 CCR5 或 pcDNA-恒河猴 CCR5 或不用质粒短暂地转染 293 细胞。转染的细胞（每个孔 10^5 个）与指定浓度的 R/T FLSC-IgG1 在 37°C 保温 1 小时。使用藻红蛋白结合的山羊抗人 IgG 检测结合的 R/T FLSC-IgG1，并用 FACS 分析。图 20 显示了 R/T FLSC-IgG1 与人和恒河猴的 CCR5 都能结合。

表达 CCR5 的狗胸腺细胞（CF2Th-R5）（ 10^5 ）与 3 μ g/ml 的 R/T FLSC-IgG1 和指定浓度的趋化因子在 37°C 保温 1 小时。使用藻红蛋白结合的山羊抗人 IgG 检测结合的 R/T FLSC-IgG1，并用 FACS 分析。RANTES 是一种 CCR5 特异性的趋化因子，正如预期一样与 R/T FLSC-IgG1 竞争受体。SDF 是一种 CXCR4 特异性趋化因子，被用做对照。图 21 进一步证明了 R/T FLSC-IgG1 可以被用做一种筛选工具，以确定能够阻断 HIV-1 通过其辅助受体 CCR5 感染细胞的试剂。

所有在此引证的参考文献被引为本发明的参考，为所教导的一切和为所有目的。应该理解的是，本发明结合其细节已经在此描述，上面的描述只是为了说明，对本发明的范围没有限制，本发明的范围由附带的权利要求书的范围所界定。其它的特点、优点和修改也落在权利要求书的范围之内。

 序列表

- <110> 马里兰大学生物技术研究所
(UNIVERSITY OF MARYLAND BIOTECHNOLOGY INSTITUTE)
- <120> 病毒外壳蛋白/受体嵌合体及使用方法
(VIRUS COAT PROTEIN/RECEPTOR CHIMERAS AND METHODS OF USE)
- <130> SCT040688-47
- <140> US 09/934,060
- <141> 2001-08-21
- <150> US 09/684,026
- <151> 2000-10-06
- <150> US 60/158,321
- <151> 1999-10-08
- <160> 33
- <170> PatentIn version 3.1
- <210> 1
- <211> 2159
- <212> DNA
- <213> 人工序列
- <220>
- <223> 合成结构 (Synthesized construct)
- <400> 1
atgcccatgg ggtctctgca accgctggcc acctgtacc tgctggggat gctggtoct

60

tctgcctcg gaaacgccga ggagaagctg tgggtgaccg tgtactacgg cgtgcccgtg	120
tggaaggagg ccaccaccac cctgttctgc gccagcgacc gcaaggccta cgacaccgag	180
gtgcacaacg tgtgggccac ccacgcctgc gtgcccaccg accccaaccc ccaggaggtg	240
gagctgaaga acgtgaccga gaacttcaac atgtggaaga acaacatggt ggagcagatg	300
cacgaggaca tcatcagcct gtgggaccag agcctgaagc cctgcgtgaa gctgaccccc	360
ctgtgcgtga ccctgaactg caccgacctg cgcaacgcc acaacggcaa cgacaccaac	420
accactagta gcagccgagg catggtgggc ggcggcgaga tgaagaactg cagcttcaac	480
atcaccacca acatccgagg caaggtgcag aaggagtacg ccctgttcta caagctggac	540
atgccccca tcgacaaca cagcaacaac cgctaccgcc tgatcagctg caacaccagc	600
gtgatcacc aggctgccc caaggtgagc ttcgagccca tccccatcca ctactgccc	660
cccgcggct tcgccatcct gaagtgcagg gacaagaagt tcaacggcaa gggcccctgc	720
accaacgtga gcaccgtgca gtgcaccac gccatccgcc ccgtggtgag caccagctg	780
ctgctgaacg gcagcctggc cgaggaggag gtggtgatcc gcagcgcaa cttcgccgac	840
aacgccaagg tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc	900
aacaacaaca cccgcaagtc catccacatc ggccccggcc ggccttcta caccaccggc	960
gagatcatcg ggcacatccg ccaggccac tgcaacctga gccgcgcaa gtggaacgac	1020
accctgaaca agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc	1080
aagcacagca gggcgggcga ccccgagatc gtgaccaca gcttcaattg cggcgggcag	1140
ttcttctact gcaacagcac ccagctgttc aacagcacct ggaacgtgac cgaggagagc	1200
aacaacaccg tggagaaca caccatcacc ctgccctgcc gcatcaagca gatcatcaac	1260
atgtggcagg aggtgggccc cgccatgtac gccccccca tccgggcca gatccgctgc	1320
agttchaaca tcaccggcct gctgctgacc cgcgacggcg gccccgagga caacaagacc	1380
gaggtgttcc gccccggcgg cggcgacatg cgcgacaact ggcgcagcga gctgtacaag	1440
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggcca ggcgcgctg	1500
gtgcagcgcg agaagcgtgg atcctctggt ggcgggtggct cgggctccgg aggaggtggg	1560
tcgggtggcg ggcggcccgc taagaaagtg gtgctgggca aaaaagggga tacagtggaa	1620
ctgacctgta cagcttcca gaagaagagc atacaattcc actggaaaaa ctccaaccag	1680
ataaagattc tgggaaatca gggctccttc ttaactaaag gtccatcaa gctgaatgat	1740
cgcgctgact caagaagaag cctttgggac caaggaaact tccccctgat catcaagaat	1800
cttaagatag aagactcaga tacttacatc tgtgaagtgg aggaccagaa ggaggaggtg	1860
caattgctag tgttcggatt gactgccaac tctgacacc acctgcttca ggggcagagc	1920

```

ctgaccctga ccttggagag cccccctggt agtagcccct cagtgcaatg taggagtcca 1980
aggggtaaaa acatacaggg ggggaagacc ctctccgtgt ctcagctgga gctccaggat 2040
agtggcacct ggacatgcac tgtcttgcag aaccagaaga aggtggagtt caaaatagac 2100
atcgtgggtgc tagctgaaca aaaactcatc tcagaagagg atctgtaata tgtttaaac 2159

```

<210> 2

<211> 720

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> MISC_特征

<222> (716)..(716)

<223> Xaa 可为任何氨基酸

<220>

<221> MISC_特征

<222> (719)..(719)

<223> Xaa 可为任何氨基酸

<400> 2

```

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
1           5           10          15

```

```

Met Leu Val Ala Ser Cys Leu Gly Asn Ala Glu Glu Lys Leu Trp Val
          20           25          30

```

```

Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
          35           40          45

```

```

Phe Cys Ala Ser Asp Arg Lys Ala Tyr Asp Thr Glu Val His Asn Val
          50           55          60

```

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val
 65 70 75 80

Glu Leu Lys Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met
 85 90 95

Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
 100 105 110

Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr
 115 120 125

Asp Leu Arg Asn Ala Thr Asn Gly Asn Asp Thr Asn Thr Thr Ser Ser
 130 135 140

Ser Arg Gly Met Val Gly Gly Gly Glu Met Lys Asn Cys Ser Phe Asn
 145 150 155 160

Ile Thr Thr Asn Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Leu Phe
 165 170 175

Tyr Lys Leu Asp Ile Ala Pro Ile Asp Asn Asn Ser Asn Asn Arg Tyr
 180 185 190

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys
 195 200 205

Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
 210 215 220

Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Lys Gly Pro Cys
 225 230 235 240

Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
 245 250 255

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val
 260 265 270

Ile Arg Ser Ala Asn Phe Ala Asp Asn Ala Lys Val Ile Ile Val Gln
 275 280 285

Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr
 290 295 300

Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
 305 310 315 320

Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser Arg Ala
 325 330 335

Lys Trp Asn Asp Thr Leu Asn Lys Ile Val Ile Lys Leu Arg Glu Gln
 340 345 350

Phe Gly Asn Lys Thr Ile Val Phe Lys His Ser Ser Gly Gly Asp Pro
 355 360 365

Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
 370 375 380

Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Val Thr Glu Glu Ser
 385 390 395 400

Asn Asn Thr Val Glu Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys
 405 410 415

Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro
 420 425 430

Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu
 435 440 445

Leu Thr Arg Asp Gly Gly Pro Glu Asp Asn Lys Thr Glu Val Phe Arg
 450 455 460

Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys
 465 470 475 480

Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala
 485 490 495

Lys Arg Arg Val Val Gln Arg Glu Lys Arg Gly Ser Ser Gly Gly Gly
 500 505 510

Gly Ser Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Ala Ala Ala Lys
 515 520 525

Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr
 530 535 540

Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln
545 550 555 560

Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser
565 570 575

Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly
580 585 590

Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr
595 600 605

Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val
610 615 620

Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln Ser
625 630 635 640

Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln
645 650 655

Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser
660 665 670

Val Ser Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr Val
675 680 685

Leu Gln Asn Gln Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val Leu
690 695 700

Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Xaa Tyr Val Xaa Thr
705 710 715 720

<210> 3

<211> 2159

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 3

atgcccatgg ggtctctgca accgctggcc accttgtagc tgctggggat gctggctcgt

60

tectgcctcg gaaacgccga ggagaagctg tgggtgaccg tgtactacgg cgtgcccgtg	120
tgggaaggagg ccaccaccac cctgttctgc gccagcgacc gcaaggccta cgacaccgag	180
gtgcacaacg tgtggggccac ccacgcctgc gtgcccaccg accccaaccc ccaggaggtg	240
gagctgaaga acgtgaccga gaacttcaac atgtggaaga acaacatggt ggagcagatg	300
cacgaggaca tcatcagcct gtgggaccag agcctgaagc cctgcgtgaa gctgaccccc	360
ctgtgcgtga ccctgaactg caccgacctg cgcaacgcc acaacggcaa cgacaccaac	420
accactagta gcagccgagg catggtgggc ggcggcgaga tgaagaactg cagcttcaac	480
atcaccacca acatccgagg caaggtgcag aaggagtacg ccctgttcta caagctggac	540
atcgccccc tgcacaaca cagcaacaac cgctaccgcc tgatcagctg caacaccagc	600
gtgatcacc aggcctgccc caaggtgagc ttcgagccca tccccatcca ctactgccc	660
cccgcggct tgcctatcct gaagtgaag gacaagaagt tcaacggcaa gggcccctgc	720
accaacgtga gcaccgtgca gtgcaccac ggcattccgc ccgtggtgag caccagctg	780
ctgctgaacg gcagcctggc cgaggaggag gtggtgatcc gcagcgcaa cttcgccgac	840
aacgccaagg tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc	900
aacaacaaca cccgcaagtc catccacatc ggccccggcc ggccttcta caccaccggc	960
gagatcatcg gcgacatccg ccaggccac tgcaacctga gccgcgcaa gtggaacgac	1020
accctgaaca agatcgtgat caagctgccc gagcagttcg gcaacaagac catcgtgttc	1080
aagcacagca gcggcggcga ccccagatc gtgaccaca gttcaattg cggcggcgag	1140
ttcttctact gcaacagcac ccagctgttc aacagcacct ggaacgtgac cgaggagagc	1200
aacaacaccg tggagaacaa caccatcacc ctgccctgcc gcatcaagca gatcatcaac	1260
atgtggcagg aggtgggccc cgccatgtac gccccccca tccgcgcca gatccgctgc	1320
agttchaaca tcaccggcct gctgctgacc cgcgacggcg gccccgagga caacaagacc	1380
gaggtgttcc gccccggcg cggcgacatg cgcgacaact ggcgagcga gctgtacaag	1440
tacaaggtgg tgaagatcga gccctgggc gtggcccca ccaaggcaa gcgccgctg	1500
gtgcagcgcg agaagaccgg atcctctggt ggcggtggct cgggctccgg aggaggtggg	1560
tccgggtggcg gcgcggccc taagaaagt gtgctgggca aaaaagggga tacagtggaa	1620
ctgacctgta cagcttcca gaagaagagc atacaattcc actggaaaaa ctccaaccag	1680
ataaagattc tgggaaatca gggctccttc ttaactaaag gtccatcaa gctgaatgat	1740
cgcgctgact caagaagaag cctttgggac caaggaaact tccccctgat catcaagaat	1800
cttaagatag aagactcaga tacttacatc tgtgaagtgg aggaccagaa ggaggaggtg	1860
caattgctag tgttcggatt gactgccaac tctgacacc accctgcttca ggggcagagc	1920

ctgaccctga ccttggagag cccccctggt agtagcccct cagtgcaatg taggagtcca 1980
 aggggtaaaa acatacaggg ggggaagacc ctctccgtgt ctcagctgga gctccaggat 2040
 agtggcacct ggacatgcac tgtcttgcag aaccagaaga aggtggagtt caaaatagac 2100
 atcgtggtgc tagctgaaca aaaactcatc tcagaagagg atctgtaata tgtttaaac 2159

<210> 4

<211> 720

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> MISC_特征

<222> (716)..(716)

<223> Xaa 可为任何氨基酸

<220>

<221> MISC_特征

<222> (719)..(719)

<223> Xaa 可为任何氨基酸

<400> 4

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1 5 10 15

Met Leu Val Ala Ser Cys Leu Gly Asn Ala Glu Glu Lys Leu Trp Val
 20 25 30

Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
 35 40 45

Phe Cys Ala Ser Asp Arg Lys Ala Tyr Asp Thr Glu Val His Asn Val
 50 55 60

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val
 65 70 75 80

Glu Leu Lys Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met
 85 90 95

Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
 100 105 110

Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr
 115 120 125

Asp Leu Arg Asn Ala Thr Asn Gly Asn Asp Thr Asn Thr Thr Ser Ser
 130 135 140

Ser Arg Gly Met Val Gly Gly Gly Glu Met Lys Asn Cys Ser Phe Asn
 145 150 155 160

Ile Thr Thr Asn Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Leu Phe
 165 170 175

Tyr Lys Leu Asp Ile Ala Pro Ile Asp Asn Asn Ser Asn Asn Arg Tyr
 180 185 190

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys
 195 200 205

Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
 210 215 220

Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Lys Gly Pro Cys
 225 230 235 240

Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
 245 250 255

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val
 260 265 270

Ile Arg Ser Ala Asn Phe Ala Asp Asn Ala Lys Val Ile Ile Val Gln
 275 280 285

Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr
 290 295 300

Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
 305 310 315 320

Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser Arg Ala
 325 330 335

Lys Trp Asn Asp Thr Leu Asn Lys Ile Val Ile Lys Leu Arg Glu Gln
 340 345 350

Phe Gly Asn Lys Thr Ile Val Phe Lys His Ser Ser Gly Gly Asp Pro
 355 360 365

Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
 370 375 380

Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Val Thr Glu Glu Ser
 385 390 395 400

Asn Asn Thr Val Glu Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys
 405 410 415

Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro
 420 425 430

Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu
 435 440 445

Leu Thr Arg Asp Gly Gly Pro Glu Asp Asn Lys Thr Glu Val Phe Arg
 450 455 460

Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys
 465 470 475 480

Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala
 485 490 495

Lys Arg Arg Val Val Gln Arg Glu Lys Thr Gly Ser Ser Gly Gly Gly
 500 505 510

Gly Ser Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Ala Ala Ala Lys
 515 520 525

Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr
 530 535 540

Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln
545 550 555 560

Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser
565 570 575

Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly
580 585 590

Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr
595 600 605

Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val
610 615 620

Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln Ser
625 630 635 640

Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln
645 650 655

Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser
660 665 670

Val Ser Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr Val
675 680 685

Leu Gln Asn Gln Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val Leu
690 695 700

Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Xaa Tyr Val Xaa Thr
705 710 715 720

<210> 5

<211> 1668

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 5

atgcccatgg ggtctctgca accgctggcc accttgtagc tgctggggat gctggtcgct

60

tcctgcctcg gaaacgccga ggagaagctg tgggtgaccg tgtactacgg cgtgcccgctg	120
tggaaggagg ccaccaccac cctgttctgc gccagcgacc gcaaggccta cgacaccgag	180
gtgcacaacg tgtgggccac ccacgcctgc gtgcccaccg accccaaccc ccaggaggtg	240
gagctgaaga acgtgaccga gaacttcaac atgtggaaga acaacatggt ggagcagatg	300
cacgaggaca tcatcagcct gtgggaccag agcctgaagc cctgcgtgaa gctgaccccc	360
ctgtgcgtga ccctgaactg caccgacctg cgcaacgcc acaacggcaa cgacaccaac	420
accactagta gcagcccgcg catggtgggc ggcggcgaga tgaagaactg cagcttcaac	480
atcaccacca acatcccgcg caaggtgcag aaggagtacg ccctgttcta caagctggac	540
atcgccccca tcgacaacaa cagcaacaac cgctaccgcc tgatcagctg caacaccagc	600
gtgatcacc aggcctgccc caaggtgagc ttcgagccca tccccatcca ctactgcgcc	660
cccgcggct tcgccatcct gaagtgcaag gacaagaagt tcaacggcaa gggcccctgc	720
accaacgtga gcaccgtgca gtgcaccac ggcacccgcc ccgtggtgag caccagctg	780
ctgctgaacg gcagcctggc cgaggaggag gtggtgatcc gcagcgcaa cttcgccgac	840
aacgccaaagg tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc	900
aacaacaaca cccgcaagtc catccacatc ggcgccggcc gcgccttcta caccaccggc	960
gagatcatcg gcgacatccg ccaggcccac tgcaacctga gccgcgcaa gtggaacgac	1020
acctgaaca agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc	1080
aagcacagca gcggcggcga ccccagatc gtgaccaca gcttcaattg cggcgggcag	1140
ttcttctact gcaacagcac ccagctgttc aacagcacct ggaacgtgac cgaggagagc	1200
aacaacaccg tggagaacaa caccatcacc ctgccctgcc gcatcaagca gatcatcaac	1260
atgtggcagg aggtgggccc cgccatgtac gccccccca tccgcggcca gatccgctgc	1320
agttchaaca tcaccggcct gctgctgacc cgcgacggcg gccccgagga caacaagacc	1380
gaggtgttcc gccccggcgg cggcgacatg cgcgacaact ggcgcagcga gctgtacaag	1440
tacaaggtgg tgaagatcga gccctgggc gtggccccca ccaaggccaa gcgcccgtg	1500
gtgcagcgcg agaagaccgg atcctctggt ggcgggtggct cgggctccgg aggaggtggg	1560
tcgggtggcg gcgcgccgc ttgcaacctg gcccgctgcc agctgcgctg caagagcctg	1620
ggcctgctgg gcaagtgcgc cggcagcttc tgccgctgcg gcccctaa	1668

<210> 6

<211> 556

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> MISC_特征

<222> (556)..(556)

<223> Xaa 可为任何氨基酸

<400> 6

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
1 5 10 15

Met Leu Val Ala Ser Cys Leu Gly Asn Ala Glu Glu Lys Leu Trp Val
20 25 30

Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
35 40 45

Phe Cys Ala Ser Asp Arg Lys Ala Tyr Asp Thr Glu Val His Asn Val
50 55 60

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val
65 70 75 80

Glu Leu Lys Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met
85 90 95

Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
100 105 110

Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr
115 120 125

Asp Leu Arg Asn Ala Thr Asn Gly Asn Asp Thr Asn Thr Thr Ser Ser
130 135 140

Ser Arg Gly Met Val Gly Gly Gly Glu Met Lys Asn Cys Ser Phe Asn
145 150 155 160

Ile Thr Thr Asn Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Leu Phe
 165 170 175

Tyr Lys Leu Asp Ile Ala Pro Ile Asp Asn Asn Ser Asn Asn Arg Tyr
 180 185 190

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys
 195 200 205

Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
 210 215 220

Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Lys Gly Pro Cys
 225 230 235 240

Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
 245 250 255

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val
 260 265 270

Ile Arg Ser Ala Asn Phe Ala Asp Asn Ala Lys Val Ile Ile Val Gln
 275 280 285

Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr
 290 295 300

Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
 305 310 315 320

Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser Arg Ala
 325 330 335

Lys Trp Asn Asp Thr Leu Asn Lys Ile Val Ile Lys Leu Arg Glu Gln
 340 345 350

Phe Gly Asn Lys Thr Ile Val Phe Lys His Ser Ser Gly Gly Asp Pro
 355 360 365

Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
 370 375 380

Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Val Thr Glu Glu Ser
 385 390 395 400

Asn Asn Thr Val Glu Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys

	405		410		415
Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro	420		425		430
Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu	435		440		445
Leu Thr Arg Asp Gly Gly Pro Glu Asp Asn Lys Thr Glu Val Phe Arg	450		455		460
Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys	465		470		475
					480
Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala	485		490		495
Lys Arg Arg Val Val Gln Arg Glu Lys Thr Gly Ser Ser Gly Gly Gly	500		505		510
Gly Ser Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Ala Ala Ala Cys	515		520		525
Asn Leu Ala Arg Cys Gln Leu Arg Cys Lys Ser Leu Gly Leu Leu Gly	530		535		540
Lys Cys Ala Gly Ser Phe Cys Ala Cys Gly Pro Xaa	545		550		555

<210> 7

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (synthesized construct)

<220>

<221> 引物_结合

<222> (1)..(39)

<223>

<400> 7
 gggggtacca tgcccatggg gtctctgcaa ccgctggcc 39

<210> 8

<211> 66

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (synthesized construct)

<400> 8
 ggggtccggag cccgagccac cgccaccaga ggatccacgc ttctcgcgct gcaccacgcg 60
 gcgctt 66

<210> 9

<211> 69

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> 引物_结合

<222> (1)..(69)

<223>

<400> 9
 ggggtccggag gaggtgggtc ggggtggcggc gcggcccgcta agaaagtggc gctgggcaaa 60
 aaaggggat 69

<210> 10

<211> 76

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> 引物_结合

<222> (1)..(76)

<223>

<400> 10
 ggggtttaaa cttattacag atcctcttct gagatgagtt ttgttcagct agcaccacga 60
 tgtctatttt gaactc 76

<210> 11

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> misc_特征

<222> (1)..(21)

<223> 间隔区 (Spacer)

<400> 11
 gsggggsgs ggggsgggaa a 21

<210> 12

<211> 1769

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 12

```

atgcccattgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct      60
tcctgcctcg gaaagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag      120
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg      180
acccccctgt gcgtgacctt gggcgcgggc gagatgaaga actgcagctt caacatcggc      240
gcgggccgcc tgatcagctg caacaccagc gtgatcacc aggctgccc caaggtagc      300
ttcgagccca tccccatcca ctactgcgcc cccgccggct tcgccatcct gaagtgaag      360
gacaagaagt tcaacggcaa gggcccctgc accaacgtga gcaccgtgca gtgcacccac      420
ggcatccgcc ccgtggtgag caccagctg ctgctgaacg gcagcctggc cgaggaggag      480
gtggtgatcc gcagcgccaa cttcgccgac aacgccaaag tgatcatcgt gcagctgaac      540
gagagcgtgg agatcaactg caccgcccc aacaacaaca cccgcaagtc catccacatc      600
ggccccggcc gcgccttcta caccaccggc gagatcatcg gcgacatccg ccaggccccac      660
tgcaacctga gccgcgccaa gtggaacgac accctgaaca agatcgtgat caagctgcgc      720
gagcagttcg gcaacaagac catcgtgttc aagcacagca gcggcggcga ccccgagatc      780
gtgaccaca gcttcaattg cggcgcgag ttcttctact gcaacagcac ccagctgttc      840
aacagcacct ggaacgtgac cgaggagagc aacaacaccg tggagaacaa caccatcacc      900
ctgccctgcc gcatcaagca gatcatcaac atgtggcagg aggtgggccc cgccatgtac      960
gccccccca tccgcgcca gatccgctgc agttchaaca tcaccggcct gctgctgacc     1020
cgcgacggcg gccccgagga caacaagacc gaggtgttcc gccccggcgg cggcgacatg     1080
cgcgacaact ggcgcagcga gctgtacaag tacaaggtgg tgaagatcgg atcctctggt     1140
ggcgggtggc cgggctccgg aggaggtggg tcgggtggcg gcgcgccgc taagaaagtg     1200
gtgctgggca aaaaagggga tacagtggaa ctgacctgta cagcttcca gaagaagagc     1260
atacaattcc actggaaaaa ctccaaccag ataaagattc tgggaaatca gggctccttc     1320
ttaactaaag gtccatccaa gctgaatgat cgcgctgact caagaagaag cctttgggac     1380
caaggaaact tccccctgat catcaagaat cttaagatag aagactcaga tacttacatc     1440
tgtgaagtgg aggaccagaa ggaggagtg caattgctag tgttcggatt gactgccaac     1500
tctgacacc accctgttca ggggcagagc ctgaccctga ccttggagag cccccctggt     1560

```

```

agtagcccct cagtgcaatg taggagtcca aggggtaaaa acatacaggg ggggaagacc 1620
ctctccgtgt ctcagctgga gctccaggat agtggcacct ggacatgcac tgtcttgcag 1680
aaccagaaga aggtggagtt caaaatagac atcgtggtgc tagctgaaca aaaactcatc 1740
tcagaagagg atctgtaata tgtttaaac 1769

```

<210> 13

<211> 590

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> MISC_特征

<222> (586)..(586)

<223> Xaa 可为任何氨基酸

<220>

<221> MISC_特征

<222> (589)..(589)

<223> Xaa 可为任何氨基酸

<400> 13

```

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
1           5           10           15

```

```

Met Leu Val Ala Ser Cys Leu Gly Lys Asn Val Thr Glu Asn Phe Asn
20           25           30

```

```

Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser
35           40           45

```

```

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
50           55           60

```

Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr
305 310 315 320

Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly
325 330 335

Leu Leu Leu Thr Arg Asp Gly Gly Pro Glu Asp Asn Lys Thr Glu Val
340 345 350

Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu
355 360 365

Tyr Lys Tyr Lys Val Val Lys Ile Gly Ser Ser Gly Gly Gly Gly Ser
370 375 380

Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Ala Ala Ala Lys Lys Val
385 390 395 400

Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser
405 410 415

Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys
420 425 430

Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu
435 440 445

Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe
450 455 460

Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile
465 470 475 480

Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly
485 490 495

Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr
500 505 510

Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg
515 520 525

Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser
530 535 540

Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln

<220>

<223> 合成结构 (Synthesized construct)

<400> 15
 gggggtacca tgcccatggg gtctctgcaa ccgctggcca ccttgtaacct gctgggggatg 60
 ctggtcgctt cctgcctcgg aaagaacgtg accgagaact tcaacatgtg g 111

<210> 16

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 16
 gggggatccg atcttcacca ccttgatcct gtacagctc 39

<210> 17

<211> 75

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 17
 ctgtgcgtga ccctgggagc ggccgagatg aagaactgca gcttcaacat cggcgcgggc 60
 cgctgatca gctgc 75

<210> 18

<211> 75

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 18

gcagctgac aggcggcccg cgccgatgtt gaagctgcag ttcttcatct cgcccgcgcc 60

cagggtcacg cacag 75

<210> 19

<211> 87

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 19

tgcaacctgg cccgctgcca gctgcgctgc aagagcctgg gcctgctggg caagtgcgcc 60

ggcagcttct gcgctgcgg cccctaa 87

<210> 20

<211> 28

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 20

Cys Asn Leu Ala Arg Cys Gln Leu Arg Cys Lys Ser Leu Gly Leu Leu
1 5 10 15

Gly Lys Cys Ala Gly Ser Phe Cys Ala Cys Gly Pro
20 25

<210> 21

<211> 102

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 21
gcgggccgctt gcaacctggc ccgctgccag ctgcgctgca agagcctggg cctgctgggc 60
aagtgcgccc gcagcttctg cgctgcggc ccctaagaat tc 102

<210> 22

<211> 102

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 22
gaattcttag gggccgcagg cgcagaagct gccggcgcac ttgccagca ggcccaggct 60
cttgacgcgc agctggcagc gggccagggt gcaagcggcc gc 102

<210> 23

<211> 1518

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 23
atgcccattg ggtctctgca accgctggcc accttgtacc tgctggggat gctggctcgt 60
tcctgcctcg gaaacgccga ggagaagctg tgggtgaccg tgtactacgg cgtgcccgtg 120
tggaaggagg ccaccaccac cctgttctgc gccagcgacc gcaaggccta cgacaccgag 180
gtgcacaacg tgtggggccac ccacgcctgc gtgcccaccg accccaacc ccaggaggctg 240
gagctgaaga acgtgaccga gaacttcaac atgtggaaga acaacatggt ggagcagatg 300
cacgaggaca tcatcagcct gtgggaccag agcctgaagc cctgcgtgaa gctgaccccc 360

```

ctgtgcgtga ccctgaactg caccgacctg cgcaacgcc acaacggcaa cgacaccaac 420
accactagta gcagccgcgg catggtgggc ggcggcgaga tgaagaactg cagcttcaac 480
atcaccacca acatccgcgg caaggtgcag aaggagtacg cctgtttcta caagctggac 540
atcgccccca tcgacaacaa cagcaacaac cgctaccgcc tgatcagctg caacaccagc 600
gtgatcacc caggcctgcc caaggtgagc ttcgagccca tccccatcca ctactgcgcc 660
cccgcggct tcgccatcct gaagtgcag gacaagaagt tcaacggcaa gggcccctgc 720
accaactgta gcaccgtgca gtgcacccac ggcacccgcc ccgtggtgag caccagctg 780
ctgctgaacg gcagcctggc cgaggaggag gtggtgatcc gcagcgccaa cttcgccgac 840
aacgccaagg tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc 900
aacaacaaca cccgcaagtc catccacatc ggccccggcc ggccttcta caccaccggc 960
gagatcatcg gcgacatccg ccaggccac tgcaacctga gccgcgccaa gtggaacgac 1020
accctgaaca agatcgtgat caagctgcgc gagcagttcg gcaacaagac catcgtgttc 1080
aagcacagca gcggcggcga ccccgagatc gtgaccaca gttcaattg cggcggcgag 1140
ttcttctact gcaacagcac ccagctgttc aacagcacct ggaacgtgac cgaggagagc 1200
aacaacaccg tggagaacaa caccatcacc ctgccctgcc gcatcaagca gatcatcaac 1260
atgtggcagg aggtgggccc cgccatgtac gccccccca tccgcgccca gatccgctgc 1320
agttchaaca tcaccggcct gctgctgacc cgcgacggcg gccccgagga caacaagacc 1380
gaggtgttcc gccccggcgg cggcgacatg cgcgacaact ggcgcagcga gctgtacaag 1440
tacaaggtgg tgaagatcga gcccctgggc gtggccccca ccaaggccaa gcgccgctg 1500
gtgcagcgcg agaagcgt 1518

```

<210> 24

<211> 506

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 24

```

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
1           5           10           15

```

Met Leu Val Ala Ser Cys Leu Gly Asn Ala Glu Glu Lys Leu Trp Val
 20 25 30

Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
 35 40 45

Phe Cys Ala Ser Asp Arg Lys Ala Tyr Asp Thr Glu Val His Asn Val
 50 55 60

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val
 65 70 75 80

Glu Leu Lys Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met
 85 90 95

Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
 100 105 110

Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr
 115 120 125

Asp Leu Arg Asn Ala Thr Asn Gly Asn Asp Thr Asn Thr Thr Ser Ser
 130 135 140

Ser Arg Gly Met Val Gly Gly Gly Glu Met Lys Asn Cys Ser Phe Asn
 145 150 155 160

Ile Thr Thr Asn Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Leu Phe
 165 170 175

Tyr Lys Leu Asp Ile Ala Pro Ile Asp Asn Asn Ser Asn Asn Arg Tyr
 180 185 190

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys
 195 200 205

Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
 210 215 220

Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Lys Gly Pro Cys
 225 230 235 240

Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
 245 250 255

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val
 260 265 270

Ile Arg Ser Ala Asn Phe Ala Asp Asn Ala Lys Val Ile Ile Val Gln
 275 280 285

Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr
 290 295 300

Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly
 305 310 315 320

Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser Arg Ala
 325 330 335

Lys Trp Asn Asp Thr Leu Asn Lys Ile Val Ile Lys Leu Arg Glu Gln
 340 345 350

Phe Gly Asn Lys Thr Ile Val Phe Lys His Ser Ser Gly Gly Asp Pro
 355 360 365

Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys
 370 375 380

Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Val Thr Glu Glu Ser
 385 390 395 400

Asn Asn Thr Val Glu Asn Asn Thr Ile Thr Leu Pro Cys Arg Ile Lys
 405 410 415

Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr Ala Pro
 420 425 430

Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu
 435 440 445

Leu Thr Arg Asp Gly Gly Pro Glu Asp Asn Lys Thr Glu Val Phe Arg
 450 455 460

Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys
 465 470 475 480

Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala
 485 490 495

Lys Arg Arg Val Val Gln Arg Glu Lys Arg

500

505

<210> 25

<211> 534

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 25
aagaaagtgg tgctgggcaa aaaaggggat acagtggaac tgacctgtac agcttcccag 60
aagaagagca tacaattcca ctggaaaaac tccaaccaga taaagattct gggaaatcag 120
ggctccttct taactaaagg tccatccaag ctgaatgata gcgctgactc aagaagaagc 180
ctttgggacc aaggaaactt cccctgatc atcaagaatc ttaagataga agactcagat 240
acttacatct gtgaagtgga ggaccagaag gaggaggtgc aattgctagt gttcggattg 300
actgccaaact ctgacacca cctgcttcag gggcagagcc tgacctgac cttggagagc 360
ccccctggta gtagccctc agtgcaatgt aggagtccaa ggggtaaaaa catacagggg 420
gggaagacc tctccgtgtc tcagctggag ctccaggata gtggcacctg gacatgcact 480
gtcttgacaga accagaagaa ggtggagttc aaaatagaca tcgtggtgct agct 534

<210> 26

<211> 178

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 26

Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys
1 5 10 15

Thr Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn
20 25 30

Gln Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro
35 40 45

Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln
50 55 60

Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp
65 70 75 80

Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu
85 90 95

Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln
100 105 110

Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val
115 120 125

Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu
130 135 140

Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr
145 150 155 160

Val Leu Gln Asn Gln Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val
165 170 175

Leu Ala

<210> 27

<211> 1128

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 27

atgcccatgg ggtctctgca accgctggcc accttgtagc tgctggggat gctggctgct 60

tcctgcctcg gaaagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgggtggag 120

cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 180

```

acccccctgt gcgtgaccct gggcgcgggc gagatgaaga actgcagctt caacatcggc 240
gcggggccgcc tgatcagctg caacaccagc gtgatcaccc aggctgccc caaggtgagc 300
ttcgagccca tccccatcca ctactgcgcc cccgcccggc tcgccatcct gaagtgcaag 360
gacaagaagt tcaacggcaa gggcccctgc accaacgtga gcaccgtgca gtgcacccac 420
ggcatccgcc ccgtggtgag caccagctg ctgctgaacg gcagcctggc cgaggaggag 480
gtggtgatcc gcagcgccaa cttcgccgac aacgccaaagg tgatcatcgt gcagctgaac 540
gagagcgtgg agatcaactg caccgcccc aacaacaaca cccgcaagtc catccacatc 600
ggccccggcc gcgcctteta caccaccggc gagatcatcg gcgacatccg ccaggccccac 660
tgcaacctga gccgcgcca gtggaacgac accctgaaca agatcgtgat caagctgcgc 720
gagcagttcg gcaacaagac catcgtgttc aagcacagca gcggcggcga ccccgagatc 780
gtgaccaca gcttcaattg cggcgggcag ttcttctact gcaacagcac ccagctgttc 840
aacagcacct ggaacgtgac cgaggagagc aacaacaccg tggagaaca caccatcacc 900
ctgccctgcc gcatcaagca gatcatcaac atgtggcagg aggtgggccg cgccatgtac 960
gccccccca tccgcgcca gatccgctgc agttchaaca tcaccggcct gctgctgacc 1020
cgcgacggcg gccccgagga caacaagacc gaggtgttcc gccccggcgg cggcgacatg 1080
cgcgacaact ggcgcagcga gctgtacaag tacaaggtgg tgaagatc 1128

```

<210> 28

<211> 376

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 28

```

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
1           5           10           15

```

```

Met Leu Val Ala Ser Cys Leu Gly Lys Asn Val Thr Glu Asn Phe Asn
20           25           30

```

```

Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser
35           40           45

```

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
50 55 60

Val Thr Leu Gly Ala Gly Glu Met Lys Asn Cys Ser Phe Asn Ile Gly
65 70 75 80

Ala Gly Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys
85 90 95

Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala
100 105 110

Gly Phe Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Lys Gly
115 120 125

Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro
130 135 140

Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu
145 150 155 160

Val Val Ile Arg Ser Ala Asn Phe Ala Asp Asn Ala Lys Val Ile Ile
165 170 175

Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn
180 185 190

Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr
195 200 205

Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Leu Ser
210 215 220

Arg Ala Lys Trp Asn Asp Thr Leu Asn Lys Ile Val Ile Lys Leu Arg
225 230 235 240

Glu Gln Phe Gly Asn Lys Thr Ile Val Phe Lys His Ser Ser Gly Gly
245 250 255

Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe
260 265 270

Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Val Thr Glu
275 280 285

Glu Ser Asn Asn Thr Val Glu Asn Asn Thr Ile Thr Leu Pro Cys Arg
290 295 300

Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Met Tyr
305 310 315 320

Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly
325 330 335

Leu Leu Leu Thr Arg Asp Gly Gly Pro Glu Asp Asn Lys Thr Glu Val
340 345 350

Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu
355 360 365

Tyr Lys Tyr Lys Val Val Lys Ile
370 375

<210> 29

<211> 1518

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 29
atgcccattgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtcgct 60
tcctgcctcg gaaacgccga ggagaagctg tgggtgaccg tgtactacgg cgtgcccgtg 120
tggaaggagg ccaccaccac cctgttctgc gccagcgacc gcaaggccta cgacaccgag 180
gtgcacaacg tgtgggccac ccacgcctgc gtgcccaccg accccaacc ccaggagggtg 240
gagctgaaga acgtgaccga gaacttcaac atgtggaaga acaacatggt ggagcagatg 300
cacgaggaca tcatcagcct gtgggaccag agcctgaagc cctgcgtgaa gctgaccccc 360
ctgtgcgtga ccctgaactg caccgacctg cgcaacgcc ccaacggcaa cgacaccaac 420
accactagta gcagccgcgg catggtgggc ggcggcgaga tgaagaactg cagcttcaac 480
atcaccacca acatccgcgg caaggtgcag aaggagtacg ccctgttcta caagctggac 540
atgccccca tcgacaacaa cagcaacaac cgctaccgcc tgatcagctg caacaccagc 600
gtgatcacc aggcctgccc caaggtgagc ttcgagccca tcccatcca ctactgcgcc 660

```

ccccccggct tgcctatcct gaagtgcaag gacaagaagt tcaacggcaa gggcccctgc 720
accaacgtga gcaccgtgca gtgcacccac ggcattccgcc ccgtggtgag caccagctg 780
ctgctgaacg gcagcctggc cgaggaggag gtggtgatcc gcagcgccaa ctccgcccac 840
aacgccaagg tgatcatcgt gcagctgaac gagagcgtgg agatcaactg caccgcccc 900
aacaacaaca cccgcaagtc catccacatc ggccccggcc gcgccttcta caccaccggc 960
gagatcatcg gcgacatccg ccaggcccac tgcaacctga gccgcgccaa gtggaacgac 1020
accctgaaca agatcgtgat caagctgctc gagcagttcg gcaacaagac catcgtgttc 1080
aagcacagca gcggcggcga ccccgagatc gtgaccaca gcttcaattg cggcggcgag 1140
ttcttctact gcaacagcac ccagctgttc aacagcacct ggaacgtgac cgaggagagc 1200
aacaacaccg tggagaacaa caccatcacc ctgcctgcc gcatcaagca gatcatcaac 1260
atgtggcagg aggtggggcg cgccatgtac gccccccca tccgaggcca gatccgctgc 1320
agttchaaca tcaccggcct gctgctgacc cgcgacggcg gccccgagga caacaagacc 1380
gaggtgttcc gccccggcgg cggcgacatg cgcgacaact ggcgcagcga gctgtacaag 1440
tacaagtggt tgaagatcga gccctggggc gtggccccca ccaaggccaa gcgcccgtg 1500
gtgcagcgcg agaagacc 1518

```

<210> 30

<211> 506

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 30

```

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
1           5           10           15

```

```

Met Leu Val Ala Ser Cys Leu Gly Asn Ala Glu Glu Lys Leu Trp Val
20           25           30

```

```

Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
35           40           45

```

```

Phe Cys Ala Ser Asp Arg Lys Ala Tyr Asp Thr Glu Val His Asn Val
50           55           60

```

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val
65 70 75 80

Glu Leu Lys Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met
85 90 95

Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
100 105 110

Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr
115 120 125

Asp Leu Arg Asn Ala Thr Asn Gly Asn Asp Thr Asn Thr Thr Ser Ser
130 135 140

Ser Arg Gly Met Val Gly Gly Gly Glu Met Lys Asn Cys Ser Phe Asn
145 150 155 160

Ile Thr Thr Asn Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Leu Phe
165 170 175

Tyr Lys Leu Asp Ile Ala Pro Ile Asp Asn Asn Ser Asn Asn Arg Tyr
180 185 190

Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys
195 200 205

Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe
210 215 220

Ala Ile Leu Lys Cys Lys Asp Lys Lys Phe Asn Gly Lys Gly Pro Cys
225 230 235 240

Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val
245 250 255

Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val
260 265 270

Ile Arg Ser Ala Asn Phe Ala Asp Asn Ala Lys Val Ile Ile Val Gln
275 280 285

Leu Asn Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr
290 295 300

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

<210> 31

<211> 776

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 31

```

atgcccattgg ggtctctgca accgctggcc accttgtacc tgctggggat gctggtegct      60
tccgtgctag cggatcccga ggagcccaaa tcttgtgaca aaactcacac atgcccaccg      120
tgcccagcac ctgaactcct ggggggaccg tcagtcttcc tcttcccccc aaaacccaag      180
gacaccctca tgatctcccg gacccctgag gtcacatgcg tggtggtgga cgtgagccac      240
gaagaccctg aggtcaagtt caactggtac gtggacggcg tggaggtgca taatgccaaag      300
acaaagccgc gggaggagca gtacaacagc acgtaccggg tggtcagcgt cctcaccgtc      360
ctgcaccagg actggctgaa tggcaaggag tacaagtgca aggtctcaa caaagccctc      420
ccagccccca tcgagaaaac catctccaaa gccaaagggc agccccgaga accacaggtg      480
tacaccctgc ccccatcccg ggatgagctg accaagaacc aggtcagcct gacctgcctg      540
gtcaaaggct tctatcccag cgacatcgcc gtggagtggg agagcaatgg gcagccggag      600
aacaactaca agaccacgcc tcccgtgctg gactccgacg gctccttctt cctctacagc      660
aagctcaccg tggacaagag caggtggcag caggggaacg tcttctcatg ctccgtgatg      720
catgaggctc tgcacaacca ctacacgcag aagagcctct ccctgtctcc gggtaa      776

```

<210> 32

<211> 259

<212> PRT

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<220>

<221> MISC_特征

<222> (259)..(259)

<223> Xaa 可为任何氨基酸

<400> 32

Met Pro Met Gly Ser Leu Gln Pro Leu Ala Thr Leu Tyr Leu Leu Gly
 1 5 10 15

Met Leu Val Ala Ser Val Leu Ala Asp Pro Glu Glu Pro Lys Ser Cys
 20 25 30

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
 35 40 45

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
 50 55 60

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
 65 70 75 80

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 85 90 95

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 100 105 110

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 115 120 125

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 130 135 140

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 145 150 155 160

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 165 170 175

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 180 185 190

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 195 200 205

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 210 215 220

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
 225 230 235 240

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

245

250

255

Pro Gly Xaa

<210> 33

<211> 63

<212> DNA

<213> 人工序列

<220>

<223> 合成结构 (Synthesized construct)

<400> 33

ggatcctctg gtggcgggtg ctcgggctcc ggaggaggtg ggtcgggtgg cggcgcggcc 60

gct 63

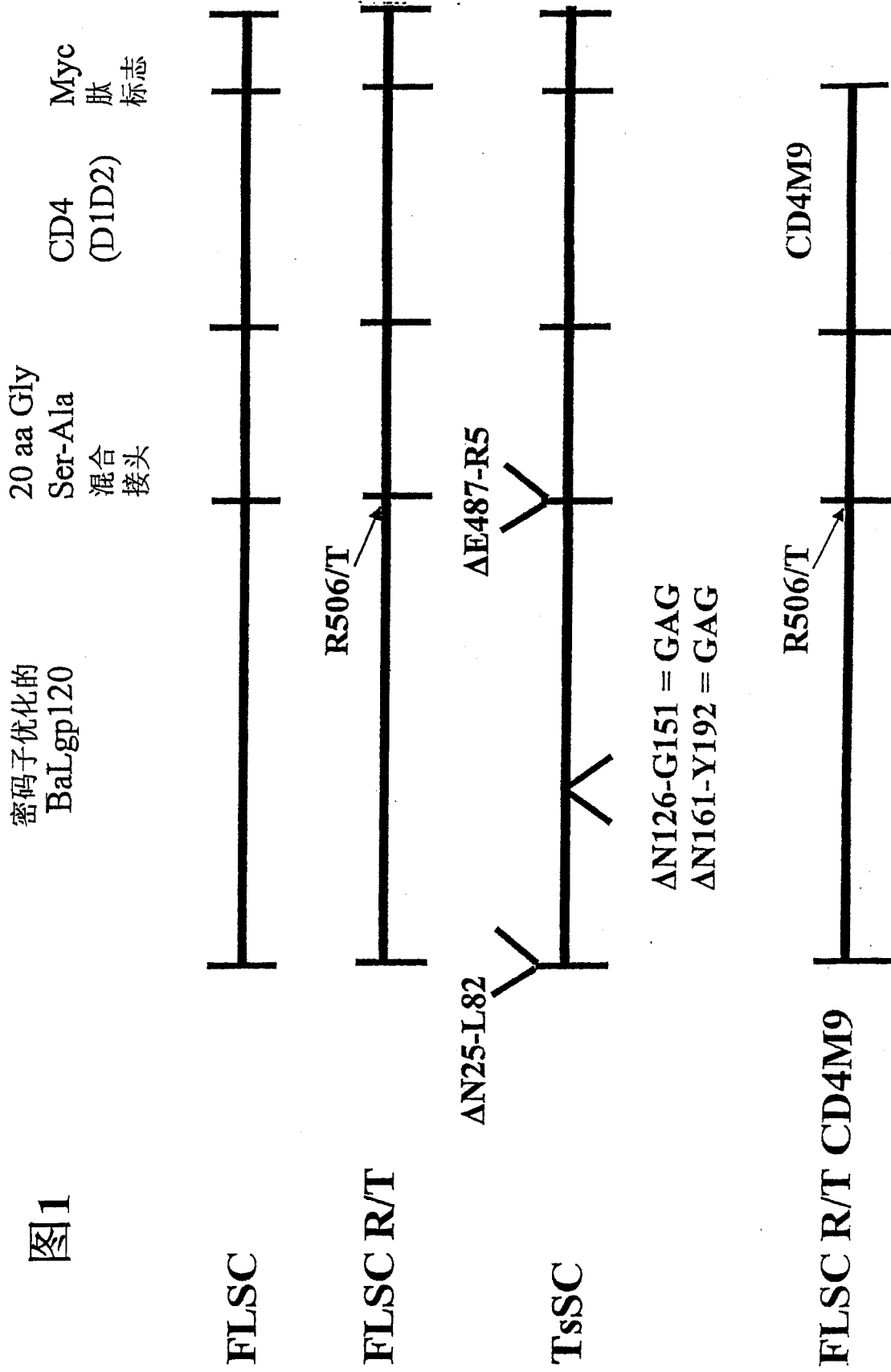


图1

图2

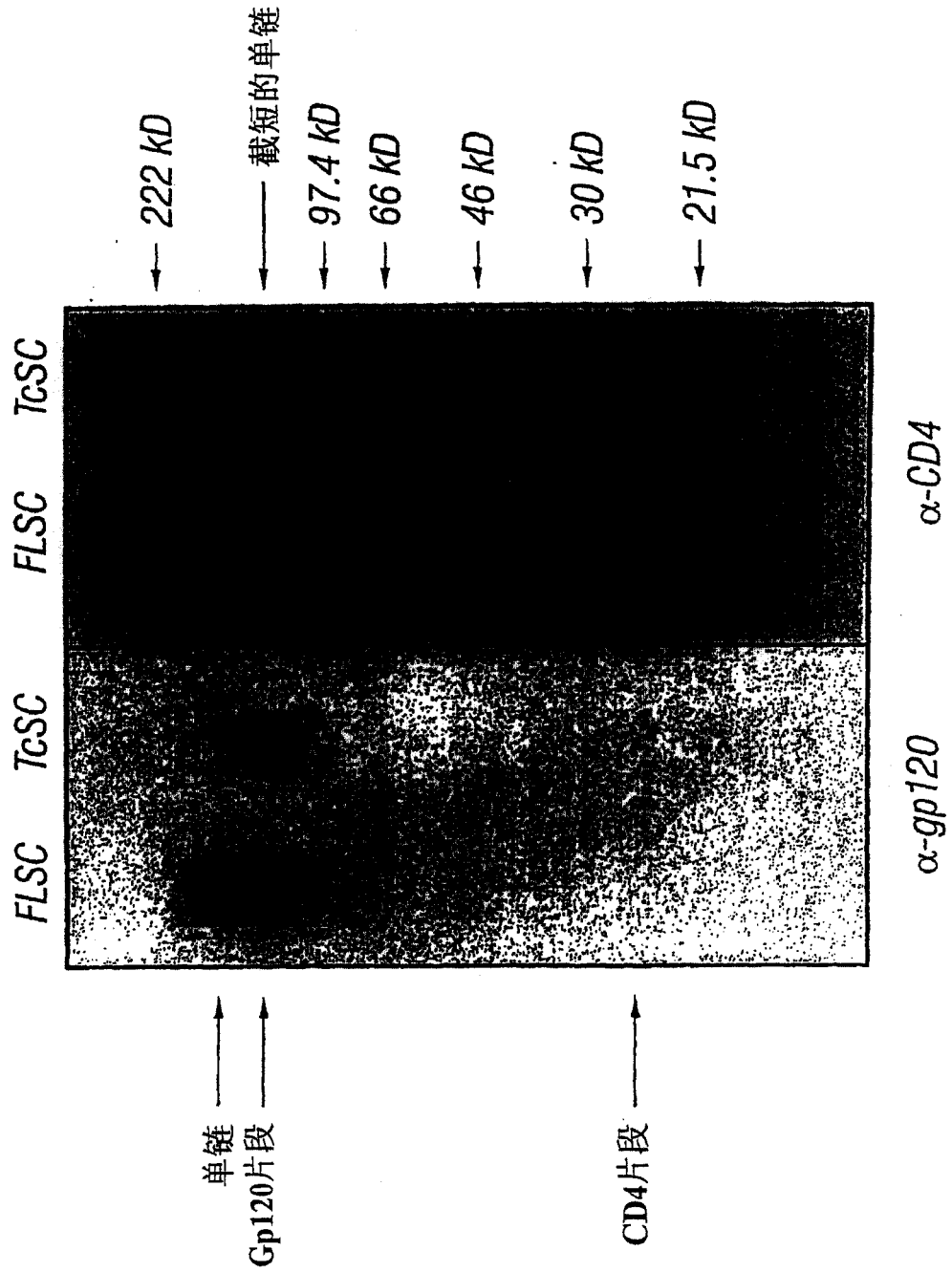


图3

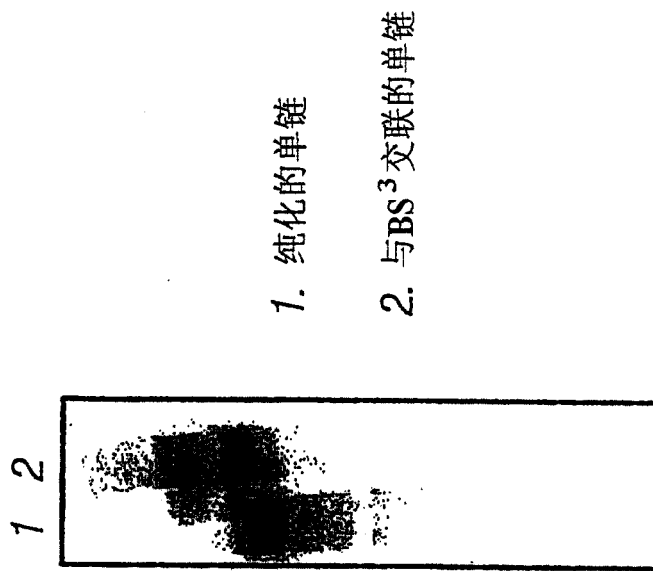
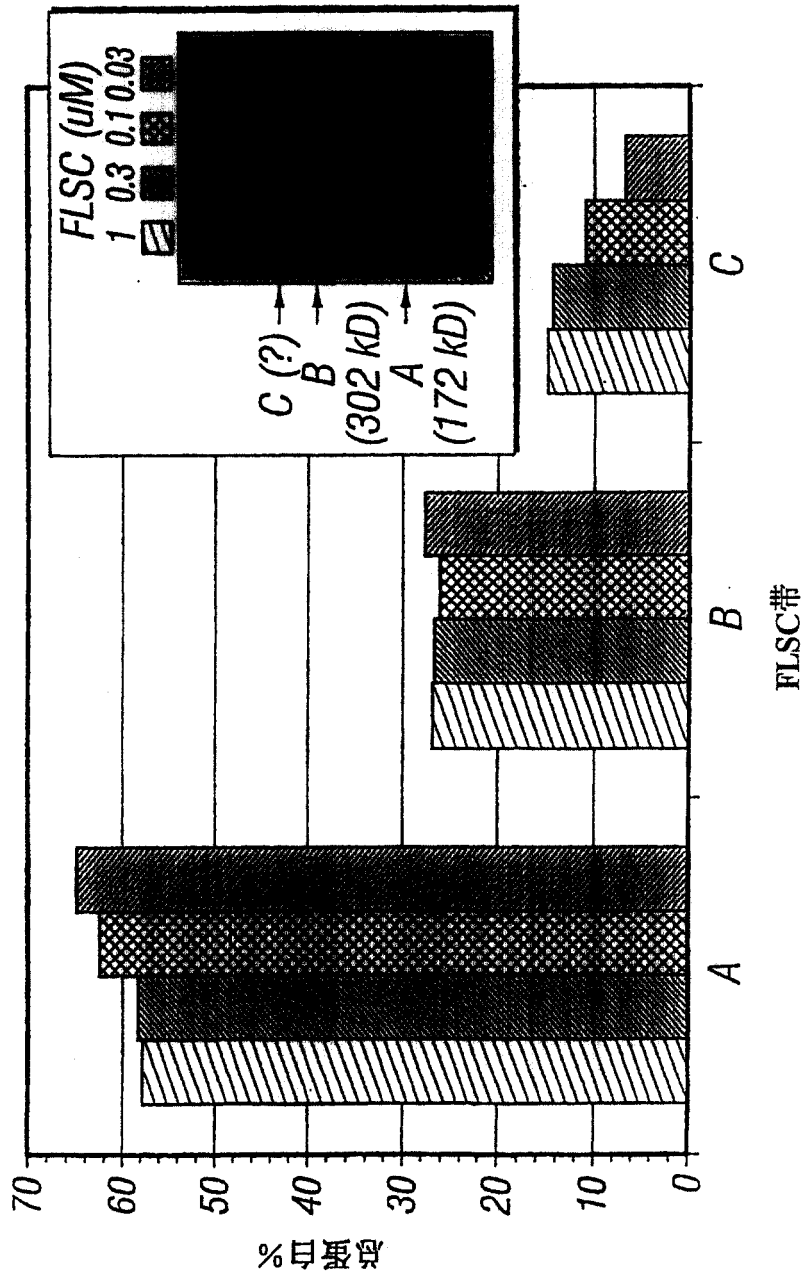


图4



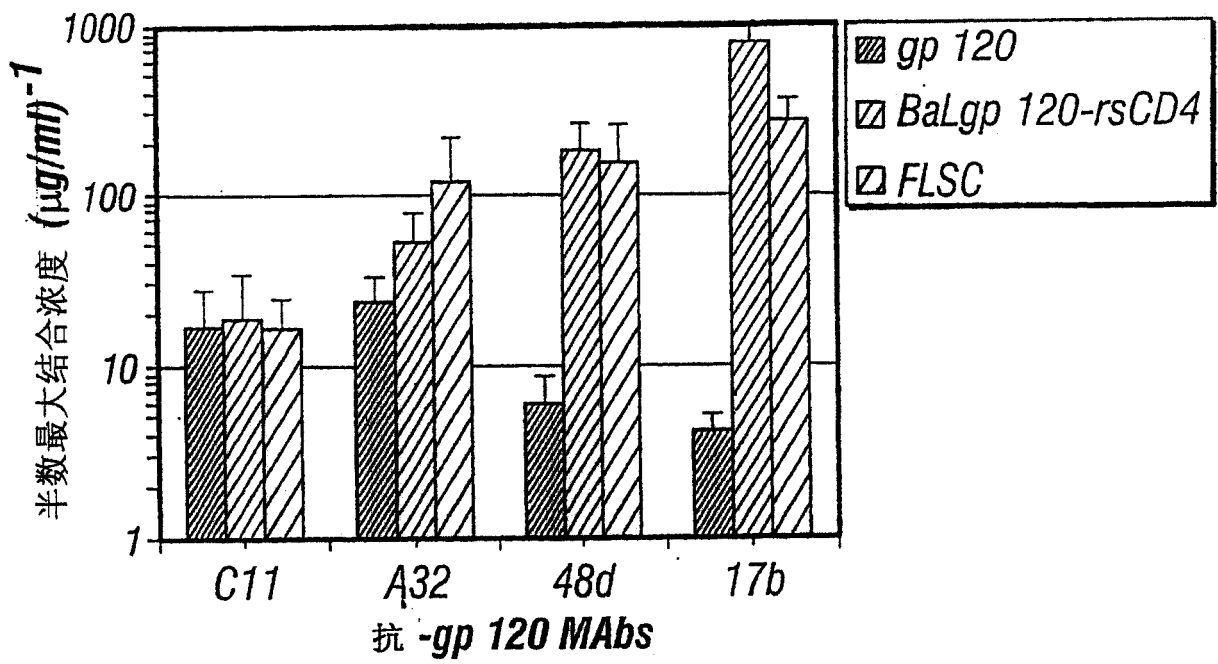


图5A

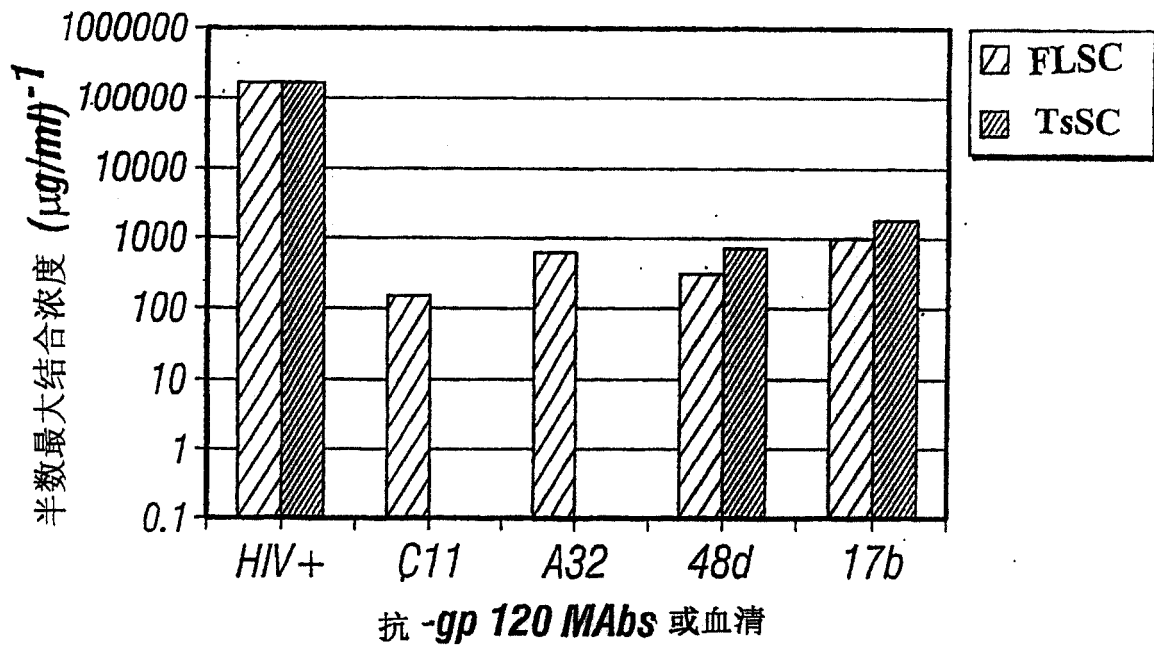


图5B

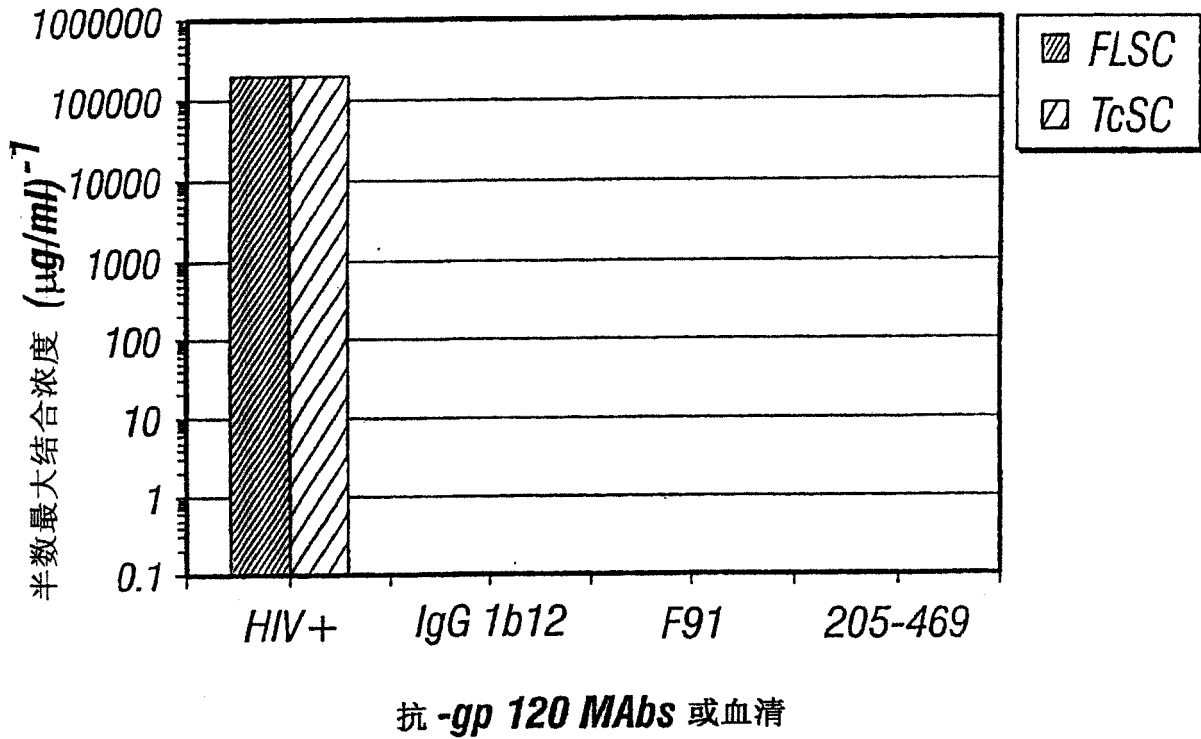


图5C

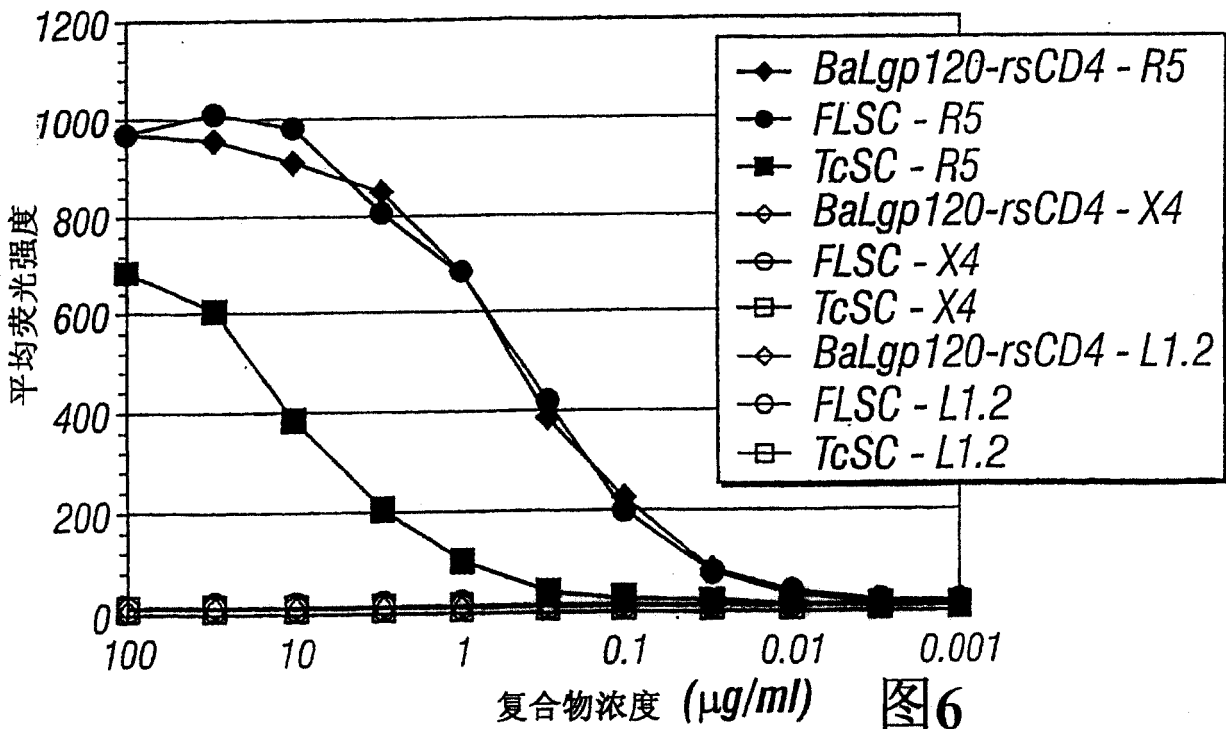


图6

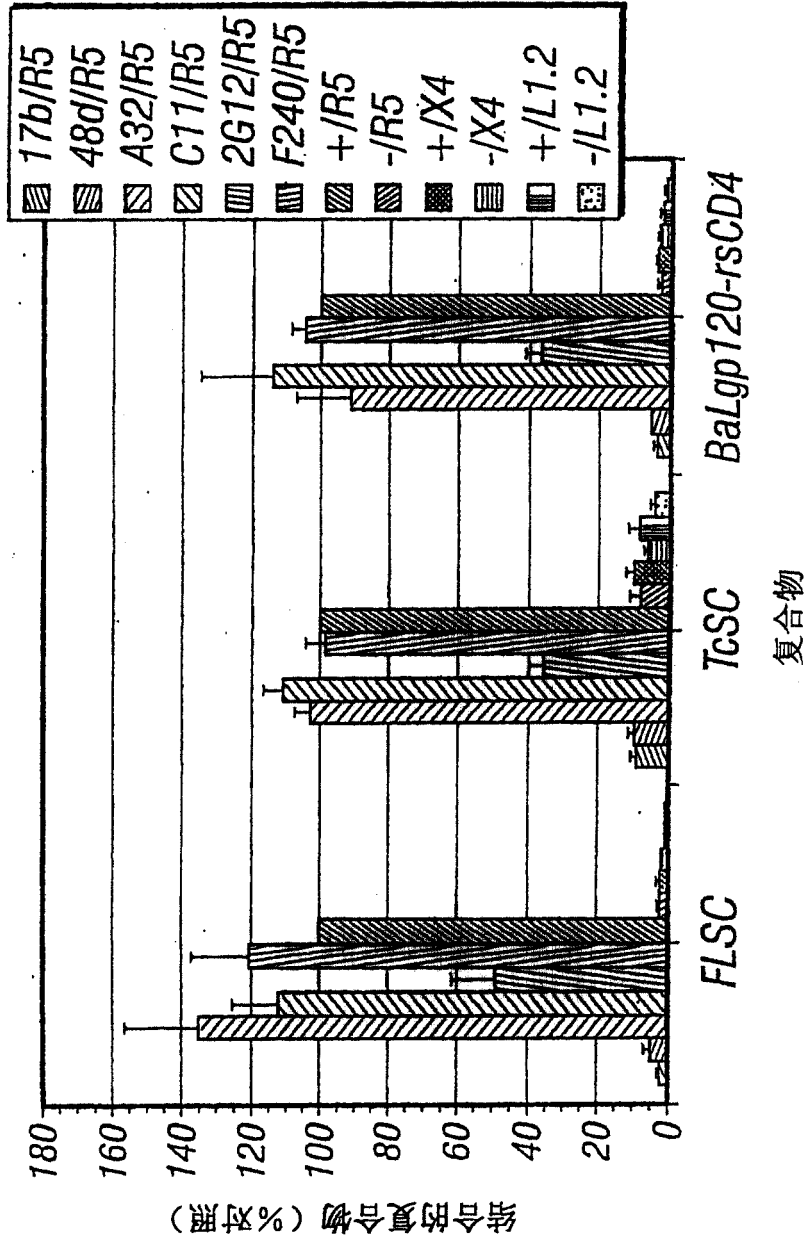


图7

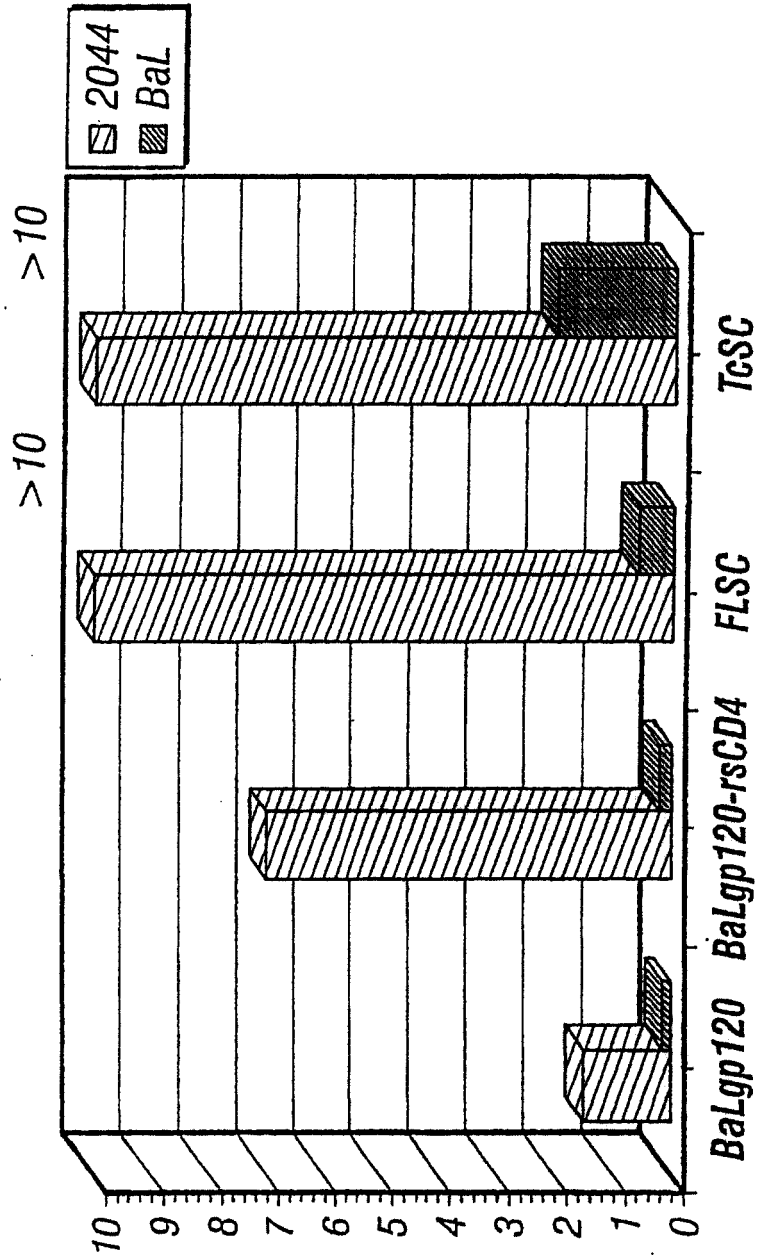


图8

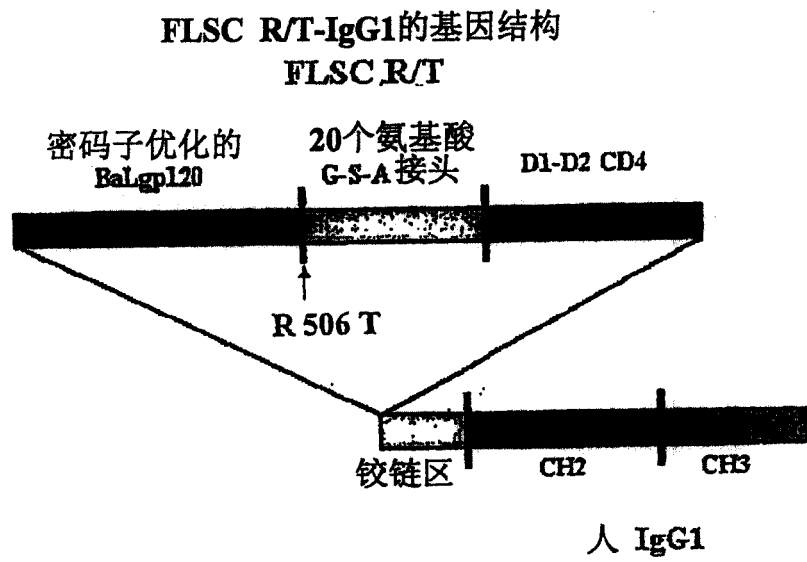


图9

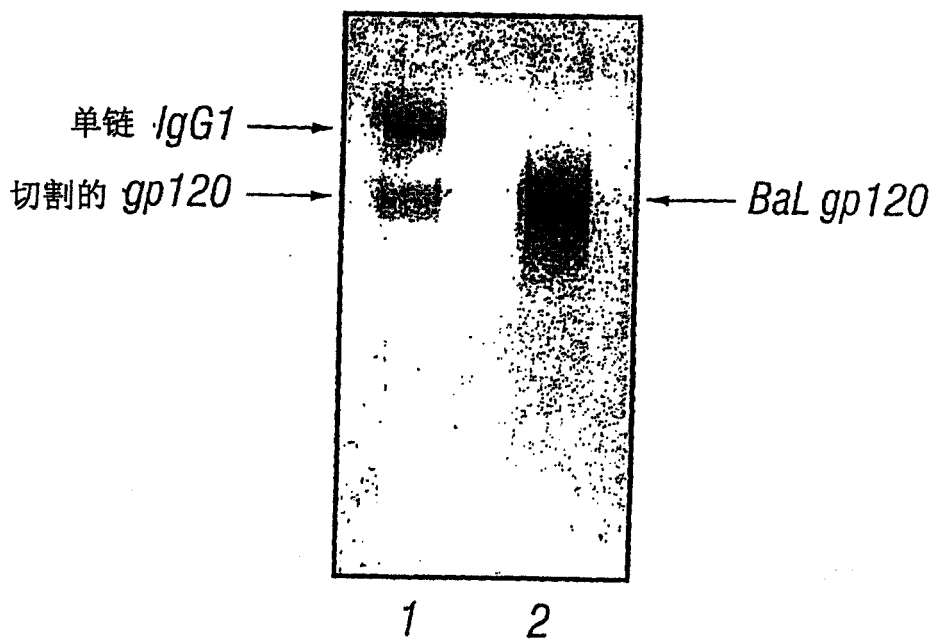


图10

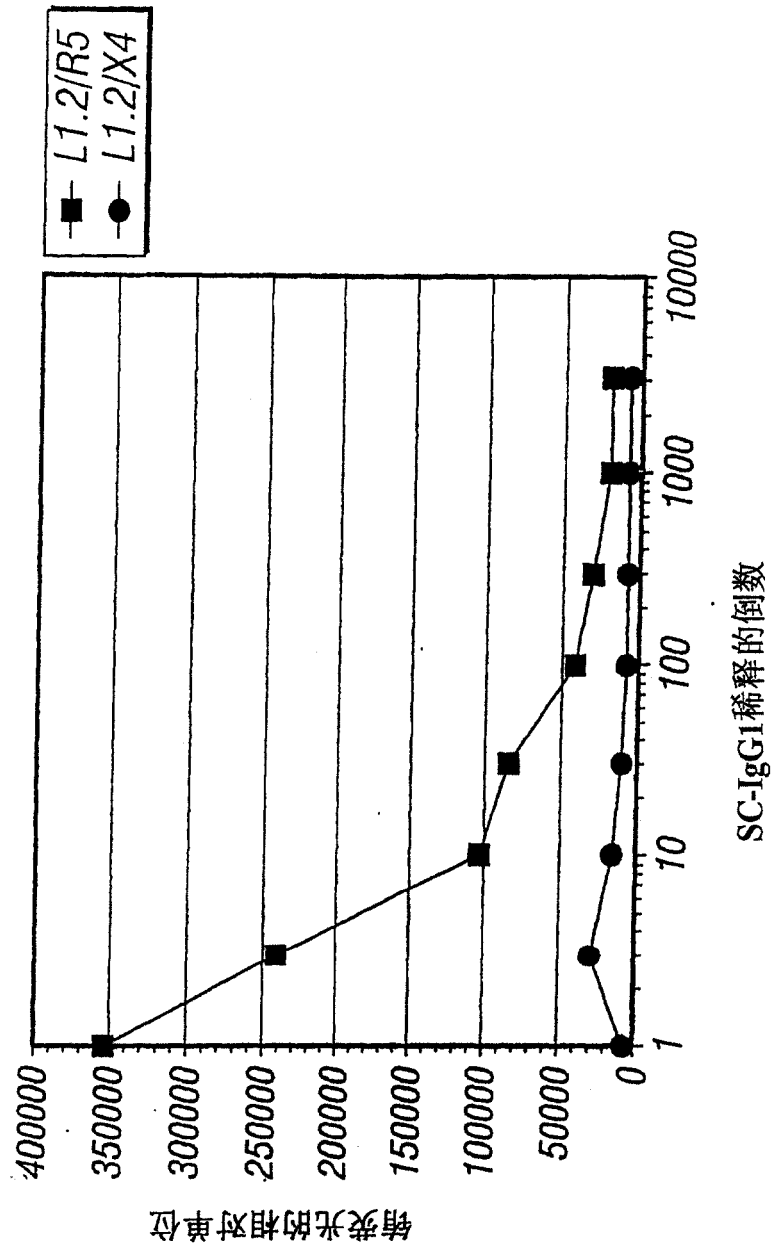


图11

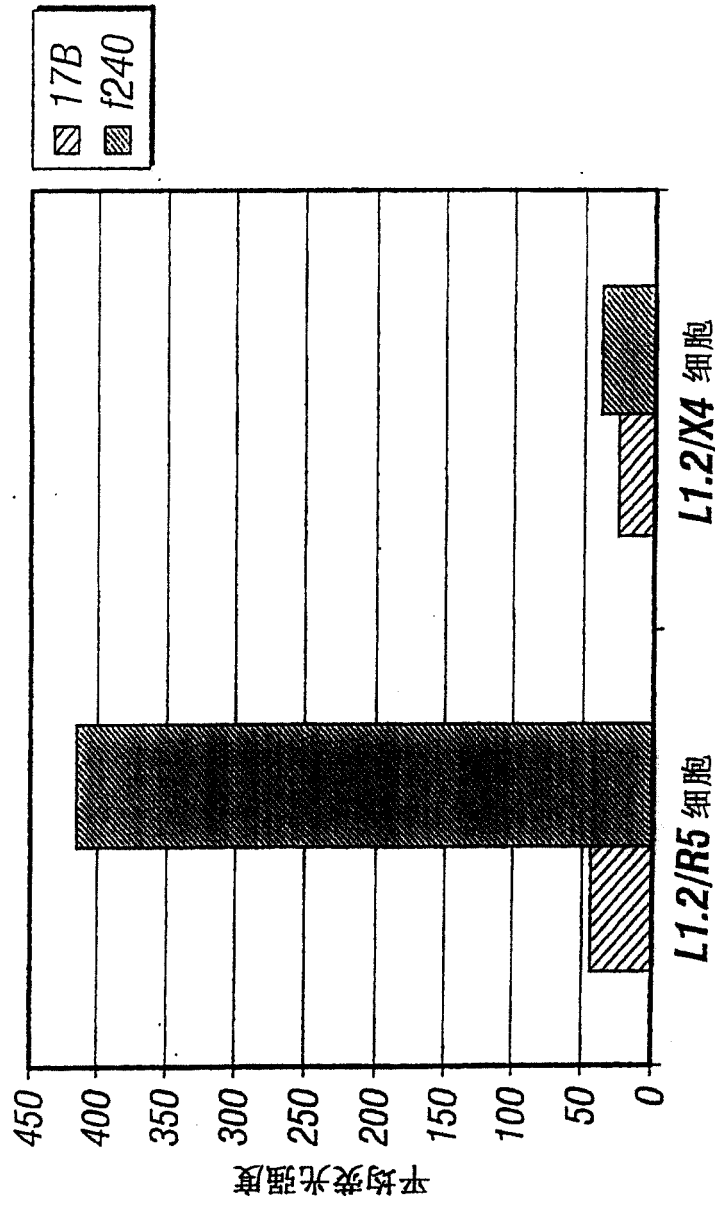
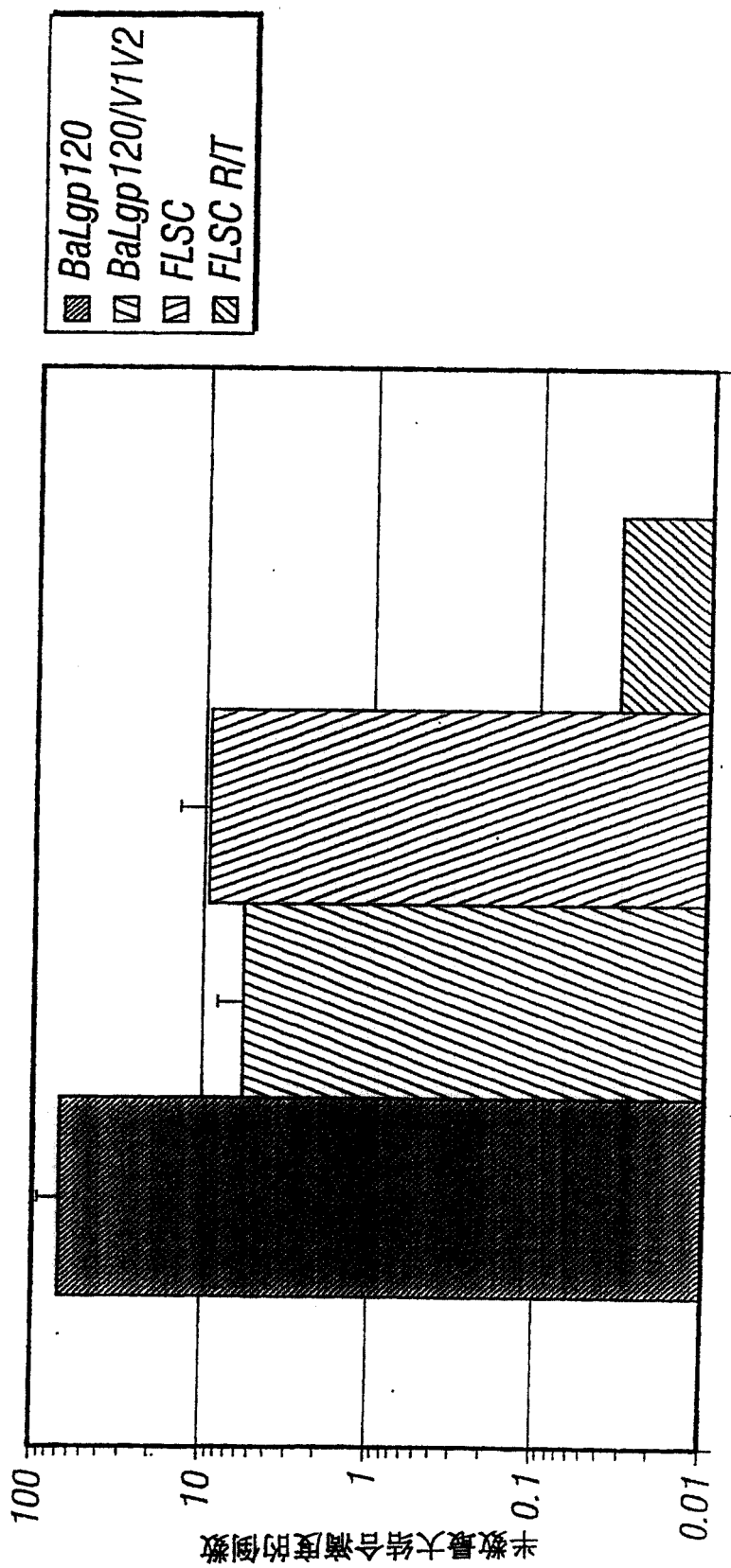


图12



复合物

图13

FLSC部分的免疫印迹

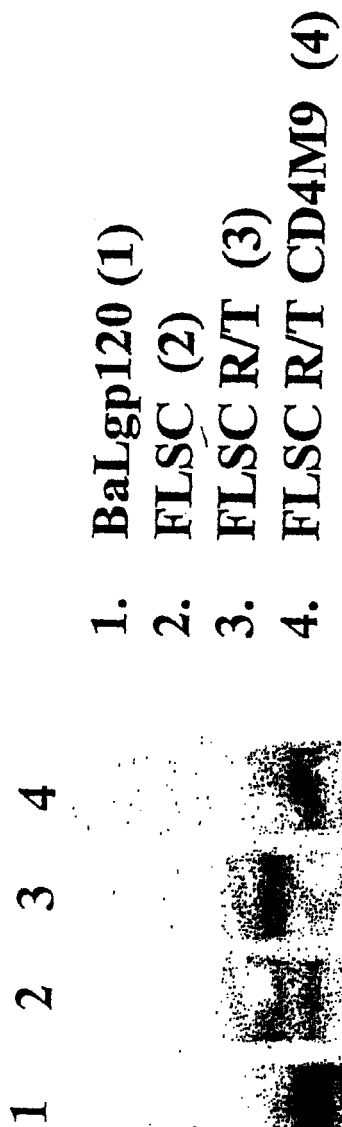


图14

图15

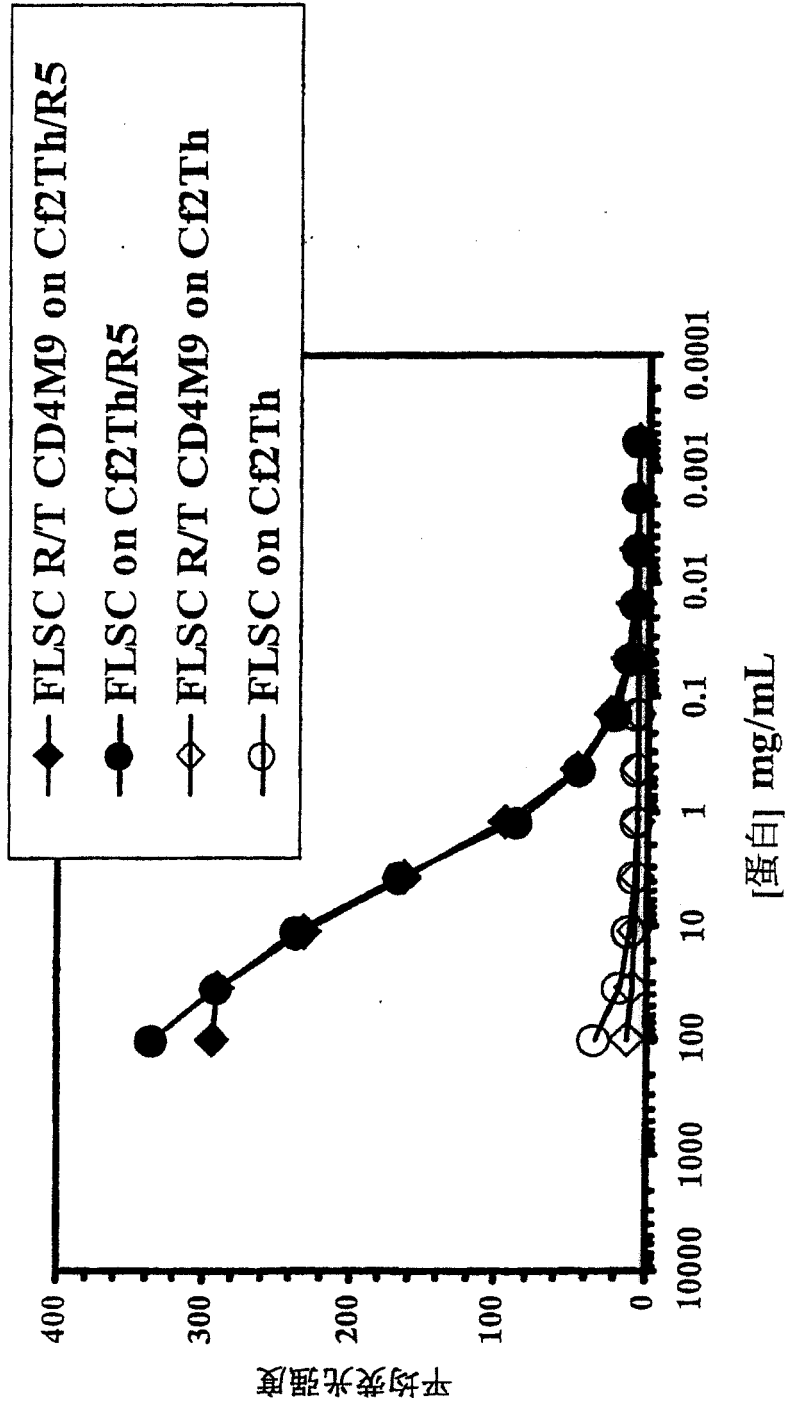
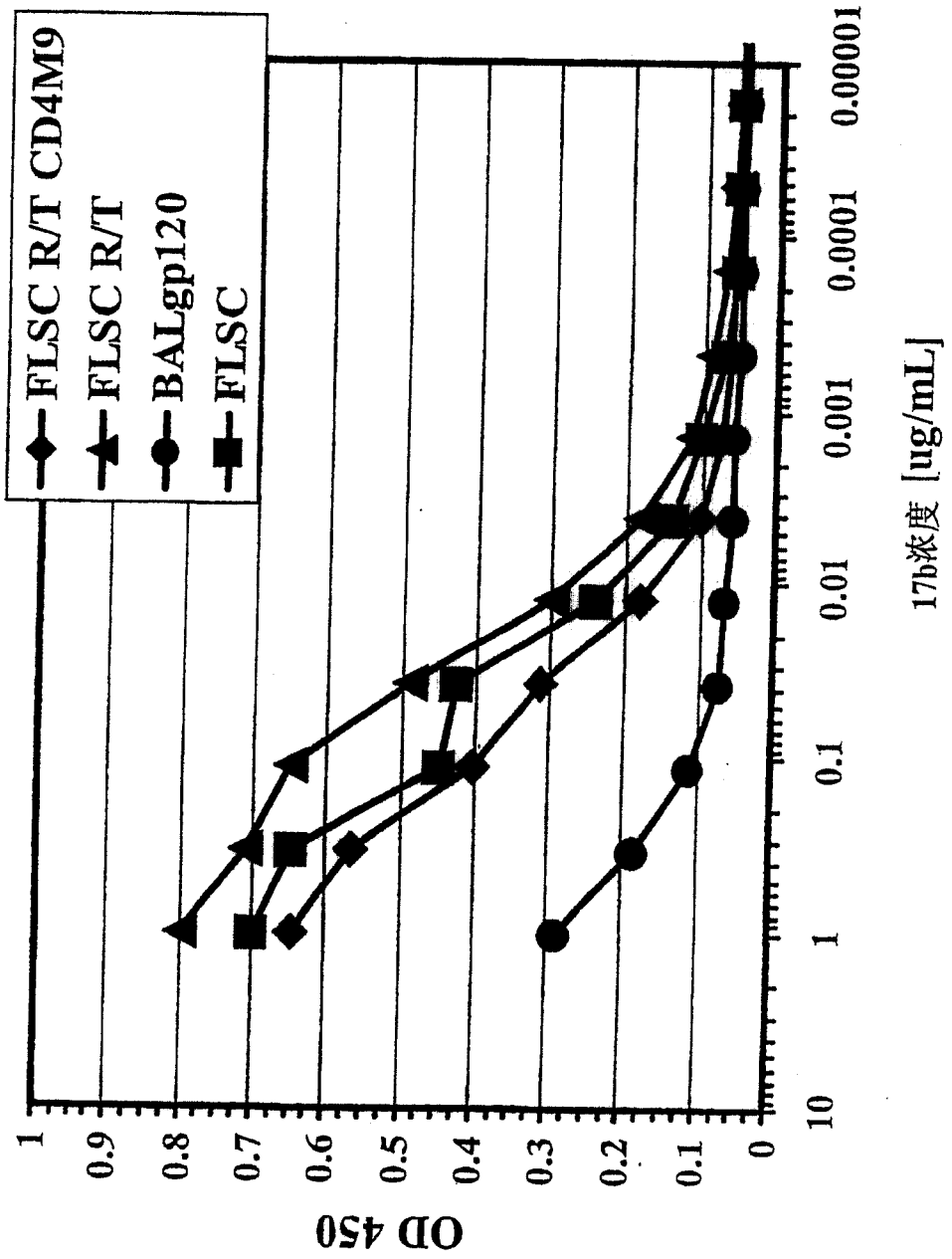


图16



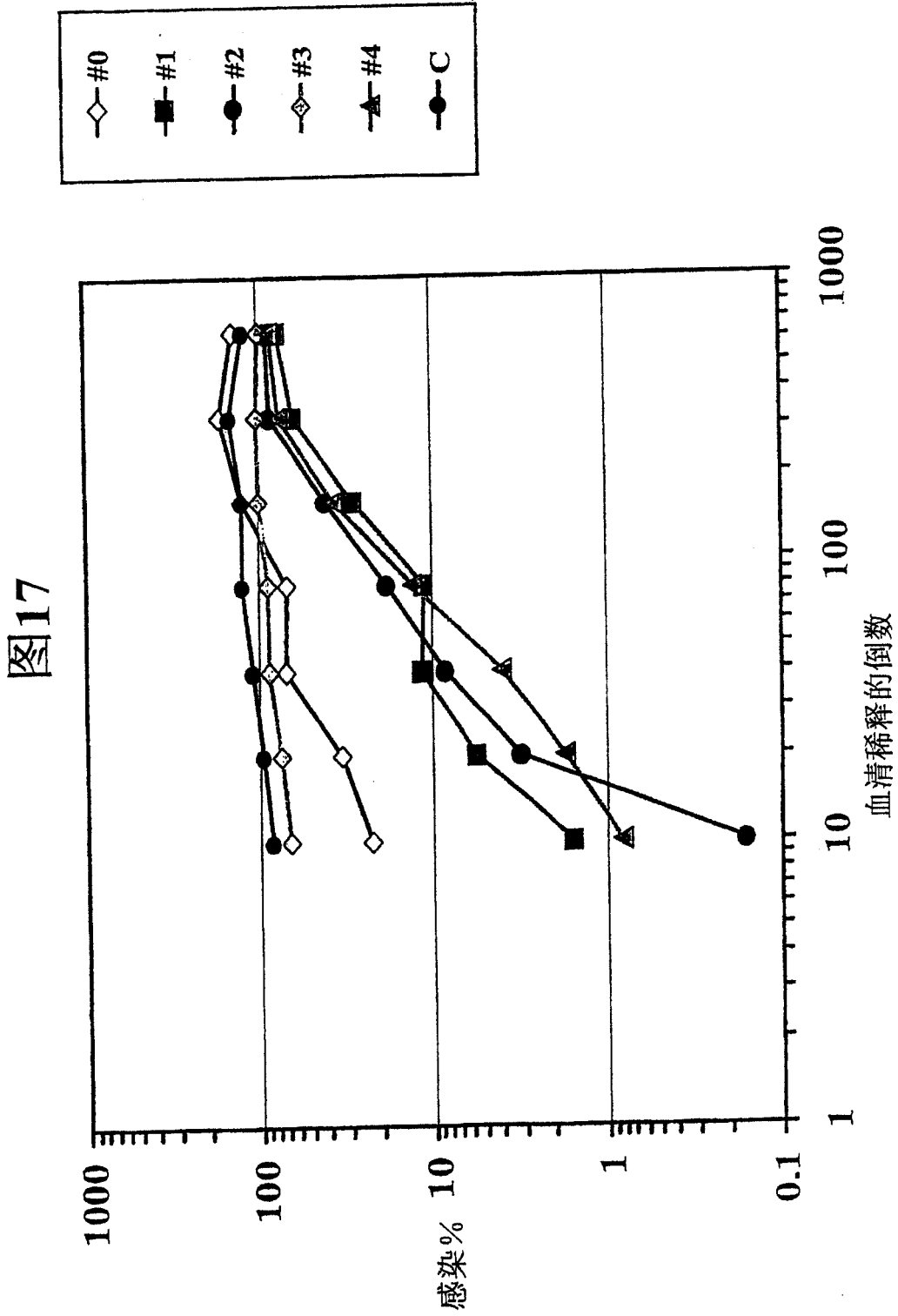
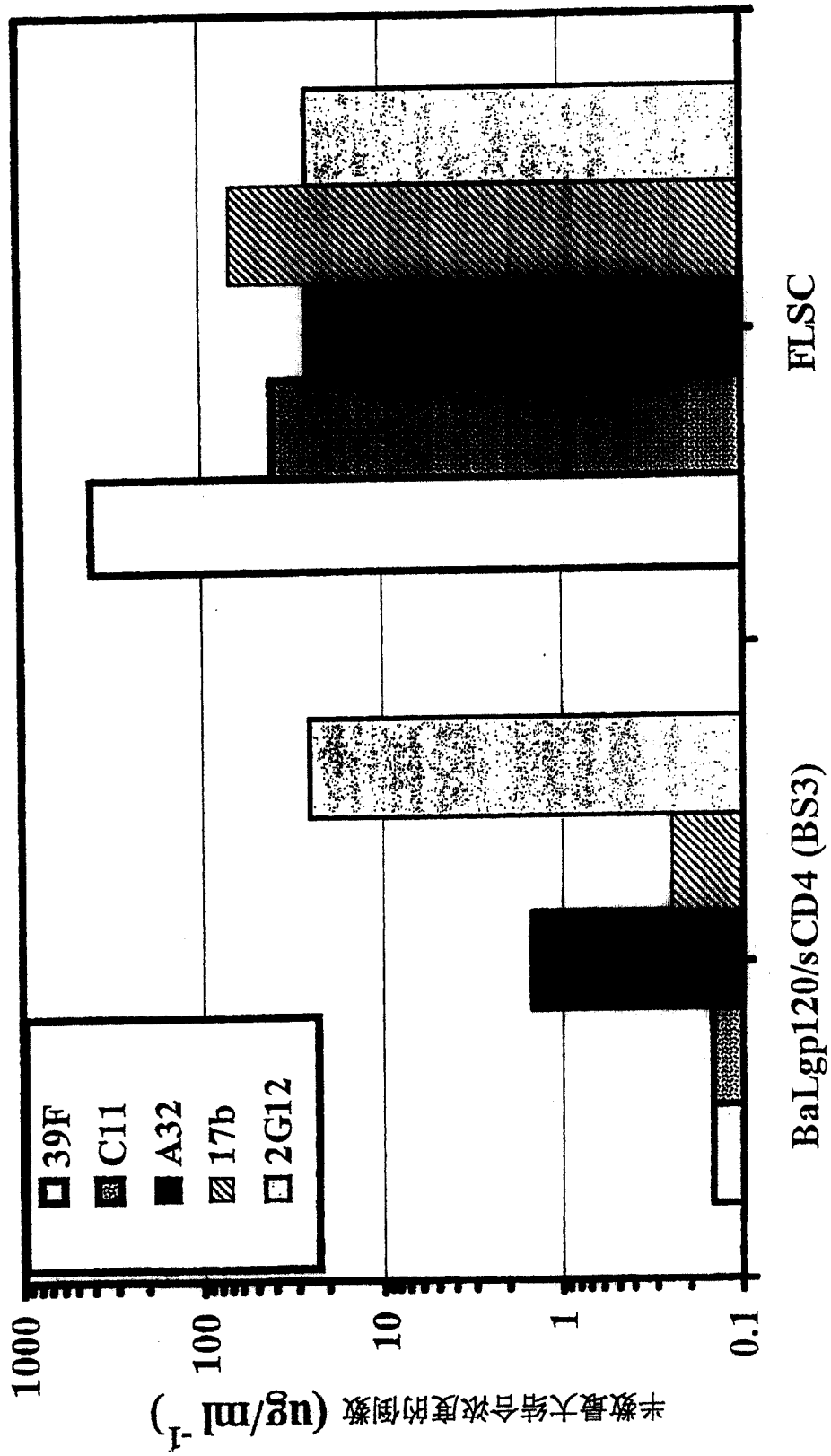
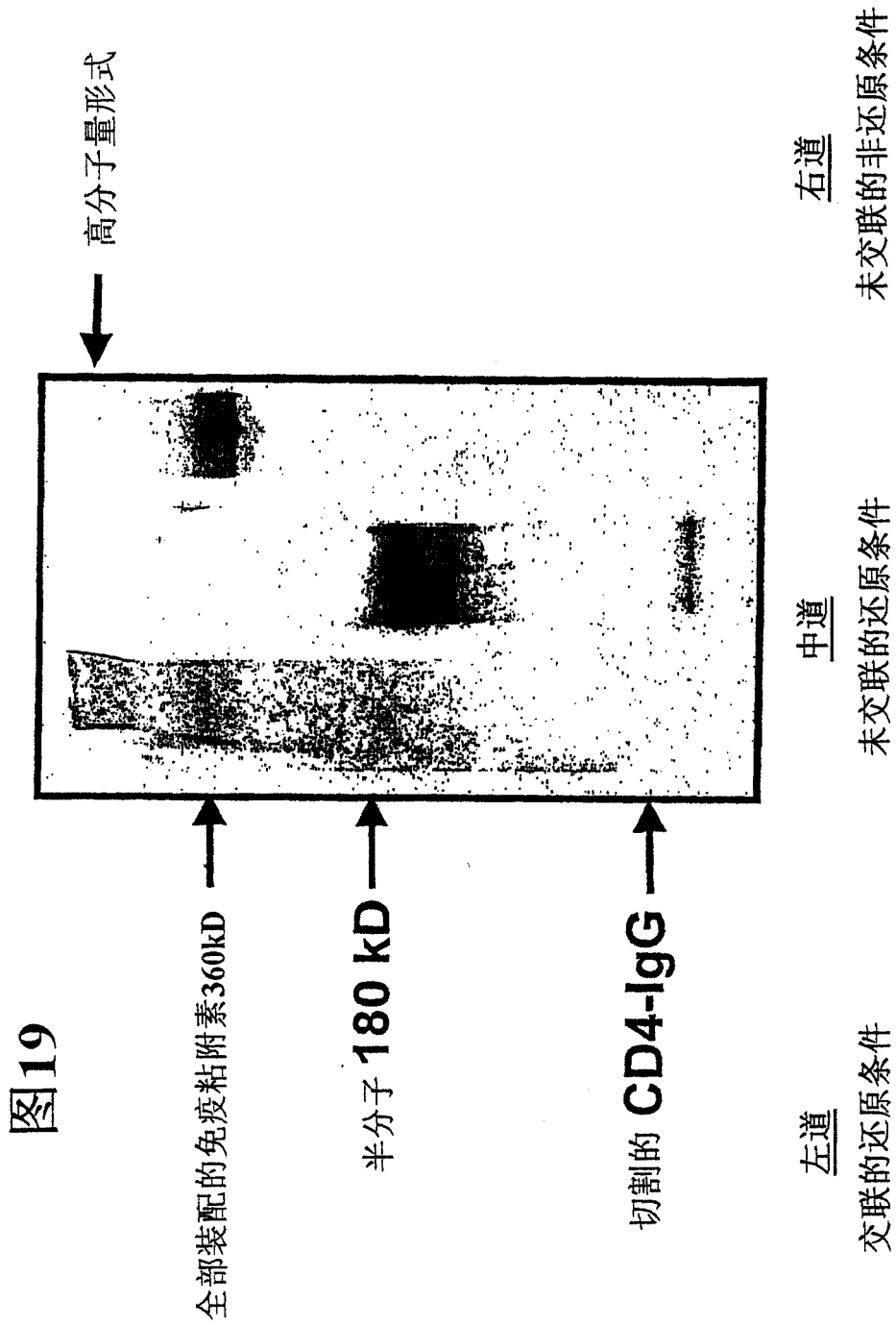


图18





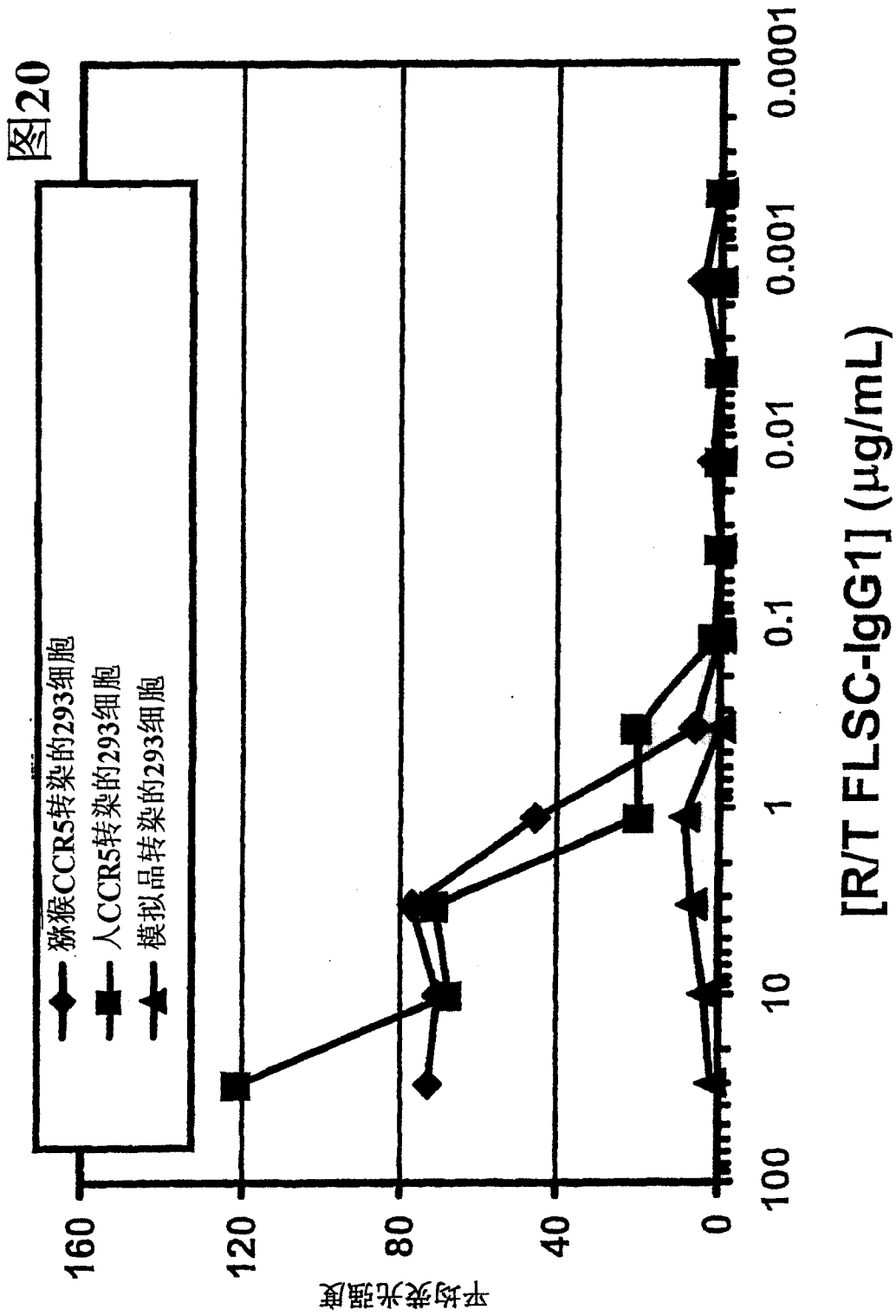
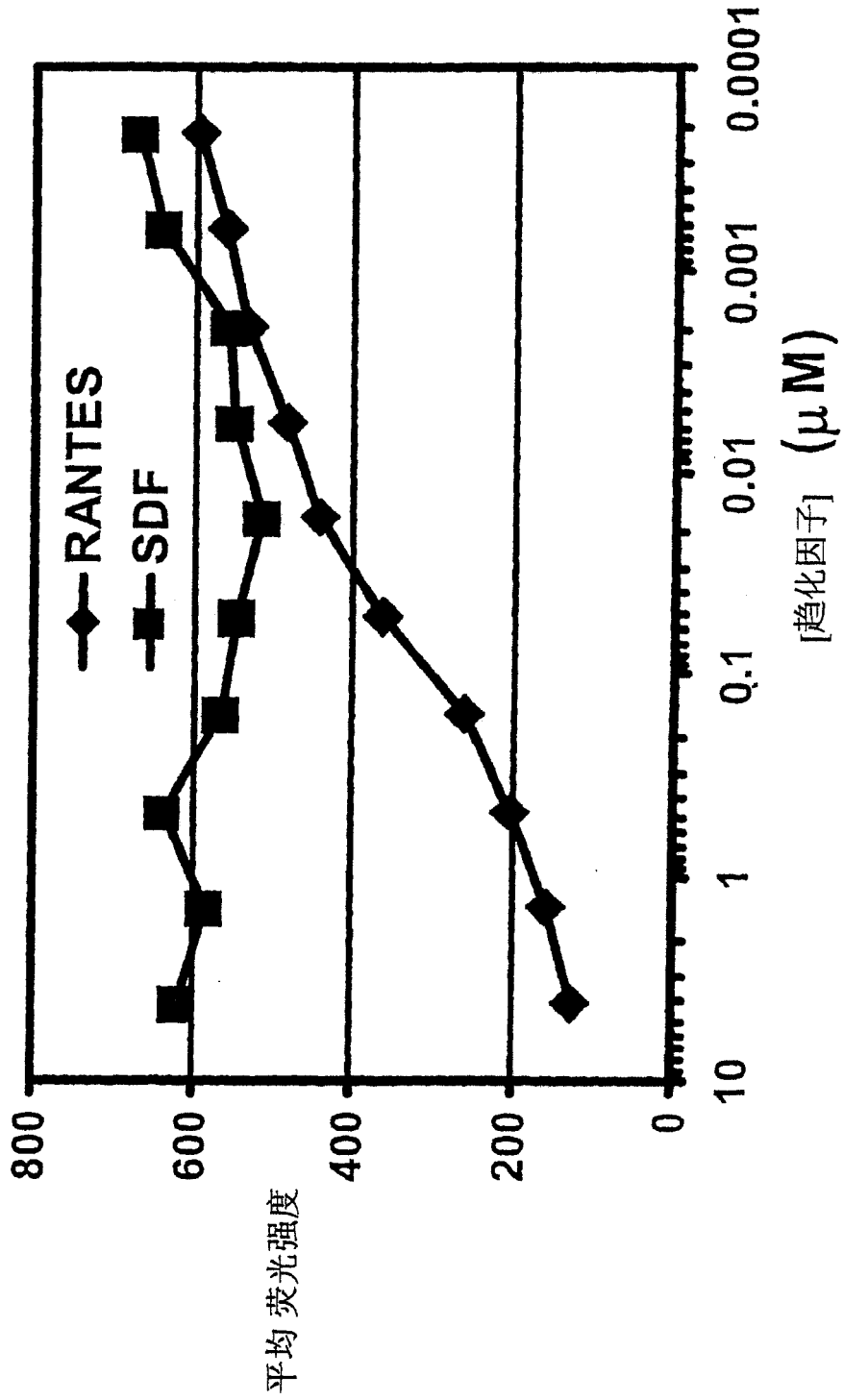


图21 R/T FLSC-IgG1与CCR5的结合受到RANTES的竞争性抑制



专利名称(译)	病毒外壳蛋白/受体嵌合体及使用方法		
公开(公告)号	CN100463694C	公开(公告)日	2009-02-25
申请号	CN02819350.4	申请日	2002-08-21
[标]发明人	安托尼L德维科 蒂莫西R福茨 罗伯特G图什坎		
发明人	安托尼·L·德维科 蒂莫西·R·福茨 罗伯特·G·图什坎		
IPC分类号	A61K39/21 A61P31/18 G01N33/50 A61K31/7088 A61K38/00 A61K39/00 A61K39/395 A61K48/00 A61P31/12 C07K14/005 C07K14/155 C07K14/16 C07K14/705 C07K14/73 C07K16/08 C07K16/10 C07K16/28 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/62 C12Q1/02 C12Q1/70 G01N33/15 G01N33/53 G01N33/569		
CPC分类号	C07K16/1063 C07K2317/34 C07K2319/00 C12N2740/16122 C07K2319/735 C07K14/70514 C07K2319/ /32 C12N2740/15022 G01N33/56983 C12N15/62 C07K2319/30 C07K14/005 G01N33/56988 C07K16/ /2812 C07K2319/40 G01N2500/00 A61K38/00 C07K2319/41 A61K39/00		
代理人(译)	杨青		
审查员(译)	苏林		
优先权	09/934060 2001-08-21 US		
其他公开文献	CN1561229A		
外部链接	Espacenet SIPO		

摘要(译)

本发明涉及含有病毒外壳序列和受体序列的嵌合分子，这些序列彼此之间能够相互作用，形成一种能够结合辅助受体的复合物。这样的嵌合分子因此表现出受体-外壳蛋白复合物特征的功能性质，可以用做试剂，其由于占据了细胞上存在的辅助受体而抑制病毒对细胞的感染。具体来说，嵌合的多肽包括免疫缺陷病毒的外膜多肽，例如HN、SN、FIV、FeL V、FPV和疱疹病毒的外膜多肽。适合于在嵌合多肽中使用的受体序列包括例如CD4 DID2和CD4M9序列。

受体 (结合亚基)	病毒 (科)	参考文献
免疫球蛋白类分子		
VCAM-1	EMC-D (小 RNA 病毒科)	Huber (1994)
(CAM-1) (第一个结构域)	大部分的 HRV, CAV 13, 18 和 21 (小 RNA 病毒科)	Celomo 等 (1986) Greve 等 (1989) Staunton 等 (1989) Tomassini 等 (1989)
PVR (第一个结构域)	脊髓灰质炎病毒 (小 RNA 病毒科) (1989)	Koike 等 (1990) Mendelsohn 等 (1989)
CD4 (第一个结构域)	HIV-1, 2; SIV (Lentiviridae) 人疱疹病毒 7	Dalglish 等 (1984) Klatzmain 等 (1984) Lusso 等 (1994)
CEA, 几个成员 (第一个结构域)	小鼠肝炎病毒 (冠状病毒科)	Williams 等 (1978)
MHC I	赛姆利基森林病毒 (Semliki forest virus) (披盖病毒科) Fectate dehydrogenase 病毒 小鼠细胞肥大病毒 (疱疹病毒科) SV-40	Hetenius 等 (1978) Oldstone 等 (1980) Inada 和 Mims (1984) Wykes 等 (1984) Brau 等 (1992)
MHC II	髓鞘脱落病毒 (Lentiviridae)	Dalziel 等 (1991)
整合蛋白		
VLA-2 (α-链)	埃可病毒 1.8 (小 RNA 病毒科)	Bergelson 等 (1992, 1993)
(RGD-结合蛋白)	FMDV (小 RNA 病毒科)	Fox 等 (1989) Mason 等 (1994)
αvβ3 (vitronectin)	CAV9, 埃可病毒 1.8 (小 RNA 病毒科)	Roivainen 等 (1994)
转运蛋白		
磷酸转运蛋白 Analogen	长臂猿白血病病毒 (反转录病毒科) Amphotropic murine (反转录病毒科)	Johann 等 (1992) Miller 等 (1994)
氨基酸阳离子转运蛋白	嗜紫性的鼠白血病病毒 (反转录病毒科)	Albritton 等 (1989)
信号受体		
LDL 受体蛋白家族	小部分的 HRV (小 RNA 病毒科) 乌白血性增生肉瘤病毒 A 亚类 (科未知)	Hofer 等 (1994) Bates 等 (1993) Connolly 等 (1994)
乙酰胆碱受体 (α-1)	狂犬病毒 (rhabdoviridae) 牛痘病毒 (痘病毒科)	Leniz 等 (1990) Marsh 和 Eppstein (1987)