



(12) 发明专利申请

(10) 申请公布号 CN 104619338 A

(43) 申请公布日 2015. 05. 13

(21) 申请号 201380032027. 2

C07K 14/005(2006. 01)

(22) 申请日 2013. 06. 17

(30) 优先权数据

61/661, 050 2012. 06. 18 US

(85) PCT国际申请进入国家阶段日

2014. 12. 17

(86) PCT国际申请的申请数据

PCT/EP2013/062553 2013. 06. 17

(87) PCT国际申请的公布数据

W02013/189901 EN 2013. 12. 27

(71) 申请人 诺华股份有限公司

地址 瑞士巴塞尔

(72) 发明人 A·卡尔菲 A·德伊 A·卡萨

I·斯里瓦斯塔瓦

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 余颖

(51) Int. Cl.

A61K 39/21(2006. 01)

权利要求书2页 说明书18页

序列表30页 附图11页

(54) 发明名称

稳定化的GP120

(57) 摘要

本发明提供了一种分离的多肽,其包含稳定于暴露已被CD4结合的表位和CD4结合位点表位的构象的HIV gp120多肽或可溶性gp140多肽。本发明还提供了治疗和预防HIV感染的免疫原性组合物和方法。

1. 一种分离的多肽,所述多肽包含 HIV gp120 多肽或可溶性 gp140 多肽,所述 HIV gp120 多肽或可溶性 gp140 多肽稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象。

2. 如权利要求 1 所述的分离的多肽,所述多肽通过内部结构域层之间的层间关联而稳定。

3. 如权利要求 1 或 2 所述的分离的多肽,所述多肽通过层 1 和层 2 之间的层间关联而稳定。

4. 如前述权利要求中任一项所述的分离的多肽,所述多肽通过层 2 和层 3 之间的层间关联而稳定。

5. 如前述权利要求中任一项所述的分离的多肽,所述层间关联是由二硫键形成的。

6. 如权利要求 5 所述的分离的多肽,所述二硫键在半胱氨酸残基之间形成,所述半胱氨酸残基所处位置相当于以下一对或多对氨基酸:SEQ ID NO:2 中的 V59C 和 S109C、V95C 和 W465C、V95C 和 R462C、V95C 和 L469C 和 / 或 H99C 和 R462C。

7. 如前述权利要求中任一项所述的分离的多肽,其中稳定化的 HIV gp120 多肽或可溶性 gp140 多肽包含与 SEQ ID NO:2 中所列内部结构域序列存在至少 90% 相同性的氨基酸序列且包含一对或多对非天然半胱氨酸残基,所述非天然半胱氨酸残基位于相当于 SEQ ID NO:2 中的 V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462 的位置。

8. 如前述权利要求中任一项所述的分离的多肽,所述稳定化的 HIV gp120 多肽或可溶性 gp140 多肽包含与 SEQ ID NO:2 中所列序列存在至少 90% 相同性的氨基酸序列且包含一对或多对非天然半胱氨酸残基,所述非天然半胱氨酸残基位于相当于 SEQ ID NO:2 中的 V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462 的位置。

9. 如前述权利要求中任一项所述的分离的多肽,所述稳定化的 HIV gp120 多肽或可溶性 gp140 多肽包含一对或多对非天然编码的半胱氨酸对,所述非天然编码的半胱氨酸对位于相当于 SEQ ID NO:2 中的 W90 和 E268、I103 和 Q413 的位置。

10. 如权利要求 7 或 8 所述的分离的多肽,所述稳定化的 HIV gp120 多肽包含 SEQ ID NO:4、6、8、10、12、14、16、18、20 或 22 中所列氨基酸序列。

11. 前述权利要求中任一项所述的分离的多肽的免疫原性片段,所述免疫原性片段包含并暴露 HIV gp120 上已被 CD4 结合的表位和 CD4 结合位点表位。

12. 一种融合蛋白,所述融合蛋白包含权利要求 1 至 10 中任一项所述的分离的多肽或者权利要求 11 所述的免疫原性片段和其他多肽。

13. 如权利要求 12 所述的融合蛋白,所述其他多肽是来自流感病毒的 HA 多肽。

14. 一种包含 3 个亚基的三聚体多肽,各亚基独立地选自权利要求 1 至 10 中任一项所述的分离的多肽、权利要求 11 所述的免疫原性片段以及权利要求 12 或 13 所述的融合蛋白。

15. 一种分离的多核苷酸,所述分离的多核苷酸编码权利要求 1 至 10 中任一项所述的分离的多肽、权利要求 11 所述的免疫原性片段或者权利要求 12 或 13 所述的融合蛋白。

16. 如权利要求 15 所述的多核苷酸,所述多核苷酸包含与 SEQ ID NO:3、5、7、9、11、13、15、17、19 或 21 所列序列具有至少 90% 相同性的核酸序列。

17. 一种免疫原性组合物,所述免疫原性组合物包含权利要求 1 至 10 中任一项所述的

分离的多肽、权利要求 11 所述的免疫原性片段、权利要求 12 或 13 所述的融合蛋白、权利要求 14 所述的三聚体和 / 或权利要求 15 或 16 所述的多核苷酸。

18. 如权利要求 17 所述的免疫原性组合物,所述免疫原性组合物还包含佐剂。

19. 用作疫苗的权利要求 17 或 18 所述的免疫原性组合物,所述疫苗用于治疗或预防 HIV 的感染和 / 或 AIDS。

20. 一种用于在对象中生成免疫应答的方法,所述方法包括向所述对象给予权利要求 1 至 10 中任一项所述的分离的多肽、权利要求 11 所述的免疫原性片段、权利要求 12 或 13 所述的融合蛋白、权利要求 14 所述的三聚体、权利要求 15 或 16 所述的多核苷酸或者权利要求 17 或 18 所述的免疫原性组合物。

21. 用于治疗或预防对象中 HIV 感染和 / 或 AIDS 的权利要求 1 至 10 中任一项所述的分离的多肽、权利要求 11 所述的免疫原性片段、权利要求 12 或 13 所述的融合蛋白、权利要求 14 所述的三聚体、权利要求 15 或 16 所述的多核苷酸和 / 或权利要求 17 或 18 所述的免疫原性组合物。

22. 一种用于治疗或预防对象中 HIV 感染和 / 或 AIDS 的方法,所述方法包括向所述对象给予权利要求 1 至 10 中任一项所述的分离的多肽、权利要求 11 所述的免疫原性片段、权利要求 12 或 13 所述的融合蛋白、权利要求 14 所述的三聚体、权利要求 15 或 16 所述的多核苷酸和 / 或权利要求 17 或 18 所述的免疫原性组合物。

23. 如权利要求 21 所述的分离的多肽、免疫原性片段、融合蛋白、三聚体、多核苷酸或免疫原性组合物或者如权利要求 20 或 22 所述的方法,所述对象是人。

稳定化的 GP120

[0001] 本专利申请要求 2012 年 6 月 18 日提交的美国临时专利申请 61/661,050 号的优先权,其全部内容通过引用纳入本文。

技术领域

[0002] 本发明的涉及 HIV 领域,且具体是 HIV 疫苗领域。

背景技术

[0003] HIV/AIDS 大流行已经席卷全球,截至 2007 年底已有超过 6000 万人感染 HIV-1 [1]。虽然已动用的大量资源以对抗该大流行,但仍没有可用的针对 HIV-1 的疫苗。虽然人体会针对初始的 HIV-1 感染生成强免疫应答,但许多这类抗体无法中和病毒 [2]。一个原因是存在多重水平的防护,其保护病毒对抗中和抗体。这包括高度的序列变异性、碳水化合物遮蔽、多聚物形成导致的表位闭塞和构象遮蔽 [3]。

[0004] 虽然存在高度的变异性,但病毒需要维持保守的能够识别和结合细胞受体的位点。这些 HIV-1gp120 中保守的受体结合位点,包括 CD4 结合位点和 CD4 诱导的(共受体)结合位点,分别包括 CD4 结合位点表位和已被 CD4 结合的表位,是对抗体介导的中和而言易受攻击的靶标并因此是有吸引力的疫苗生产靶标。然而,这些保守位点和这些位点上的表位是构象遮蔽的并因此无法接触中和抗体。

[0005] 过去所进行的尝试试图通过稳定内部和外部结构域关联来稳定 gp120 使之形成暴露蛋白保守区域的构象。例如,先前试图稳定 gp120 使之形成结合受体的构象的结构指导的尝试利用了多种方法,包括二硫桥键和空腔填充突变 [4],以及在稳定化前除去免疫优势区域(如 V1/V2 和 V3 环) [5]。通过在使用半胱氨酸残基进行靶向取代后形成二硫桥键,可获得最大程度的稳定化 [6 和 7]。然而,虽然这些二硫键稳定了 gp120 的结构使之形成暴露 gp120 中已被 CD4 结合的表位的构象,但其无法暴露 CD4 结合位点表位,如通过 CD4 结合位点抗体的亲和性所测量的那样。除此之外,策略还涉及除去同样不利的可变环,因为可变环涉及 gp120 的四级结构的调控 [8]。

发明内容

[0006] 公开的 gp120 结构中缺失了 HIV-1gp120 的内部结构域的结构 [4, 9 和 10]。然而,在 CD4 配体结合后 gp120 中发生的构象改变过程中,内部结构域的层的相互作用是重要的。本发明涉及 HIV-1gp120 和可溶性 HIV-1gp140 的内部结构域中层间相互作用的稳定化。

[0007] HIV 结构

[0008] 新揭示的 gp120 的内部结构域的结构显示,在结合 CD4 的状态中,内部结构域被组织成 3 层,由层 1、2 和 3 组成,从 β -三明治结构投影至靶细胞膜(图 1) [11-13]。内部结构域由对应于来自 SF162 毒株的 HIV gp120 (SEQ ID NO:2) 的氨基酸 1-248 和 458-497 的氨基酸组成。层 1 由对应于来自 SF162 毒株的 HIVgp120 的氨基酸 42-76 的氨基酸组成。层 2 由对应于来自 SF162 毒株的 HIV gp120 的氨基酸 90-117 和 192-223 的氨基酸组成。层 3

由对应于来自 SF162 毒株的 HIV gp120 的氨基酸 239-249 和 459-470 的氨基酸组成。内部结构域被鉴定为在多肽从未配体化至结合 CD4 的状态转变时经历剧烈构象变化的 gp120 的区域 [14, 15]。CD4 结合后发生的内部结构域中的构象变化导致 HIV 病毒与靶细胞的成功融合所必需的两个下游事件的发生。这些事件是：

[0009] 1. 初始弱 CD4 接触的稳定化, 以及

[0010] 2. 暴露 gp120 上的保守位点用于共受体结合, 即已被 CD4 结合的表位的暴露。

[0011] 在结合 CD4 的状态中, gp120 的内部结构域的三层通过特异性层间相互作用结合在一起并稳定 gp120 的结合 CD4 的构象。

[0012] 层 1 和 2 之间的相互作用通过多个残基介导, 所述残基已通过定点诱变进行鉴定 [11, 12]。层 1 和 2 之间的这些相互作用形成了独特的结构, 与“衣领”类似, 围绕内部结构域并稳定该区域 [8]。这些层间相互作用的破坏引起 gp120 的结合 CD4 的构象的去稳定化, 导致与 CD4 的亲合性降低 [11]。类似地, 降低内部结构域的层之间相互作用的突变还阻止了 gp120 与小分子 CD4 模拟物 (如 NBD-556, 其需要 gp120 的结合 CD4 的构象以进行高亲和性相互作用) 的结合 [11, 16]。抗体 (如识别 gp120 上 CD4 诱导的保守结合位点的 17b、48d 和 412d [17, 18 和 27]) 的结合也因为阻止内部结构域中层间相互作用的突变而大幅降低 [11]。通常, 降低的层间相互作用导致保守结合位点 (即 CD4 和共受体结合位点) 暴露的减少, 并因此降低了识别这些位点的配体的结合。

[0013] 与上文中公开的方法相反, 本发明人发现, 内部结构域中层间关联的稳定化诱导 gp120 形成已被 CD4 结合的表位和 CD4 结合位点表位 (也称作 CD4 诱导的或 CD4i 表位) 都暴露的构象。因此, 本发明提供了:

[0014] • 包含稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象的 HIV gp120 多肽或可溶性 gp140 多肽的分离的多肽

[0015] • 包含 HIV gp120 的已被 CD4 结合的表位和 CD4 结合位点表位的稳定化的 HIV gp120 多肽或可溶性 gp140 多肽的免疫原性片段

[0016] • 包含三个亚基的三聚体多肽, 各亚基独立地选自特定分离的多肽, 所述分离的多肽包含稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象的 HIV gp120 多肽、稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象的可溶性 gp140 多肽、包含 HIV gp120 的已被 CD4 结合的表位和 CD4 结合位点表位的稳定化的 HIV gp120 多肽和包含 HIV gp120 的已被 CD4 结合的表位和 CD4 结合位点表位的稳定化的可溶性 gp140 多肽的免疫原性片段

[0017] • 编码上文所列多肽的多核苷酸

[0018] • 包含本发明的多肽、免疫原性片段、融合蛋白和 / 或多核苷酸的免疫原性组合物

[0019] • 在对象中产生免疫应答的方法

[0020] • 治疗或预防 HIV 感染的方法

[0021] 稳定化的 HIV gp120 和可溶性 gp140 多肽

[0022] 本发明提供了一种分离的多肽, 其包含稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象的 HIV gp120 多肽或可溶性 gp140 多肽。“稳定于同时暴露结合 CD 的和 CD4 结合位点表位的构象”指甚至在不存在 CD4 配体的情况下, HIV gp120 或可溶性 gp140 多肽的内部结构域仍形成结合 CD4 的构象。因此, 在野生型 gp120 或可溶性 gp140

中仅在结合 CD4 时暴露的构象遮蔽表位在本发明的稳定化的 gp120 或可溶性 gp140 多肽中持续暴露。此外,本发明的稳定化的 gp120 或可溶性 gp140 多肽在不存在 CD4 的情况下形成结合 CD4 的构象,并因此 CD4 结合位点表位也可与已被 CD4 结合的表位同时暴露。

[0023] 根据本发明,包含稳定于同时暴露结合 CD 的和 CD4 结合位点表位的构象的 HIV gp120 多肽或可溶性 gp140 多肽的分离的多肽是特异性结合以下物质的多肽

[0024] (i) 抗 -gp120CD4 结合位点抗体,即特异性针对 gp120 中 CD4 结合位点的抗体,

[0025] (ii) 特异性针对 gp120 中 CD4- 诱导的保守结合位点的抗体,以及

[0026] (iii) 可选的,其他抗 -gp120 或抗 -可溶性 gp140 抗体。

[0027] “特异性结合”指抗体以明显强于 BSA 的亲合性结合本发明的多肽。优选地,本发明的多肽的亲合性比 BSA 强至少 100 倍、 10^3 倍、 10^4 倍、 10^5 倍、 10^6 倍等。

[0028] 典型的抗 -gp120 抗体是本领域熟知的且包括例如 B12[19]、F105[20]、17b[17]、48d[27]、412d[21]、B13[22] 和 2G12[23],以及可获自 ATCC 的抗体 46-2(CRL-2186)、46-4(CRL-2178)、46-5(CRL-2184)、55-2(CRL-2155)、55-36(CRL-2153)、55-6(CRL-2185) 和 55-83(CRL-2395)。抗 -gp120CD4 结合位点 (α -gp120-CD4BS) 抗体包括单克隆抗体 B12、F105、JL413[24]、1795[25]、448-D(ATCC HB-10895)、558-D(ATCC HB-10894) 和 559/64-D(ATCC HB-10893)[26]。特异性针对 gp120 中 CD4- 诱导的保守结合位点 (α -gp120CD4i) 的抗体包括参考文献 27 中描述的 17b、48d、412d 和 23e (2.3E)。这些 α -gp120CD4i 抗体可获自 NIH AIDS 研究和参比试剂程序 (<https://www.aidsreagent.org/>)。这些抗体的序列已经公开 [28],且核苷酸序列可获自 IGBLAST(<http://www.ncbi.nlm.nih.gov/igblast/>) 或 ImMunoGeneTics 信息系统 (IMGT) (<http://imgt.cines.fr/>) 数据库。

[0029] 在一个实施方式中, gp120 多肽或可溶性 gp140 多肽被一种或多种内部结构域的层之间的层间关联所稳定。在一个特定实施方式中,稳定化关联是由一对非天然编码的半胱氨酸残基之间的二硫键产生的。

[0030] 稳定化关联可在层 1 和 2、2 和 3 或层 1 和 2 以及 2 和 3 之间产生。

[0031] 为在两个半胱氨酸残基之间形成二硫桥键,侧链必须相距正确的距离并相对于彼此处于正确的方向。对位于内部结构域的层 1、2 和 3 的界面上的所有残基的检测鉴定到了能够被成对的半胱氨酸残基靶向取代用于形成二硫键的残基。合适残基的鉴定基于 gp120 的晶体结构 [13]。如果残基对预期具有 $3.9\text{\AA}$ 至 $5.8\text{\AA}$ 的 C β -C β 距离,则认为其是合适的。

[0032] 基于此,以下残基对经鉴定用于被半胱氨酸取代以在层之间形成稳定化的二硫桥键。V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462(残基编号根据 SF162 毒株)。可对与来自 SF162 毒株的 gp120 多肽中上文所列氨基酸对相当的来自其他 HIV-1 毒株的 gp120 中的氨基酸位置进行鉴定,例如通过将来自第二毒株的 gp120 多肽序列与 SF162gp120 多肽序列比对。图 5 显示了来自 HIV-1 的 SF162 与 Hxb2 毒株的 gp120 之间的比对。Hxb2gp120 中对应的残基如灰色编号中所示。用于被半胱氨酸取代以形成稳定化的二硫桥键的对应残基对是 Hxb2 中的 V65 和 S115、V101 和 W479、V101 和 R476、V101 和 L483 以及 H105 和 R476。

[0033] 类似地,可通过将 gp120 多肽序列与 gp140 序列进行比对来鉴定可溶性 gp140 多肽中对应的残基。SF162 可溶性 gp140 氨基酸序列示于 SEQ ID NO:23。

[0034] 因此,本发明提供了一种分离的多肽,其包含稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象的 HIV gp120 多肽或可溶性 gp140 多肽,其中通过内部结构域中层间二硫键实现稳定化且该二硫键是由位于与野生型 SF162HIV gp120 多肽中一对或多对以下对相当位置的半胱氨酸残基形成的:V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462。

[0035] 在一个特定实施方式中,稳定化的 HIV gp120 多肽或可溶性 gp140 多肽包含具有与 SEQ ID NO:2 中所列内部结构域序列存在至少 a% 相同性的氨基酸序列的内部结构域,其中 a 是选自 60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 或更大的数值,且包含一对或多对非天然产生的半胱氨酸残基,所述半胱氨酸残基的位置与 SEQ ID NO:2 中 V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462 的位置相当。

[0036] 在另一个实施方式中,稳定化的 HIV gp120 多肽或可溶性 gp140 多肽包含具有与 SEQ ID NO:2 中所列序列存在至少 b% 相同性的氨基酸序列,其中 b 是选自 60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、99.5% 或 99.9% 或更大的数值,且包含一对或多对非天然产生的半胱氨酸残基,所述半胱氨酸残基的位置与 SEQ ID NO:2 中 V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462 的位置相当。

[0037] 在另一个实施方式中,稳定化的 HIV gp120 多肽包含 SEQ ID NO:4、6、8、10、12、14、16、18、20 或 22 中所列氨基酸序列。

[0038] 本发明的稳定化的 HIV gp120 多肽或可溶性 gp140 多肽还可包含本领域已知的其他结构域间和层间稳定化。例如,本发明的稳定化的 HIV gp120 多肽可包含参考文献 4 中所述的空腔填充突变或者参考文献 6 和 7 中所述的二硫键。具体而言,本发明的稳定化的 HIV gp120 多肽还可由一对或多对非天然编码的半胱氨酸对之间的二硫键稳定,所述半胱氨酸对的位置对应于 SEQ ID NO:2 中的 W90 和 E268、I103 和 Q413。

[0039] 本发明还提供了本发明的多肽的免疫原性片段,其中所述免疫原性片段包含并同时暴露 HIV gp120 的已被 CD4 结合的表位和 CD4 结合位点表位。所述免疫原性片段的长度可以是至少 300、350、400、450、475、480、485、490、495 或 497 个氨基酸。在一个特定实施方式中,所述免疫原性片段包含位于与野生型 SF162HIV gp120 多肽中一对或多对以下对相当位置的半胱氨酸残基:V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469 和 / 或 H99 和 R462。

[0040] 与 SEQ ID NO:2 或 SEQ ID NO:23 相比,本发明的多肽可包括一个或多个(如 1、2、3、4、5、6、7、8、9、10 个等)保守性氨基酸取代以及半胱氨酸残基,其位于在一对或多对以下对处:V59 和 S109、V95 和 W465、V95 和 R462、V95 和 L469、H99 和 R462。与 SEQ ID NO:4、6、8、10、12、14、16、18、20 或 22 相比,本发明的多肽可包括一个或多个(如 1、2、3、4、5、6、7、8、9、10 个等)保守性氨基酸取代。保守性氨基酸取代被定义为使用具有相关侧链的一个氨基酸取代另一个氨基酸,前提是氨基酸取代不降低 gp120 多肽的稳定化,即该多肽仍稳定于同时暴露已被 CD4 结合的表位和 CD4 结合位点表位的构象。遗传编码的氨基酸通常分为四类:(1) 酸性,即天冬氨酸、谷氨酸;(2) 碱性,即赖氨酸、精氨酸、组氨酸;(3) 非极性,即丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸;和(4) 不带电的极性氨基酸,即甘氨酸、天冬酰胺、谷氨酰胺、胱氨酸、丝氨酸、苏氨酸、酪氨酸。有时将苯丙氨酸、色氨酸和酪氨酸一起归类为芳族氨基酸。通常,这些家族中的单个氨基酸的取代不

会对生物活性产生重要影响。此外,相对于 SEQ ID NO:2、4、6、8、10、12、14、16、18、20、22 或 23,所述多肽可含有一个或多个(如 1、2、3、4、5、6、7、8、9、10 个等)单个氨基酸的缺失。另外,相对于 SEQ ID NO:2、4、6、8、10、12、14、16、18、20、22 或 23,所述多肽可包含一个或多个(如 1、2、3、4、5、6、7、8、9、10 个等)插入(如各 1、2、3、4 或 5 个氨基酸)。在通过二硫键进行稳定的实施方式中,缺失、插入和取代可发生在在不干扰二硫键的任意位置上。

[0041] 可以多种方式制备本发明多肽,例如化学合成(全部或部分),用蛋白酶消化较长多肽,由 RNA 翻译,由细胞培养物纯化(如通过重组表达),由生物体本身制备(如细菌培养后,或直接来自患者)等。产生长度小于 40 个氨基酸的肽的优选方法包括体外化学合成[29, 30]。尤其优选固相肽合成,例如基于 tBoc 或 Fmoc[31] 化学的方法。也可部分或完全利用酶促合成[32]。作为化学合成的替代方式,可利用生物合成,例如,可通过翻译产生多肽。这一过程可以在体外或体内进行。生物学方法通常仅限于产生基于 L-氨基酸的多肽,但可通过操作翻译机器(如氨基酰基 tRNA 分子的翻译机器)引入 D-氨基酸(或其它非天然氨基酸,如碘化酪氨酸或甲基苯丙氨酸、叠氮基高丙氨酸等)[33]。然而,包含 D-氨基酸时,优选使用化学合成。本发明多肽的 C 末端和 / 或 N 末端上可具有共价修饰。

[0042] 本发明的多肽优选以纯或基本纯,即基本不含其它多肽(如不含天然产生的多肽)、特别是不含其它 HIV 或宿主细胞多肽的形式提供,并且本发明的多肽通常为至少约 50% 纯(以重量计),通常至少约 90% 纯,即组合物中少于约 50% 且更优选少于约 10% (如 5% 或更少)是其它表达的多肽。

[0043] 本发明的多肽可连接于固体支持物。本发明的多肽可包含可检测标记(如放射性或荧光标记,或生物素标记)。

[0044] 术语“多肽”指任何长度的氨基酸聚合物且包括其他修饰和变体中的糖蛋白。HIV gp120 和或可溶性 gp140 是糖蛋白且本发明的多肽优选是糖基化的。氨基酸聚合物可以是直链或支链的,其可包含修饰的氨基酸并可被非氨基酸打断。该术语也包括天然修饰或通过介入修饰的氨基酸聚合物;例如,二硫键形成、其他糖基化、部分或完全去糖基化、脂化、乙酰化、磷酸化或任何其它操作或修饰,如与标记组分偶联。该定义还包括,例如,含有一个或多个氨基酸类似物(包括例如,非天然氨基酸等),以及本领域已知的其它修饰的多肽。多肽可以单链或结合链的形式产生。本发明的多肽可以是天然或非天然糖基化的(即该多肽的糖基化模式不同于相应天然产生多肽的糖基化模式)。本发明的多肽可以是分离或纯化的。

[0045] 本发明提供含有序列 -X-Y- 或 -Y-X- 的多肽,其中 : -X- 是上文所定义的氨基酸序列,即本发明的多肽,而 -Y- 不是上文所定义的序列,即本发明提供融合蛋白。在一个特定实施方式中, -Y- 是流感血凝素多肽或辅助本发明多肽功能的任何合适的蛋白或生物分子。

[0046] 本发明还提供了多聚体形式的本发明的多肽。具体而言,本发明提供了由 3 个亚基构成的三聚体,其中各亚基是本发明的多肽、融合蛋白或免疫原性片段。在一个实施方式中,所述三聚体包含至少两个相同的亚基。在一个特定实施方式中,全部三个亚基都是相同的。在一个替代性实施方式中,全部三个亚基是不同的。

[0047] 本发明提供了生产本发明的多肽的方法,该方法包括在诱导多肽表达的条件下培养本发明的宿主细胞的步骤。

[0048] 本发明提供了生产本发明的多肽的方法,其中使用化学方式部分或整体地合成所

述多肽。

[0049] 核酸

[0050] 本发明还提供了编码本发明的多肽的核酸。本发明提供了包含与编码本发明的多肽的核苷酸序列存在序列相同性的核苷酸序列的核酸。这类核酸包括使用替代密码子编码相同氨基酸的那些核酸。

[0051] 本发明还提供了可与这些核酸杂交的核酸。可以在不同“严谨性”的条件下进行杂交反应。增加杂交反应严谨性的条件在本领域众所周知且已发表。相关条件的示例包括（按严谨性提高的顺序）：25℃、37℃、50℃、55℃和 68℃的孵育温度；10x SSC、6x SSC、1x SSC、0.1x SSC 的缓冲液浓度（其中 SSC 是 0.15M NaCl 和 15mM 柠檬酸盐缓冲液）和使用其它缓冲液体系的等同条件；0%、25%、50%和 75%的甲酰胺浓度；5 分钟至 24 小时的孵育时间；1 个、2 个或更多个洗涤步骤；1、2 或 15 分钟的洗涤孵育时间；和 6x SSC、1x SSC、0.1x SSC 或去离子水的洗涤溶液。杂交技术及其优化方法为本领域熟知 [例如参见参考文献 34]。

[0052] 本发明包括含有这些序列的互补序列的核酸（例如，用于反义或探针，或用作引物）。

[0053] 本发明的核酸可采取各种形式（如单链、双链、载体、引物、探针、标记形式等）。本发明的核酸可以是环状或分支的，但通常是线性的。除非另有说明或要求，利用核酸的本发明任何实施方式可利用双链形式和构成该双链形式的两条互补单链形式中的每条链。引物、探针以及反义核酸通常是单链。

[0054] 本发明的核酸优选以纯或基本纯的形式提供，即基本不含其它核酸（如不含天然产生的核酸），特别是不含其它肺炎球菌或宿主细胞核酸，通常其至少约 50% 纯（以重量计），且通常至少约 90% 纯。本发明的核酸优选为肺炎球菌核酸。

[0055] 可以多种方式制备本发明的核酸，例如，完全或部分化学合成（例如 DNA 的亚磷酸胺合成）、利用核酸酶（例如限制性酶）消化较长核酸、连接较短核酸或核苷酸（例如使用连接酶或聚合酶）、由基因组或 cDNA 文库制备等。

[0056] 本发明的核酸可连接于固体支持物（如珠、平板、滤纸、膜、玻片、微阵列支持物、树脂等）。可用，例如放射性或荧光标记、或生物素标记标记本发明的核酸。在准备将核酸用于检测技术时，例如核酸是引物或探针时，这特别有用。

[0057] 术语“核酸”通常包括任何长度的核苷酸聚合形式，其包含脱氧核糖核苷酸、核糖核苷酸和 / 或其类似物。它包括 DNA、RNA、DNA/RNA 杂交体。它也包括 DNA 或 RNA 类似物，如含有修饰主链（如肽核酸（PNA）或硫代磷酸酯）或修饰碱基的类似物。因此，本发明包括 mRNA、tRNA、rRNA、核酶、DNA、cDNA、重组核酸、分枝核酸、质粒、载体、探针、引物等。本发明核酸采取 RNA 形式时，它可能具有或不具有 5' 帽。

[0058] 本发明核酸可以是载体的一部分，即设计用于转导 / 转染一种或多种细胞类型的核酸构建物的一部分。载体可以是，例如，设计用于分离、增殖和复制所插入核苷酸的“克隆载体”，设计用于在宿主细胞中表达核苷酸序列的“表达载体”，设计用于产生重组病毒或病毒样颗粒的“病毒载体”，或具有一种以上载体类型的属性的“穿梭载体”。优选的载体是质粒。“宿主细胞”包括单个细胞或细胞培养物，其可以是或已经是外源核酸的受体。宿主细胞包括单个宿主细胞的后代，由于天然、偶然或有意的突变和 / 或改变，这些后代与原始亲

本细胞（在形态或总 DNA 互补方面）不必完全相同。宿主细胞包括使用本发明的核酸体内或体外转染或感染的细胞。

[0059] 应理解，当核酸是 DNA 时，RNA 序列中的“U”被 DNA 中的“T”所替代。相似地，应理解当核酸是 RNA 时，DNA 序列中的“T”被 RNA 中的“U”所替代。

[0060] 涉及核酸时，术语“互补”或“互补的”指沃森－克里克碱基配对。因此，C 的互补物是 G，G 的互补物是 C，A 的互补物是 T（或 U），T（或 U）的互补物是 A。也能使用如 I（嘌呤肌苷）的碱基，例如与嘧啶（C 或 T）互补。

[0061] 本发明的核酸可用于，例如：在体外或体内生产多肽；作为用于检测生物样品中核酸的杂交探针；生成额外的核酸拷贝；生成核酶或反义寡核苷酸；作为单链 DNA 引物或探针；或作为形成三链体的寡核苷酸。

[0062] 本发明提供了生产本发明的核酸的方法，其中使用化学方式部分或完整地合成所述核酸。

[0063] 本发明提供了包含本发明的核苷酸序列的载体（如克隆或表达载体）和使用这类载体转化的宿主细胞。

[0064] 免疫原性组合物

[0065] 本发明还提供了免疫原性组合物。本发明的免疫原性组合物包含本发明的多肽、其免疫原性片段、本发明的融合蛋白、本发明的三聚体多肽、本发明的多核苷酸和 / 或其组合，其在本文中称作抗原。该免疫原性组合物可用作疫苗。这些疫苗可以是预防性（即预防感染）或治疗性（即治疗感染）疫苗，但一般是预防性疫苗。

[0066] 通常，免疫原性组合物包含其他组分以使其在药学上可接受。除抗原外，它们通常还包含其它组分，例如它们一般包含一种或多种药物载体和 / 或赋形剂。对这类组分的详细讨论参见参考文献 35。

[0067] 免疫原性组合物通常以水性形式给予哺乳动物。然而在给药前，该组合物可以是非水性形式。例如，虽然一些疫苗制备成水性形式然后以水性形式填装、分配和给药，但其他疫苗在制备过程中冻干，并在使用时重建成水性形式。因此，本发明组合物可以是经干燥的，例如冻干制剂。

[0068] 该免疫原性组合物可包含防腐剂，如硫柳汞或 2- 苯氧乙醇。然而，疫苗优选应基本不含（即少于 5 μ g/ml）含汞物质，如不含硫柳汞。更优选无汞的疫苗。特别优选不含防腐剂的疫苗。

[0069] 为了控制张力，优选包含生理盐（如钠盐）。优选氯化钠（NaCl），它的浓度可以是 1 至 20mg/ml，例如约 10 $\pm$ 2mg/ml NaCl。可存在的其它盐包括氯化钾、磷酸二氢钾、无水磷酸氢二钠、氯化镁、氯化钙等。

[0070] 免疫原性组合物的渗透压通常为 200mOsm/kg-400mOsm/kg，优选 240-360mOsm/kg，更优选落入 290-310mOsm/kg 的范围。

[0071] 免疫原性组合物可包含一种或多种缓冲剂。典型的缓冲剂包括：磷酸盐缓冲剂；Tris 缓冲剂；硼酸盐缓冲剂；琥珀酸盐缓冲剂；组氨酸缓冲剂（尤其是含氢氧化铝佐剂时）；或柠檬酸盐缓冲剂。包含的缓冲剂的浓度一般为 5 至 20mM 的范围。

[0072] 组合物的 pH 通常为 5.0 至 8.1，更常为 6.0 至 8.0，例如 6.5 至 7.5，或者 7.0 至 7.8。

[0073] 该组合物优选无菌。该组合物优选无热原,如每剂量含有小于 1EU(内毒素单位,标准量度),优选每剂量小于 0.1EU。该组合物优选不含谷蛋白。

[0074] 所述组合物可含有一次免疫的物质,或者可含有多次免疫的物质(即“多剂量”试剂盒)。多剂量配置优选含有防腐剂。作为多剂量免疫原性组合物中包含防腐剂的替代方案(或补充方案),所述免疫原性组合物可包含在装有利于取出物质的无菌接头的容器中。

[0075] 本发明的免疫原性组合物也可包含一种或多种免疫调节剂。优选地,一种或多种免疫调节剂包括一种或多种佐剂,例如两种、三种、四种或更多种佐剂,例如参考文献 36 和 37 中所述(例如,佐剂包含一种或多种铝盐,或包含亚微米水包油乳液)。

[0076] 可将所述免疫原性组合物制备为液体溶液或悬浮液形式的注射剂。也可制备适合在注射前溶解或悬浮于液体载剂的固体形式(如冻干组合物或喷雾冻干组合物)。所述免疫原性组合物可制备成口服给药制剂,如片剂或胶囊,喷雾剂,或糖浆剂(任选调味)。所述免疫原性组合物可以制成鼻部或眼部给药制剂,例如滴剂。可将所述免疫原性组合物设计为试剂盒的形式,从而在临对患者给予之前重建合并的组合物。此类试剂盒可包含一种或多种液体形式的抗原以及一种或多种冻干抗原。

[0077] 临用前制备组合物(例如,以冻干形式提供组分时)并以试剂盒形式提供组合物时,该试剂盒可包括两个药瓶,或者可包括一个填充好的注射器和一个药瓶,注射器内容物用于在注射前重激活药瓶内容物。

[0078] 用作疫苗的免疫原性组合物包含免疫有效量的抗原,以及需要的任何其它组分。“免疫有效量”是指以单一剂量或一系列剂量的部分给予个体的对治疗或预防有效的量。该量取决于待治疗个体的健康和身体状况、年龄、待治疗个体的分类组(如非人灵长动物、灵长动物等)、个体的免疫系统合成抗体的能力、所需的保护程度、疫苗制剂、治疗医生对医学情况的评估和其它相关因素。预计所述量将落入可通过常规试验可确定的较宽范围内。

[0079] 治疗方法

[0080] 可将本发明的分离的多肽、其免疫原性片段、本发明的融合蛋白、本发明的三聚体多肽、本发明的多核苷酸、本发明的免疫原性组合物和/或其组合用于在对象中产生免疫应答的方法中以及治疗和预防对象中 HIV 感染和/或 AIDS 的方法中。在一个优选实施方式中,所述对象是人。

[0081] 用于在对象中产生免疫应答的方法包括向所述对象给予本发明的分离的多肽、其免疫原性片段、本发明的融合蛋白、本发明的三聚体多肽、本发明的多核苷酸、本发明的免疫原性组合物和/或其组合。

[0082] 本发明的方法和用途所引起的免疫应答通常包括抗体应答,优选保护性抗体应答。用于评估 HIV 疫苗接种后抗体应答、中和能力和保护的方法是本领域熟知的,例如参考文献 38、39、40、41 和 42 中所述。

[0083] 本发明还提供了治疗或预防对象中 HIV 感染和/或 AIDS 的方法,包括向所述对象给予本发明的分离的多肽、其免疫原性片段、本发明的融合蛋白、本发明的三聚体多肽、本发明的多核苷酸、本发明的免疫原性组合物和/或其组合。

[0084] 可以多种方式给予本发明免疫原性组合物。最优选的免疫途径是肌肉内注射(如注射到上肢或下肢),但其它可用的途径包括皮下注射、鼻内[43-45]、皮内[46, 47]、口服[48]、经皮、透皮[49]等。皮内和鼻内途径很吸引人。皮内给药可能涉及微注射装置,例如

使用约 1.5mm 长的针头。

[0085] 可通过单剂量方案或多剂量方案进行治疗。多剂量可用于初次免疫方案和 / 或加强的免疫方案。在多剂量方案中,可通过相同或不同的途径(如肠胃外初次和粘膜加强,粘膜初次和肠胃外加强等)给予多个剂量。给予一个以上剂量(一般是两个剂量)特别可用于免疫未曾免疫接触的患者,例如之前从未接受过 HIV 疫苗的人。一般以至少 1 周(例如约 2 周、约 3 周、约 4 周、约 6 周、约 8 周、约 10 周、约 12 周、约 16 周等)的间隔给予多个剂量。

[0086] 可将本发明的免疫原性组合物与抗逆转录病毒化合物(具体是对 HIV 有活性的抗逆转录病毒化合物)基本同时给予患者(在对健康护理专业人员进行的同一用药咨询或访问期间)。

[0087] 概述

[0088] 术语“包括”涵盖“包含”以及“由……组成”,例如,“包括”X 的组合物可以仅由 X 组成或可包含其它物质,例如 X+Y。

[0089] 术语“基本上”不排除“完全”,如“基本上不含”Y 的组合物可能完全不含 Y。如有需要,“基本上”一词可从本发明的定义中省略。

[0090] 与数值 x 相关的术语“约”是可任选的,并且表示,例如 $x \pm 10\%$ 。

[0091] 优选通过 MPSRCH 程序(牛津分子科技公司(Oxford Molecular))执行的 Smith-Waterman 同源性搜索算法,利用仿射缺口搜索测定多肽序列之间的相同性,其中参数为缺口开放罚分 = 12 且缺口延伸罚分 = 1。

[0092] 附图简要说明

[0093] 图 1A: 来自 HXBc2 毒株的 HIV gp120 的内部结构域的结构及其组织成层 1、2 和 3。

[0094] B: 显示相关层和结构域的来自 SF162 毒株的 HIV-1gp120 的序列(SEQ ID NO:2)。内部结构域为黑色文本且外部结构域为灰色文本。形成层 1 的氨基酸带单线下划线,形成层 2 的氨基酸带双线下划线,且形成层 3 的氨基酸带虚线下划线。

[0095] 图 2 半胱氨酸取代的和天然的 gp120 结构的比较

[0096] 图 3 基于来自 HXBc2 毒株的 gp120 结构的层 1 和 2 二硫键桥连的靶标。

[0097] 图 4 基于来自 HXBc2 毒株的 gp120 结构的层 2 和 3 二硫键桥连的靶标。

[0098] 图 5Hxb2-gp120(结构信息的来源)和 SF162(参比序列)比对中显示的半胱氨酸取代的位点。靶向半胱氨酸取代和二硫键桥连的位点以序列编号上加框的形式表示。

[0099] 图 6 哺乳动物细胞中野生型和半胱氨酸取代的 gp120 的表达的比较。将野生型、单突变和双突变的编码 SF162gp120 的质粒转染至 293T 细胞中,并在培养 48 小时后收获并裂解细胞。在还原或非还原条件下将等量的细胞裂解物在 SDS-PAGE 上运行并使用抗 gp120 抗体(2G12)通过 western 印迹检测裂解物中的 HIV-gp120。

[0100] 图 7A: 典型的考马斯染色 PAGE 凝胶,显示纯化的 SF162gp120 和二硫键稳定化的 SF162L1-SS-L2gp120(SEQ ID NO:4),显示蛋白纯度 >90%。

[0101] B: 典型的考马斯染色 PAGE 凝胶,显示纯化的 SF162gp140 和二硫键稳定化的 SF162L1-SS-L2gp140(V59C 和 S109C),显示蛋白纯度 >90%。

[0102] 图 8 测定 gp120 和 gp120L1-SS-L2 与受体(可溶性 CD4, sCD4)、mAb b12(其结合 CD4 结合位点)和 mAb 17b(以 CD4 诱导的方式在 CD4 结合后结合的抗体)的结合亲和性的

表面等离子共振 (SPR) 分析。显示了对各配体注射的多种浓度 gp120 蛋白的实验数据, 以及 1:1 拟合曲线。

[0103] 图 9A: 抗体结合研究, 第 1 次免疫后 2 周 (2wp1)、第 2 次免疫后 2 周 (2wp2)、第 3 次免疫后 2 周 (2wp3)、第 3 次免疫后 4 周 (4wp3) 和第 3 次免疫后 8 周 (8wp3) 时间点

[0104] B: 亲和力指数, 全部 4 种免疫原

[0105] 图 10T2M-b1 试验中假病毒 (交叉亚型) 的中和。A - gp120 ;B - gp120L1-SS-L2 ; C - gp140 ;D - gp140L1-SS-L2。

[0106] 图 11 由 A - gp120 ;B - gp120L1-SS-L2 ;C - gp140 ;D - gp140L1-SS-L2 产生的抗体的表位特异性。

[0107] 实施本发明的方式

[0108] 用于稳定层间关联的策略

[0109] 半胱氨酸突变和二硫键桥连位点的鉴定

[0110] 为在两个半胱氨酸残基之间形成二硫桥键, 侧链必须相距适当的距离并相对于彼此处于正确的方向。我们检查了位于层 1、2 和 3 的界面上的所有残基并试图鉴定可被靶向用于使用半胱氨酸进行取代以形成二硫键的残基。我们使用了 gp120 的晶体结构 [13] 和以下标准以进行该选择:

[0111] a. 先前解析的 gp120 结构中形成二硫桥键的半胱氨酸的 C β (碳 - β) 原子之间的平均距离是约 4.22Å, 在 3.9Å - 4.7Å 的范围内。因此, 我们对 C β -C β 距离落入该范围内的相互作用层上的残基进行了筛选。

[0112] b. 我们还检查了其中通过使用半胱氨酸残基进行取代导入了人工二硫桥键的 gp120 的晶体结构 [6]。在该结构中, 两对残基 (W90-E268) 和 (I103-Q413) 被取代为半胱氨酸 (W90C-E268C) 和 (I103C-Q413C)。天然的和半胱氨酸取代的残基的 C β 距离分别为 4.1Å 和 4.3Å。此外, 半胱氨酸取代的残基中二硫桥键正确地形成, 且这些二硫桥键连接的 gp120 的结合 CD4 的结构无法与天然 gp120-CD4 结合结构区分 (图 2)。

[0113] 基于这些标准, 我们鉴定到了多对残基, 如果被取代为半胱氨酸, 这些残基对就可在层间形成二硫桥键 (图 3、4 和 5)。

[0114] 半胱氨酸取代和靶向二硫键插入的位点的选择

[0115] 如上文所述, 可使用可用的结构信息通过观察选择位于层的界面且距离相近 (约 4-5 埃) 的残基, 所述残基如果被半胱氨酸取代就会形成二硫键 [13]。我们还使用了软件 Disulfide-by-Design 以进一步了解推测距离足够接近、如果被半胱氨酸取代就会形成二硫键的残基。

[0116] 用于半胱氨酸取代的定点诱变

[0117] 一旦我们确定了用于半胱氨酸取代的残基, 我们根据快速改变 (Quick Change) 定点诱变 (司查塔基公司 (Stratagene)) 设计互补引物。野生型 SF162gp120 编码质粒用作模板, 并在各轮诱变后选择克隆并进行测序以确保存在正确的突变且不存在非预期的突变。进一步扩增正确的克隆用于大规模质粒 DNA 制备。

[0118] 层 1 和层 2 桥连的靶标

[0119] 经选择用于半胱氨酸取代以桥连层 1 和 2 的靶标是来自层 1 的缬氨酸 59 和来自层 2 的丝氨酸 109 (图 3 和 5)。该键桥在本文中称作 L1-SS-L2。新的结合 CD4 的 gp120 结构

中这些残基之间的 C β 原子距离是 3.7Å, 这是用于二硫桥连的理想距离。使用 Stratagen® Quick Change 实验方案通过定点诱变将这两个残基突变为半胱氨酸, 并通过 DNA 测序进行确认 (表 1)。通过小规模转染和 western 印迹确认, 在将这些残基突变为半胱氨酸相关的 gp120 表达或加工过程中不存在不良影响 (图 6)。

[0120] 表 1

[0121]

HXBc2中的 序列编号	SF162中的 序列编号	L1 - L2	L2 - L3	L1 - L2 - L3	(对照1) L1和 L3	(对照2) W90C - V268C, I103C - Q413C 和 W90C - V268C - I103C - Q413C
1 V65C	V59C ✓	V59C + S109C ✓		✓ L1 - L2 + V95C + W465C ✓ +R462C ✓ +L469C	V59 + W465C	
2 V101C	V95C ✓		V95C + W465C ✓ V95C + R462C ✓ V95C + L469C ✓ H99C + R462C ✓	✓ L1 - L2 + V95C + S109C		
3 S115C	S109C ✓	S109C + V59C ✓		✓ L1 - L2 + V95C + W465C		
4 W479C	W465C ✓		W465C + V95C ✓			
5 H105C	H99C ✓					
6 R478C	R462C ✓					
7 L483C	L469C ✓					
8 W96C	W90C ✓					
9 V275C	E268C ✓					
10 I109C	I103C ✓					✓ W90C - V268C ✓ I103C - Q413C
11 Q428C	Q413C ✓					✓ W90C - E268C - I103C

✓ 经测序或正在测序的克隆

[0122] 层 2 和层 3 桥连的靶标

[0123] 基于与上文所述类似的标准, 我们鉴定到了位于可通过二硫键桥连的层 2 和 3 上的多个残基。这些残基和 C β 原子之间的距离如下: V95-W465: 4.5Å、V95-R462: 5.3Å、V95-L469: 5.6Å 和 H99-R462: 5.8Å。残基 V95 和 W465 是半胱氨酸突变和层 2 和 3 之间二硫键桥连的主要靶标 (图 4 和 5)。选择的残基是突变的且确认了存在突变且不存在来自突变的不良影响, 如上文所述 (表 1 和图 6)。

[0124] 组合诱变和多个二硫键的插入

[0125] 为实现进一步的稳定化, 我们继续进行如表 1 所示和下文所述的组合诱变。包括:

[0126] a. 使用 V59C 和 S109C 以及 V95C 和 W465C 之间二硫键的层 1-2 和层 2-3 双重桥连的组合

[0127] b. 层 1-2 桥连与先前公开的、二硫键桥连的 gp120 结构 [6] (W90C 和 E268C ; I103C 和 Q413C) 的组合, 使用 V59C 和 S109C 之间的二硫键。

[0128] c. 层 2-3 桥连与先前公开的、二硫键桥连的 gp120 结构 [6] (W90C 和 E268C ; I103C 和 Q413C) 的组合, 其联合 V95C 和 W465C、W90C 和 E268C 或 I103C 和 Q413C 处的二硫键。

[0129] 蛋白表达与纯化

[0130] 重组 HIV-1 包膜 (Env) 糖蛋白 gp120 和 gp140 源自 B 亚型 CCR5 热带株 HIV-1SF162 并通过 HEK293T 细胞的转染生产。环 1 和 2 二硫键稳定化的 gp120 (gp120L1-SS-L2) 和 gp140 (gp140L1-SS-L2) 也源自 SF162 并在 HEK293T 细胞中生产。使用三步纯化方法纯化全

部四种糖蛋白,所述三步纯化方法涉及雪花莲-琼脂糖(GNA)亲和柱、阳离子交换 DEAE 柱和最后的陶瓷羟磷灰石(CHAP)柱,如 Srivastava 等所述 [52]。随后通过 SDS-PAGE(用于分析纯化水平)和(对抗 SF162gp140 多克隆兔血清)具有特异反应性的免疫印迹分析纯化的糖蛋白。纯化的糖蛋白是同质的(>95%的 gp120 单体;>80%的 gp140 三聚体),纯度>98%。使用 Endosafe®筒和 Endosafe®-PTSTM 分光光度计(查尔斯河实验室国际公司(Charles River Laboratories International, Inc.)),马萨诸塞州威尔明顿)测量糖蛋白中的内毒素水平,并发现该水平是 ≤ 0.05 EU/免疫剂量。

[0131] 还生产并纯化了用于表位作图目的的其他 SF162gp120 变体(如 gp120 Δ V3、gp120 Δ V1V2、gp120D368R(CD4 结合位点突变体)、gp120I420R(CD4i 位点突变体)),如上文和先前所述 [50]。从蛋白的考马斯染色 SDS PAGE 和 Western 印迹估计蛋白的纯度。

[0132] 图 7A 显示了纯化的 SF162gp120(泳道 1)和二硫键稳定化的 SF162L1-SS-L2gp120(SEQ ID NO:4)(泳道 2)的考马斯染色的 SDS-PAGE,显示蛋白纯度>90%。箭头显示两种 120kDa 蛋白。图 7B 显示了纯化的 SF162gp140(泳道 3)和二硫键稳定化的 SF162L1-SS-L2gp140(V59C 和 S109C)(泳道 4)的考马斯染色的 SDS-PAGE,显示蛋白纯度>90%(B 图)。箭头显示两种 140kDa 蛋白。

[0133] MW 指分子量标记物;显示了标记物中各条带的分子量(以 kDa 单位)。

[0134] 通过表面等离子共振(SPR)试验对蛋白的表征

[0135] 通过 SPR 测试了来自各突变体的纯化的蛋白的结构编号和构象固定,所述 SPR 测量与特定配体结合的动力学。为测量野生型 gp120 与二硫键稳定化的 gp120(gp120L1-SS-L2)的结合亲和性,我们使用了基于 SPR 的 BIAcore 3000。通过胺偶联将 200RU 的 sCD4 或 mAb(b12 或 17b)直接固定在 CM5 传感器芯片上。随后以 80 μ l/分钟注射不同浓度的野生型或二硫键稳定化的 gp120。结合分析在 25°C 下进行,使用 HBS-EP 缓冲液作为运行缓冲液。随后将实验曲线拟合至 1:1 朗格缪尔结合模型 BIAevaluation 软件 3.2(BIAcore 公司)。1:1 拟合后得出的结合速率 K_{on} 、解离速率 K_{off} 和解离常数 K_d 示于表 2。

[0136] 结果

[0137] 为鉴定通过 gp120 的结构构象上的二硫键固定层间移动的效果,通过表面等离子共振(SPR)试验纯化并表征来自野生型(未变化)和层 1-2 二硫键交联突变体(V59C 和 S109C;L1-SS-L2)的蛋白。在该试验中,与 gp120 的不同构象结合的配体被捕获至 CM5 芯片上(参见方法),并比较其与野生型或突变的可溶性 gp120 的结合。与野生型相比,突变体 gp120 与 17b 抗体(一种选择性结合 gp120 的结合 CD4 的构象的 CD4 诱导型抗体)的结合显著升高。由二硫键对层 1 和 2 之间相互作用的稳定化导致与 17b 结合的亲和性升高了超过 5 倍。有趣的是,该亲和性方面的升高几乎全部来源于结合速率的升高(参见表 2),其确认了结合 CD4 的构象中 gp120 的固定。然而,单独由层 1 和 2 之间的二硫键稳定化的 gp120 中仍存在大量灵活性,这在与可溶性 CD4 结合后尤为明显。

[0138] 表面等离子共振(SPR)分析也用于测量 gp120 和 gp120L1-SS-L2 与受体(可溶性 CD4;sCD4)、mAb b12(其结合 CD4 结合位点)以及 mAb 17b 之间的结合亲和性(图 8)。直线表示对各配体注射的多种浓度 gp120 蛋白的实验数据;曲线表示 1:1 拟合曲线。两条曲线的几乎完全重叠验证了与 1:1 结合模型拟合(Chi²,表 2)的良好性。

[0139] 表 2

[0140]

配体 - 分析物	K_{on} (1/Ms)	K_{off} (1/s)	K_d (M)
Gp120 - sCD4	$3.21e^4$	$7.84e^{-4}$	$2.45e^{-8}$
Gp120L1-SS-L2 - sCD4	$8.02e^3$	$7.46e^{-4}$	$9.31e^{-8}$
Gp120 - b12	$1.65e^5$	$7.12e^{-3}$	$4.48e^{-8}$
Gp120L1-SS-L2 - b12	$2.91e^4$	$8.85e^{-3}$	$3.04e^{-7}$
Gp120(+sCD4) - 17b	$2.04e^5$	$5.35e^{-5}$	$2.62e^{-10}$
Gp120L1-SS-L2 - 17b	$1.05e^6$	$4.64e^{-4}$	$4.42e^{-10}$
Gp120L1-SS-L2(+sCD4) - 17b	$3.46e^5$	$3.33e^{-5}$	$9.62e^{-11}$

[0141] (wt/ 未修饰的) gp120 和二硫键稳定化的 gp120L1-SS-L2 的结合的比较动力学分析如下所示：

[0142] 1. 与 CD4 的结合：gp120 以 $24.5nM(2.48e^{-8}M)$ 的 K_d (解离常数) 结合可溶性 CD4(sCD4)，而稳定化的 gp120(gp120L1-SS-L2) 以低 4 倍的亲和性结合 ($K_d = 93.1nM/9.31e^{-8}M$)。该 4 倍差异完全是由于二硫键稳定化的 gp120 与 sCD4 之间较低的结合速率 (K_{on})，该低结合速率是由将蛋白“锁定”在结合 CD4 的构象而导致的。[比较第 2 和第 3 排]。

[0143] 2. 与 b12 的结合：在 gp120 和 gp120L1-SS-L2 之间的 K_d 方面，我们观察到类似的趋势，但具有更大的 ≤ 7 倍的倍数差异。[比较第 4 和第 5 排]。

[0144] 3. 与 17b(+/-sCD4) 的结合：我们知道，在 sCD4 不存在的情况下，gp120 不能有效地结合 mAb 17b。然而，在 sCD4 存在的情况下 [第 6 排]，gp120 以 $0.26nM(2.62e^{-10}M)$ 的亚纳摩尔级亲和性结合。重要和关键的是，在 sCD4 不存在的情况下，二硫键稳定化的 gp120(gp120L1-SS-L2) 结合 mAb 17b 的亲和性与结合 CD4 的状态 (gp120+CD4) 的 gp120 ($0.44nM(4.42e^{-10}M)$) 类似 [第 7 排]。将 sCD4 添加至稳定化的 gp120 后，亲和性提高了 20 倍并显示 $96pM$ 的解离常数 [$9.62e^{-11}M$, 第 7 排]。

[0145] 为实现结合 CD4 的构象中的 gp120 的完全稳定化，正开发层 2 和 3 之间进一步稳定化的突变。

[0146] 候选免疫原的分析

[0147] 一旦生产了各进一步稳定化的多肽，则评估其蛋白表达、正确折叠和配体结合的水平。将这些与先前稳定化的结构和为增强保守结合位点暴露而研发的其他经修饰的 gp120 比较。实现 gp120 上保守结合位点稳定暴露的那些蛋白将在小规模动物试验中进一步作为候选免疫原进行分析。

[0148] 佐剂 -Carbopol971P NF+MF59[51]

[0149] Carbopol® 971P NF (在该研究中称作 Carbopol 971P) 购自卢比作公司

(Lubrizol) (粉末状) 并随后在无菌调节下重悬于水中以生成 0.5% 均匀、低粘度的悬浮液。将该悬浮液储存于 4℃ 待用。制备 gp140 蛋白与 0.5% (w/v) Carbopol971P (pH $\geq$ 3.0) 的 1:1 (v/v) 混合物用于所有体外评估。对于动物中给药, 首先制备 gp140 与 0.5% (w/v) Carbopol971P 的 1:1 (v/v) 混合物, 并在添加等体积的 MF59 前就将 gp140 蛋白 -Carbopol971P 复合体孵育 30 分钟, 从而使给予的 Carbopol971P 悬浮液的终浓度保持在 0.125% (w/v)。对于所有用于体内研究的 Carbopol971P 悬浮液, 使用 Endosafe® 筒和 Endosafe®-PTSTM 分光光度计 (查尔斯河实验室国际公司 (Charles River Laboratories International, Inc.)), 马萨诸塞州威尔明顿) 测量内毒素水平。

[0150] 兔免疫

[0151] 免疫研究在加利福尼亚州纳帕的乔斯曼有限责任公司 (Josman LLC) 进行, 这是一家通过 USDA 批准 (第 93-R-0260 号) 且具有来自 NIH 的公共卫生服务 (PHS) 保险 (第 A3404-01 号) 的研究机构。该研究中使用了 3 组新西兰白兔 (每组 5 只年轻成年雌性)。使用添加了单独的 MF59、单独的 Carbopol971P 或者 Carbopol971P 加 MF59 佐剂的 SF162gp140 蛋白对兔进行免疫。在第 0、4、12 和 24 周时, 在臀中肌肉内给予 4 次免疫 (每次免疫 2 个位点)。各免疫的蛋白剂量是 50 μ g。由第 1 次免疫 (预采血) 之前和每次免疫后的各个时间点 (2wp2、2wp3、2wp4、4wp4 和 15wp4 放血) 上收集的血液制备血清样品并且分析结合和中和。

[0152] 对于 0.125% (w/v) Carbopol971P 的注射相关的局部反应原性的评估, 在给予剂量前、紧接在免疫后和免疫后 24 小时和 48 小时肉眼观察皮肤上注射位点处是否有水肿和红斑。紧接在免疫后和免疫后 24 小时和 48 小时进行任何明显临床迹象的一般观察。还在研究开始前、各免疫前和各免疫后 24 小时和 48 小时记录体重。

[0153] 根据动物福利法、用于实验动物护理和使用的 ILAR 指南以及所有适用的地方、州和联邦法律和法规中所列的动物人道护理和使用的要求, 该研究得到了诺华公司的动物护理和使用委员会的全面批准 (批准号 09NVD 044.3.3.09)。

[0154] 结果

[0155] 通过使用 Carbopol 和作为佐剂的 MF59 进行的仅使用蛋白 (25 μ g) 的免疫, 我们进行了兔研究。野生型 gp120 和 gp140 分别与 gp120L1-SS-L2 (SEQ ID NO:4) 和 gp140L1-SS-L2 (V59C 和 S109C) 比较。免疫按下表 3 所示进行。

[0156] 表 3

[0157]

组 (n=5)	免疫原	剂量/免疫	周	佐剂
1	gp120	25 μ g	0, 4, 12	Carbopol971P + MF59

[0158]

2	gp120 L1-SS-L2	25 μ g	0, 4, 12	Carbopol971P + MF59
3	gp140	25 μ g	0, 4, 12	Carbopol971P + MF59
4	gp140 L1-SS-L2	25 μ g	0, 4, 12	Carbopol971P + MF59

[0159] 在第 0、4 周和 12 周的时间点使用所示免疫原和佐剂对兔进行免疫。

[0160] 免疫后,我们在第 1 次免疫后 2 周 (2wp1)、第 2 次免疫后 2 周 (2wp2)、第 3 次免疫后 2 周 (2wp3)、4wp3 和 8wp3 时间点评估了结合的 Ab。我们观察到,全部 4 种免疫原都引发了 $>10e+5-6$ GMT 的 2wp2/2wp3 免疫。结果见图 9A。

[0161] 我们还观察到,全部 4 种免疫原都在 5 个时间点处产生类似水平的亲合力,例外是 gp140L1-SS-L2,其在 2wp3(最终免疫后 2 周)时生成亲合力显著较高的抗体。结果见图 9B。

[0162] HIV-1 病毒中和

[0163] 在 TZM-bl 试验中,我们还分析了用于假病毒(交叉亚型)中和的 2wp3 血清。在 TZM-bl 细胞中使用标准化良好的使用假病毒的试验和荧光素酶报告基因试验测量病毒中和效价 [John C. Kappes 博士、Xiaoyun Wu 博士和北卡罗来纳州达勒姆的詹西木公司 (Tranzyme, Inc.)],如前人所述 [46, 47]。简言之,向稀释的血清样品中加入总共 200TCID₅₀ 假病毒 / 孔并且在 37°C 下孵育 1 小时。孵育后,在含有 DEAE- 葡聚糖的基质中加入 10,000 个细胞 / 孔并且在 37°C 下孵育 48 小时。DEAE- 葡聚糖的终浓度是 10 μ g/ml。48 小时孵育后,将 100 μ l 的细胞转移至 96 孔黑色固体板 (克斯塔公司 (Costar)),用于如供应商 (普洛麦格公司 (Promega)) 所述使用 Bright Glo 底物溶液测量发光。中和效价指扣除背景相对发光单位 (RLU) 后 RLU 比病毒对照孔减少 50% 的稀释度。通过用表达质粒共转染 293T 细胞来制备 HIV-1 包膜假病毒,所述表达质粒含有来自一组 HIV-1 分离物的全长分子克隆的 gp160 包膜基因,所述包膜基因通过使用 FuGENE-6HD (罗氏应用科学 (Roche Applied Sciences),印第安纳州印第安纳波利斯) 与包膜缺陷型 HIV-1 主链载体 (pSG3 Δ env) 连接,如前人所报道 [Montefiori, D. C., Measuring HIV neutralization in a luciferase reporter gene assay (在荧光素酶报告基因试验中测量 HIV 中和), HIV Protocols (《HIV 实验方案》),第二版, G. V. K. Vinayaka R. Prasad 编,卷 485, 2009, 休曼出版社 (Humana Press), 第 395-405 页]。48 小时后,通过 0.45 μ m 过滤器过滤含有假病毒的细胞培养物上清液并在 -80°C 下储存备用。

[0164] 我们观察到,全部 4 种免疫原都引发了针对 Tier 1a 病毒 [SHIV-Bal-P4 (嵌合、亚型 B)、MN. 3 (亚型 B)、SF162. LS (亚型 B) 和 MW965. 26 (亚型 C)] 的包膜特异性中和抗体。重要的是,二硫键稳定化的 gp120 和 gp140 都产生高 >0.5 log 的同源中和效价 (即针对 SF162 病毒)。结果见图 10。

[0165] 表位特异性

[0166] 我们使用了 2wp3 血清来评估使用 4 种免疫原免疫时生成的抗体的表位特异性。我们研究的 4 种表位是 V3 环、V1V2 环、CD4 结合位点 (CD4BS) 和 CD4 诱导型 (CD4i ;结合 CD4 的) 位点。通过使用 SF162gp140 蛋白的标准 ELISA 试验对来自动物的血清中包膜特异性总抗体效价进行定量,所述动物使用添加 MF59 佐剂的 gp140 蛋白、添加 Carbopol971P 佐剂的 gp140 蛋白或者添加 Carbopol971P 加 MF59 佐剂的 gp140 蛋白免疫,如前人所述 [52]。使用他处所述的硫氰酸铵 (NH_4SCN) 置换 ELISA 进行抗体亲和指数的测定 [52]。

[0167] 分析后,我们观察到 4 种免疫原所生成的抗体质量方面的显著差异。gp120 引发了高水平的针对 V3 (约 50–55%) 和 V1V2 表位的抗体,而针对更保守的 CD4BS 和 CD4i 位点的抗体则较低 (10–20%) (图 11A)。然而,二硫键稳定化的 gp120 生成了显著较低水平的针对 V3 (小于 40%) 和 V1V2 (15–20%) 区域的抗体和相对较高水平的针对 CD4BS (大于等于 30%;在 gp120 免疫的臂中大于约 20%) 和 CD4 位点的抗体 (大于 30%;在 gp120 免疫的臂中大于 10%) (图 11B)。

[0168] 使用 gp140 的免疫引发了小于 40% 的 V3 活性抗体和小于等于 20% 的 V1V2 活性抗体。针对保守的 CD4BS 的抗体大于等于 20% 且针对 CD4i 位点的抗体为约 10% (图 11C)。然而,使用二硫键稳定化的 gp140 免疫时,2wp3 显示出类似的抗 V3 抗体水平 (小于 40%) 但显著较高的针对 V1V2 (30–35%)、CD4BS (30–35%) 和 CD4i 位点 (小于等于 40%) 的抗体水平 (图 11D)。

[0169] 总之,二硫键稳定化的 gp120 和 gp140 生成显著较高的针对 CD4i 位点的抗体,且二硫键稳定化的 gp140 引发针对所有物质的所有表位的最“平衡”的应答。

[0170] 初免 – 加强方案或更强力的佐剂的使用可有助于提高应答的宽度或总体质量。

[0171] 应理解,仅以举例的方式描述了本发明,在本发明的范围和构思内可对之进行修改。

[0172] 参考文献

[0173] [1]UNAIDS. 2008. UNAIDS Report on the global HIV/AIDS epidemic-2008 (UNAIDS 关于全球 HIV/AIDS 流行的报告 -2008). Report. UNAIDS.

[0174] [2]Johannes F 等 (2009)Nature 458:636–640.

[0175] [3]Kwong D 等 (2002)Nature 420:678–682.

[0176] [4]Kwong D 等 (2000)Structure 8:1329 – 1339.

[0177] [5]Dey B 等 (2009)PLoS Pathogens 5:e1000445

[0178] [6]Zhou T 等 (2007)Nature. 445(7129):732–7.

[0179] [7]Chen L 等 (2009)Science. 2009326(5956):1123–7.

[0180] [8]Kwon 等 (2012)PNAS 109(15)5663–5668 ;2012 年 3 月 26 日印刷前提前发表, doi:10.1073/pnas.1112391109

[0181] [9]Kwong D 等 (1998)Nature. 393, 648–659.

[0182] [10]Chih-chin Huang 等 (2005). Science 310:1025–1028.

[0183] [11]Kassa A 等 (2009). J Virol. 83:8364–78.

[0184] [12]Finzi A 等 (2010). Mol. Cell. 37:656–67.

[0185] [13]Pancera M 等 (2010). Proc Natl Acad Sci 107:1166–71

[0186] [14]Chen B 等 (2005). Nature 433:834 – 841.

- [0187] [15]Kwong P (2005)Nature. 433(7028):834-41.
- [0188] [16]Kassa A 等 (2009). J. Virol. 83:4476-4488.
- [0189] [17]Thali, M 等 (1993)J. Virol. 67:3978-3988
- [0190] [18]Choe, H 等 (2003)Cell 114:161-170.
- [0191] [19]Burton DR 等 (1994)Science 266(5187):1024-7.
- [0192] [20]Posner MR 等 (1991)J Immunol. 146(12):4325-32
- [0193] [21]Huang CC 等 (2007)Science 317(5846):1930-4
- [0194] [22]Moore 等 (1994)J. Virol., 68:469-484
- [0195] [23]Calarese DA 等 (2003)Science. 300(5628):2065-71
- [0196] [24]Karle S 等 (2004). AIDS 18(2):329-331.
- [0197] [25]McKeating J 等 (1992)AIDS Res. Hum. Retroviruses. 8:451-459
- [0198] [26] 美国专利 6, 241, 986
- [0199] [27]Xiang S 等 (2002)AIDS Res. Hum. Retrovir. 18:1207-1217
- [0200] [28]Huang 等 (2004)PNAS;101:2706-2711
- [0201] [29]Bodanszky(1993)Principles of Peptide Synthesis(ISBN:0387564314).
- [0202] [30]Fields 等 (1997)Meth Enzymol 289:Solid-Phase Peptide Synthesis. ISBN:0121821900.
- [0203] [31]Chan 和 White(2000)Fmoc Solid Phase Peptide Synthesis. ISBN:0199637245.
- [0204] [32]Kullmann(1987)Enzymatic Peptide Synthesis. ISBN:0849368413.
- [0205] [33]Ibba(1996)Biotechnol Genet Eng Rev 13:197-216
- [0206] [34] 美国专利 5, 707, 829
- [0207] [35]Remington:The Science and Practice of Pharmacy(《雷明顿:药物科学与实践》)(Gennaro, 2000 ;第 20 版, ISBN:0683306472)
- [0208] [36]Vaccine Design:The Subunit and Adjuvant Approach(《疫苗设计:亚基和佐剂方法》)(Powell 和 Newman 编) 普莱南出版社 (Plenum Press)1995 (ISBN 0-306-44867-X).
- [0210] [37]Vaccine Adjuvants:Preparation Methods and Research Protocols(《疫苗佐剂:制备方法和研究方案》)(分子药物方法丛书的第 42 卷). ISBN:1-59259-083-7. O' Hagan 编.
- [0211] [38]Wei, X 等 (2003)Nature 422:307-312
- [0212] [39]Montefiori, C(2004)Current Protocols in Immunology - 约翰威利公司 (John Wiley&Sons)
- [0213] [40]Barnett 等 (2010)J. Virol. 第 5975-5985 页, 卷 84, 第 12 期.
- [0214] [41]Burke B 等 (2009)Virology 387(1):147-56
- [0215] [42]Li 等 (2005)J. Virol. 79:10108-10125
- [0216] [43]Greenbaum 等 (2004)Vaccine 22:2566-77.
- [0217] [44]Zurbriggen 等 (2003)Expert Rev Vaccines 2:295-304.
- [0218] [45]Piascik(2003)J Am Pharm Assoc(Wash DC). 43:728-30.

- [0219] [46]Halperin 等 (1979)Am J Public Health 69:1247-50.
- [0220] [47]Herbert 等 (1979)J Infect Dis 140:234-8.
- [0221] [48]Mann 等 (2004)Vaccine 22:2425-9.
- [0222] [49]Chen 等 (2003)Vaccine 21:2830-6.
- [0223] [50]AK Dey 等, PLoS One, 2012. 7(1):p. e30233
- [0224] [51]AK Dey 等, Vaccine. 2012 年 4 月 5 日 ;30(17):2749-59. Epub 2012 年 2 月 22 日)
- [0225] [52]Srivastava, I. K. 等, J Virol, 2002. 76(6): 第 2835-47 页

[0001]

<110> 诺华股份有限公司 (NOVARTIS AG)

<120> 稳定化的GP120

<130> PAT054316-WO-PCT

<140> PCT/EP2013/062553

<141> 2013年6月17日

<150> 61/661, 050

<151> 2012年6月18日

<160> 23

<170> SeqWin99, version 1.02

<210> 1

<211> 1494

<212> DNA

<213> HIV-1

<400> 1

```

atggatgcaa tgaagagagg gcctctgtgt gtgtctgtgc tgtgtggagc agtcttcgtt 60
tcgccacgcg ccgtggagaa gctgtgggtg accgtgtact acggcgtgcc cgtgtggaag 120
gaggccacca ccacctgttt ctgtgccagc gacgccaaag cctacgacac cgaggtgcac 180
aacgtgtggg ccaccacgcg ctgtgtgcc accgacccca acceccagga gatcgtgtctg 240
gagaacgtga ccgagaactt caacatgttg aagacaaca tgggtggagca gatgcacgag 300
gacatcatca gctgtggga ecagagcctg aagccctgcg tgaagctgac cccctgtgtc 360
ttagccctgc actgcacca cclgaagaa gccaccaaca ccaagagcag caactggaa 420
gagatggacc gggcgagat caagaactgc agcttcaagg tgaaccaag catccgcaac 480
aagatgcaga aggagtacgc cctgtttctac aagctggacg tgggtcccat cgacaacgac 540
aacaccagct acaagctgat caactgcaac accagcgtga tcaccaggc ctgcccacag 600
gtgagctteg agcccatccc cactcactac tgcgcccccg ccggttctgc cactctgaag 660
tgcaacgcga agaagtlcaa cggcagcgcc ccttgcacca acgtgagcac cgtgcagtc 720
accacggcga tccgcccggt ggtgagcacc cagctgtgtc tgaacggcag cctggccgag 780
gaggggcgtg tgatccgcag cgagaacttc accgacaacg ccaagaccat catcgtgcag 840
ctgaaggaga gcttgagat caactgcacc cgcaccaaca acaacacccg caagagcatc 900
accatcgccc ccggcccgcc ctctacgccc accggcgaca tcttcggcga catccgcag 960
gcccactgca acatcagcgg cgagaagtgg aacaacaccc tgaagcagat cgtgaccaag 1020
ctgcaggccc agttcggcaa caagaccatc gtgttcaagc agagcagcgg cggcgacccc 1080
gagatcgtga tgcacagctt caactgcggc ggcgagttct tctactgcaa cagcaccag 1140
ctgttcaaca gcacctggaa caaacccatc ggccccaaca acaccaacgg caccatcac 1200
ctgccctgcc gcatcaagca gatcataaac cgttggcagg aggtgggcaa ggccatgtac 1260
gcccccccca tccggcgcca gatccgtgc agcagcaaea tcaccgccc gctgtgtgac 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgcccg cgcgggcgac 1380
atgcgcgaca actggcgag cgagctgtac aagtacaagg tggtagaat cgagcccctg 1440
gacgtggccc ccaccaaggc caagcgccgc gtagtgacg gggagaagcg ctac 1494

```

<210> 2

<211> 497

<212> PRT

<213> HIV-1

<220>

<221> 结构域

<222> 1-248, 458-497

<223> 内部结构域

<400> 2

```

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1           5           10          15

```

```

Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
20          25          30

```

```

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

```

```

Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
50          55          60

```

```

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

```

```

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

```

[0002]

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

[0003]

Arg

<210> 3
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V59C和S109C

<400> 3
 atggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcgtt 60
 tgcgccagcg ccgtggagaa gctgtgggtg accgtgtact acggcgigcc cgtgtggag 120
 gagggcacca ccacctgttt ctgcgcagc gacgccaagg cctacgacac cgagtggcac 180
 aacgtgtggg ccaccacgcg ctgcgtgccc accgaccca acccccagga gatcgtgctg 240
 gagaacgtga ccgagaactt caacatglgg aagaacaaca tggltggagca galgcacgag 300
 gacateatca gcctgtggga ccagtgcctg aagccctgcg tgaagctgac cccctgtgc 360
 gtgaccctgc actgcacca cctgaagaac gccaccaaca ccaagagcag caactggaag 420
 gagatggacc gggcgagat caagaactgc agcttcaagg tgaccaccag catccgcaac 480
 aagatgcaga aggagtacgc cctgttctac aagctggagc tgggtgccc atgacaacgac 540
 aacaccagct acaagctgat caactgcaac accagcgtga tcaccagagc ctgccccag 600
 gtgagcttgc agcccatccc catccactac tgcgcgcccg ccggttgcg catcctgaag 660
 tgcacgcgaa agaagttcaa cggcagtgcc cctgcacca acgtgagcag cgtgcagtcg 720
 acccagcgca tccgcccggt ggtgagcacc cagctgctgc tgaacggcag cctggccgag 780
 gagggcggtg tgatecgcag cgagaacttc accgacaacg ccaagaccat catcgtgcag 840
 ctgaaggaga gcgtggagat caactgcacc cgcgcccaaca acaacacccg caagagcate 900
 accatcgccc ccggccgcgc cttctacgcc accggcgaca tcatcgccga catccgccag 960
 gcccaactga acatcagcgg cgagaagtgg aacaacaacc tgaagcagat cgtgaccaag 1020
 ctgcaggccc agttcggcaa caagaccatc gtgttcaagc agagcagcgg cggcgacccc 1080
 gagatcgtga tgcacagctt caactgcgga ggcgagttct tctactgcaa cagcaaccag 1140
 ctgttcaaca gcacctggaa caacaccate ggccccaaca acaccaacgg caccatcacc 1200
 ctgccttgcg gcatacagca galcalcaac cgtlggcagg agglgggcaa ggccatgtac 1260
 gcccccccga tccggcgcca gatccgtgc agcagcaaca tcacggcct gctgctgacc 1320
 cgcgacggcg gcaagsagat cagcaacacc accgagatct tccgccccgg cggcgcgac 1380
 atgcgcgaca actggcgag cgagctgtac aagtacaagg tggtagaat cgagccctg 1440
 ggctgtgccc ccaccaagcg caagcgccgc gtggtgcagc gcgagaagcg ctaa 1494

<210> 4
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V59C 和 S109C

<400> 4
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
 Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
 20 25 30
 Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
 35 40 45
 Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Cys His Asn Val Trp Ala
 50 55 60
 Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
 65 70 75 80
 Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu
 85 90 95
 Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Cys Leu Lys Pro
 100 105 110
 Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Leu
 115 120 125
 Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
 130 135 140
 Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr Thr Ser Ile Arg Asn
 145 150 155 160

[0004]

Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
 165 170 175
 Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
 180 185 190
 Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
 195 200 205
 His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
 210 215 220
 Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys
 225 230 235 240
 Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
 245 250 255
 Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser Glu Asn Phe Thr Asp
 260 265 270
 Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu Ser Val Glu Ile Asn
 275 280 285
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Thr Ile Gly Pro
 290 295 300
 Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
 305 310 315 320
 Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn Asn Thr Leu Lys Gln
 325 330 335
 Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn Lys Thr Ile Val Phe
 340 345 350
 Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn
 355 360 365
 Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser
 370 375 380
 Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr Asn Gly Thr Ile Thr
 385 390 395 400
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg Trp Gln Glu Val Gly
 405 410 415
 Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser
 420 425 430
 Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Lys Glu Ile Ser
 435 440 445
 Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn
 450 455 460
 Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Gln Pro Leu
 465 470 475 480
 Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys
 485 490 495

Arg

<210> 5
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C 和 W465C

<400> 5

[0005]

```

atggatgcaa tgaagagagg gctetgctgt gtgetgctgc tgtgtggagc agtettctgtt 60
tcgcccagcg ccglggagaa gctgtgggag accgigtact accggcgtgcc cgtgtlgaag 120
gaggccacca ccacctgttt ctgcgccagc gacggcaagg cctacgacac cgaggtgcac 180
aacgtgtgtg ccaccacgc ctgcgtgccc accgaaccca acccccagga gatcgtgctg 240
gagaacgtga ccgagaactt caacatgtgg aagaacacaa tgltygagca gatgcacgag 300
gacatcatca gctgttggga ccagagcctg aagccctgcg tgaagctgac cccctgttgc 360
gtgaccctgc actgcaccaa cctgaagaac gccaccaaca ccaagagcag caactggaag 420
gagatggacc gcggcgagat caagaactgc agcttcaagg tgaccaccag catccgcaac 480
aagatgcaga agaggtacgc cctgtttctac aagctggacg tgggtgccat cgacaacgac 540
aacaccagct acaagctgat caactgcac accagcgiga tcaccaggc ctgcccacag 600
gtgagcttcg agcccatccc catccaetac tgcgcccccg cgggttgcg catcctgaag 660
tgcaacgaca agaaglteaa eggcagcggc cctlgcacea acgtgagcag cglgcagtgc 720
acccaaggca tccgcccggt ggtgagcacc cagctgctgc tgaacggcag cctggccgag 780
gaggcggtgg tgatccgcag cgagaacttc accgacaacg ccaagaccat catcgtgcag 840
ctgaaggaga gctggagat caactgcacc cgcaccaaca acaacacccg caagagcatt 900
accatcgccc ccggccgggc cttctacgcc accggcgaca tcctcggcga catccgccag 960
gcccactgca acatcagcgg cgagaagtgg aacaacaccc tgaagcagat cgtgaccaag 1020
ctgcaggccc agttcggcaa caagaccatc gtgttcaagg agagcagcgg cggcgacccc 1080
gagatcgtga tgcacagctt caactgcagg gccgagttct tctactgcaa cagcaccacg 1140
ctgttcaaca gcacctggaa caacaccatc ggccccaaca acaccaacgg caccatcacc 1200
ctgcctctgc gcatcaagca gatcatcaac cgtctggcagg aggtgggcaa ggcatgtac 1260
gcccccccca tcccgcgcca gatccgtgc agcagcaaca tcaccggcct gctgcigacc 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgccccgg cggcgcgagc 1380
atgcgcgaca actgycgcag cgagctgtac aagtacaagg tggatgaagat cgagccctctg 1440
ggcgtggccc ccaccaaggc caagcgccgc gtgggtgcagc gcgagaagcg cttaa 1494

```

<210> 6
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C 和 W465C

<400> 6
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
 Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
 20 25 30
 Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
 35 40 45
 Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
 50 55 60
 Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
 65 70 75 80
 Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Cys Glu
 85 90 95
 Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
 100 105 110
 Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Leu
 115 120 125
 Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
 130 135 140
 Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr Thr Ser Ile Arg Asn
 145 150 155 160
 Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
 165 170 175
 Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
 180 185 190
 Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
 195 200 205
 His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
 210 215 220

[0006]

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

Arg

<210> 7
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 Y95C 和 R462C

<400> 7
 atggatgcaa tgaagagagg gctctgctgt gtgetgetge tgtgtggagc agtcttcgtt 60
 icgcccagcg ccgtggagaa gcgtgggtg accgtgtact acggcgtgcc cgtgtggaag 120
 gaggccacca ccacctgttt ctgcgccagc gacgccaagg cctacgacac cgaggtgcac 180
 aacgtgtggg ccaccacgc ctgcgtgcc accgaacca acccccagga gatcgtgctg 240
 gagaacgtga ccgagaactt caacatgtgg aagaacaaca tgtgygagca gatgcacgag 300
 gacatcatca gcctgtggga ccagagcctg aagccctgeg tgaagctgac cccctgtgce 360
 gtgacctgc actgcacca cctgaagaac gccaccnaca ccaagagcag caactggaag 420
 gagatggacc gcggcgagat caagaactgc agcttcaagg tgaccaccag catccgcaac 480
 aagatcaga aggagtacgc cctgttttac aagctggacg tggigcccat cgacaacgac 540
 aacaccagct acaagctgat caactgcaac accagcgtga tcaccagge ctgcccgaag 600
 gtgagcttgc agcccatccc catccactac tgcgcccccg ceggttccg catcctgaag 660
 tgaacagaca agangttaa cggcagcgge cctgcacca acgtgagca cgtgcagtgc 720

[0007]

```

accacaggca tccgcccggt ggtgagcacc cagctgctgc tgaacggcag cctggccgag 780
gagggcggtg tgatccgcag cgagaaactc accgacaacg ccaagaccat catcgtgcag 840
ctgaaggaga gcgtggagat caactgcacc cgcccacaaca acaacacccg caagagcacc 900
accatcggcc ccggcccgccg cttctacgcc accggcgaca tcacggcgga catccgccag 960
gcccactgca acatcagcgg cgagaagtgg aacaacaccc tgaagcagat cgtgaccaag 1020
ctgcaggccc agttcggcaa caagaccatc gtgttcaagc agagcagcgg cgcgaccccc 1080
gagalegtga lgcacagcct caactgcggc ggcgagttct tctactgcaa cagcaccacc 1140
ctgttcaaca gcacctggaa caacaccatc gcccacaaca acaccaacgg caccatcacc 1200
ctgccctgcc gcataagca gatacatcac cgctggcagg aggtgggcaa ggccatgtac 1260
gcccccccca tccggggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgcccgcg cgcgcgac 1380
atgtcygaca actggcgag cgagctgtac aagtacaagg tggtaagat cgagccctg 1440
ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg ctaa 1494

```

<210> 8
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C 和 R462C

<400> 8
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
 Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
 20 25 30
 Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
 35 40 45
 Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
 50 55 60
 Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
 65 70 75 80
 Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Cys Glu
 85 90 95
 Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
 100 105 110
 Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Leu
 115 120 125
 Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
 130 135 140
 Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr Thr Ser Ile Arg Asn
 145 150 155 160
 Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
 165 170 175
 Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
 180 185 190
 Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
 195 200 205
 His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
 210 215 220
 Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys
 225 230 235 240
 Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
 245 250 255
 Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser Glu Asn Phe Thr Asp
 260 265 270
 Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu Ser Val Glu Ile Asn
 275 280 285

[0008]

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

Arg

<210> 9
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C 和 L469C

<400> 9
 atggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcgtt 60
 tgcgccagcg ccgtggagaa gctgtgggtg accgtgtact acggcgtgcc cgtgtggaag 120
 gaggccacca ccacctgtt ctgcgccagc gacgccaaag cctacgacac cgagggtcac 180
 aacgtgtggg ccacccacgc ctgcgtgcc accgaccca acccccagga gatcgtgtgt 240
 gagaacgtga ccgagaactt caacatgttg aagaacaaca tgtgygagca gatgcacgag 300
 gacatcatca gectgtggga ccagagcctg aagccctgcg tgaagctgac cccctgtgc 360
 glgacctgc acigcaccaa cctgaagaac gccaccaaca ccaagagcag caactggaa 420
 gagatggacc gcggcgagat caagaactgc agcttaagg tgaccaccag catccgcaac 480
 aagatgcaga aggagtacgc cctgtttctac aagctggagc tgggtcccat cgacaacgac 540
 aacaccagct acaagctgat caactgcaac accagcgtga tcaccaggc ctgcccacaag 600
 gtgagcttcg agcccatccc catccactac tgcgcccccg ccgcttcgc catcctgaag 660
 tgcacgaca agaagttaaa cggcagcggc cctgcacca acgtgagcac cgtgcagtgc 720
 acccacggca tcgcccctg ggtgagcacc cagctgctgc tgaacggcag cctggccgag 780
 gagggcgtgg tgatccgag cgagaacttc accgacaacg ccaagaccat catcgtgcag 840
 ctgaaggaga gcgtggagat caactgcacc cgcgccaaaca acaacaccg caagagcatc 900
 accatcggtc ccggccgcgc ctctacgcc accggcgaca tcacggcgga catccgccag 960
 gccactgca acalcagcgg cgaguagtg aacaacatcc tgaagcagat cglgaccaag 1020
 ctgcaggccc agttcgcaaa caagaccatc gtgttcagc agagcagcgg cggcgacccc 1080
 gagatcgtga tgcacagctt caactgcggc ggcgagttct tctactgcaa cagcaccacg 1140
 ctgttcacaa gcacctggaa caacaccate ggccccaaaca acaccaacgg caccatcacc 1200
 ctgccctgcc gcataagca gatacatcaac cgtggcagg aggtgggcaa ggccatgtac 1260
 gcccccacca tcgcgggcca gacccgtgc agcagcaaca tcaccggtc gtgtgtgacc 1320
 cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgccccgg cggcgcgac 1380
 atgcgcgaca actggcgag cgagtgytac aagtacaagg tggtaagat cgagccctc 1440

[0009]

ggcgtggccc ccactaaggc caagcgcgc gtgggcagc gcgagaagcg ctaa 1494

<210> 10
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C 和 L469C

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

[0010]

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

Arg

<210> 11
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 H99C 和 R462C

<400> 11
 atggatgcac tgaagagagg gctctgctgf gtgctgctgc tgtgiggagc agtcttctgt 60
 tcgcccagcg ccgtggagaa gctgtgggtg accgtgtact acgacgtgcc cgtgtggaag 120
 gagggcaacta ccacctgttt ctggcccgag gacgccaagg cctacgacac cgaggtgcac 180
 aacgtgtggg ccacccacgc ctgctgtccc accgacccca acccccagga gatcgtgctg 240
 gagaacgtga ccgagaactt caacatgttg aagaacaaca tgggtggagca gatgtgcgag 300
 gacatcatca gctgtgtgga ccagagcccg aagccctgcg tgaagctgac cccctgtgac 360
 gtgacctgc actgcacca cctgaagaac gccaccaaca ccaagagcag caactggaag 420
 gagaagacc gcggcgagat caagaactgc agcttcaagg tgaccaccag cctctgcgac 480
 aagatgcaga agagtagcgc cctgttctac aagctggagc tgggtgccat cgacaacgac 540
 aacaccagct acaagctgat caactgcaac accagcgtga tcaccacggc ctgcccacag 600
 gtgagcttgc agcccatccc catccactac tgcgcctccg ccgcttctgc catcctgag 660
 tgaacagaca agaagttcaa cggcagcggc cctgcacca acgtgagcac cgtgcagtgc 720
 acccagcgca tccgcccctg ggtgagcacc cagctgtctc tgaacggcag cctggccgag 780
 gagggcgttg tgatccgcag cgagaacttc accgacacac ccaagaccat catcgtgcag 840
 ctgaaggaga gcgtggagat caactgcacc cgcaccaaca acaacaccgc caagagcacc 900
 accatcgccc ccggccgcgc ctcttaccgc accggcgaca tcctcggcga catccgcgag 960
 gccactgca acatcagcgg cgagaagtgg aacaacaccc tgaagcagat cgtgaccaag 1020
 ctgcaggccc agtccgcaaa caagaccata gtgttcaagg agagcagcgg ccgacacccc 1080
 gagatcgtga tgcacagctt caactgcggc ggcgagttct tctactgcaa cagcaccacg 1140
 ctgttcaaca gcacctggaa caacaccatc ggcaccaaaa acaccaacgg caccatcacc 1200
 ctgcccctgc gattcagcaa gateatcaac cgtgtgcagg aggtgggcaa gaccatgtac 1260
 gcccccacca tccgcccga gatccgtgc agcagcaaca tcaccggcct gctgctgacc 1320
 cgcagcggcg gcaaggagat cagcaacacc accgagatct tccgcccgg ccgcccggac 1380
 atgtgygaca actggcgag cgagctgtac aagtacaagg tgggtgaagat cgagccctg 1440
 ggcgtggccc ccaccaaggc caagcgcgc gtggtgcagc gcgagaagcg ctaa 1494

<210> 12
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 H99C 和 R462C

<400> 12
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly

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

[0012]

Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser
 420 425 430
 Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Lys Glu Ile Ser
 435 440 445
 Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Cys Asp Asn
 450 455 460
 Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
 465 470 475 480
 Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys
 485 490 495

Arg

<210> 13
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V59C S109C 和 V95C W465C

<400> 13
 atgsatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcgtt 60
 tcgccagcgg ccgtggagaa gctgtgggtg accgtgtact accgctgtcc cgtgtggaag 120
 gagccacca ccacccgtgt ctgcgccagc gacgccaaagg cctacgacac cgagtgtccac 180
 aacgtgtggg ccacccacgc ctgcgtgccc accgacccca accccagga gatcgtgctg 240
 gagaacgtga ccgagaactt caacatgtgg aagaacaaca tgtgygagca gatgcacgag 300
 gacatcatca gctgtgggga ccagtgcttg aagccctgcg tgaagctgac cccctgtgc 360
 gtgacccgtc actgcaccaa cctgaagaac gccaccaaca ccaagagcag caactggaag 420
 gagatggacc gggcgagat caagaactgc agcttcaagg tgaccaaccag caccgcaac 480
 aagatgcaga aggagtacgc cctgttctac aagctggacg tggatgcccac cgacaacgac 540
 aacaccagct acaagctgat caactgcaac accagcgtga tcaccaggc ctgcccacag 600
 gtgagcttcg agcccatccc caccacttac tgcgccccg ccgcttccg caccctgaag 660
 tgcacacgaa agaagttcaa eggcagcggc ccctgacca accgtgagc cgtgcagtgc 720
 acccagggca tccgtccgtt ggtgagcacc cagctgtgtc tgaacggcag cctggccgag 780
 gaggcgctgg tgatccgag cgagaacttc accgacaacg ccaagaccat cctcgtgcag 840
 ctgaaggaga gegtggagat caactgcacc cgcaccaaca acaacaccg caagagcacc 900
 accatcgccc ccggccgcgc ctctacgcc accggcgaca tcacggcgca caccgccag 960
 gccacactga acatcagcgg cgagaagtgg aacaacacc tgaagcagat cgtgaccaag 1020
 ctgcaggccc agtctggcaa caagaccalc gtttcaage agagcagcgg cggcgacccc 1080
 gagatcgiga tgcacagctt caactcgggc ggcagttct tctactgcaa cagcaccag 1140
 ctgttcaaca gcacctggaa caacaccalc ggcaccaaca acaccaacgg caccatcacc 1200
 ctgcccctgc gcataagca gatcatcaac cgttggcagg aggtgggcaa ggccatgtac 1260
 gcccccacca tcgcgggcaa gatccgctgc agcagcaaca tcaccggcct gctgctgacc 1320
 cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgccccgg cggcgggcag 1380
 atgcgcgaca actgycgcag cgagctgtac aagtaesagg tggtaagat cggcccctg 1440
 ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg ctaa 1494

<210> 14
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V59C S109C 和 V95C W465C

<400> 14
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
 Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
 20 25 30
 Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
 35 40 45
 Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Cys His Asn Val Trp Ala
 50 55 60
 Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu

[0013]

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

[0014]

Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys
485 490 495

Arg

<210> 15
<211> 1494
<212> DNA
<213> 人工序列

<220>
<223> HIV-1 gp120 V59C S109C 和 W90C E268C

<400> 15
atggatgcaa tgaagagagg gctctgctgt gtgetgetge tgtgtggagc agtccttcgtt 60
tcgcccagcg ccgtggagaa gctgtgggtg accgtgtact acggcgtgcc cgtgtggaag 120
gaggccacca ccacctgfl clcgccagc gacgccaagg cctacgacac cgagtgcac 180
aacgtgtggg ccacccacgc ctgctgtccc accgaccca acccccagga gatcgtgctg 240
gagaacgtga ccgagaactt caacatgtgy aagaacaaca tgggtggagca gatgeacgag 300
gacatcatca gectgtggga ccagtgcctg aagccctgcg tgaagctgac cccctgtgtc 360
gtgaccttgc actgacccaa cctgaagaac gccaccaaca ccaagagcag caactggaag 420
gagatggacc gcggcgagat caagaactgc agcttcaagg tgaccaccag catccgcaac 480
aagatgcaga agaggtacgc cctgttctac aagctggacg tgggtgcccac cgacaacgac 540
aacaccagct acaagctgat caactgcaac accagcgtga tcaccagggc ctgcccgaag 600
gtgagcttgc agccctcccc cctccactac tgcgcccccg ccgcttgcgc catcctgaag 660
tgcacacgaa agaatgtcaa cggcagcggc cctgcacca acgtgagcac cgtgcagtgc 720
accacacgca tcgcccctgt ggtgagcacc cagctgtctg tgaacggcag cctggccgag 780
gaggcgctgg tgateccgag ctgyaacttc accgacaacg ccaagaccat catcgtgcag 840
ctgaaggaga gcgtggagat caactgcacc cgccccaaaca acaacaccgc caagagcacc 900
accatcgggc cggcgccgcg ctctctacgc accggcgaca tcctcgccga catccgccag 960
gcccactgca acatcagcgg cgagaagtgg aacaaacccc tgaageagat cgtgaccaag 1020
ctgcagggcc agttcgccaa caagaccatc gtgttcaagg agagcagcgg cggcgacccc 1080
gagatcgtga tgcaacgctt caactgcggc ggagagttct tctactgcaa cagcaccag 1140
ctgttcaaca gcacctggaa caacaccatc ggccccaaaca acaccaacgg caccatcacc 1200
ctgcctgtgc gcatcaagca gataatcaac cgctggcagg aggtgggcaa ggccatgtac 1260
gcccccccca tcgcgcccca gctccgtgc agcagcaaca tcaccggcct gctgtgacc 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tcgccccgg cgcgcgagac 1380
atgcgcgaca actggcgag cgagctgtac aagtacaagg tggtaagat cgagcccttg 1440
agcgtggccc ccaccaagge caagcgccgc gtggtgcagc gcgagaagcg ctaa 1494

<210> 16
<211> 497
<212> PRT
<213> HIV-1 gp120 V59C S109C 和 W90C E268C

<400> 16
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15
Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
20 25 30
Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
35 40 45
Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Cys His Asn Val Trp Ala
50 55 60
Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
65 70 75 80
Glu Asn Val Thr Glu Asn Phe Asn Met Cys Lys Asn Asn Met Val Glu
85 90 95
Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Cys Leu Lys Pro
100 105 110
Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Leu
115 120 125
Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
130 135 140
Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr Thr Ser Ile Arg Asn

[0015]

```

145              150              155              160
Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
165              170
Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
180              185              190
Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
195              200              205
His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
210              215              220
Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys
225              230              235              240
Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
245              250              255
Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser Cys Asn Phe Thr Asp
260              265              270
Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu Ser Val Glu Ile Asn
275              280              285
Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Thr Ile Gly Pro
290              295              300
Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
305              310              315              320
Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn Asn Thr Leu Lys Gln
325              330              335
Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn Lys Thr Ile Val Phe
340              345              350
Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn
355              360              365
Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser
370              375              380
Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr Asn Gly Thr Ile Thr
385              390              395              400
Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg Trp Gln Glu Val Gly
405              410              415
Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser
420              425              430
Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Lys Glu Ile Ser
435              440              445
Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn
450              455              460
Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
465              470              475              480
Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys
485              490              495

```

Arg

<210> 17
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V59C S109C 和 I103C Q413C

[0016]

```

<400> 17
atggatgcaa tgaagagagg gctctgctgt gtgtgtgtgc tgtgtggagc agtcttctgt 60
tcgcccagcg cegtggagaa cgtgtgggtg accgtgtact acggcgtgce cgtgtggaag 120
gaggccacca ccacctgttt ctgcgccage gacgccaagg cctacgacac cgagtgccac 180
aacgtgtggg ccacccacgc ctgcgtgccc accgacccca acccccagga gatcgtgtgt 240
gagaacgtga ccgagaactt caacatgtgg aagaacaaca tggtagagca gatgacgag 300
gacatctgya gctgttggga ccagtgcctg aagccctgag tgaagctgac cccctgtgtc 360
gtgacctgac actgcaccaa cctgaagaac gccaccaaca caaagagcag caactggang 420
gagatggacc gctgcgagat caagaactgc agcttcaagg tgaccaccag cctccgaaac 480
aagatgcaga aggagtacgc cctgtttctac aagctggacg tggtagccat cgacaacgac 540
aacaccagct acaagctgat caactgcaac accagcgtga tcacctaggc ctgcccacag 600
gtgagcttcc agcccattcc catccactac tgcgcccccg ccggtttcgc catcctgaag 660
tgcaacgaca agaagttaaa cggcagcggc cctgcacca acgtgagcac cgtgcagtgc 720
accacaggca tccgccccgt ggtgagcacc cagctgtgtc tgaacggcag cctggccgag 780
gaggcggtgg tgatccgag cgagaacttc accgacaacg ccaagaccat catcgtgcag 840
ctgaaggaga gcgtggagat caactgcacc cgcaccaaca acaacaccgc caagagcacc 900
accatggccc ccggccgcgc ctctacgccc accggcgaca tcatcgcgca catccgcag 960
gcccactgca acatcagcgg cgagaagtgg acaacacccc tgaagcagat cgtgaccaag 1020
ctgcaggccc agttcgagca caagaccata gtgttcaagg agacagcgg ccgacgcccc 1080
gagatcgtga tgcacagett caactgcgga ggaggttct tctactgcaa cagcaccacg 1140
ctgttcaaca gcacctggaa caacaccatc ggcccacaaca acaaccaagg caccatcacc 1200
ctgcctgtcc gcacaaagca gatcatcaac cgtgtgtgag agtggggcaa ggcatgtac 1260
gcccccccca tccggcgcca gatcctgtgc agcagcaaca tcaccggcct gctgtgtacc 1320
cggcagcgcc gcaaggagat cagcaaacac accgagatct tccgccccgg cggcggcgac 1380
atgcgcgaca atcgcgcgag cgagctgtac nagtaacagg tggtagagat cgagccctgt 1440
ggcgtggccc ccaccaggca caagcgccgc gggtagcagc gcgagaagcg ctac 1494

```

<210> 18

<211> 497

<212> PRT

<213> HIV-1 gp120 V59C S109C 和 I103C Q413C

<400> 18

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
20 25 30

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

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

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

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

Gln Met His Glu Asp Ile Cys Ser Leu Trp Asp Gln Cys Leu Lys Pro
100 105 110

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

Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
130 135 140

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

Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
165 170 175

Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
180 185 190

Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
195 200 205

His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
210 215 220

Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys

[0017]

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

Arg

<210> 19
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C W465C 和 W90C E268C

<400> 19
 atggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtctttggt 60
 tgcgccagcg ccgtggagaa gctgtgggtg accgtgtact acggcgtagc cgtgtggaag 120
 gaggccacca ccacctgttt ctgcgccagc gacgccaagg cctacgacac cgaggtagac 180
 aacgtgtggg ccacceacgc ctgcgtgcc accgaccca acccccagga gatcgtgctg 240
 gagaacgtga ccgagaactt caacaigtgy aagaacaaac tgtgygagca gatgcacgag 300
 gacatcataa gctgtggga ccagagcctg aagccctgcg tgaagctgac cccctgtgc 360
 gtgaccctgc actgcaccaa cctgaagaac gccaccaaca ccaagagcag caactggaag 420
 gagatggacc gcgagagat caagaactgc agcttcaagg tgaccaccag catccgcaac 480
 aagatgcaga aggagtagc cctgttctac aagctggacg tgggtcccat cgacaacgac 540
 aacaccagct acaagctgat caactgcaac accagcgtga tcaccaggc ctgccccang 600
 gtgagcttcg agcccatccc catccactac tgcgccccg ccggtcttcg catcctgaag 660
 tgaacgaca agaagttaaa cggcagcggc cctgcacca acgtgagcac cgtgcagtgc 720
 acccaaggaa tccgcccgt ggtagcacc cagctgtgc tgaacggcag cctggcagag 780
 gagggcgtgg tgatccgag ctgyaacttc accgacaacg ccaagaccat catcgtgcag 840

[0018]

```

cigaaggaga gegtggagat caactgcacc cgcaccaaca acaacacccg caagagcatt 900
accatcgccc ccggccgggc ctcttaccgc accggcgaca tcacggcgga caccggccag 960
gcccactgca acatcagcgg cgagaagigg aacaacaccc tgaagcagat cgtgaccaag 1020
ctgcaggccc agttcgccaa caagaccatc gtgttcaage agagcagcgg cggcgacccc 1080
gagatcgtga tgcacagctt caactgcggc ggcgagttct tctactgcaa cagcaccacg 1140
ctgttcaaca gcaccigga caacaccalc ggccccaaca acaccaacgg caccatcacc 1200
ctgccctgcc gcatcaagea gatcatcaac cgtggcgagg aggtgggcaa ggccatgtac 1260
gcccccccca tccggggcca gatccgctgc agcagcaaca tcaccggcct gctgctgacc 1320
cgcgacggcg gcaaggagat cagcaaacac accgagatct tccgcgccgg cggcgcgagc 1380
atgcgcgaca actgycgcag cgagctgtac aagtcacagg tggtagaagat cggcccccg 1440
ggcgtggccc ccaccaaggc caagcgcgcg gtggtgcagc gcgagaagcg ctaa 1494

```

<210> 20
 <211> 497
 <212> PRT
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C W465C 和 W90C E268C

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

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

Arg

<210> 21
 <211> 1494
 <212> DNA
 <213> 人工序列

<220>
 <223> HIV-1 gp120 V95C W465C 和 1103C Q413C

<400> 21
 atggatgcaa tgaagagagg gctctgctgt gtctgtctgc tgtgtggagc agtettcgtt 60
 tcgcccagcg ccgtggagaa gctgtgggtg accgtgtact acggcgtgcc cgtgtggaag 120
 gaggccacca ccacctgttt ctgcgccagc gacgccaagg cctacgacac cgagggtcac 180
 aacgtgtggg ccaccacagc ctgcgtgccc accgaccca acccccagga gatcgtgctg 240
 gagaacgtga ccgagaactt caacatgttg aagaacaaca tttgygagca gatgcacgag 300
 gacatctgya gctgtggga ccagagcctg aagccctgcg tgaagctgac cccctgtgc 360
 gtgacctgc actgcaccaa cctgaagaac gccaccaaca ccaagagcag caactggaag 420
 gagatggacc gcggcgagat caagaactgc agcttcaagg tgaccacag catccgaac 480
 aagatgcaga aggaglacgc cctgttctac aagctggacg tgggtgccal cgacaacgac 540
 aacaccagct acaactgat caactgcaac accagcgtga tcaccaggc ctgcccacag 600
 gtgagcttgc agccatecc caccactac tgcgccccg ccggttgcg cactctgaag 660
 tgcaacgaca agaagttaaa cggcagcggc cctgcacca acgtgagcag cgtgcagtgc 720
 acccaccgca tcgccccgt ggtgagcacc cagctgtgc tgaacggcag cctggccgag 780
 gagggcgttg tgatccgcag cgagaacttc accgacaacg ccaagaccat cactgtcag 840
 ctgaaggaga gcgtggagat caactgcacc cgccccaa acaacaccg caagagcacc 900
 accatcggcc ccggccgccc ctctacgcc accggcgaca tcctggcga catccgccag 960
 gccactgca acatcagcg cgagaagtgg aacaacacc tgaagcagat cgtgaccaag 1020
 ctgcaggccc agttcgccaa caagaccatc gtgttcagc agagcagcgg cggcgacccc 1080
 gagatcgtga tgcacagctt caactgcggc ggcgagttct tctactgcaa cagcaaccag 1140
 ctgttcaaca gcaactggaa caacaccatc ggcaccaaca acaccaacgg caccatcacc 1200
 ctgccccgac gcatcaagca gatcatcaac cgtgtgtgy aggtgggcaa ggccatgtac 1260
 gcccccacca tcgcgggcca gatccgtgc agcagcaaca tcaccggcct gctgctgacc 1320
 cgcgacggcg gaaggagat cagcaacacc accgagatct tcgccccgg cggcgccgac 1380
 atgcgcgaca actgycgcag cgagctgtac aagtacaagg tggtaagat egagccctg 1440
 ggcgtggccc ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cttaa 1494

[0020]

<210> 22
 <211> 497
 <212> PRT
 <213> 人工序列

 <220>
 <223> HIV-1 gp120 V95C W465C 和 I103C Q413C

 <400> 22
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
 Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
 20 25 30
 Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
 35 40 45
 Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
 50 55 60
 Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
 65 70 75 80
 Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Cys Glu
 85 90 95
 Gln Met His Glu Asp Ile Cys Ser Leu Trp Asp Gln Ser Leu Lys Pro
 100 105 110
 Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Leu
 115 120 125
 Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
 130 135 140
 Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr Thr Ser Ile Arg Asn
 145 150 155 160
 Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
 165 170 175
 Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
 180 185 190
 Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
 195 200 205
 His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
 210 215 220
 Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys
 225 230 235 240
 Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
 245 250 255
 Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser Glu Asn Phe Thr Asp
 260 265 270
 Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu Ser Val Glu Ile Asn
 275 280 285
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Thr Ile Gly Pro
 290 295 300
 Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
 305 310 315 320
 Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn Asn Thr Leu Lys Gln
 325 330 335
 Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn Lys Thr Ile Val Phe
 340 345 350
 Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn

[0021]

```

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

<210> 23
<211> 668
<212> PRT
<213> HIV-1

<400> 23
Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1          5          10          15
Ala Val Phe Val Ser Pro Ser Ala Val Glu Lys Leu Trp Val Thr Val
20          25          30
Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys
35          40          45
Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
50          55          60
Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
65          70          75          80
Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu
85          90          95
Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
100          105          110
Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Leu
115          120          125
Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp Lys Glu Met Asp Arg
130          135          140
Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr Thr Ser Ile Arg Asn
145          150          155          160
Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro
165          170          175
Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile Asn Cys Asn Thr Ser
180          185          190
Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile
195          200          205
His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys
210          215          220

```

[0022]

Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys
 225 230 235 240
 Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
 245 250 255
 Ser Leu Ala Glu Gly Val Val Ile Arg Ser Glu Asn Phe Thr Asp
 260 265 270
 Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu Ser Val Glu Ile Asn
 275 280 285
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Thr Ile Gly Pro
 290 295 300
 Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
 305 310 315 320
 Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn Asn Thr Leu Lys Gln
 325 330 335
 Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn Lys Thr Ile Val Phe
 340 345 350
 Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn
 355 360 365
 Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser
 370 375 380
 Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr Asn Gly Thr Ile Thr
 385 390 395 400
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg Trp Gln Glu Val Gly
 405 410 415
 Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser Ser
 420 425 430
 Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Lys Glu Ile Ser
 435 440 445
 Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn
 450 455 460
 Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu
 465 470 475 480
 Gly Val Ala Pro Thr Lys Ala Ile Ser Ser Val Val Gln Ser Glu Lys
 485 490 495
 Ser Ala Val Thr Leu Gly Ala Met Phe Leu Gly Phe Leu Gly Ala Ala
 500 505 510
 Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu Thr Val Gln Ala Arg
 515 520 525
 Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala
 530 535 540
 Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys
 545 550 555 560
 Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln
 565 570 575
 Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr
 580 585 590
 Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Asp Gln Ile
 595 600 605
 Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile Asp Asn Tyr
 610 615 620

[0023]

Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu
625 630 635 640

Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp
645 650 655

Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr Ile
660 665

<210> 24
<211> 510
<212> PRT
<213> HIV-1

<220>
<221> misc feature
<223> 来自图5的HXB2

<400> 24

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg
1 5 10 15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu
20 25 30

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

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
50 55 60

Val His Asn Val Trp Ala Thr His Ala Gly Val Pro Thr Asp Pro Asn
65 70 75 80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp
85 90 95

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

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser
115 120 125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser
130 135 140

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

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

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr
180 185 190

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

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

[0024]

```

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

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

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

Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Glu Leu
                275                280                285

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

Lys Lys Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile
305                310                315                320

Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala
                325                330                335

Lys Trp Asn Ala Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln
                340                345                350

Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp
                355                360                365

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

Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp
385                390                395                400

Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu
                405                410                415

Pro Cys Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val Gly Lys
                420                425                430

Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn
                435                440                445

Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Gly
                450                455                460

Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg
465                470                475                480

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

Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg
                500                505                510

<210> 25
<211> 512
<212> PRT

```

[0025]

<213> HIV-1

<220>
<221> misc_feature
<223> 来自图5的共有序列

<220>
<221> misc_feature
<222> (5)..(14)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (17)..(17)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (20)..(20)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (31)..(31)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (74)..(74)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (84)..(84)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (87)..(87)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (99)..(99)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (128)..(128)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (130)..(130)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (133)..(133)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (137)..(137)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (141)..(141)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (144)..(148)
<223> 非共有氨基酸

[0026]

<220>
<221> misc_feature
<222> (150)..(151)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (160)..(162)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (167)..(167)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (169)..(169)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (175)..(175)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (181)..(182)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (188)..(188)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (192)..(192)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (194)..(195)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (230)..(230)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (232)..(232)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (236)..(236)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (269)..(269)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (275)..(275)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (289)..(290)
<223> 非共有氨基酸

<220>

[0027]

<221> misc_feature
<222> (306)..(306)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (308)..(308)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (310)..(311)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (318)..(320)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (322)..(322)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (324)..(324)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (326)..(327)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (336)..(337)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (341)..(341)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (347)..(348)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (351)..(352)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (357)..(357)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (361)..(361)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (374)..(374)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (397)..(397)
<223> 非共有氨基酸

<220>
<221> misc_feature

[0028]

```

<222> (399)..(402)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (404)..(404)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (406)..(406)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (410)..(410)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (412)..(413)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (424)..(424)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (427)..(427)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (441)..(441)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (461)..(461)
<223> 非共有氨基酸

<220>
<221> misc_feature
<222> (466)..(467)
<223> 非共有氨基酸

<400> 2

Met Arg Val Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Arg
1 5 10 15

Xaa Gly Thr Xaa Leu Leu Gly Met Leu Met Ile Cys Ser Ala Xaa Glu
20 25 30

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

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
50 55 60

Val His Asn Val Trp Ala Thr His Ala Xaa Val Pro Thr Asp Pro Asn
65 70 75 80

Pro Gln Glu Xaa Val Leu Xaa Asn Val Thr Glu Asn Phe Asn Met Trp
85 90 95

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

```

[0029]

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

[0030]

Leu Pro Cys Arg Ile Lys Gln Xaa Ile Asn Xaa Trp Gln Glu Val Gly
420 425 430

Lys Ala Met Tyr Ala Pro Pro Ile Xaa Gly Gln Ile Arg Cys Ser Ser
435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Xaa Xaa Xaa Xaa
450 455 460

Asn Xaa Xaa Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn
465 470 475 480

Trp Arg Ser Glu Leu Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly
485 490 495

Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg
500 505 510

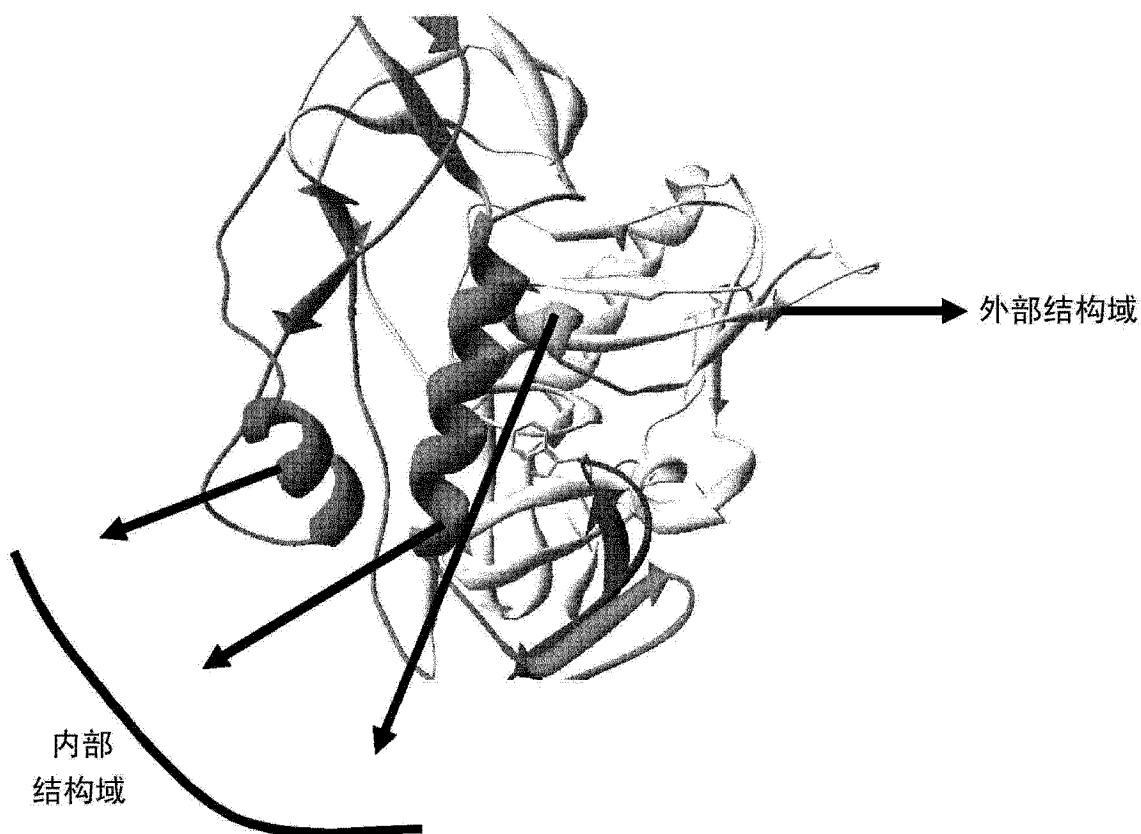


图 1A

MDAMKRGLCCVLLLCGAVFVSPSAVEKLVWTVVYGVVPVWKEATTTLFCASDAKAYDTEVH
NVWATHACVPTDNPQEIVLENVTFENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLC
 VTLHCTNLKNATNTKSSNWKEMDRGEIKNCSFKVTTsirNKMqKEYALFYKLDVVPIDND
 NTSYKLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFN_{MD}SGPCTNVSTVQC
THGIRPVVSTQLLNGSLAEEGVVIRSENFTDnAKTIIVQLKESVEINCTRPNNNTRKSI
 TIGPGRAFyATGDIIGDIRQAHCNISGEKWNNTLKQIVTKLQAQFGNKTIVFKQSSGGDP
 EIVMHSFNCGGEFFYCNSTOLFNSTWNNTIGPNNTNGTITLPCRIKQIINRWQEVGKAMY
 APPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGDMRDNWRSELYKYKVVKIEPL
 GVAPTKAKRRVVQREKR

图 1B

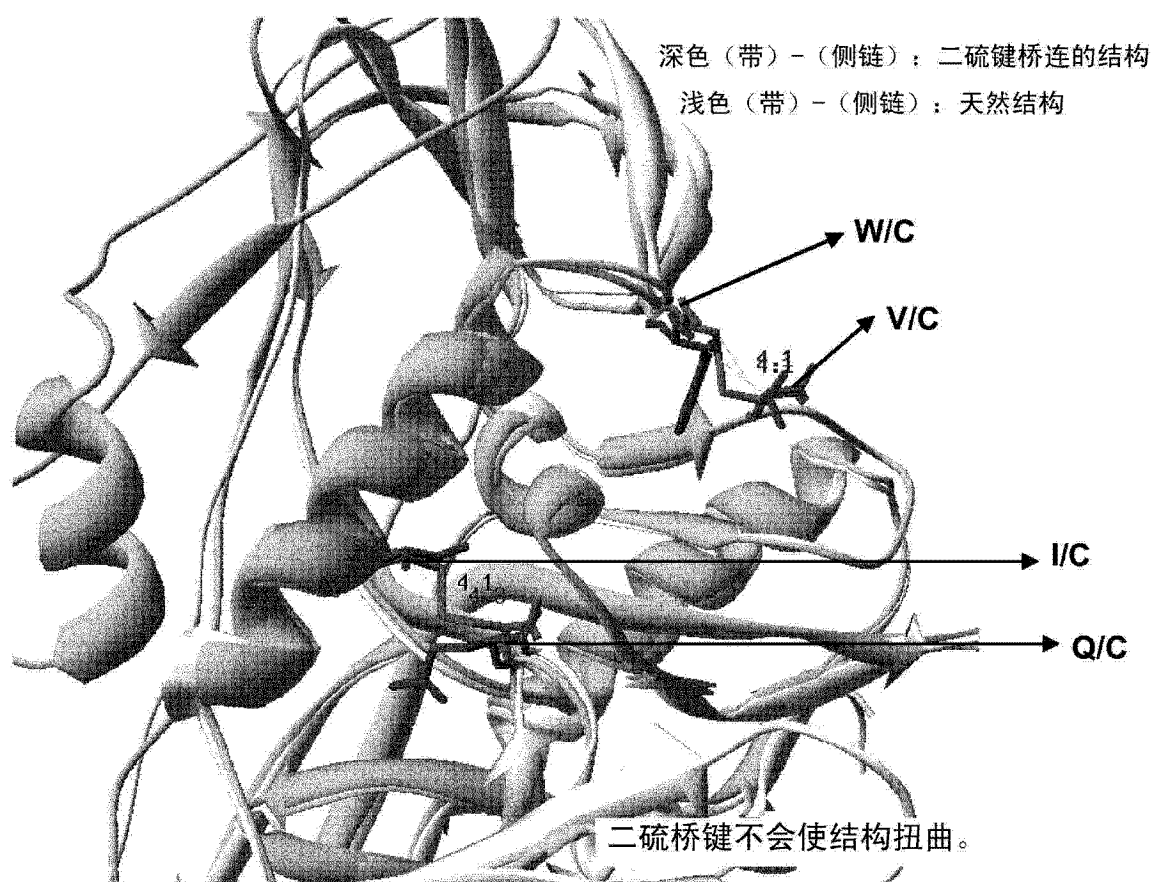


图 2

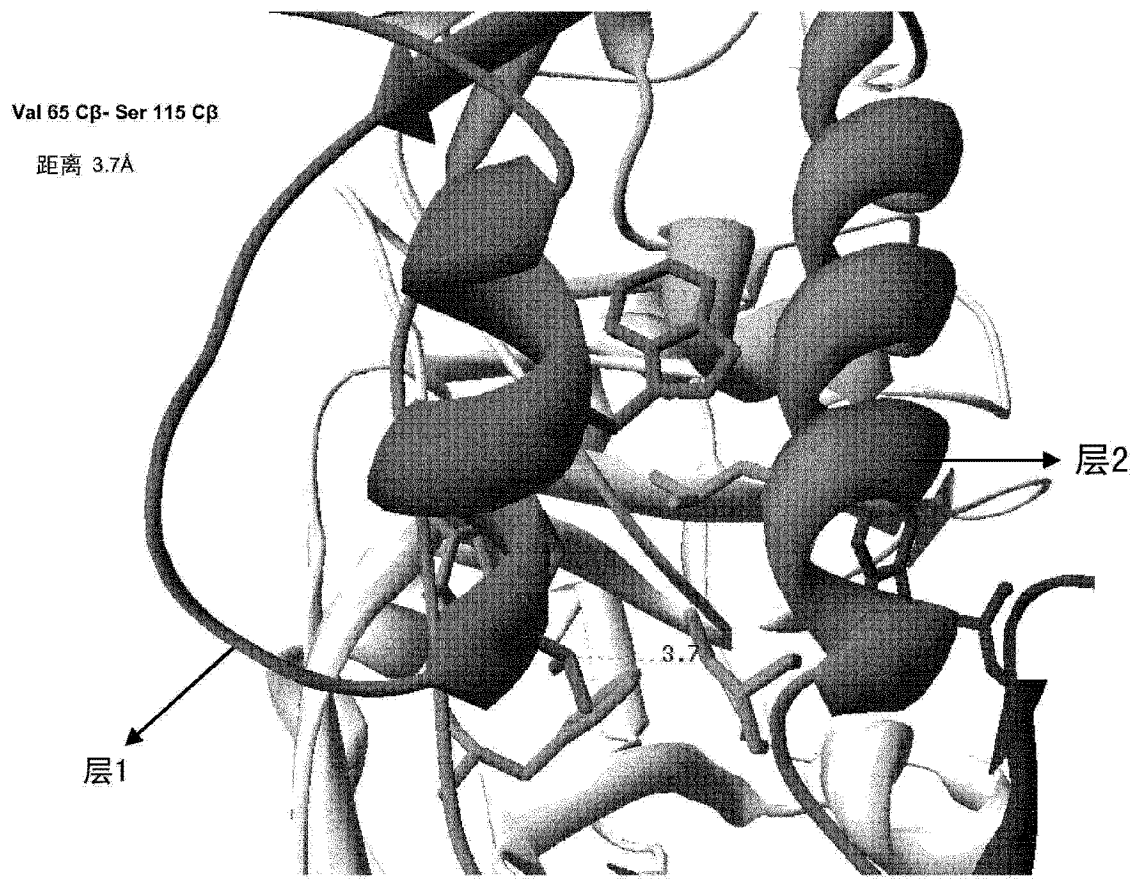


图 3

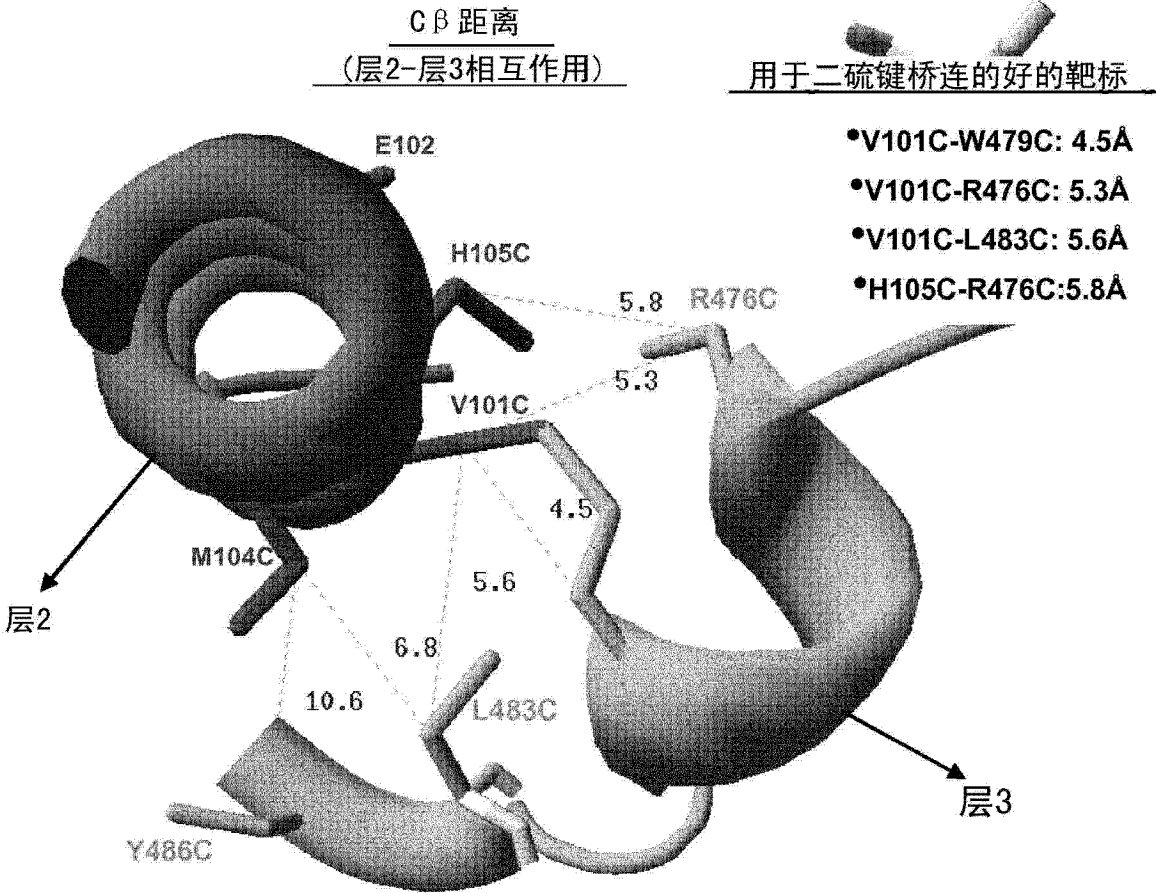


图 4

SF162 MRVKGIRKN- YQHLWRGGTL LLGMLMICS A VEKLWVTVYY GVPVWKEATT TLFCASDAKA YDTEV V65C(V59C)
 HxB2 MRVKEKYQHL WRWGWRWGIM LLGMLMICS A TEKLWVTVYY GVPVWKEATT TLFCASDAKA YDTEV
 共有 MRVK-----WR-GT- LLGMLMICS A -EKLWVTVYY GVPVWKEATT TLFCASDAKA YDTEV

V101C(V95C) H105C(H99C) S115C(S109C)
 SF162 HNVWA THACVPTDPN PQEIVLENT ENFNMMKNM VEQMHEDIIS LWDQSLKPCV KLIPLCVTLH
 HxB2 HNVWA THAGVPTDPN PQEVVLNVT ENFNMMKNM VEQMHEDIIS LWDQSLKPCV KLIPLCVSLK
 共有 HNVWA THA-VPTDPN PQE-VL-NVT ENFNMMKN-M VEQMHEDIIS LWDQSLKPCV KLIPLCV-L-

SF162 CTNLKNATNT KSSNWKE-MD RGEIKNCSFK VITSIRNKMQ KEYALFYKLD VVPIDNDNIS YKLIN
 HxB2 CTDLKNDTNT NSSSGRMIME KGEIKNCSFN ISTSIRGKVQ KEYAFFYKLD IIPIDNDTIS YLTIS
 共有 CT-LKN-TNT -SS-----M- -GEIKNCSF- --TSIR-K-Q KEYA-FYKLD ---PIDND-TS Y-L--

SF162 CNTSV ITOACPKVSF EPIPIHYCAP AGFAILKCN KKFNGSGPCT NVSTVQCTHG IRPVVSTQLL
 HxB2 CNTSV ITOACPKVSF EPIPIHYCAP AGFAILKCN KKFNGSGPCT NVSTVQCTHG IRPVVSTQLL
 共有 CNTSV ITOACPKVSF EPIPIHYCAP AGFAILKCN- K-FNG-GPCT NVSTVQCTHG IRPVVSTQLL

SF162 LNSGLAEEGV VIRSENFTDN AKTIIVOLKE SVEINCTRPN NNTKRSITI- -GPGRFYAT GDIIG
 HxB2 LNSGLAEEGV VIRSVNFTDN AKTIIVOLNT SVEINCTRPN NNTKKIRIQ RGPGRFVTI GKI-G
 共有 LNSGLAEE-V VIRS-NFTDN AKTIIVOL-- SVEINCTRPN NNTK-I-I- -GPGRF--- G-I-G

SF162 DIRQA HCNISGERKN NTLKQIVTKL QAQFGN-KTI VFKQSSGGDP EIVMHSFNCG GEFFYCNSIQ
 HxB2 NMRQA HCNISRAKWN ATLKQIASKL REQFGNNKTI IFKQSSGGDP EIVTHSFNCG GEFFYCNSIQ
 共有 --RQA HCNIS--KWN -TLKQI--KL --QFGN-KTI -FKQSSGGDP EIV-HSFNCG GEFFYCNSIQ

SF162 LFNSTWNN-- --TIGPNNTN G--TITLPCR IKQIINRWQE VGKAMYAPPI RGQIRCSSNI TGLLL
 HxB2 LFNSTWFNST WSTEGSNNTG GSDTITLPCR IKQFINRWQE VGKAMYAPPI SGQIRCSSNI TGLLL
 共有 LFNSTW-N-- --T-G-NNT- G--TITLPCR IKQ-IN-WQE VGKAMYAPPI -GQIRCSSNI TGLLL

W479C(W465C) R475C(R462C) L483C(L469C)
 SF162 TRDGG KEISNTTEIF RPPGGDMRDN WRSELYKYKV VKIEPLGVAP TKAKRRVVQR EKR
 HxB2 TRDGG NN-NGSEIF RPPGGDMRDN WRSELYKYKV VKIEPLGVAP TKAKRRVVQR EKR
 共有 TRDGG ----N--EIF RPPGGDMRDN WRSELYKYKV VKIEPLGVAP TKAKRRVVQR EKR

Seq # in HxB2(Seq # in SF162)

图 5

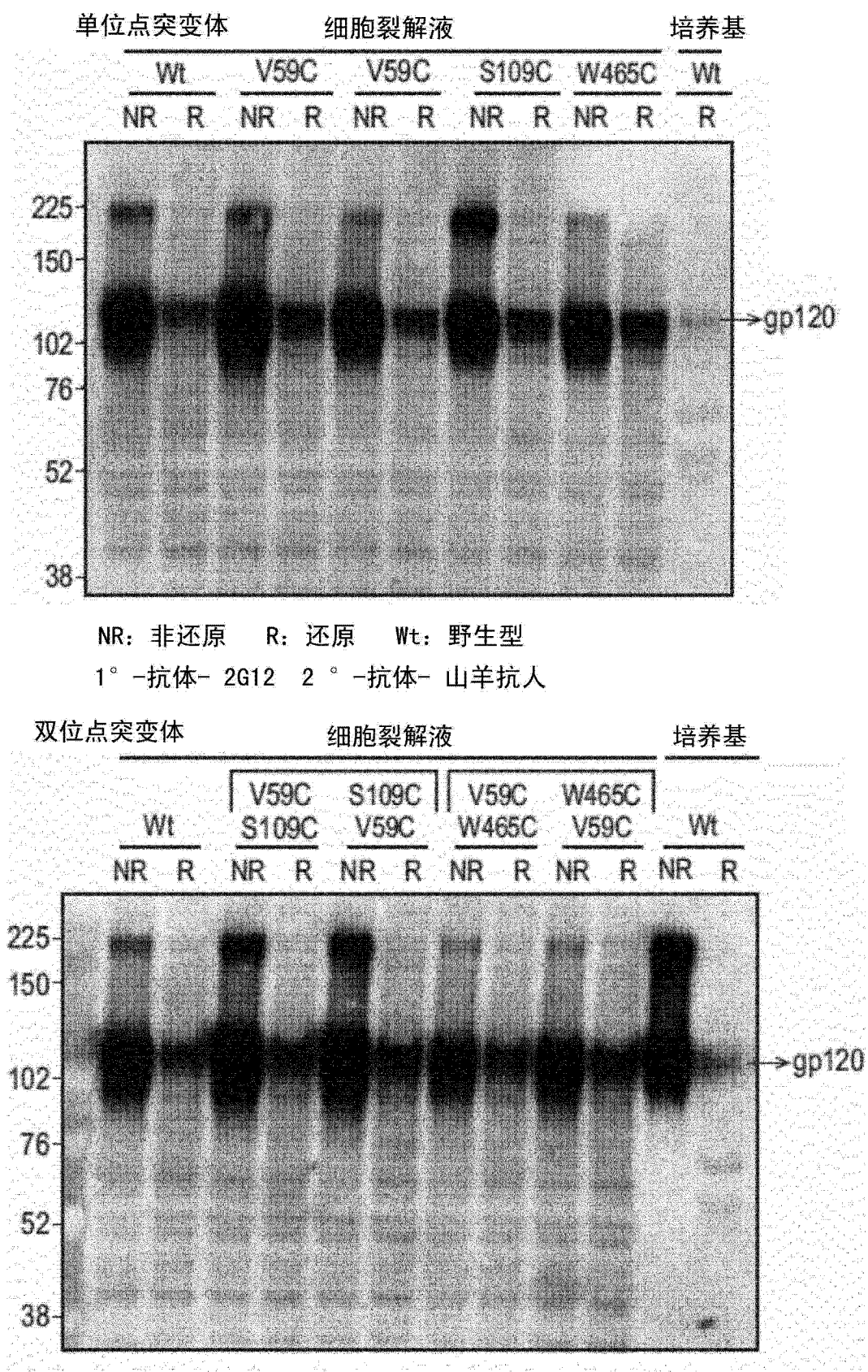


图 6

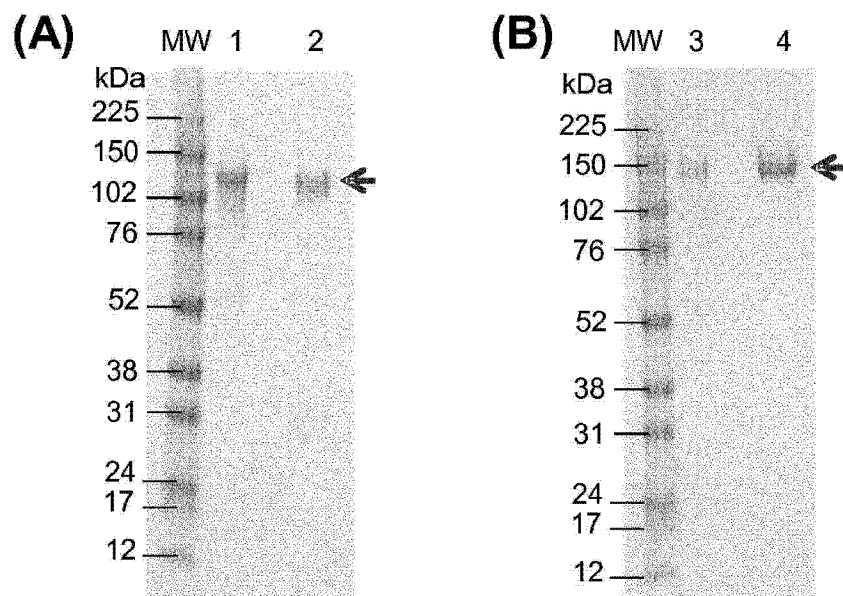


图 7

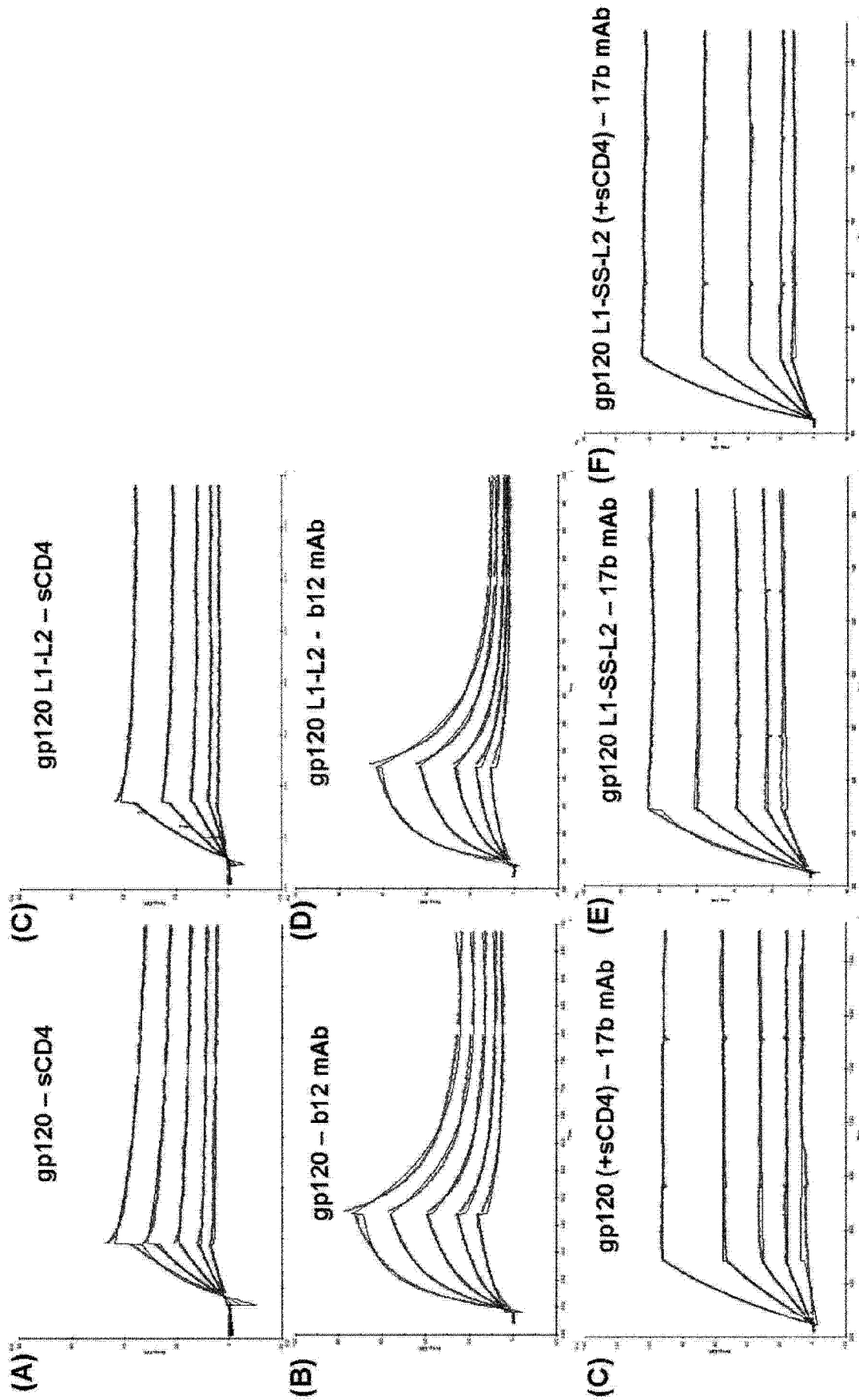


图 8

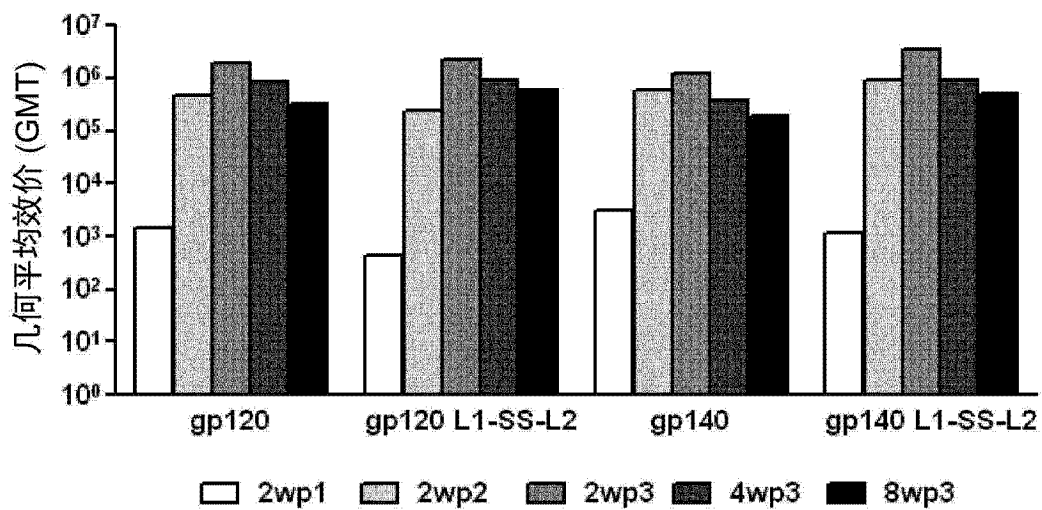
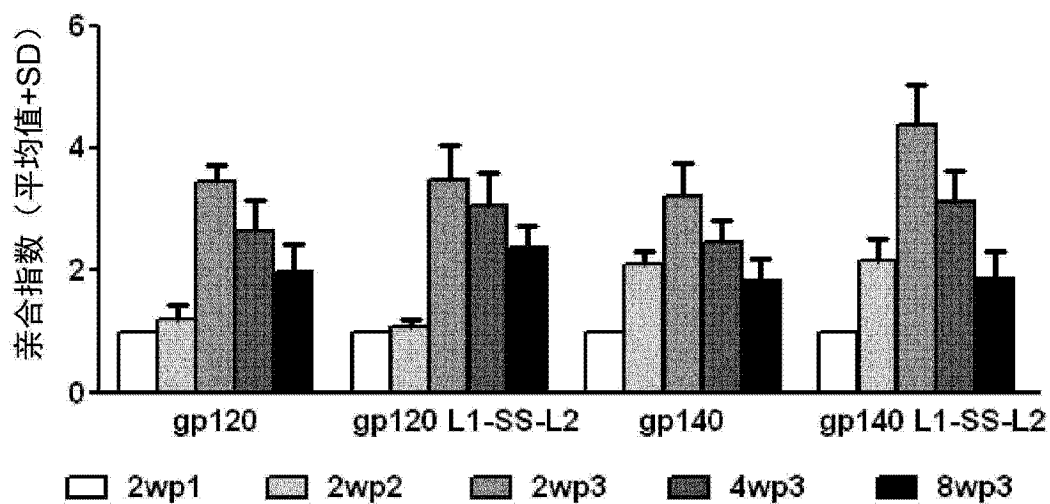
A**B**

图 9

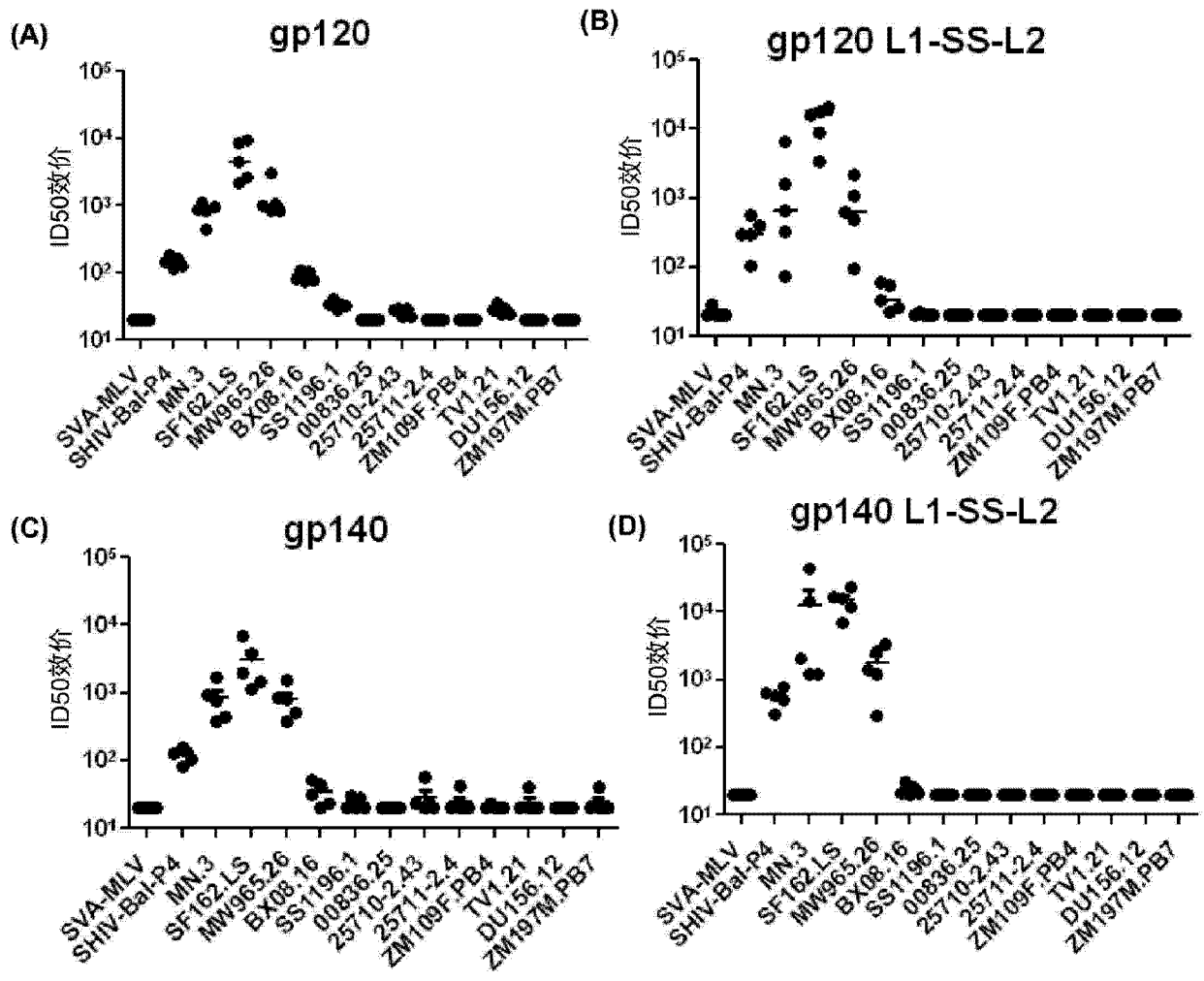


图 10

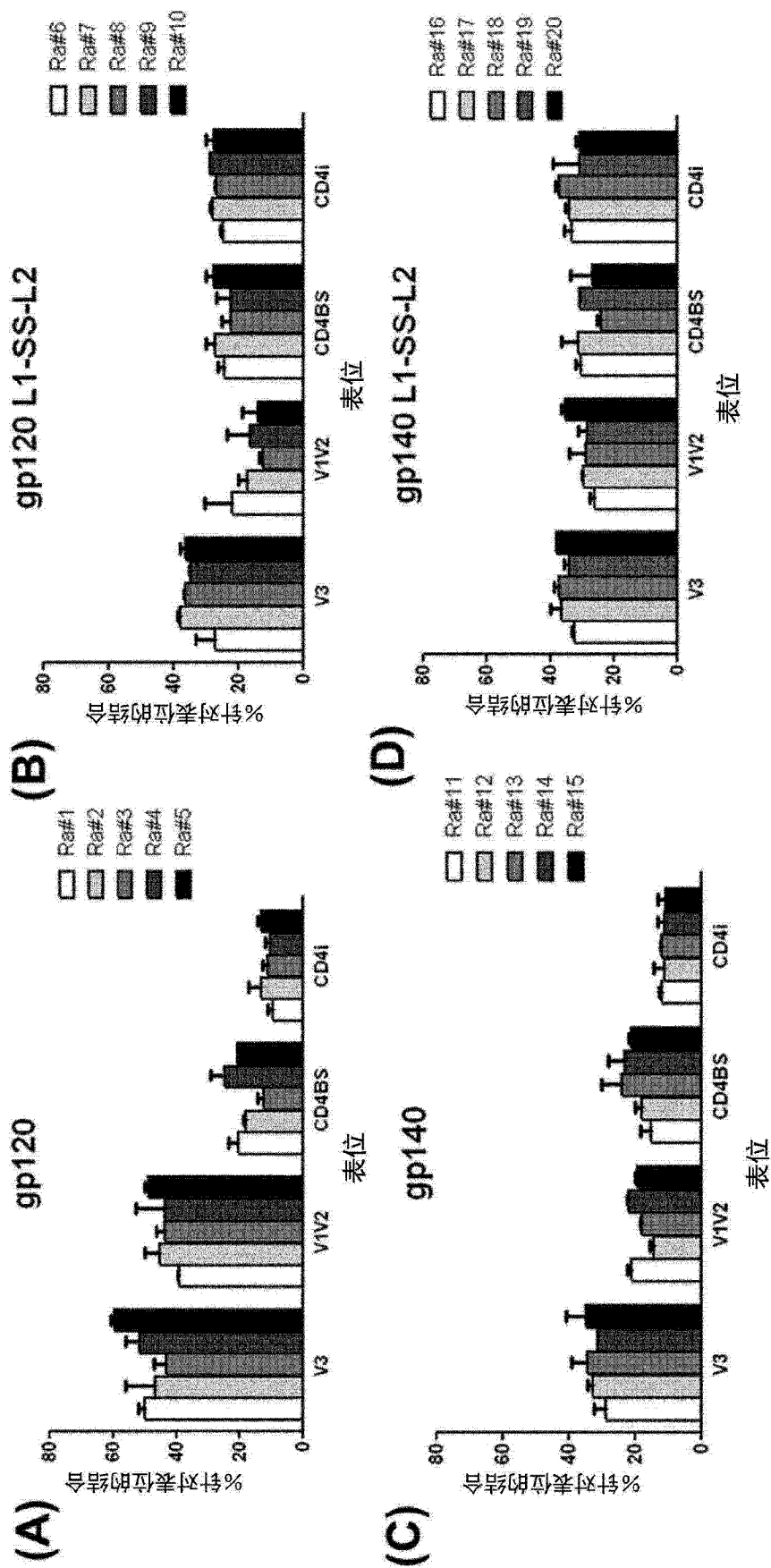


图 11