



(12) 发明专利

(10) 授权公告号 CN 102164952 B

(45) 授权公告日 2016. 05. 25

(21) 申请号 200980138400. 6

A61K 39/04(2006. 01)

(22) 申请日 2009. 07. 24

(56) 对比文件

(30) 优先权数据

61/083720 2008. 07. 25 US

US 6613881 B1, 2003. 09. 02,

CA 2290774 A1, 1998. 11. 26,

WO 0162893 A2, 2001. 08. 30,

CA 2326598 A1, 1999. 10. 14,

(85) PCT国际申请进入国家阶段日

2011. 03. 25

审查员 李翠莹

(86) PCT国际申请的申请数据

PCT/EP2009/059580 2009. 07. 24

(87) PCT国际申请的公布数据

W02010/010177 EN 2010. 01. 28

(73) 专利权人 葛兰素史密丝克莱恩生物有限公
司

地址 比利时里克森萨特

专利权人 葛兰素集团有限公司

(72) 发明人 J·布朗 P·梅滕斯 D·墨菲

(74) 专利代理机构 中国专利代理(香港)有限公
司 72001

代理人 李进 刘健

(51) Int. Cl.

C07K 14/35(2006. 01)

权利要求书1页 说明书60页

序列表106页 附图13页

(54) 发明名称

新型组合物和方法

(57) 摘要

本发明涉及用于治疗或预防潜伏性TB的多肽,其包含:(i)Rv3616c蛋白序列;(ii)Rv3616c蛋白序列的变体;或(iii)Rv3616c蛋白序列的免疫原性片段。在其它方面,本发明涉及用于治疗或预防潜伏性结核病的相关多核苷酸、融合蛋白和方法。

1. 分离多肽在制备用于治疗潜伏性结核病的药物中的用途,所述分离多肽:
 - (i) 由SEQ ID NO: 1的Rv3616c蛋白序列组成;或
 - (ii) 包含Rv3616c蛋白序列的免疫原性片段,所述片段选自SEQ ID NO: 127、128、130、131-133、135、143-148或150-156。
2. 根据权利要求1的用途,其中所述多肽由SEQ ID NO: 1的Rv3616c蛋白序列组成。
3. 根据权利要求1的用途,其中所述多肽包含Rv3616c蛋白序列的免疫原性片段,所述片段选自SEQ ID NO: 127、128、130、131-133、135、143-148或150-156。
4. 根据权利要求3的用途,其中所述多肽包含Rv3616c蛋白序列的免疫原性片段,所述片段选自SEQ ID NO: 127、128、147、148或152。
5. 根据权利要求1-4中任一项的在制备用于预防结核病再活化的药物中的用途。
6. 根据权利要求1-4中任一项的在制备用于延迟结核病再活化的药物中的用途。
7. 分离多核苷酸在制备用于治疗潜伏性结核病的药物中的用途,所述分离多核苷酸包含编码以下多肽的核酸序列,所述多肽:
 - (i) 由SEQ ID NO: 1的Rv3616c蛋白序列组成;或
 - (ii) 包含Rv3616c蛋白序列的免疫原性片段,所述片段选自SEQ ID NO: 127、128、130、131-133、135、143-148或150-156。
8. 根据权利要求7的用途,其中所述多肽由SEQ ID NO: 1的Rv3616c蛋白序列组成。
9. 根据权利要求7的用途,其中所述多肽包含Rv3616c蛋白序列的免疫原性片段,所述片段选自SEQ ID NO: 127、128、130、131-133、135、143-148或150-156。
10. 根据权利要求9的用途,其中所述多肽包含Rv3616c蛋白序列的免疫原性片段,所述片段选自SEQ ID NO: 127、128、147、148或152。
11. 根据权利要求7-10中任一项的在制备用于预防结核病再活化的药物中的用途。
12. 根据权利要求7-10中任一项的在制备用于延迟结核病再活化的药物中的用途。

新型组合物和方法

发明领域

[0001] 本发明涉及用于治疗或预防结核病、尤其用于治疗或预防潜伏性结核病和预防或延迟结核病再活化的多肽和多核苷酸(以及相关方法)。本发明还涉及包含所述多肽和多核苷酸的药物组合物和免疫原性组合物,以及用于诊断结核病(尤其潜伏性结核病)的方法。

[0002] 发明背景

[0003] 结核病(TB)为由结核分枝杆菌(*Mycobacterium tuberculosis*)和其它分枝杆菌感染引起的慢性传染病。其为发展中国家的主要疾病,以及世界发达地区日益增加的问题。据信超过20亿人感染了TB杆菌,并且每年约920万新TB病例和170万人死亡。10% TB杆菌感染者将形成活动性TB,每年每个患活动性TB的人传染平均10-15个其它人。虽然全球每年发病率已达高峰,但由于人口增长,死亡和病例数仍然上升(世界卫生组织Tuberculosis Facts 2008)。

[0004] 结核分枝杆菌通过呼吸途径感染个体。肺泡巨噬细胞吞噬该细菌,但其能通过抑制吞噬体与酸性溶酶体融合来存活和增殖。接着发生涉及CD4+和CD8+T细胞的复杂免疫应答,最终导致肉芽肿形成。结核分枝杆菌成功作为病原体的关键为以下事实:该隔离但未根除的细菌可长期存留,使得个体易于以后形成活动性TB。

[0005] 在感染后的最初几年,少于5%的感染个体形成活动性TB。肉芽肿可存留数十年并认为其包括处于休眠期、缺乏氧气和营养物的活结核分枝杆菌。然而,最近有人提出处于休眠期细菌的大多数位于全身分布的非巨噬细胞细胞类型中(Locht等, *Expert Opin. Biol. Ther.* 2007 7(11):1665-1677)。活动性TB的形成发生在当宿主天然免疫力与病原体之间的平衡改变时,例如由于抑制免疫事件(Anderson P *Trends in Microbiology* 2007 15(1):7-13; Ehlers S *Infection* 2009 37(2):87-95)。

[0006] 也已提出描述潜伏性TB与活动性TB之间的平衡的动态假说(Cardana P-J *Inflammation & Allergy-Drug Targets* 2006 6:27-39; Cardana P-J *Infection* 2009 37(2):80-86)。

[0007] 尽管感染可在相当长期内无症状,但是活动性疾病最通常表现为肺的急性炎症,导致疲劳、体重减轻、发热和持续咳嗽。如果不加以治疗,则通常导致严重并发症和死亡。

[0008] 通常可使用长期抗生素治疗来控制结核病,但这种治疗并不足以防止该疾病的蔓延。感染个体可为无症状的,但持续若干时间有传染性。此外,尽管治疗方案的顺应性为关键的,但是患者行为难以监测。某些患者并不完成疗程,这可导致无效治疗和产生耐药性。

[0009] 多重耐药TB(MDR-TB)为对一线药物无反应的一种形式。所有TB病例中的5%为MDR-TB,估计每年出现490,000个新MDR-TB病例。广泛耐药性TB(XDR-TB)当在MDR-TB基础上形成的对二线药物的耐受性时出现。据估计每年出现40,000个几乎不能治疗的XDR-TB新病例(世界卫生组织Tuberculosis Facts 2008)。

[0010] 即使完成抗生素治疗的整个疗程,结核分枝杆菌的感染也可能不能从感染个体中根除,并可遗留为可再活化的潜伏性感染。

[0011] 为了控制结核病的传播,该疾病的有效疫苗和准确早期诊断是最为重要的。

[0012] 通常使用结核菌素皮试实现对潜伏性TB感染的诊断,该测试涉及真皮内暴露至结核菌素纯蛋白衍生物(PPD)。抗原特异性T细胞应答导致在注射后48-72小时于注射部位产生可测量的硬结,这表示对分枝杆菌抗原的暴露。然而,灵敏度和特异性为该测试的一个问题,并且已接种BCG的个体并不能总是容易地与感染个体区分(鉴于BCG并不保护潜伏性感染这个事实,这尤其重要)。通常,已接受BCG但未被结核分枝杆菌感染的个体显示小于10mm直径的PPD反应,而具有大于10mm直径PPD反应的人被认为已受到结核分枝杆菌感染。然而,该规则并不适用于由于HIV感染而免疫抑制的个体,HIV感染可导致小于10mm直径的PPD反应);或其中由非结核分枝细菌感染的人可显示大于10mm直径PPD反应的地方流行国家。

[0013] 最近几年的进展已见证体外T细胞基测定法的开发,所述测定法基于干扰素- γ 释放并使用比PPD更特异于结核分枝杆菌的抗原,即ESAT-6和CFP-10。这些高特异性测试似乎是至少与结核菌素皮试同样灵敏,并且还显示较低的BCG接种所致的交叉反应性。参见Pai M等Expert Rev.Mol.Diagn.2006 6(3):413-422关于潜伏性TB诊断的最新综述。然而,由于ESAT-6/CFP-10为早期抗原,所以基于ESAT-6/CFP-10的测定法仅可在最近感染的人中表现最佳。因此,特异关联潜伏性结核病的新抗原的鉴定可有助于开发更灵敏的测定法,其可检测长期潜伏性感染。

[0014] 仍对治疗和预防结核病、尤其治疗和预防潜伏性TB以及预防TB的再活化的有效策略存在需要。

[0015] 发明简述

[0016] 本发明主要涉及Rv3616c作为潜伏性TB抗原的鉴定,在预防和治疗潜伏性TB以及预防或延迟TB再活化中的相关方法和用途。

[0017] 本发明提供用作治疗或预防潜伏性TB的药物的多肽,其包含:

[0018] (i)Rv3616c蛋白序列;

[0019] (ii)Rv3616c蛋白序列的变体;或者

[0020] (iii)Rv3616c蛋白序列的免疫原性片段。

[0021] 本发明另一个方面涉及用于治疗或预防潜伏性TB的方法,其包括给予安全且有效量的以下多肽至有需要的受试者,所述多肽包含:

[0022] (i)Rv3616c蛋白序列;

[0023] (ii)Rv3616c蛋白序列的变体;或者

[0024] (iii)Rv3616c蛋白序列的免疫原性片段;

[0025] 其中所述多肽诱导免疫应答,尤其针对结核分枝杆菌的免疫应答。

[0026] 以下多肽在制备用于治疗或预防潜伏性TB的药物中的用途表示本发明的另一个方面,所述多肽包含:

[0027] (i)Rv3616c蛋白序列;

[0028] (ii)Rv3616c蛋白序列的变体;或者

[0029] (iii)Rv3616c蛋白序列的免疫原性片段。

[0030] 还提供用作治疗或预防潜伏性TB药物的多核苷酸,其包含编码以下多肽的核酸序列,所述多肽包含:

[0031] (i)Rv3616c蛋白序列;

[0032] (ii)Rv3616c蛋白序列的变体;或者

[0033] (iii)Rv3616c蛋白序列的免疫原性片段。

[0034] 本发明另一个方面涉及用于治疗或预防潜伏性TB的方法,其包括给予安全且有效量的以下多核苷酸至有需要的受试者,所述多核苷酸包含编码以下多肽的核酸序列,所述多肽包含:

[0035] (i)Rv3616c蛋白序列;

[0036] (ii)Rv3616c蛋白序列的变体;或者

[0037] (iii)Rv3616c蛋白序列的免疫原性片段;

[0038] 其中所述多核苷酸诱导免疫应答,尤其针对结核分枝杆菌的免疫应答。

[0039] 以下多核苷酸在制备用于治疗或预防潜伏性TB的药物中的用途表示本发明的另一个方面,所述多核苷酸包含编码以下多肽的核酸序列,所述多肽包含:

[0040] (i)Rv3616c蛋白序列;

[0041] (ii)Rv3616c蛋白序列的变体;或者

[0042] (iii)Rv3616c蛋白序列的免疫原性片段。

[0043] 此外,提供用作治疗或预防潜伏性TB的药物的药物组合物,其包含:

[0044] (a)以下多肽,其包含:

[0045] (i)Rv3616c蛋白序列;

[0046] (ii)Rv3616c蛋白序列的变体;或

[0047] (iii)Rv3616c蛋白序列的免疫原性片段;或者

[0048] (b)包含编码(a)多肽的核酸序列的多核苷酸;

[0049] 以及

[0050] (c)药学上可接受的载体或赋形剂。

[0051] 此外,提供用作治疗或预防潜伏性TB的药物的免疫原性组合物,其包含:

[0052] (a)以下多肽,其包含:

[0053] (i)Rv3616c蛋白序列;

[0054] (ii)Rv3616c蛋白序列的变体;或

[0055] (iii)Rv3616c蛋白序列的免疫原性片段;或者

[0056] (b)包含编码(a)多肽的核酸序列的多核苷酸;

[0057] 以及

[0058] (c)非特异性免疫应答增强剂。

[0059] 另外提供用于诊断潜伏性TB的特异结合至以下多肽的抗体或其片段,所述多肽包含:

[0060] (i)Rv3616c蛋白序列;

[0061] (ii)Rv3616c蛋白序列的变体;或者

[0062] (iii)Rv3616c蛋白序列的免疫原性片段。

[0063] (例如用于诊断潜伏性结核病的方法,其包括测定来自测试受试者的生物学样品中的特异结合至本发明多肽的抗体或其片段的存在。)

[0064] 还提供用于治疗潜伏性TB的方法,其包括以下步骤:

[0065] (i)鉴定具有潜伏性TB感染的受试者(例如通过PPD或T细胞基测定法);以及

[0066] (ii)给予所述受试者安全且有效量的如本文所述的多肽或多核苷酸(例如以药物

组合物或免疫原性组合物形式)。

[0067] 还提供本发明多肽在制备用于鉴定测试受试者内的潜伏性TB的诊断试剂盒中的用途。

[0068] 在一个实施方案中,接受多肽、多核苷酸或组合物的受试者可具有活动性结核病(例如结核分枝杆菌的活动性感染)。在第二个实施方案中,受试者可具有潜伏性结核病(例如结核分枝杆菌的休眠性感染)。在第三个实施方案中,受试者可无结核病(例如没有结核分枝杆菌感染)。

[0069] 接受多肽、多核苷酸或组合物的受试者先前可已接受结核病接种(例如针对结核分枝杆菌感染接种),例如已用卡介苗(*Bacillus Calmette-Guerin*, BCG)接种。或者,接受本发明多肽、多核苷酸或组合物的受试者先前可没有接受0结核病接种(例如未针对结核分枝杆菌感染接种),例如未用卡介苗(BCG)接种。

[0070] 附图简述

[0071] 图1:Rv3616c肽全长序列比对。

[0072] 图2:对Rv3616c肽的PBMC反应。

[0073] 图3:在第21天(即第二次免疫接种(immunisation)7天后),来自表达IFN- γ 和/或IL-2和/或TNF- α 细胞因子的免疫CB6F1小鼠的CD4和CD8细胞百分比。

[0074] 图4:在免疫CB6F1小鼠中抗原特异性CD4反应在第21天(即第二次免疫接种7天后)的细胞因子概况。

[0075] 图5:在免疫CB6F1小鼠中抗原特异性CD8反应在第21天(即第二次免疫接种7天后)的细胞因子概况。

[0076] 图6:在第35天(即第三次免疫接种7天后),来自表达IFN- γ 和/或IL-2和/或TNF- α 细胞因子的免疫CB6F1小鼠的CD4和CD8细胞百分比。

[0077] 图7:在免疫CB6F1小鼠中抗原特异性CD4反应在第35天(即第三次免疫接种7天后)的细胞因子概况。

[0078] 图8:在免疫CB6F1小鼠中抗原特异性CD8反应在第35天(即第三次免疫接种7天后)的细胞因子概况。

[0079] 图9:在第21天(即第二次免疫接种7天后),来自表达IFN- γ 和/或IL-2和/或TNF- α 细胞因子的免疫C57BL/6小鼠的CD4和CD8细胞百分比。

[0080] 图10:在免疫C57BL/6小鼠中抗原特异性CD4反应在第21天(即第二次免疫接种7天后)的细胞因子概况。

[0081] 图11:在第35天(即第三次免疫接种7天后),来自表达IFN- γ 和/或IL-2和/或TNF- α 细胞因子的免疫C57BL/6小鼠的CD4和CD8细胞百分比。

[0082] 图12:在免疫C57BL/6小鼠中抗原特异性CD4反应在35天(即第三次免疫接种7天后)的细胞因子概况。

[0083] 图13:在免疫C57BL/6小鼠中抗原特异性CD8反应在35天(即第三次免疫接种7天后)的细胞因子概况。

[0084] 图14:无感染(naive)和潜伏性感染的人中的抗原特异性CD4 T细胞应答。

[0085] 所列序列简述

[0086] SEQ ID No:1:来自结核分枝杆菌H37Rv菌株的Rv3616c的多肽序列。

- [0087] SEQ ID No:2:来自结核分枝杆菌H37Rv菌株的Rv3616c的多核苷酸序列。
- [0088] SEQ ID No:3:来自结核分枝杆菌CDC1551菌株的Rv3616c的多肽序列。
- [0089] SEQ ID No:4:来自结核分枝杆菌F11菌株的Rv3616c的多肽序列。
- [0090] SEQ ID No:5:来自结核分枝杆菌Haarlem A菌株的Rv3616c的多肽序列。
- [0091] SEQ ID No:6:来自结核分枝杆菌C菌株的Rv3616c的多肽序列。
- [0092] SEQ ID No:7:来自BCG的Rv3616c的多肽序列。
- [0093] SEQ ID No:8:Mtb8.4的多肽序列。
- [0094] SEQ ID No:9:Mtb9.8的多肽序列。
- [0095] SEQ ID No:10:Mtb9.9的多肽序列。
- [0096] SEQ ID No:11:Ra12的多肽序列。
- [0097] SEQ ID No:12:Ra35的多肽序列。
- [0098] SEQ ID No:13:TbH9的多肽序列。
- [0099] SEQ ID No:14:Mtb41的多肽序列。
- [0100] SEQ ID No:15:ESAT-6的多肽序列。
- [0101] SEQ ID No:16:Ag85A的多肽序列。
- [0102] SEQ ID No:17:Ag85B的多肽序列。
- [0103] SEQ ID No:18: α -晶体蛋白的多肽序列。
- [0104] SEQ ID No:19:MPT64的多肽序列。
- [0105] SEQ ID No:20:Mtb32A的多肽序列。
- [0106] SEQ ID No:21:Ser/Ala突变的成熟Mtb32A的多肽序列。
- [0107] SEQ ID No:22:TB10.4的多肽序列。
- [0108] SEQ ID No:23:Mtb72f的多肽序列。
- [0109] SEQ ID No:24:M72的多肽序列。
- [0110] SEQ ID No:25:Mtb71f的多肽序列。
- [0111] SEQ ID No:26:M92融合体的多肽序列。
- [0112] SEQ ID No:27:M103融合体的多肽序列。
- [0113] SEQ ID No:28:M114融合体的多肽序列。
- [0114] SEQ ID No:29:推定的人CD4细胞表位1。
- [0115] SEQ ID No:30:推定的人CD4细胞表位2。
- [0116] SEQ ID No:31:推定的人CD4细胞表位3。
- [0117] SEQ ID No:32:推定的人CD4细胞表位4。
- [0118] SEQ ID No:33:推定的人CD4细胞表位5。
- [0119] SEQ ID No:34:推定的人CD4细胞表位6。
- [0120] SEQ ID No:35:推定的人CD4细胞表位7。
- [0121] SEQ ID No:36:推定的人CD4细胞表位8。
- [0122] SEQ ID No:37:推定的人CD4细胞表位9。
- [0123] SEQ ID No:38:推定的人CD4细胞表位10。
- [0124] SEQ ID No:39:推定的人CD4细胞表位11。
- [0125] SEQ ID No:40:推定的人CD4细胞表位12。

- [0126] SEQ ID No:41:推定的人CD4细胞表位13。
- [0127] SEQ ID No:42:推定的人CD4细胞表位14。
- [0128] SEQ ID No:43:推定的人CD4细胞表位15。
- [0129] SEQ ID No:44:推定的人CD4细胞表位16。
- [0130] SEQ ID No:45:推定的人CD4细胞表位17。
- [0131] SEQ ID No:46:推定的人CD4细胞表位18。
- [0132] SEQ ID No:47:推定的人CD4细胞表位19。
- [0133] SEQ ID No:48:推定的人CD8细胞表位1。
- [0134] SEQ ID No:49:推定的人CD8细胞表位2。
- [0135] SEQ ID No:50:推定的人CD8细胞表位3。
- [0136] SEQ ID No:51:推定的人CD8细胞表位4。
- [0137] SEQ ID No:52:推定的人CD8细胞表位5。
- [0138] SEQ ID No:53:推定的人CD8细胞表位6。
- [0139] SEQ ID No:54:推定的人CD8细胞表位7。
- [0140] SEQ ID No:55:推定的人CD8细胞表位8。
- [0141] SEQ ID No:56:推定的人CD8细胞表位9。
- [0142] SEQ ID No:57:推定的人CD8细胞表位10。
- [0143] SEQ ID No:58:推定的人CD8细胞表位11。
- [0144] SEQ ID No:59:推定的人CD8细胞表位12。
- [0145] SEQ ID No:60:推定的人CD8细胞表位13。
- [0146] SEQ ID No:61:推定的人CD8细胞表位14。
- [0147] SEQ ID No:62:推定的人CD8细胞表位15。
- [0148] SEQ ID No:63:推定的人CD8细胞表位16。
- [0149] SEQ ID No:64:推定的人CD8细胞表位17。
- [0150] SEQ ID No:65:推定的人CD8细胞表位18。
- [0151] SEQ ID No:66:推定的人CD8细胞表位19。
- [0152] SEQ ID No:67:推定的人CD8细胞表位20。
- [0153] SEQ ID No:68:推定的人CD8细胞表位21。
- [0154] SEQ ID No:69:推定的人CD8细胞表位22。
- [0155] SEQ ID No:70:推定的人CD8细胞表位23。
- [0156] SEQ ID No:71:推定的人CD8细胞表位24。
- [0157] SEQ ID No:72:推定的人CD8细胞表位25。
- [0158] SEQ ID No:73:推定的人CD8细胞表位26。
- [0159] SEQ ID No:74:推定的人CD8细胞表位27。
- [0160] SEQ ID No:75:推定的人CD8细胞表位28。
- [0161] SEQ ID No:76:推定的人CD8细胞表位29。
- [0162] SEQ ID No:77:推定的人CD8细胞表位30。
- [0163] SEQ ID No:78:推定的人CD8细胞表位31。
- [0164] SEQ ID No:79:推定的人CD8细胞表位32。

- [0165] SEQ ID No:80:推定的人CD8细胞表位33。
- [0166] SEQ ID No:81:推定的人CD8细胞表位34。
- [0167] SEQ ID No:82:推定的人CD8细胞表位35。
- [0168] SEQ ID No:83:推定的人CD8细胞表位36。
- [0169] SEQ ID No:84:推定的人CD8细胞表位37。
- [0170] SEQ ID No:85:推定的人CD8细胞表位38。
- [0171] SEQ ID No:86:推定的人CD8细胞表位39。
- [0172] SEQ ID No:87:推定的人CD8细胞表位40。
- [0173] SEQ ID No:88:推定的人CD8细胞表位41。
- [0174] SEQ ID No:89:推定的人CD8细胞表位42。
- [0175] SEQ ID No:90:推定的人CD8细胞表位43。
- [0176] SEQ ID No:91:推定的人CD8细胞表位44。
- [0177] SEQ ID No:92:推定的人CD8细胞表位45。
- [0178] SEQ ID No:93:推定的人CD8细胞表位46。
- [0179] SEQ ID No:94:推定的人CD8细胞表位47。
- [0180] SEQ ID No:95:推定的人CD8细胞表位48。
- [0181] SEQ ID No:96:推定的人CD8细胞表位49。
- [0182] SEQ ID No:97:推定的人CD8细胞表位50。
- [0183] SEQ ID No:98:推定的人CD8细胞表位51。
- [0184] SEQ ID No:99:推定的人CD8细胞表位52。
- [0185] SEQ ID No:100:推定的人CD8细胞表位53。
- [0186] SEQ ID No:101:推定的人CD8细胞表位54。
- [0187] SEQ ID No:102:推定的人CD8细胞表位55。
- [0188] SEQ ID No:103:推定的人CD8细胞表位56。
- [0189] SEQ ID No:104:推定的人CD8细胞表位57。
- [0190] SEQ ID No:105:推定的人CD8细胞表位58。
- [0191] SEQ ID No:106:推定的人CD8细胞表位59。
- [0192] SEQ ID No:107:推定的人CD8细胞表位60。
- [0193] SEQ ID No:108:推定的人CD8细胞表位61。
- [0194] SEQ ID No:109:推定的人CD8细胞表位62。
- [0195] SEQ ID No:110:推定的人CD8细胞表位63。
- [0196] SEQ ID No:111:推定的人CD8细胞表位64。
- [0197] SEQ ID No:112:推定的人CD8细胞表位65。
- [0198] SEQ ID No:113:推定的人CD8细胞表位66。
- [0199] SEQ ID No:114:推定的人CD8细胞表位67。
- [0200] SEQ ID No:115:推定的人CD8细胞表位68。
- [0201] SEQ ID No:116:推定的人CD8细胞表位69。
- [0202] SEQ ID No:117:推定的人CD8细胞表位70。
- [0203] SEQ ID No:118:推定的人CD8细胞表位71。

- [0204] SEQ ID No:119:推定的人CD8细胞表位72。
- [0205] SEQ ID No:120:推定的人CD8细胞表位73。
- [0206] SEQ ID No:121:推定的人CD8细胞表位74。
- [0207] SEQ ID No:122:推定的人CD8细胞表位75。
- [0208] SEQ ID No:123:推定的人CD8细胞表位76。
- [0209] SEQ ID No:124:推定的人CD8细胞表位77。
- [0210] SEQ ID No:125:推定的人CD8细胞表位78。
- [0211] SEQ ID No:126:推定的人CD8细胞表位79。
- [0212] SEQ ID No:127:肽1。
- [0213] SEQ ID No:128:肽2。
- [0214] SEQ ID No:129:肽3。
- [0215] SEQ ID No:130:肽4。
- [0216] SEQ ID No:131:肽5。
- [0217] SEQ ID No:132:肽6。
- [0218] SEQ ID No:133:肽7。
- [0219] SEQ ID No:134:肽8。
- [0220] SEQ ID No:135:肽9。
- [0221] SEQ ID No:136:肽10。
- [0222] SEQ ID No:137:肽11。
- [0223] SEQ ID No:138:肽12。
- [0224] SEQ ID No:139:肽13。
- [0225] SEQ ID No:140:肽14。
- [0226] SEQ ID No:141:肽15。
- [0227] SEQ ID No:142:肽16。
- [0228] SEQ ID No:143:肽17。
- [0229] SEQ ID No:144:肽18。
- [0230] SEQ ID No:145:肽19。
- [0231] SEQ ID No:146:肽20。
- [0232] SEQ ID No:147:肽21。
- [0233] SEQ ID No:148:肽22。
- [0234] SEQ ID No:149:肽23。
- [0235] SEQ ID No:150:肽24。
- [0236] SEQ ID No:151:肽25。
- [0237] SEQ ID No:152:肽26。
- [0238] SEQ ID No:153:肽27。
- [0239] SEQ ID No:154:肽28。
- [0240] SEQ ID No:155:肽29。
- [0241] SEQ ID No:156:肽30。
- [0242] SEQ ID No:157:来自结核分枝杆菌H37Rv菌株的Rv1753c的多肽序列。

[0243] SEQ ID No:158:来自结核分枝杆菌H37Rv菌株的Rv2386c的多肽序列。

[0244] SEQ ID No:159:来自结核分枝杆菌H37Rv菌株的Rv2707c的多肽序列。

[0245] 发明详述

[0246] 目前,活细菌接种为用于诱导保护性免疫的最有效方法。用于此目的的最常用的分枝杆菌为卡介苗(BCG),牛分枝杆菌(*M.bovis*)的无毒株,其于60多年之前开发出。然而,BCG的安全性和有效性为争议的来源——尽管在儿童中针对严重疾病表现提供保护,但是BCG并不在成人生命中预防潜伏性TB的建立或肺部疾病的再活化。此外,某些国家例如美国并不用该制剂接种公众。

[0247] 几乎所有的目前在临床开发中的新一代TB疫苗已设计为暴露前疫苗(pre-exposure vaccine)。它们包括亚单位疫苗,其尤其有效于加强由先前BCG接种诱导的免疫;以及改进的活分枝杆菌疫苗,其旨在用更有效和/或更安全的菌株代替BCG。尽管这些疫苗旨在促进对感染的抗性,但是它们作为暴露后或治疗性疫苗在潜伏性TB情况中很可能较不有效(Lin MY等Endocrine, Metabolic & Immune Disorders-Drug Targets 2008;15-29)。

[0248] 在分枝杆菌感染早期期间强烈表达的数种蛋白在动物接种模型中已表现出提供强保护效能。然而,用感染初期强烈表达的抗原接种可能不能为处理感染晚期提供最优免疫应答。感染晚期的充分控制可能需要特异于在该时间表达的特定抗原的T细胞。

[0249] 直接靶向持久休眠细菌的暴露后疫苗(post-exposure vaccine)可能有助于针对TB再活化提供保护,从而增强TB控制或甚至能够清除该感染。因此,靶向潜伏性TB的疫苗可显著且经济地降低全球TB感染率。

[0250] 基于晚期抗原的亚单位疫苗也可与早期抗原联合使用以提供多阶段疫苗。或者,晚期抗原可用于补充和改进BCG接种(通过加强BCG反应或通过开发改进的重组BCG菌株)。

[0251] Rv3616c,也称为Mtb40或HTCC1,先前已显示涉及与结核病有关的免疫应答(参见,例如W098/53075)。Al-Attayah等Clin.Exp.Immunol.2004 138:139-144已表明Rv3616c被肺结核患者良好识别(通过PMBC增殖和IFN- γ 产生)。Mustafa等Infect.Immun.2006 74(8):4566-4572已研究牛分枝杆菌感染且BCG接种的牛对Rv3616c的识别。

[0252] 最近,已基于全基因组结核分枝杆菌基因组的生物信息学分析(Zvi等BMC Medical Genetics 2008 1:18)和基于在活动性和潜伏性感染个体中差异表达蛋白的测试(Schuck SD等PLoS ONE 2009 4(5):e5590)提出一系列结核分枝杆菌候选疫苗。

[0253] 虽然已表明巨噬细胞充当分枝杆菌免疫的主要效应物,但是T细胞为所述免疫的主导诱导物。T细胞针对结核病提供保护的关键作用通过在人免疫缺陷病毒感染的个体中CD4+T细胞的相关耗竭所致的TB再活化的速率增加来阐明。此外,已表明在对结核分枝杆菌的初次免疫应答最盛时获得的CD4+T细胞的过继转移赋予T细胞缺陷型小鼠针对结核分枝杆菌的保护(Orme等J.Exp.Med.1983 158:74-83)。

[0254] 已表明分枝杆菌反应性CD4+T细胞为 γ -干扰素(IFN- γ)的强有力的产生者,还已表明IFN- γ 引发小鼠中的巨噬细胞的抗-分枝杆菌作用(Flynn等J.Exp.Med.1993 178:2249-2254)。尽管IFN- γ 在人类中的作用较不明确,但是研究已表明1,25-二羟基-维生素D3单独或者与IFN- γ 或肿瘤坏死因子- α 联合,都活化人巨噬细胞抑制结核分枝杆菌感染。此外,已知IFN- γ 刺激人巨噬细胞产生1,25-二羟基-维生素D3。同样,已表明白介素-12

(IL-12)在刺激抗结核分枝杆菌感染中起作用。有关结核分枝杆菌感染免疫学的综述参见Chan和Kaufmann, *Tuberculosis: Pathogenesis, Protection and Control* (Bloom编著, 1994)、*Tuberculosis* (第2版, Rom和Garay编著, 2003)和Harrison的 *Principles of Internal Medicine*, 第150章, 第953-966页(第16版, Braunwald等编著, 2005)。

[0255] 本发明主要涉及Rv3616c作为与潜伏性TB关联的TB抗原的鉴定, 在预防和治疗潜伏性TB以及预防或延迟TB再活化中的相关方法和用途。

[0256] 因此, 本发明提供用于治疗或预防潜伏性TB的Rv3616c蛋白、其变体或其免疫原性片段或者编码所述蛋白、变体或片段的多核苷酸。合适的是, 所述用途可具体为治疗潜伏性TB。或者, 所述用途可为预防或延迟TB再活化(尤其是延迟TB再活化, 例如延迟数月、数年或甚至无限期地延迟)。

[0257] 术语“结核分枝杆菌复合群菌种”包括传统认为引起疾病结核病的分枝杆菌以及在免疫缺陷患者例如AIDS患者中引起结核病和肺疾病的环境和机会性分枝杆菌, 例如结核分枝杆菌、牛分枝杆菌或非洲分枝杆菌(*M. africanum*)、BCG、鸟分枝杆菌(*M. avium*)、胞内分枝杆菌(*M. intracellulare*)、隐藏分枝杆菌(*M. celatum*)、日内瓦分枝杆菌(*M. genavense*)、嗜血分枝杆菌(*M. haemophilum*)、堪萨斯分枝杆菌(*M. kansasii*)、猿分枝杆菌(*M. simiae*)、母牛分枝杆菌(*M. vaccae*)、偶发分枝杆菌(*M. fortuitum*)和瘰癧分枝杆菌(*M. scrofulaceum*)(参见例如Harrison的 *Principles of Internal Medicine*, 第150章, 第953-966页(第16版, Braunwald等编著, 2005)。本发明尤其针对结核分枝杆菌感染。

[0258] 术语“活动性感染”是指表现出疾病症状和/或损伤(与所表现出的疾病症状相称的)的感染(例如结核分枝杆菌感染)。

[0259] 术语“非活动性感染”“休眠性感染”或“潜伏性感染”是指未表现出疾病症状和/或损伤(与未表现出疾病症状相称的)的感染(例如结核分枝杆菌感染)。

[0260] 术语“原发性结核病”是指直接在感染(例如结核分枝杆菌感染)后的临床疾病(疾病症状的表现)。参见Harrison的 *Principles of Internal Medicine*, 第150章, 第953-966页(第16版, Braunwald等编著, 2005)。

[0261] 术语“继发性结核病”或“原发后结核病(*postprimary tuberculosis*)”是指休眠性、非活动性或潜伏性感染(例如结核分枝杆菌感染)的再活化。参见Harrison的 *Principles of Internal Medicine*, 第150章, 第953-966页(第16版, Braunwald等编著, 2005)。

[0262] 术语“结核病再活化(*tuberculosis reactivation*)”是指感染(例如在结核菌素皮试中, 合适地在体外基于T细胞的测定中)测试呈阳性、但没有明显疾病症状的个体在以后表现出疾病症状。阳性诊断测试表明所述个体已被感染, 然而, 所述个体先前可表现出或未表现出活动性疾病症状, 所述症状已被充分治疗以使结核病进入非活动性或潜伏性状态。应认识到的是, 用于预防、延迟或治疗结核病再活化的方法可在表现疾病活动性症状的个体中开始。

[0263] 术语“耐药性”结核病是指以下感染(例如结核分枝杆菌感染): 其中一种或多种有效治疗结核病的所谓“前线”化疗药(例如异烟肼、利福平、乙胺丁醇、链霉素和吡嗪酰胺)不能抑制或杀伤感染菌株(即对所述化疗药有抗性)。

[0264] 术语“多重耐药性”结核病是指以下感染(例如结核分枝杆菌感染): 其中感染菌株

对两种以上有效治疗结核病的“前线”化疗药具有抗性。

[0265] “化疗药”是指本领域已知并用于治疗结核病(例如结核分枝杆菌感染)的药物。用于治疗结核病的例示性药物包括但不限于阿米卡星、氨基水杨酸、卷曲霉素、环丝氨酸、乙胺丁醇、乙硫异烟胺、异烟肼、卡那霉素、吡嗪酰胺、利福霉素(即利福平、利福喷汀和利福布汀)、链霉素、氧氟沙星、环丙沙星、克拉霉素、阿奇霉素和氟喹诺酮类。用于治疗非耐药性结核病的“一线”或“前线”化疗药包括异烟肼、利福平、乙胺丁醇、链霉素和吡嗪酰胺。用于治疗已显示对一种或多种“一线”药物具有耐药性的结核病的“二线”化疗药包括氧氟沙星、环丙沙星、乙硫异烟胺、氨基水杨酸、环丝氨酸、阿米卡星、卡那霉素和卷曲霉素。这类药物综述于Goodman和Gilman的The Pharmacological Basis of Therapeutics,第48章,Hardman和Limbird编著,2001。

[0266] 术语“多肽”、“肽”和“蛋白”在本文中可互换使用,是指氨基酸残基的聚合体。该术语还适用于其中一个或多个氨基酸残基是相应天然存在氨基酸的人工化学模拟物的氨基酸聚合体,以及适用于天然存在氨基酸聚合体和非天然存在氨基酸聚合体。合适的本发明多肽仅由天然存在氨基酸残基、尤其是由遗传密码编码的氨基酸组成。

[0267] 术语“氨基酸”是指天然存在和合成的氨基酸,以及以与天然存在氨基酸类似的方式起作用的氨基酸类似物和氨基酸模拟物。天然存在氨基酸为由遗传密码编码的氨基酸,以及随后被修饰的氨基酸,例如羟脯氨酸、 γ -羧基谷氨酸和O-磷酸丝氨酸。氨基酸类似物是指与天然存在氨基酸具有相同基本化学结构的化合物,即与氢、羧基、氨基和R基团连接的 α 碳,例如高丝氨酸、正亮氨酸、甲硫氨酸亚砷(methionine sulfoxide)、甲硫氨酸甲基砷(methionine methyl sulfonium)。这类类似物具有修饰R基团(例如正亮氨酸)或修饰肽主链,但保留与天然存在氨基酸相同的基本化学结构。氨基酸模拟物是指具有与氨基酸一般化学结构不同的结构、但以与天然存在氨基酸类似的方式起作用的化合物。合适的氨基酸为天然存在氨基酸或氨基酸类似物,尤其是天然存在氨基酸和特别是由遗传密码编码的氨基酸。

[0268] “核酸”是指单链或双链形式的脱氧核糖核苷酸或核糖核苷酸及其聚合体。该术语包括含有已知核苷酸类似物或修饰主链残基或键的核酸,其为合成的、天然存在的和非天然存在的,其具有与参比核酸相似的结合特性,而且其以与参比核苷酸相似的方式代谢。这类类似物的实例包括而限于硫代磷酸酯、氨基磷酸酯、甲基膦酸酯、手性-甲基膦酸酯、2'-O-甲基核糖核苷酸、肽-核酸(PNA)。合适的是,术语“核酸”是指天然存在的脱氧核糖核苷酸或核糖核苷酸及其聚合体。

[0269] 除非另有说明,否则特定核酸序列也隐含包括其保守修饰变体(例如简并密码子取代物)和互补序列,以及所明确指出的序列(合适的是,其是指明确指出的序列)。具体而言,可通过产生其中一个或多个所选(或全部)密码子的第3位被混合碱基和/或脱氧肌苷残基取代的序列,得到简并密码子取代物(Batzer等,Nucleic Acid Res.19:5081(1991); Ohtsuka等,J.Biol.Chem.260:2605-2608(1985);Rossolini等,Mol.Cell.Probes 8:91-98(1994))。术语核酸与基因、cDNA、mRNA、寡核苷酸和多核苷酸可互换使用。

[0270] 氨基酸在本文中可用它们公知的三字母符号表示,或者可用IUPAC-IUB生物化学命名委员会(Biochemical Nomenclature Commission)所推荐的一字母符号来表示。同样,核苷酸也可用它们公知的单字母码来表示。

[0271] 本文所用术语“Rv3616c蛋白序列”意指SEQ ID No:1提供的多肽序列或其来自以下分枝杆菌的同源物,所述分枝杆菌为结核分枝杆菌复合群菌种例如结核分枝杆菌、牛分枝杆菌或非洲分枝杆菌,或者为环境或机会性的并在免疫缺损宿主(例如AIDS患者)中引起机会性感染例如肺部感染的分枝杆菌例如BCG、鸟分枝杆菌、胞内分枝杆菌、隐藏分枝杆菌、日内瓦分枝杆菌、嗜血分枝杆菌、堪萨斯分枝杆菌、猿分枝杆菌、母牛分枝杆菌、偶发分枝杆菌和瘰癧分枝杆菌(参见,例如Harrison的Principles of Internal Medicine,第150章,第953-966页(第16版,Braunwald等编著,2005)。

[0272] 为在接种宿主之间确保高有效率(efficacy rate),疫苗的组分应在临床上重要的菌株间充分保守。合适的是,Rv3616c蛋白来源于结核分枝杆菌H37Rv(即SEQ ID No:1提供的多肽序列)或其来自另一种结核分枝杆菌菌株(例如CDC1551、F11、Haarlem A和C菌株)的同源物。具有耐药性(例如MDR或尤其XDR)的结核分枝杆菌菌株为Rv3616c蛋白序列的特别有价值的基础。引人关注的菌株包括:

[0273] CDC1551-可传播且有毒力的菌株

[0274] Haarlem家族(例如Haarlem A)-在密集人群中发现的耐药性菌株。结核分枝杆菌菌株的Haarlem家族成员已在世界许多地方发现。该家族的第一个代表在荷兰Haarlem发现。

[0275] KZN4207-来自南非KwaZulu-Natal患者的药物敏感性分离群

[0276] KZN1435-来自南非KwaZulu-Natal患者的多重耐药性(MDR)分离群

[0277] KZN605-来自南非KwaZulu-Natal患者的广泛耐药性(XDR)分离群

[0278] C-在纽约市高度传播。在一个研究中发现该菌株在注射药物使用者中更为常见且对活性氮中间体具有抗性(Friedman等J.Infect.Dis.1997 176(2):478-84)

[0279] 94_M4241A-在1994年于旧金山分离自出生于中国的患者。先前通过基因组缺失分析来分析该菌株(Gagneux等,PNAS 2006 103(8):2869-2873)。

[0280] 02_1987-在2002年于旧金山分离自出生于韩国的患者。先前通过基因组缺失分析来分析该菌株(Gagneux等,PNAS 2006 103(8):2869-2873)。

[0281] T92-在1999年于旧金山分离自出生于菲律宾的患者。该菌株公布于Hirsh等PNAS 2004 101:4871-4876)。

[0282] T85-在1998年于旧金山分离自出生于中国的患者。该菌株公布于Hirsh等PNAS 2004 101:4871-4876)。

[0283] EAS054-在1993年于旧金山分离自出生于印度的患者。先前通过基因组缺失分析来分析该菌株(Gagneux等.,PNAS 2006 103(8):2869-2873)。

[0284] Gagneux等,PNAS 2006 103(8):2869-2873和Herbert等Infect.Immun.2007 75(12):5798-5805对已知存在的结核分枝杆菌菌株的范围提供了有价值的背景。

[0285] 最合适的是,Rv3616c蛋白选自SEQ ID No:1和3-7、尤其是SEQ ID No:1和3-6(例如SEQ ID No:1)提供的多肽序列。

[0286] 特别引人关注的多核苷酸为包含以下序列(例如由其组成)的多核苷酸,所述序列编码:

[0287] (i)Rv3616c蛋白序列;

[0288] (ii)Rv3616c蛋白序列的变体;或

[0289] (iii)Rv3616c蛋白序列的免疫原性片段。

[0290] 多核苷酸适宜包含编码Rv3616c蛋白的免疫原性片段的SEQ ID NO:2变体或SEQ ID NO:2片段(例如由其组成)。

[0291] 组合

[0292] 用于本发明中的Rv3616c相关多肽还可包含设计以增强其免疫原性或在其它方面改进这些抗原的其它组分。例如,可通过向抗原的一端添加一段组氨酸残基(公知为his-标记)以有利于多肽抗原的分离改进。

[0293] 术语“his-标记”是指一串插入参比序列内的组氨酸残基,通常6个残基。为使参比序列活性的破坏最小化,his-标记通常在N端插入,通常紧接起始甲硫氨酸残基之后,或者在C端插入。它们通常与天然序列异源,但由于它们通过提高蛋白与固定化金属亲和层析树脂(IMAC)的结合来促进分离而将其插入。一般而言,从引起针对参比蛋白的所需免疫应答的角度来看,his-标记存在与否是不重要的。然而,为避免针对his-标记本身的不利反应的风险,据认为最好减少his-标记的长度例如至4个以下的残基,尤其是2个残基(或者完全不使用his-标记)。

[0294] 为提高所引发的免疫应答的强度和/或幅度,组合物、多肽和核酸可包含多拷贝的本发明抗原和/或来自分枝杆菌(尤其是结核分枝杆菌)的其它异源多肽(或其编码多核苷酸)。

[0295] 本领域技术人员认识到的是,当联合使用多种组分时,确切存在形式可变化。例如,Rv3616c组分和其它拷贝的抗原或其它异源抗原组分可存在为:

[0296] (1)两种独立的多肽组分;

[0297] (2)包含两种多肽组分的融合蛋白;

[0298] (3)一种多肽组分和一种多核苷酸组分;

[0299] (4)两种独立的多核苷酸组分;

[0300] (5)编码两种独立多肽组分的单个多核苷酸;或者

[0301] (6)编码包含两种多肽组分的融合蛋白的单个多核苷酸。

[0302] 该灵活性同样适用于联合使用三种或更多种组分的情形。然而,为了方便,理想的是,通常当存在许多组分时将它们包含在单个融合蛋白或编码单个融合蛋白的多核苷酸内。在本发明的一个实施方案中,所有抗原组分作为多肽提供(例如在单个融合蛋白内)。在本发明的备选实施方案中,所有抗原组分作为多核苷酸提供(例如单个多核苷酸,例如编码单个融合蛋白的单个多核苷酸)。

[0303] 术语“异源的”当用于提及核酸的部分时,是指所述核酸包含彼此在自然界不存在相同关系的两个或更多个亚序列。例如,核酸通常是重组产生的、具有两个或更多个来自排列以得到新功能核酸的无关基因的序列,例如一个来源的启动子与另一来源的编码区。同样,异源蛋白是指所述蛋白包含彼此在自然界不存在相同关系的两个或更多个亚序列(例如融合蛋白)。

[0304] “融合多肽”或“融合蛋白”是指具有至少两个直接或通过氨基酸接头共价连接的异源多肽(例如至少两个分枝杆菌多肽)的蛋白。形成融合蛋白的多肽通常是C端与N端连接,但它们也可以是C端与C端、N端与N端、或N端与C端连接。融合蛋白的多肽可呈任何顺序。该术语也指构成融合蛋白的抗原的保守修饰变体、多态变体、等位基因、突变体、免疫原性

片段和种间同源物。结核分枝杆菌抗原描述于Cole等,Nature 393:537(1998)中,该文献公开了完整的结核分枝杆菌基因组。可例如使用本文所述的序列比较算法或本领域技术人员已知的其它方法(例如杂交测定法和抗体结合测定法),来鉴定来自其它分枝杆菌的、对应于结核分枝杆菌抗原的抗原。

[0305] 术语“融合的”是指融合蛋白内的两个多肽间的共价连接。所述多肽通常通过肽键直接相互连接或通过氨基酸接头连接。任选的是,所述肽可通过本领域技术人员已知的非肽共价键连接。

[0306] 可与Rv3616c组合的例示性结核分枝杆菌抗原包括以下抗原中的一种或多种(例如1-5种,例如1-3种,尤其是1种)(例如(i)至(xi)中的一种或多种):

[0307] (i)Mtb8.4(也称为DPV和Rv1174c),其多肽序列描述于W097/09428的SEQ ID No:102(cDNA描述于SEQ ID No:101中)和Coler等Journal of Immunology 1998 161:2356-2364中。尤其引人关注的是成熟Mtb8.4序列,其不存在前导信号肽(即来自W097/09428 SEQ ID No:102的氨基酸残基15-96)。Mtb8.4的全长多肽序列示于SEQ ID No:8;

[0308] (ii)Mtb9.8(也称为MSL和Rv0287),其多肽序列描述于W098/53075的SEQ ID No:109(MSL片段公开于W098/53075的SEQ ID No:110-124,SEQ ID No:119和120尤其引人关注)以及Coler等Vaccine 2009 27:223-233(具体地,活性片段示于其中的图2中)中。Mtb9.8的全长多肽序列示于SEQ ID No:9;

[0309] (iii)Mtb9.9(也称为Mtb9.9A、MTI、MTI-A和Rv1793),其多肽序列描述于W098/53075的SEQ ID No:19和Alderson等Journal of Experimental Medicine 2000 7:551-559(MTI片段公开于W098/53075的SEQ ID No:17和51-66,SEQ ID No:17、51、52、53、56和62-65尤其引人关注)中。许多MTI多肽变体描述于W098/53075 SEQ ID No:21、23、25、27、29和31以及Alderson等Journal of Experimental Medicine 2000 7:551-559中。Mtb9.9的全长多肽序列示于SEQ ID No:10;

[0310] (iv)Ra12(也称为Mtb32A C端抗原),其多肽序列描述于W001/98460的SEQ ID No:10和Skeiky等Journal of Immunology 2004 172:7618-7682中。Ra12的全长多肽序列示于SEQ ID No:11;

[0311] (v)Ra35(也称为Mtb32A N端抗原),其多肽序列描述于W001/98460的SEQ ID No:8和Skeiky等Journal of Immunology 2004 172:7618-7682中。Ra35的全长多肽序列示于SEQ ID No:12;

[0312] (vi)TbH9(也称为Mtb39、Mtb39A、TbH9FL和Rv1196),其多肽序列描述于W097/09428的SEQ ID No:107以及Dillon等Infection and Immunity 1999 67(6):2941-2950和Skeiky等Journal of Immunology 2004 172:7618-7682中。TbH9的全长多肽序列示于SEQ ID No:13;

[0313] (vii)Mtb41(也称为MTCC2和Rv0915c),其多肽序列描述于W098/53075的SEQ ID No:142(cDNA描述于SEQ ID No:140中)和Skeiky等Journal of Immunology 2000 165:7140-7149。Mtb41的全长多肽序列示于SEQ ID No:14;

[0314] (viii)ESAT-6(也称为esxA和Rv3875),其多肽序列描述于W097/09428的SEQ ID No:103(cDNA描述于SEQ ID No:104中)和Sorensen等Infection and Immunity 1995 63(5):1710-1717中。ESAT-6的全长多肽序列示于SEQ ID No:15;

[0315] (ix)Ag85复合抗原(例如Ag85A,也称为fbpA和Rv3804c;或Ag85B,也称为fbpB和Rv1886c),其论述于例如Content等Infection and Immunity 1991 59:3205-3212和Huygen等Nature Medicine 1996 2(8):893-898中。Ag85A的全长多肽序列示于SEQ ID No:16(残基43-338(即缺少信号肽)的成熟蛋白尤其引人关注)。Ag85B的全长多肽序列示于SEQ ID No:17(残基41-325(即缺少信号肽)的成熟蛋白尤其引人关注);

[0316] (x) α -晶体蛋白(也称为hspX和Rv2031c),其描述于Verbon等Journal of Bacteriology 1992 174:1352-1359和Friscia等Clinical and Experimental Immunology 1995 102:53-57(特别引人关注的为对应于残基71-91、21-40、91-110和111-130的片段)中。 α -晶体蛋白的全长多肽序列示于SEQ ID No:18;

[0317] (xi)Mpt64(也称为Rv1980c),其描述于Roche等Scandinavian Journal of Immunology 1996 43:662-670中。MPT64的全长多肽序列示于SEQ ID No:19(残基24-228(即缺少信号肽)的成熟蛋白尤其引人关注);

[0318] (xii)Mtb32A,其多肽序列描述于W001/98460的SEQ ID No:2(全长)和SEQ ID No:4(成熟)的残基8-330中,尤其是具有至少一个催化三联体突变(例如催化丝氨酸残基,其可例如突变为丙氨酸)的变体。Mtb32A的全长多肽序列示于SEQ ID No:20。具有Ser/Ala突变的Mtb32A的成熟形式示于SEQ ID No:21;

[0319] (xiii)TB10.4, TB10.4的全长多肽序列示于SEQ ID No:22;

[0320] (xiv)Rv1753c,来自结核分枝杆菌H37Rv的Rv1753c的全长多肽序列示于SEQ ID No:157;

[0321] (xv)Rv2386c,来自结核分枝杆菌H37Rv的Rv2386c的全长多肽序列示于SEQ ID No:158;和/或

[0322] (xvi)Rv2707c,来自结核分枝杆菌H37Rv的Rv2707c的全长多肽序列示于SEQ ID No:159。

[0323] 或它们的组合,例如(例如(a)-(g)的组合):

[0324] (a)Ra12、TbH9和Ra35组分的例如以融合蛋白(例如Mtb72f)形式的组合。Mtb72f的多肽序列描述于W02006/117240的SEQ ID No:6(cDNA描述于SEQ ID No:5中)和Skeiky等Journal of Immunology 2004 172:7618-7682(其中其掺入了任选的His-标记以有助于纯化,当用于本发明时,合适的是Mtb72f不存在所述任选的组氨酸残基)中。Mtb72f的多肽序列示于SEQ ID No:23;

[0325] (b)Ra12、TbH9和Ser/Ala突变Ra35(即其中催化丝氨酸残基已被丙氨酸置换)组分的例如以融合蛋白(例如M72)形式的组合。M72的多肽序列描述于W02006/117240的SEQ ID No:4(cDNA描述于SEQ ID No:3中),其中其掺入了任选的双组氨酸以有助于制备,当用于本发明时,M72也可掺入双组氨酸,但合适的是M72不存在所述任选的双组氨酸(即W02006/117240 SEQ ID No:4的残基4-725尤其引人关注)。M72的多肽序列示于SEQ ID No:24;

[0326] (c)Mtb8.4、Mtb9.8、Mtb9.9和Mtb41组分的例如以融合蛋白(例如Mtb71f)形式的组合。Mtb71f的多肽序列描述于W099/051748的SEQ ID No:16(cDNA描述于SEQ ID No:15中),其中其掺入了任选的His-标记以有助于纯化,当用于本发明时,合适的是Mtb71f对应于W099/051748 SEQ ID NO:16的氨基酸残基9-710。Mtb71f的多肽序列示于SEQ ID No:25;

[0327] (d)Mtb72f或M72(合适的是,没有任选的组氨酸残基以有助于表达)与Mtb9.8和

Mtb9.9的例如以融合蛋白的组合。M72-Mtb9.9-Mtb9.8融合体的多肽序列示于SEQ ID No: 26(M92融合体),当用于本发明时,M72-Mtb9.9-Mtb9.8融合体可任选在起始甲硫氨酸残基后面掺入双组氨酸以有助于制备;

[0328] (e)Mtb72f或M72(合适的是,没有任选的组氨酸残基以有助于表达)与Ag85B的例如以融合蛋白(例如Mtb103f)的组合。Mtb103f的多肽序列描述于W003/070187的SEQ ID No:18(cDNA描述于SEQ ID No:10中),其中其掺入了任选的His-标记以有助于纯化,当用于本发明时,合适的是Mtb103f对应于W003/070187 SEQ ID No:18的氨基酸残基8-1016。同样尤其引人关注的是M103,即在Ra35组分中掺入Ser/Ala突变的Mtb103f,当用于本发明时,合适的是M103对应于W003/070187 SEQ ID No:18的氨基酸残基8-1016,其中在第710位的Ser残基已用Ala置换。M103的多肽序列示于SEQ ID No:27,当用于本发明时,M72-Mtb9.9-Mtb9.8融合体可任选在起始甲硫氨酸残基后掺入双组氨酸以有助于制备;

[0329] (f)Mtb72f或M72(合适的是,没有任选的组氨酸残基以有助于表达)与Mtb41的例如以融合蛋白(例如Mtb114f)的组合。Mtb114f的多肽序列描述于W003/070187的SEQ ID No:16(cDNA描述于SEQ ID No:9),其中其掺入了任选的His-标记以有助于纯化,当用于本发明时,合适的是Mtb114f对应于W003/070187 SEQ ID No:16的氨基酸残基8-1154。同样尤其引人关注的是M114,即在Ra35组分中掺入Ser/Ala突变的Mtb114f,当用于本发明时,合适的是M114对应于来自W003/070187的SEQ ID No:16的氨基酸残基8-1154,其中在第710位的Ser残基已用Ala置换。M114的多肽序列示于SEQ ID No:28,当用于本发明时,M72-Mtb9.9-Mtb9.8融合体可任选在起始甲硫氨酸残基后掺入双组氨酸以有助于制备;

[0330] (g)Ag85B与ESAT-6组分的组合,例如以在Doherty等Journal of Infectious Diseases 2004 190:2146-2153所述融合体的组合;和/或

[0331] (h)Ag85B与TB10.4组分的组合,例如以在Dietrich等Journal of Immunology 2005 174(10):6332-6339 190:2146-2153所述融合体的组合。

[0332] Rv3616c组分与Rv1753c组分的组合尤其引人关注。显而易见的是,这种组合可任选包含其它另外的抗原组分(例如M72组分)。

[0333] 另一种引人关注的组合包含Rv3616c组分和M72组分。

[0334] 另一种引人关注的组合包含Rv3616c组分和Rv2386c组分。

[0335] 其它引人关注的组合包括包含Rv3616c组分与Rv2707c组分的组合。

[0336] 另外的引人关注的组合包含Rv3616c组分和 α -晶体蛋白组分。

[0337] 本领域技术人员应认识到的是,组合无需依赖于上述(i)-(xvi)和(a)-(h)中描述的特定序列,以及所述序列的保守修饰变体(例如具有至少70%同一性,例如至少80%同一性、尤其是至少90%同一性和特别是至少95%同一性)或免疫原性片段(例如全长抗原的至少20%,例如该抗原的至少50%、尤其是至少70%和特别是至少80%)可用于获得相同的实施效果。

[0338] 上述独立抗原序列中的每一个还公开于Cole等Nature 1998 393:537-544和Camus Microbiology 2002 148:2967-2973中。结核分枝杆菌H37Rv的基因组为公众可获得的,例如在Welcome Trust Sanger Institute网页(www.sanger.ac.uk/Projects/M_tuberculosis/)和其它地方。

[0339] 许多上述抗原还公开于美国专利申请号08/523,435、08/523,436、08/658,800、

08/659,683、08/818,111、08/818,112、08/942,341、08/942,578、08/858,998、08/859,381、09/056,556、09/072,596、09/072,967、09/073,009、09/073,010、09/223,040、09/287,849和PCT专利申请PCT/US98/10407、PCT/US98/10514、PCT/US99/03265、PCT/US99/03268、PCT/US99/07717、W097/09428和W097/09429、W098/16645、W098/16646,上述任一参考文献都通过引用结合于本文中。

[0340] 本发明组合物、多肽和核酸还可包含来自其它来源的其它多肽。例如,本发明组合物和融合蛋白可包含多肽或编码多肽的核酸,其中所述多肽增强抗原的表达,例如NS1、流感病毒蛋白的表达(参见,例如W099/40188和W093/04175)。本发明核酸可根据所选物种例如人类(在体内表达的情况下)或特定细菌(在多肽制备的情况下)偏好的密码子来进行改造。

[0341] Rv3616c组分还可与一种或多种有效抗结核病(例如结核分枝杆菌感染)的化疗药一起给予。这类化疗药的实例包括但不限于阿米卡星、氨基水杨酸、卷曲霉素、环丝氨酸、乙胺丁醇、乙硫异烟胺、异烟肼、卡那霉素、吡嗪酰胺、利福霉素(即利福平、利福喷汀和利福布汀)、链霉素、氧氟沙星、环丙沙星、克拉霉素、阿奇霉素和氟喹诺酮类。这类化学疗法由治疗医师判断使用优选药物组合来确定。用于治疗非耐药性结核病(例如结核分枝杆菌感染)的“一线”化疗药包括异烟肼、利福平、乙胺丁醇、链霉素和吡嗪酰胺。用于治疗已显示对一种或多种“一线”药物具有耐药性的结核病(例如结核分枝杆菌感染)的“二线”化疗药包括氧氟沙星、环丙沙星、乙硫异烟胺、氨基水杨酸、环丝氨酸、阿米卡星、卡那霉素和卷曲霉素。

[0342] 常规化疗药通常在相对长时期(大约9个月)内给予。常规化疗药与本发明Rv3616c组分的给药的组合能够缩短化疗治疗期(例如至8个月、7个月、6个月、5个月、4个月、3个月或更短)而不会降低有效性。

[0343] 尤其引人关注的是将Rv3616c组分与卡介苗(BCG)联用,例如以重组表达Rv3616c(或如本文所述的其变体或片段)的修饰BCG形式。或者,Rv3616c组分可通过同时给药或通过加强先前的BCG接种,用于增强受试者对BCG接种的反应。当用于增强受试者对BCG接种的反应时,Rv3616c组分显然可以多肽或多核苷酸的形式提供(任选与如上文所述的其它抗原组分一起)。

[0344] 本领域技术人员应认识到的是,组分的组合无需一起给予,并可如下施用:单独或联合;同时、序贯或在短期内;通过相同途径或通过不同途径。然而,为了方便,通常理想(当给药方案可配伍时)的是,作为单个组合物给予组分的组合。

[0345] 通常将本发明多肽、多核苷酸和组合物给予人,但也有效于包括家养哺乳动物(例如狗、猫、兔、大鼠、小鼠、豚鼠、仓鼠、栗鼠)和农业哺乳动物(例如牛、猪、绵羊、山羊、马)在内的其它哺乳动物。

[0346] 免疫原性片段

[0347] T细胞表位为短连续氨基酸段,其被T细胞(例如CD4⁺或CD8⁺T细胞)识别。T细胞表位的鉴定可通过表位作图实验实现,表位作图实验为本领域技术人员所熟知(参见,例如Paul,Fundamental Immunology,第三版,243-247(1993);Beißbarth等Bioinformatics 2005 21(增刊1):i29-i37)。

[0348] 或者,可使用实施例中论述的方法预测表位。

[0349] 由于T细胞应答在结核病中的决定性参与,所以显而易见的是,全长Rv3616c多肽

的包含至少一个T细胞表位的片段具有免疫原性,并且可有助于免疫保护。这类片段在本文中称为免疫原性片段。

[0350] 本发明免疫原性片段通常包含来自全长多肽序列的至少9个连续氨基酸(例如至少10个),例如至少12个连续氨基酸(例如至少15个或至少20个连续氨基酸),尤其是至少50个连续氨基酸,例如至少100个连续氨基酸(例如至少200个连续氨基酸)。合适的是,免疫原性片段为全长多肽序列长度的至少20%,例如至少50%、至少70%或至少80%。

[0351] 应理解的是,在多样远系群体例如人类中,不同的HLA型意指特定表位可能不被该群体的所有成员识别。因此,为使对多肽的识别水平和免疫应答规模最大化,通常需要免疫原性片段包含来自全长序列的多个表位(适宜全部表位)。

[0352] Rv3616c蛋白的可用的具体片断包括含至少一个CD4+表位、适宜至少两个CD4+表位和尤其所有CD4+表位的片段(例如在实施例和SEQ ID No:29-47中描述的表位,尤其是与多个HLA等位基因关联的表位,例如与2、3、4、5或更多个等位基因关联的表位)。

[0353] Rv3616c蛋白的可用的其它片断包括含至少一个CD8表位、适宜至少两个CD8表位和尤其所有CD8表位的片段(例如在实施例和SEQ ID No:48-126中描述的表位,尤其是与多个HLA等位基因关联的表位,例如与2、3、4、5或更多个等位基因关联的表位)。

[0354] 当使用全长多肽的单个片段时,当其引起以下反应时认为这类片段是免疫原性的:在用特定抗原体外再刺激(restimulation)PBMC或全血的测定(例如再刺激数小时至高达两周的时期,例如高达1天、1天至1周或1-2周)中,所述反应为参比序列活性的至少20%、适宜至少50%和尤其至少75%(例如至少90%),所述测定通过淋巴组织增生、在培养上清液中细胞因子的产生(由ELISA、CBA等测量)测量细胞的活化,或者通过胞内和胞外染色(例如使用特异于免疫标志物(例如CD3、CD4、CD8、IL2、TNF α 、IFN γ 、CD40L、CD69等)的抗体)接着用流式细胞仪分析,来表征T和B细胞应答。合适的是,当其引起以下反应时认为片段是免疫原性的:在T细胞增殖和/或IFN- γ 产生测定中,所述反应为参比序列活性的至少20%、适宜至少50%和尤其至少75%(例如至少90%)。

[0355] 在某些情况下,可使用全长多肽的多个片段(其可重叠或不重叠并且可覆盖或不覆盖整个全长序列)以得到与全长序列本身等效的生物反应。例如,组合中的如上所述的至少2个免疫原性片段(例如3、4或5个),在PBMC或全血的体外再刺激测定(例如T细胞增殖和/或IFN- γ 产生测定)中提供参比序列活性的至少50%、适宜至少75%和尤其至少90%。

[0356] SEQ ID No:127-156的肽为特别引人关注的片段(尤其是SEQ ID No:127-133和143-156的肽)。

[0357] 变体

[0358] “变体”或“保守修饰变体”适用于氨基酸序列和核酸序列两者。对于特定核酸序列而言,保守修饰变体是指编码相同或基本相同氨基酸序列的核酸,或者在核酸并不编码氨基酸序列时,是指基本相同的序列。

[0359] 由于遗传密码的简并性,大量功能相同的核酸编码任何既定蛋白。例如,密码子GCA、GCC、GCG和GCU都编码氨基酸丙氨酸。因此,在被密码子指定为丙氨酸的每个位置,可将该密码子改为任何所述相应密码子,而不改变所编码的多肽。这样的核酸变异导致产生“沉默”或“简并”变体,其为保守修饰变异中的一种。本文编码多肽的每个核酸序列也都描述了所述核酸的每种可能的沉默变异。技术人员应认识到核酸中的各密码子(除了AUG(其通常

是甲硫氨酸的唯一密码子)和TGG(其通常是色氨酸的唯一密码子)之外)都可被改变,以得到功能相同的分子。因此,编码多肽的核酸中的每个沉默变异都隐含在每个所述序列中。

[0360] 与参比序列相比,本发明多核苷酸可包含许多沉默变异(可改变例如1-50个,例如1-25个、尤其1-5个和特别1个密码子)。与参比序列相比,本发明多核苷酸可包含许多非沉默保守变异(可改变例如1-50个,例如1-25个、尤其1-5个和特别1个密码子)。非沉默变异为导致所编码氨基酸序列改变的变异(通过氨基酸残基的取代、缺失或添加)。本领域技术人员应认识到特定多核苷酸序列可包含沉默和非沉默保守变异两者。

[0361] 对于蛋白序列的变体而言,技术人员应认识到的是,当一个或多个改变导致功能类似的氨基酸对氨基酸的取代或者不会实质影响变体生物功能的残基取代/缺失/添加时,对多肽进行的改变、添加或缺失单个氨基酸或小百分比氨基酸的单个取代、缺失或添加为“保守修饰变体”。

[0362] 提供功能类似的氨基酸的保守取代表为本领域所熟知。这类保守修饰变体作为多态变体、种间同源物和等位基因的补充(并且不排除这些变体)。

[0363] 与参比序列相比,本发明多肽可包含许多保守取代(可改变例如1-50个,例如1-25个、尤其1-10个和特别1个氨基酸残基)。一般而言,这类保守取代将落入以下指定的氨基酸分组的一个中,但在某些情况下,在不会实质影响抗原的免疫原性的条件下,其它取代是可能的。以下8个组别每一个都含通常相互之间为保守取代的氨基酸:

[0364] 1)丙氨酸(A),甘氨酸(G);

[0365] 2)天冬氨酸(D),谷氨酸(E);

[0366] 3)天冬酰胺(N),谷氨酰胺(Q);

[0367] 4)精氨酸(R),赖氨酸(K);

[0368] 5)异亮氨酸(I),亮氨酸(L),甲硫氨酸(M),缬氨酸(V);

[0369] 6)苯丙氨酸(F),酪氨酸(Y),色氨酸(W);

[0370] 7)丝氨酸(S),苏氨酸(T);和

[0371] 8)半胱氨酸(C),甲硫氨酸(M);

[0372] (参见例如Creighton,Proteins(1984))。

[0373] 合适的是,这类取代不发生在表位区,因而不会对抗原的免疫原性产生显著影响。

[0374] 与参比序列相比,蛋白变体还可包括其中插入另外的氨基酸的变体,例如这类插入可在1-10个位置(例如1-5个位置,适宜1或2个位置,尤其是1个位置)出现,并且可例如涉及在各个位置插入50个或更少的氨基酸(例如20个或更少,尤其是10个或更少,特别是5个或更少)。合适的是,这类插入不发生在表位区,因而不会对抗原的免疫原性产生显著影响。插入的一个实例包括一小段组氨酸残基(例如2-6个残基)以有助于目标抗原的表达和/或纯化。

[0375] 与参比序列相比,蛋白变体还可包括其中已缺失氨基酸的变体,例如这类缺失可在1-10个位置(例如1-5个位置,适宜1或2个位置,尤其是1个位置)出现,并且可例如涉及在各个位置缺失50个或更少的氨基酸(例如20个或更少,尤其是10个或更少,特别是5个或更少)。合适的是,这类缺失不发生在表位区,因而不会对抗原的免疫原性产生显著影响。

[0376] 技术人员应认识到特定蛋白变体可包含取代、缺失和添加(或其任何组合)。

[0377] 确定抗原表位区的方法描述且例示于实施例中。

[0378] 变体优选显示与相关参比序列的至少约70%同一性、更优选至少约80%同一性和最优选至少约90%同一性(例如至少约95%、至少约98%或至少约99%)。

[0379] 关于两个或更多个核酸或多肽序列的术语“同一的”或百分比“同一性”是指如利用以下序列比较算法中的一种或者通过人工比对和目测检查所测定,当在比较窗或指定区域内进行比较和比对最大一致性时,相同或者具有特定百分比的相同氨基酸残基或核苷酸(即在指定区域内有70%同一性、任选75%、80%、85%、90%、95%、98%或99%同一性)的两个或更多个序列或亚序列。那么这类序列就称为“基本同一的”。该定义也涉及测试序列的互补序列(compliment)。任选的是,同一性存在于长度至少约25个至约50个氨基酸或核苷酸的区域内,或任选在长度为75-100个氨基酸或核苷酸的区域内。合适的是,在对应于参比序列的完整长度的窗口内进行比较。

[0380] 对于序列比较,通常将一个序列用作参比序列,将测试序列与其进行比较。当使用序列比较算法时,将测试序列和参比序列都输入计算机中,必要时指定亚序列坐标,并且指定序列算法程序参数。可以使用缺省程序参数,或者可以指定备选参数。然后根据程序参数,序列比较算法计算出测试序列相对于参比序列的百分比序列同一性。

[0381] 本文所用的“比较窗”是指以下区段:其中在将一个序列与具有相同数目连续位置的参比序列进行最佳排列之后,可将该序列与参比序列进行比较。用于比较的序列比对方法是本领域众所周知的。可以按照诸如以下的方法等进行用于比较的序列最佳比对:局部同源性算法,Smith和Waterman,Adv. Appl. Math. 2:482(1981);同源性比对算法,Needleman和Wunsch, J. Mol. Biol. 48:443(1970);相似性搜索方法,Pearson和Lipman, Proc. Nat'l Acad. Sci. USA 85:2444(1988);这些算法的计算机工具(Wisconsin Genetics Software Package中的GAP、BESTFIT、FASTA和TFASTA, Genetics Computer Group, 575 Science Dr., Madison, WI);或者人工比对和目测检查(参见例如Current Protocols in Molecular Biology(Ausubel等编著,1995增刊))。

[0382] 有用算法的一个实例是PILEUP。PILEUP运用渐近配对比对,从一组相关序列中创建多重序列比对,以显示相互关系和百分比序列同一性。其也绘出显示聚类关系的树状图或分枝图(dendrogram),所述聚类关系用于创建所述比对。PILEUP采用Feng和Doolittle, J. Mol. Evol. 35:351-360(1987)的渐近比对的简化法。所使用的方法类似于Higgins和Sharp, CABIOS 5:151-153(1989)描述的方法。所述程序可比对多达300个序列,每个序列最大长度为5,000个核苷酸或氨基酸。所述多重比对法始于两个最为相似序列的配对比对,产生一组两个经比对的序列。然后该组序列与下一个最相关的序列或下一组经比对的序列进行比对。通过两个单独序列的配对比对的简单延伸,将两组序列进行比对。通过一系列渐近配对比对,完成最终比对。通过对序列比较区指定具体序列及其氨基酸或核苷酸坐标并通过指定程序参数,运行该程序。运用PILEUP,采用以下参数:缺省空位加权(default gap weight)(3.00)、缺省空位长度加权(default gap length weight)(0.10)和加权末端空位(weighted end gap),可将参比序列与其它测试序列进行比较,以确定百分比序列同一性关系。PILEUP可得自GCG序列分析软件包,例如7.0版本(Devereaux等, Nuc. Acids Res. 12:387-395(1984))。

[0383] 适用于测定百分比序列同一性和序列相似性的算法的另一个实例是BLAST和BLAST 2.0算法,其分别在Altschul等, Nuc. Acids Res. 25:3389-3402(1977)和Altschul

等, *J. Mol. Biol.* 215:403-410(1990)中有描述。用于进行BLAST分析的软件公众可通过国家生物技术信息中心(National Center for Biotechnology Information)(网址在 www.ncbi.nlm.nih.gov/)得到。该算法包括首先通过鉴定在查询序列中与数据库序列中相同长度的字串比对时匹配或满足某些正阈值分数T的长度W的短字串,来鉴定高分序列对(HSP)。T被称为邻近字串分数阈值(Altschul等,见上文)。这些起始邻近字串命中作为用于起始搜索以发现含有其的更长HSP的种子。只要累积比对分数可增加,字串命中则以两个方向沿每个序列延伸。对于核苷酸序列,使用参数M(对匹配残基对的奖分(reward score);总是 >0)和N(对错配残基的罚分(penalty score);总是 <0),计算累积分数。对于氨基酸序列,使用计分矩阵来计算累积分数。字串命中在每个方向的延伸当遇到下述情况时终止:累积比对分数比其获得的最大值低数量X;累积分数为零或零以下,这是由于一个或多个负计分残基比对的累加所致;或者到达任一序列的末端。BLAST算法参数W、T和X决定所述比对的灵敏度和速度。BLASTN程序(对于核苷酸序列)使用字长(W)为11,期望值(E)为10, $M=5$, $N=-4$ 作为缺省值,并且比较两条链。对于氨基酸序列而言,BLASTP程序作为缺省值使用字长(W)为3和期望值(E)为10,以及BLOSUM62计分矩阵(参见Henikoff和Henikoff, *Proc. Natl. Acad. Sci. U.S.A.* 89:10915(1989))比对(B)为50,期望值(E)为10, $M=5$, $N=-4$,并且比较两条链。

[0384] BLAST算法也进行两个序列之间相似性的统计学分析(参见例如Karlin和Altschul, *Proc. Natl. Acad. Sci. U.S.A.* 90:5873-5787(1993))。BLAST算法提供的相似性的一种衡量是最小总和概率(smallest sum probability)($P(N)$),其提供两个核苷酸序列或氨基酸序列偶然发生匹配的概率的指标。例如,如果测试核酸与参比核酸比较中的最小总和概率小于约0.2、更优选小于约0.01和最优选小于约0.001,则认为核酸与参比序列相似。

[0385] 本发明还涉及用于治疗或预防潜伏性TB的多核苷酸,其包含第一核苷酸序列,该序列在中等严格条件下(例如在高度严格条件下)与第二核苷酸序列的互补序列(complement)选择性杂交,所述第二核苷酸序列编码以下多肽,其包含:

[0386] (i) Rv3616c蛋白序列;

[0387] (ii) Rv3616c蛋白序列的变体;或

[0388] (iii) Rv3616c蛋白序列的免疫原性片段。

[0389] 措词“高度严格杂交条件”是指以下条件:在该条件下探针将杂交至通常在核酸的复杂混合物中的其靶亚序列,但不会杂交至其它序列。高度严格条件为序列依赖性的并且不同情况下不同。较长的序列在较高的温度下特异性杂交。有关核酸杂交的详细指南可参见Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, “Overview of principles of hybridization and the strategy of Nucleic Acid assays(杂交原理和核酸测定策略的综述)”(1993)。通常,将高度严格条件选择在限定的离子强度、pH下,比特定序列的热解链温度(T_m)低约5-10°C。 T_m 为(在限定的离子强度、pH和核酸浓度下)平衡时50%与靶互补的探针杂交至靶序列的温度(由于靶序列过量存在,所以在 T_m 下,50%探针在平衡时被占用)。高度严格条件是以下条件:其中pH 7.0-8.3,盐浓度低于约1.0M钠离子、通常约0.01-1.0M钠离子浓度(或其它盐),温度对短探针(例如10-50个核苷酸)至少约30°C和对长探针(例如超过50个核苷酸)至少约60°C。高度严格条件也可以通过加入去稳定剂例如甲酰胺来达到。对于选择性或特异性杂交,阳性信

号至少是背景的两倍,任选为背景杂交的10倍。

[0390] 例示性的高度严格杂交条件可为如下:50%甲酰胺,5×SSC和1%SDS,于42℃保温,或者5×SSC,1%SDS,于65℃保温,在0.2×SSC和0.1%SDS中于65℃洗涤。

[0391] 如果核酸所编码的多肽基本同一,则在高度严格条件下彼此不杂交的核酸仍然在功能上等同。这发生于例如当使用遗传密码所允许的最大密码子简并性产生核酸拷贝时。在此情况下,核酸通常在中等严格杂交条件下杂交。

[0392] 例示性的“中等严格杂交条件”包括在40%甲酰胺、1M NaCl、1%SDS的缓冲液中于37℃杂交,然后在1×SSC中于45℃洗涤。阳性杂交至少是背景的两倍。普通技术人员容易认识到可采用备选的性杂交和洗涤条件以提供类似严格性的条件。

[0393] 措词“选择性(或特异性)杂交至”是指当特定核苷酸序列存在于复杂混合物(例如总细胞或文库DNA或RNA)中时,在严格杂交条件下分子仅与该序列结合、双链化或杂交。

[0394] 在任何情况下,多肽序列的变体具有与参比序列基本相同的活性(在多核苷酸的情况下,变体多核苷酸序列编码具有与参比序列活性基本相同的多肽)。基本相同的活性意指在用特定抗原体外再刺激PBMC或全血的测定(例如再刺激数小时至高达两周的时期,例如高达1天、1天至1周或1-2周)中的参比序列活性的至少50%、适宜至少75%和尤其至少90%,所述测定通过淋巴组织增生、在培养上清液中细胞因子的产生(由ELISA、CBA等测量)测量细胞的活化,或者通过胞内和胞外染色(例如使用特异于免疫标志物(例如CD3、CD4、CD8、IL2、TNF α 、IFN γ 、CD40L、CD69等)的抗体)接着用流式细胞仪分析,来表征T和B细胞应答。合适的是,基本相同的活性意指在T细胞增殖和/或IFN- γ 产生测定中的参比序列活性的至少50%、适宜至少75%和尤其至少90%。

[0395] 多核苷酸组合物

[0396] 本文所用术语“多核苷酸”是指已分离、无特定物种总基因组DNA的分子。因此,编码多肽的多核苷酸是指含一个或多个编码序列、但基本上从获取所述多核苷酸的物种总基因组DNA中完全分离或纯化的多核苷酸区段。

[0397] 本领域技术人员应理解的是,用于本发明的多核苷酸可包括基因组序列、基因组外序列和质粒编码的序列以及较小的改造的基因区段,其表达或可适于表达蛋白、多肽、肽等。这类区段可以是天然分离的,或人工合成修饰的。

[0398] 本文所用的“分离的”是指多核苷酸基本上与其它编码序列分开并且所述多核苷酸不含大部分无关编码DNA,例如大染色体片段或其它功能基因或多肽编码区。分离核酸与其它可读框分离,所述可读框位于基因侧翼并编码除该基因外的蛋白。当然,这是指最初分离的DNA区段,并不排除以后通过人工向该区段添加的基因或编码区。

[0399] 技术人员应认识到,多核苷酸可以是单链(编码链或反义链)或双链的,并且可以是DNA(基因组、cDNA或合成的)或RNA分子。RNA分子包括HnRNA分子(其包含内含子并以一对一方式对应于DNA分子)和mRNA分子(其不包含内含子)。额外的编码或非编码序列可以(但并非必需)存在于本发明多核苷酸内,并且多核苷酸可以(但并非必需)与其它分子和/或支持材料连接。

[0400] 多核苷酸可包含天然序列(即编码分枝杆菌抗原或其部分的内源序列)或可包含所述序列的变体或者生物学等同物或功能等同物。如下文进一步描述,多核苷酸变体可含有一个或多个取代、添加、缺失和/或插入,相对于参比蛋白来说,优选使得所编码多肽的免

疫原性并未减少。对所编码多肽的免疫原性的影响通常可如本文所述进行评价。

[0401] 在另外的实施方案中,本发明提供分离的多核苷酸和多肽,其包含与本文所公开的一个或多个序列相同或互补的不同长度的连续序列段。例如,本发明提供的多核苷酸包含本文所公开的参比序列的至少约30、40、50、75、100、150、200、300、400、500或1000个或更多个连续核苷酸以及它们之间的所有中间长度。容易理解的是,在该范围之内,“中间长度”是指所给出的数值间的任何长度,例如30、31、32等;50、51、52、53等;100、101、102、103等;150、151、152、153等;包括200-500;500-1,000等的所有整数。

[0402] 此外,本领域普通技术人员应认识到的是,由于遗传密码的简并性,所以存在许多编码本文所述的多肽的核苷酸序列。这些多核苷酸中的某些具有与任何天然基因的核苷酸序列的相对低的同一性。尽管如此,因密码子使用差异而不同的多核苷酸都明确由本发明所考虑,例如为人和/或灵长类密码子选择而最优化的多核苷酸。此外,包含本文提供的多核苷酸序列的基因的等位基因也在本发明范围之内。等位基因是因一个或多个突变而改变的内源基因,所述突变为例如核苷酸的缺失、添加和/或取代。所得mRNA和蛋白可以(但并非必需)具有改变的结构或功能。等位基因可以用标准技术(例如杂交、扩增和/或数据库序列比较)进行鉴定。

[0403] 多肽鉴定和表征

[0404] 可用多种已良好建立的技术中的任一种,来鉴定、制备和/或操作多核苷酸。例如,如下文更详细描述,可通过筛选cDNA微阵列来鉴定多核苷酸。可例如按照生产商的说明书(以及基本如以下文献所述:Schena等,Proc.Natl.Acad.Sci.USA 93:10614-10619(1996)和Heller等,Proc.Natl.Acad.Sci.USA 94:2150-2155(1997)),使用Synteni微阵列(Palo Alto,CA),进行这样的筛选。或者,可以从表达本文所述蛋白的细胞(例如结核分枝杆菌细胞)制备的cDNA中扩增多核苷酸。可通过聚合酶链式反应(PCR)扩增这样的多核苷酸。对于该方法,可根据本文提供的序列来设计序列特异性引物,也可购买或合成引物。

[0405] 使用众所周知的技术,可将多核苷酸的扩增部分用于从合适文库(例如结核分枝杆菌cDNA文库)中分离出全长基因。在这样的技术中,用适合扩增的一个或多个多核苷酸探针或引物筛选文库(cDNA或基因组文库)。优选的是,文库经大小选择以包含较大的分子。也可优选将随机引物文库用于鉴定基因的5'和上游区。优选将基因组文库用于获取内含子和延伸5'序列。

[0406] 对于杂交技术,可以使用众所周知的技术标记部分序列(例如通过切口平移或末端标记³²P)。然后一般通过使含有变性菌落(或含有噬斑的菌苔)滤膜与经标记的探针杂交,筛选细菌文库或噬菌体文库(参见Sambrook等,Molecular Cloning:A Laboratory Manual (2000))。选择并扩增杂交菌落或噬斑,然后分离DNA用于进一步分析。可通过例如PCR,使用来自部分序列的引物和来自载体的引物,分析cDNA克隆以确定额外序列的量。可产生限制图谱和部分序列,以鉴定一个或多个重叠克隆。然后可使用可包括产生系列缺失克隆的标准技术,确定完整序列。然后可将所得重叠序列装配成单个连续序列。可使用众所周知的技术,通过连接合适片段产生全长cDNA分子。

[0407] 或者,存在大量的扩增技术用于根据部分cDNA序列得到全长编码序列。这些技术中,一般通过PCR进行扩增。多种市售试剂盒中的任一种都可用于进行扩增步骤。可使用例如本领域众所周知的软件来设计引物。引物优选长度为22-30个核苷酸,GC含量至少为50%

并且在约68°C-72°C的温度下退火到靶序列上。可如上所述对扩增区进行测序,并将重叠序列装配成连续序列。

[0408] 一种这样的扩增技术是反向PCR(参见Triglia等,Nucl.Acids Res.16:8186(1988)),该技术使用限制酶在基因已知区内产生片段。再通过分子内连接使片段环化,并用作使用来自已知区的反向引物(divergent primer)的PCR模板。在备选方法中,邻接部分序列的序列可通过用接头序列的引物和对已知区具有特异性的引物扩增来回溯(retireve)。扩增的序列通常经过第2轮扩增,其用同样的接头引物和对已知区具有特异性的第二引物。该方法的一个改动就是使用沿已知序列的相反方向开始延伸的两个引物,参见WO 96/38591。另一种这样的技术称为“cDNA末端快速扩增”或RACE。该技术包括使用内部引物和外部引物,其能与polyA区或载体序列杂交,以鉴定位于已知序列的5'和3'处的序列。另外的技术包括捕获PCR(Lagerstrom等,PCR Mehtods Applic.1:111-19(1991))和步移PCR(walking PCR)(Parker等,Nucl.Acids.Res.19:3055-60(1991))。使用扩增的其它方法也可用于获取全长cDNA序列。

[0409] 在某些情况下,可通过分析例如得自GenBank的表达序列标签(EST)数据库提供的序列,获取全长cDNA序列。通常可用众所周知的程序(例如NCBI BLAST搜索)进行重叠EST的搜索,并且这样的EST可用于产生连续的全长序列。也可通过分析基因组片段获取全长DNA序列。

[0410] 在宿主细胞中的多核苷酸表达

[0411] 编码多肽或者其融合蛋白或功能等同物的多核苷酸序列或其片段可用于重组DNA分子中,以在合适宿主细胞中指导多肽表达。因为遗传密码的固有简并性,所以可产生编码基本相同或功能等同氨基酸序列的其它DNA序列,而且这些序列可用于克隆和表达既定多肽。

[0412] 本领域技术人员应理解,在某些情况下产生具有非天然存在密码子的编码多肽的核苷酸序列可为有利的。例如,可以选择特定原核或真核宿主偏好的密码子,以增加蛋白表达速率或产生具有所需特性(例如半衰期比天然存在序列所产生的转录物的长)的重组RNA转录物。

[0413] 此外,可使用本领域公知的方法改造多核苷酸序列,以为了各种原因改变多肽编码序列,包括但不限于修饰基因产物的克隆、加工和/或表达的改变。例如,通过对基因片段和合成寡核苷酸的随机片段化和PCR再装配的DNA改组,可用于改造核苷酸序列。另外,位点定向诱变可用于插入新限制位点、改变糖基化模式、改变密码子偏好、产生剪接变体或引入突变等。

[0414] 可将天然、修饰或重组核酸序列与异源序列连接,以编码融合蛋白。例如,为了针对多肽活性抑制剂筛选肽文库,其可用于编码可被市售抗体识别的嵌合蛋白。也可改造融合蛋白,以在介于多肽编码序列与异源蛋白序列之间含有切割位点,使得可切割多肽并经纯化而与异源部分分开。

[0415] 可采用本领域众所周知的化学方法,全部或部分合成编码所需多肽的序列(参见Caruthers,M.H.等,Nucl.Acids Res.Symp.Ser.第215-223页(1980),Horn等,Nucl.Acids Res.Symp.Ser.第225-232页(1980))。或者,可用化学方法合成多肽氨基酸序列或其部分,来产生蛋白本身。例如,可使用各种固相技术进行肽合成(Roberge等,Science 269:202-

204(1995)),而且可例如使用ABI 431A Peptide Synthesizer(Perkin Elmer,Palo Alto,CA)来实现自动化合成。

[0416] 可通过制备型高效液相色谱(例如Creighton,Proteins,Structures and Molecular Principles(1983))或本领域可用的其它类似技术,基本纯化新合成的肽。合成肽的组成可通过氨基酸分析或测序(例如Edman降解法)得以证实。另外,多肽的氨基酸序列或其任何部分可在直接合成期间改变和/或使用化学方法与来自其它蛋白的序列或其任何部分组合,以产生变体多肽。

[0417] 为了表达所需多肽,可以将编码多肽的核苷酸序列或功能等同物插入到合适表达载体,即含有用于转录和翻译所插入编码序列的必要元件的载体中。可以用本领域技术人员众所周知的方法来构建含有目标多肽编码序列及合适的转录和翻译控制元件的表达载体。这些方法包括体外重组DNA技术、合成技术和体内遗传重组。这些技术参见Sambrook等,Molecular Cloning,A Laboratory Manual(2000)和Ausubel等,Current Protocols in Molecular Biology(每年更新)。

[0418] 各种表达载体/宿主系统可用于包含和表达多核苷酸序列。它们包括但不限于微生物,例如用重组噬菌体、质粒或粘粒DNA表达载体转化的细菌;用酵母表达载体转化的酵母菌;用病毒表达载体(例如杆状病毒)感染的昆虫细胞系统;用病毒表达载体(例如花椰菜花叶病毒,CaMV;烟草花叶病毒,TMV)或用细菌表达载体(例如Ti或pBR322质粒)转化的植物细胞系统;或动物细胞系统。

[0419] 表达载体中存在的“控制元件”或“调节序列”是载体的非翻译区——增强子、启动子、5'和3'非翻译区,它们与宿主细胞蛋白相互作用以进行转录和翻译。这类元件的强度和特异性可不同。根据所用的载体系统和宿主,可使用许多合适的转录和翻译元件,包括组成型和诱导型启动子。例如,当在细菌系统中克隆时,可使用诱导型启动子,例如PBLUESCRIPT噬菌粒(Stratagene,La Jolla,Calif.)或PSPORT1质粒(Gibco BRL,Gaithersburg,MD)的杂合lacZ启动子等。在哺乳动物细胞系统中,通常优选来自哺乳动物基因或来自哺乳动物病毒的启动子。如果需要产生含有多拷贝的多肽编码序列的细胞系,基于SV40或EBV的载体可有利地与合适的选择标记一起使用。

[0420] 在细菌系统中,可根据所表达多肽的预期用途,选择多种表达载体。例如,当需要大量多肽例如用于诱导抗体时,可使用指导容易纯化的融合蛋白高水平表达的载体。这样的载体包括但不限于多功能大肠杆菌克隆和表达载体,例如BLUESCRIPT(Stratagene),其中编码目标多肽的序列可与氨基端Met和其后7个残基 β -半乳糖苷酶的序列一起符合读框地连接到载体中,使得产生杂合蛋白;pIN载体(Van Heeke和Schuster,J.Biol.Chem.264:5503-5509(1989));等等。pGEX载体(Promega,Madison,Wis.)也可用于将外源多肽表达为与谷胱甘肽S-转移酶(GST)的融合蛋白。一般而言,这样的融合蛋白是可溶性的,并可通过吸附到谷胱甘肽-琼脂糖珠上、接着在游离谷胱甘肽存在下进行洗脱,从裂解细胞中容易地纯化该融合蛋白。可设计在这种系统中制备的蛋白,以使其包含肝素、凝血酶或因子XA蛋白酶切割位点,使得克隆的目标多肽可随意从GST部分释放出来。

[0421] 在酵母菌酿酒酵母(*Saccharomyces cerevisiae*)中,可使用多种含有组成型或诱导型启动子(例如 α 因子、醇氧化酶和PGH)的载体。其它载体含有的组成型或诱导型启动子包括GAP、PGK、GAL和ADH。有关综述可参见Ausubel等(见上文)、Grant等,Methods

Enzymol.153:516-544(1987)和Romas等,Yeast 8 423-88(1992)。

[0422] 在使用植物表达载体的情况下,多肽编码序列的表达可由任意多种启动子驱动。例如,病毒启动子(例如CaMV的35S和19S启动子)可单独使用,或者与来自TMV的 Ω 前导序列一起使用(Takamatsu,EMBO J.6:307-311(1987))。或者,可使用植物启动子例如RUBISCO的小亚基或热激启动子(Coruzzi等,EMBO J.3:1671-1680(1984);Broglie等,Science 224:838-843(1984);和Winter等,Results Probl.Cell Differ.17:85-105(1991))。可通过直接DNA转化或病原体介导的转染,将这些构建体导入植物细胞中。这类技术可参见大量普遍可得的综述(参见例如Hobbs,载于McGraw Hill Yearbook of Science and Technology,第191-196页(1992))。

[0423] 昆虫系统也可用于表达目标多肽。例如,在一个这样的系统中,将苜蓿银纹夜蛾核型多角体病毒(Autographa californica nuclear polyhedrosis virus,AcNPV)用作载体,以在草地夜蛾(Spodoptera frugiperda)细胞或粉纹夜蛾(Trichoplusia)幼虫中表达外源基因。多肽编码序列可克隆至该病毒的非必需区,例如多角体蛋白基因,并处于多角体蛋白启动子的控制之下。多肽编码序列的成功插入,将会使多角体蛋白基因失活,并产生缺乏外壳蛋白的重组病毒。然后将重组病毒用于感染例如草地夜蛾细胞或粉纹夜蛾幼虫,其中可表达目标多肽(Engelhard等,Proc.Natl.Acad.Sci.U.S.A.91:3224-3227(1994))。

[0424] 在哺乳动物宿主细胞中,多种基于病毒的表达系统为通常可得到的。例如,在将腺病毒用作表达载体的情况下,可将编码目标多肽的序列连接到由晚期启动子和三联前导序列组成的腺病毒转录/翻译复合体中。在病毒基因组的非必需E1或E3区中的插入,可用于获得在受感染的宿主细胞中能表达所述多肽的活病毒(Logan和Shenk,Proc.Natl.Acad.Sci.U.S.A.81:3655-3659(1984))。另外,转录增强子(例如劳斯肉瘤病毒(RSV)增强子)可用于增加在哺乳动物宿主细胞中的表达。采用腺病毒载体的方法和方案的综述参见Wold,Adenovirus Methods and Protocols,1998。关于腺病毒载体应用的其它参考文献可参见Adenovirus:A Medical Dictionary,Bibliography,and Annotated Research Guide to Internet References,2004。

[0425] 特定起始信号也可用于达到更有效地翻译目标多肽编码序列。这类信号包括ATG起始密码子和邻接序列。在多肽编码序列的情况下,将其起始密码子和上游序列插入到合适表达载体中,无需其它转录或翻译控制信号。然而,在仅插入编码序列或其部分的情况下,应提供外源翻译控制信号,其包括ATG起始密码子。此外,起始密码子应处于正确读框中,以确保翻译整个插入物。外源翻译元件和起始密码子可以来自不同来源,天然和合成的均可。可通过包含适于所用特定细胞系统的增强子来提高表达效率,例如文献中所述的增强子(Scharf等,Results Probl.Cell Differ.20:125-162(1994))。

[0426] 另外,可根据宿主细胞株以所需方式调节所插入序列的表达的能力或加工所表达蛋白的能力来对其进行选择。多肽的这类修饰包括但不限于乙酰化、羧化、糖基化、磷酸化、脂质化(lipidation)和酰化。可将切割蛋白的“前原(prepro)”形式的翻译后加工用于促进正确插入、折叠和/或行使功能。可选择对这类翻译后活性具有特定细胞机构(machinery)和特征性机制的不同宿主细胞(例如CHO、HeLa、MDCK、HEK293和WI38),以确保外源蛋白的正确修饰和加工。

[0427] 为了长期、高收率地制备重组蛋白,通常优选稳定表达。例如,稳定表达目标多核

苷酸的细胞系可用表达载体转化,所述表达载体可在相同或不同载体上含有病毒复制起点和/或内源表达元件和选择标记基因。导入载体后,可让细胞在富集培养基上生长1-2天,然后换到选择培养基上。选择标记的目的是赋予对选择的抗性,并且它的存在使可成功表达所导入序列的细胞生长和对其回收。稳定转化细胞的抗性克隆可使用适于该细胞类型的组织培养技术进行增殖。

[0428] 许多选择系统可用于回收转化细胞系。这些系统包括但不限于单纯疱疹病毒胸苷激酶(Wigler等,Cell 11:223-32(1977))基因和腺嘌呤磷酸核糖转移酶(Lowy等,Cell 22:817-23(1990))基因,这两者可分别用于 tk^- 或 $aprt^-$ 细胞。同样,针对抗代谢物、抗生素或除草剂的抗性也可用作选择基础;例如,赋予甲氨蝶呤抗性的 $dhfr$ (Wigler等,Proc.Natl.Acad.Sci.U.S.A.77:3567-70(1980));赋予氨基糖苷类、新霉素和G-418抗性的 npt (Colbere-Garapin等,J.Mol.Biol.150:1-14(1981));和分别赋予氯磺隆(chlorsulfuron)和膦丝菌素(phosphinotricin)乙酰转移酶抗性的 als 或 pat (Murry,见上文)。其它选择基因已有描述,例如允许细胞利用吡啶代替色氨酸的 $trpB$ 、或者允许细胞利用组氨酸(histidinol)代替组氨酸的 $hisD$ (Hartman和Mulligan,Proc.Natl.Acad.Sci.U.S.A.85:8047-51(1988))。近来,可见标记的使用得到普及,这样的标记例如花色素苷、 β -葡糖醛酸糖苷酶及其底物GUS、萤光素酶及其底物萤光素,不仅广泛用于鉴定转化体,而且广泛用于定量特定载体系统的瞬时或稳定蛋白表达的量(Rhodes等,Methods Mol.Biol.55:121-131(1995))。

[0429] 尽管标记基因表达的存在/不存在表明目标基因也存在,但它的存在和表达可能需要证实。例如,如果将编码多肽的序列插入到标记基因序列内,则可根据标记基因功能缺失,鉴定含有该序列的重组细胞。或者,标记基因可与多肽编码序列串联起来并处于一个启动子控制之下。标记基因响应诱导或选择的表达,通常也表示串联基因的表达。

[0430] 或者,可通过各种本领域技术人员已知的方法鉴定含有并表达所需多核苷酸序列的宿主细胞。这些方法包括但不限于DNA-DNA或DNA-RNA杂交和蛋白生物测定或免疫测定技术,其包括基于膜、溶液或芯片的技术,用于检测和/或定量测定核酸或蛋白。

[0431] 使用对产物具有特异性的多克隆或单克隆抗体来检测和测定多核苷酸编码产物的表达的许多方案,都为本领域已知。实例包括酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)和荧光激活细胞分选术(FACS)。对于某些应用来说,可能优选两位点的基于单克隆的免疫测定法(使用对既定多肽上的两个非干扰表位具有反应性的单克隆抗体),但也可使用竞争性结合测定法。这些测定和其它测定尤其描述于Hampton等,Serological Methods,a Laboratory Manual(1990)和Maddox等,J.Exp.Med.158:1211-1216(1983)。

[0432] 多种标记和缀合技术为本领域技术人员已知,并且可用于各种核酸和氨基酸测定。用于产生用于检测多核苷酸相关序列的经标记的杂交或PCR探针的方法包括寡核苷酸标记、切口平移、末端标记或使用标记核苷酸的PCR扩增。或者,可将序列或其任何部分克隆到载体中,用于产生mRNA探针。这类载体是本领域已知的,是市售的,并且可通过添加适当的RNA聚合酶(例如T7、T3或SP6)和标记的核苷酸,用于体外合成RNA探针。可用各种市售试剂盒实施这些方法。可使用的合适报道分子或标记包括放射性核素、酶、荧光剂、化学发光剂或生色剂及底物、辅因子、抑制剂、磁性颗粒等。

[0433] 可将用目标多核苷酸序列转化的宿主细胞在适于表达和从细胞培养物中回收蛋

白的条件下培养。重组细胞产生的蛋白可以是分泌型的或者包含在细胞内,这视所用序列和/或载体而定。本领域技术人员应理解的是,含有多核苷酸的表达载体可经设计,以包含指导所编码多肽分泌通过原核或真核细胞膜的信号序列。其它的重组构建体可用于使编码目标多肽的序列与编码多肽结构域的核苷酸序列连接起来,所述多肽结构域促进可溶性蛋白的纯化。这样的促进纯化的结构域包括但不限于允许在固定化金属上纯化的金属螯合肽例如组氨酸-色氨酸组件、允许在固定化免疫球蛋白上纯化的蛋白A结构域、以及用于FLAG延伸/亲和纯化系统(Immunex Corp., Seattle, Wash.)的结构域。可使用在纯化结构域与所编码多肽之间包含的可切割接头序列(例如对因子XA或肠激酶具有特异性的序列(Invitrogen, San Diego, Calif.)),以促进纯化。将一种这类表达载体用于以下融合蛋白的表达:所述融合蛋白包含目标多肽和位于硫氧还蛋白或肠激酶切割位点之前的核酸编码的6个组氨酸残基。组氨酸残基有助于在IMAC(固定化金属离子亲和色谱)上进行纯化,参见Porath等, *Prot. Exp. Purif.* 3:263-281(1992),而肠激酶切割位点则提供用于从融合蛋白中纯化所需多肽的方法。有关含有融合蛋白的载体的论述参见Kroll等, *DNA Cell Biol.* 12:441-453(1993))。

[0434] 体内多核苷酸递送技术

[0435] 在另外的实施方案中,将包含本发明的一个或多个多核苷酸的遗传构建体体内导入细胞内。这可用多种众所周知的方法中的任一种来完成,其中的一些方法概述如下,用于说明的目的。

[0436] 1. 腺病毒

[0437] 用于体内递送一个或多个核酸序列的优选方法中的一种包括使用腺病毒表达载体。“腺病毒表达载体”意指包括含足以执行以下功能的腺病毒序列的构建体:(a)支持构建体的包装和(b)表达已按有义或反义方向克隆在其中的多核苷酸。当然,在反义构建体的情况下,表达并不需要合成该基因产物。

[0438] 表达载体包括腺病毒的基因工程改造形式。腺病毒遗传构成的知识(36kb,线状双链DNA病毒)允许用多达7kb的外源序列备选大片段腺病毒DNA(Grunhaus和Horwitz,1992)。与反转录病毒相反,腺病毒感染宿主细胞不导致染色体整合,因为腺病毒DNA可以附加体方式复制,没有潜在基因毒性。此外,腺病毒结构稳定,在大量扩增后没有检出基因组重排。腺病毒可感染几乎所有上皮细胞,与其细胞周期阶段无关。迄今为止,腺病毒感染看来仅引起人的轻微疾病,例如急性呼吸道疾病。

[0439] 腺病毒尤其适用于作为基因传递载体,这是因为其基因组大小适中、易操作、滴度高、靶细胞范围广、侵染性高。病毒基因组两端含有100-200个碱基对反向重复序列(ITR),其是病毒DNA复制和包装所需的顺式元件。基因组早期(E)和晚期(L)区含有不同转录单位,它们按病毒DNA复制开始来划分。E1区(E1A和E1B)编码负责调节病毒基因组和少量细胞基因转录的蛋白。E2区(E2A和E2B)的表达导致合成用于病毒DNA复制的蛋白。这些蛋白涉及DNA复制、晚期基因表达和宿主细胞关闭(Renan,1990)。包括大部分病毒衣壳蛋白在内的晚期基因产物仅在由主要晚期启动子(MLP)引起的对单一初级转录物的明显加工之后才表达。MLP(位于16.8m.u.)在感染晚期特别有效,并且该启动子产生的所有mRNA都具有5'-三联前导(TPL)序列,该序列使它们成为翻译的优选mRNA。

[0440] 在目前的系统中,重组腺病毒由穿梭载体与原病毒载体之间同源重组产生。因为

两个原病毒载体之间可能重组,所以该过程可能产生野生型腺病毒。因此,从单个噬斑分离出病毒单克隆并检查其基因组结构是至关重要的。

[0441] 目前复制缺陷型腺病毒载体的产生和增殖,取决于独特的辅助细胞系(称为293),其由Ad5 DNA片段转化自人胚肾细胞,并且能组成型地表达E1蛋白(Graham等,1977)。因为E3区并非腺病毒基因组所必需的(Jones和Shenk,1978),所以目前的腺病毒载体,在293细胞辅助下,在E1、D3或这两个区携带外源DNA(Graham和Prevec,1991)。在自然界,腺病毒可包装约105%的野生型基因组(Ghosh-Choudhury等,1987),提供约2kB的额外DNA容量。与在E1和E3区可替代的约5.5kB DNA组合,目前腺病毒载体的最大容量小于7.5kB,或约载体总长的15%。超过80%腺病毒基因组保留在载体主链中,是载体带有的细胞毒性的来源。而且,缺失E1的病毒的复制缺陷是不完全的。例如,在高感染复数(MOI)下,用目前可用载体已观察到病毒基因表达的渗漏(Mulligan,1993)。

[0442] 辅助细胞系可来自人类细胞,例如人胚肾细胞、肌细胞、造血细胞或其它人胚胎间充质细胞或上皮细胞。或者,辅助细胞可来源于与人腺病毒相容的其它哺乳动物细胞。这类细胞包括例如Vero细胞或其它猴胚胎间充质细胞或上皮细胞。如上所述,目前优选的辅助细胞系是293。

[0443] Racher等(1995)公开了用于培养293细胞和增殖腺病毒的改良方法。在一种形式中,通过将单个细胞群接种到含100-200ml培养基的1升硅化摇瓶(siliconised spinner flask)(Techne, Cambridge, UK)中,使天然细胞聚集体生长。在40rpm搅拌后,用锥虫蓝评价细胞生存力。在另一种形式中,将Fibra-Cel微载体(Bibby Sterlin, Stone, UK)(5g/l)如下使用。在250ml锥形瓶中,将重悬于5ml培养基中的细胞接种物加入到载体(50ml)中,静置1-4小时,偶尔搅拌。然后将培养基用50ml新鲜培养基更换,然后振摇。为了病毒制备,让细胞生长至约80%汇合,然后更换培养基(至25%的终体积)并以MOI为0.05加入腺病毒。让培养物静置过夜,然后将体积增加至100%并再开始振摇72h。

[0444] 除了需要腺病毒载体是复制缺陷型或者至少是条件缺陷型之外,认为腺病毒载体的特性对本发明的成功实施来说并非至关重要。腺病毒可为42种不同已知血清型或A-F亚群中的任一种。腺病毒C亚群5型为获得用于本发明的条件复制缺陷型腺病毒载体的优选原料,因为腺病毒5型是人腺病毒,关于其的许多生化 and 遗传信息是已知的,而且其曾用于大多数以腺病毒为载体的构建中。

[0445] 如上所述,本发明的典型载体是复制缺陷型且没有腺病毒E1区。因此,最合宜的是将编码目标基因的多核苷酸引入到已去除E1-编码序列的位置上。然而,腺病毒序列内构建体的插入位置对本发明而言并非至关重要。编码目标基因的多核苷酸也可插入到如Karlsson等(1986)所述的E3置换型载体中取替缺失的E3区或其中辅助细胞系或辅助病毒能弥补E4缺损的E4区中。

[0446] 腺病毒容易培养和操作,体外和体内均具有广泛宿主范围。该类病毒可以高滴度获取,例如 10^9 - 10^{11} 噬斑形成单位/ml,而且是高感染性的。腺病毒生命周期无需整合到宿主细胞基因组中。腺病毒载体所递送的外源基因是附加体,因此对宿主细胞的基因毒性低。在野生型腺病毒接种研究中,未报道有副作用(Couch等,1963;Top等,1971),证明其作为体内基因转移载体具有安全性和治疗潜力。

[0447] 腺病毒载体已用于真核基因表达(Levrero等,1991;Gomez-Foix等,1992)和疫苗

开发(Grunhaus和Horwitz,1992;Graham和Prevec,1992)。近来,动物研究表明重组腺病毒可用于基因疗法(Stratford-Perricaudet和Perricaudet,1991;Stratford-Perricaudet等,1990;Rich等,1993)。将重组腺病毒给予不同组织的研究包括气管滴注(Rosenfeld等,1991;Rosenfeld等,1992)、肌注(Ragot等,1993)、外周静脉注射(Herz和Gerard,1993)和立体定向接种(stereotactic inoculation)到脑部(Le Gal La Salle等,1993)。

[0448] 腺病毒载体可来源于人腺病毒。或者,它们可来源于诸如黑猩猩等其它物种的腺病毒,其优点是病毒载体不被许多人类受试者中循环的针对人腺病毒的抗体所中和(参见例如:Tatsis N等Gene Therapy 2006 13:421-429)。

[0449] 腺病毒35型(其相对不常见,因此存在低水平的预先存在的对该载体本身的免疫)已用作某些正在开发中的结核病疫苗的递送系统(参见,例如Radosevic等Infection and Immunity 2007 75(8):4105-4115)。腺病毒35型在本发明中作为递送载体还可具有特殊价值。

[0450] 2. 反转录病毒

[0451] 反转录病毒是一类单链RNA病毒,特征在于以下能力:在感染细胞中通过反转录过程能将其RNA转变成双链DNA(Coffin,1990)。然后所得DNA稳定整合在细胞染色体中成为原病毒并指导病毒蛋白合成。所述整合导致在受体细胞及其后代中保留病毒基因序列。反转录病毒基因组含有gag、pol和env三个基因,它们分别编码衣壳蛋白、聚合酶和包膜组分。在gag基因上游发现的序列含有将基因组包装在病毒体中的信号。两个长末端重复(LTR)序列存在于病毒基因组的5'和3'端。它们含有强启动子和增强子序列,也是整合到宿主细胞基因组中所需要的(Coffin,1990)。

[0452] 为了构建反转录病毒载体,将编码一个或多个目标寡核苷酸或多核苷酸序列的核酸插入到病毒基因组内某些病毒序列的位置上,以产生复制缺陷型病毒。为了产生病毒体,构建含有gag、pol和env基因、但不含LTR和包装组分的包装细胞系(Mann等,1983)。当将含有cDNA以及反转录病毒LTR和包装序列的重组质粒导入该细胞系(通过例如磷酸钙沉淀)时,包装序列使重组质粒的RNA转录物包装到病毒颗粒中,病毒颗粒随后被分泌到培养基中(Nicolas和Rubenstein,1988;Temin,1986;Mann等,1983)。然后收集含重组反转录病毒的培养基,任选浓缩,然后用于基因转移。反转录病毒载体能够感染各种各样的细胞类型。然而,整合和稳定表达则需要宿主细胞分裂(Paskind等,1975)。

[0453] 最近,根据通过使乳糖残基经化学法添加到病毒包膜上而对反转录病毒进行化学修饰,开发了用于允许反转录病毒载体特异性靶向的新方法。该修饰使得可通过唾液酸糖蛋白受体特异性感染肝细胞。

[0454] 设计了重组反转录病毒靶向的不同方法,其中使用针对反转录病毒包膜蛋白和特异性细胞受体的生物素化抗体。经使用链霉抗生物素,该抗体通过生物素组分进行偶联(Roux等,1989)。Roux等人使用针对I类和II类主要组织相容性复合体抗原的抗体,表明携带这些表面抗原的各种人类细胞在体外可被同向性病毒感染(Roux等,1989)。

[0455] 3. 腺伴随病毒

[0456] AAV(Ridgeway,1988;Hermonat和Muzycska,1984)是细小病毒(parovirus),其被发现为腺病毒原液的污染物。其是与任何疾病无关的遍在病毒(美国人口中有85%都存在抗体)。其也被分类为依赖病毒,因为其复制依赖于诸如腺病毒等辅助病毒的存在。已经分

离出5种血清型,其中已充分表征了AAV-2。AAV具有单链线状DNA,其包裹在衣壳蛋白VP1、VP2和VP3中,形成直径20-24nm的二十面体病毒体(Muzyczka和McLaughlin,1988)。

[0457] AAV DNA长约4.7千碱基。其包含两个可读框并位于两个ITR的侧翼。AAV基因组中有两个主要基因:rep和cap.rep基因编码负责病毒复制的蛋白,而cap编码衣壳蛋白VP1-3。每个ITR形成T-形发夹结构。这些末端重复序列是对AAV的染色体整合唯一必需的顺式组分。因此,AAV可用作载体,其中所有病毒编码序列都被除去并用递送基因盒替代。根据图谱位置,鉴定并命名了p5、p19和p40三个病毒启动子。从p5和p19转录导致产生rep蛋白,而从p40转录产生衣壳蛋白(Hermonat和Muzyczka,1984)。

[0458] 有若干因素促使研究人员研究使用rAAV作为表达载体的可行性。一个因素就是对递送基因以整合到宿主染色体的要求出奇地少。需要具有145bp ITR,其仅占AAV基因组的6%。这在载体中留下装配4.5kb DNA插入物的空间。尽管这样的携带容量可阻止AAV递送大基因,但是这完全适合递送反义构建体。

[0459] 由于AAV的安全性,所以其也是递送载体的良好选择。存在相对复杂的拯救机制:为了动员rAAV,不仅需要野生型腺病毒而且需要AAV基因。而且,AAV为无致病性并与任何疾病无关。除去病毒编码序列使对病毒基因表达的免疫反应最小化,因此,rAAV不会引起炎症反应。

[0460] 4. 作为表达构建体的其它病毒载体

[0461] 其它病毒载体可用作本发明的表达构建体,用于将寡核苷酸或多核苷酸序列递送给宿主细胞。可使用来源于诸如痘苗病毒(Ridgeway,1988;Coupar等,1988)、慢病毒、脊髓灰质炎病毒和疱疹病毒等病毒的载体。也可预期源于其它痘病毒的载体例如源于禽痘的载体为有用的。它们给各种哺乳动物细胞带来一些有吸引力的特征(Friedmann,1989;Ridgeway,1988;Coupar等,1988;Horwich等,1990)。

[0462] 随着目前对缺陷型乙肝病毒的认识,获得了有关不同病毒序列的结构-功能关系的新认识。体外研究表明,病毒可保留辅助-依赖性包装和反转录的能力,即使其基因组缺失高达80%(Horwich等,1990)。这表明基因组的大部分都可用外源遗传物质替换。趋肝性和持久性(整合)对于肝定向基因转移而言是特别有吸引力的性质。Chang等(1991)将氯霉素乙酰转移酶(CAT)基因在聚合酶编码序列、表面编码序列和前表面编码序列的位置导入鸭乙肝病毒基因组中。将其与野生型病毒共转染到禽类肝癌细胞系中。用含高滴度重组病毒的培养基来感染原代小鸭肝细胞。转染后持续至少24天检测到稳定的CAT基因表达(Chang等,1991)。

[0463] 另外的“病毒”载体包括病毒样颗粒(VLP)和噬菌体。

[0464] 5. 非病毒载体

[0465] 为了实现本发明寡核苷酸或多核苷酸序列的表达,必须将表达构建体递送到细胞中。可在体外(如在用于转化细胞系的实验室方法中)或者在体内或离体(如在某些疾病状态的治疗中)完成这样的递送。如上所述,一种优选的递送机制是通过病毒感染,其中表达构建体包裹在感染性病毒颗粒之内。

[0466] 一旦将表达构建体递送到细胞中,编码所需寡核苷酸或多核苷酸序列的核酸可在不同位置定位和表达。在某些实施方案中,编码构建体的核酸可稳定整合到细胞基因组中。这样的整合可通过同源重组(基因置换)处于特定位置和方向,或者其可整合在随机的非特

异性位置上(基因增强)。在再一个实施方案中,核酸可以稳定保留在细胞中,作为单独的附加体DNA区段。这类核酸区段或“附加体”编码以下序列:其足以允许独立于或同步于宿主细胞周期的保留和复制。表达构建体如何递送给细胞以及核酸保留在细胞的何处,取决于所用表达构建体的类型。

[0467] 在本发明的某些实施方案中,包含一个或多个寡核苷酸或多核苷酸序列的表达构建体可仅由裸的重组DNA或质粒构成。构建体的转移可例如通过以物理法或化学法透过细胞膜的任何方法来进行。这尤其适用于体外转移,但也可用于体内使用。Dubensky等(1984)成功地将多瘤病毒DNA以磷酸钙沉淀形式注入成年小鼠和新生小鼠肝和脾中,所述小鼠表现出活动性病毒复制和急性感染。Benvenisty和Reshef(1986)也表明直接腹膜内注射磷酸钙沉淀的质粒,导致转染基因的表达。预期也可在体内以类似方式转移编码目标基因的DNA并表达基因产物。

[0468] 用于将裸DNA表达构建体转入细胞的本发明的另一个实施方案涉及粒子轰击。该方法依赖于使包被DNA的微粒加速到高速而让其穿过细胞膜和进入细胞而不杀死细胞的能力(Klein等,1987)。已经开发了用于加速小颗粒的一些装置。一个这样的装置依靠高压放电产生电流,这样的电流又提供原动力(Yang等,1990)。所用的微粒由诸如钨珠或金珠等生物惰性物质组成。

[0469] 已经对包括大鼠和小鼠的肝、皮肤和肌肉组织在内的所选器官进行体内轰击(Yang等,1990;Zelenin等,1991)。这可能需要经手术暴露组织或细胞,以消除枪与靶器官之间的任何中间组织,即离体处理。此外,编码特定基因的DNA可通过该方法递送,这也包括在本发明中。

[0470] 细菌也可用作递送手段(例如李斯特菌,参见W02004/11048),尤其是BCG。

[0471] 多肽组合物

[0472] 通常,用于本发明中的多肽应为分离的多肽(即与其通常与之一起存在于自然中的组分分离)。

[0473] 例如,如果天然存在的蛋白与天然系统中共同存在的某些或所有物质分离,则所述蛋白为分离的。优选这类多肽的纯度为至少约90%,更优选至少约95%,最优选至少约99%。如果多核苷酸例如被克隆到并非其天然环境的一部分的载体中时,则认为其为分离的。

[0474] 多肽可用任何众所周知的技术来制备。如上所述的DNA序列编码的重组多肽可容易利用本领域普通技术人员已知的任何表达载体从该DNA序列制备。可在已用含编码重组多肽的DNA分子的表达载体转化或转染的任何合适宿主细胞中进行表达。合适宿主细胞包括原核细胞、酵母细胞和高等真核细胞,例如哺乳动物细胞和植物细胞。优选使用的宿主细胞是大肠杆菌、酵母菌或哺乳动物细胞系例如COS或CHO。可以先用市售滤器浓缩来自合适宿主/载体系统的上清液,所述合适宿主/载体系统分泌重组蛋白或多肽至培养基中。浓缩后,可将浓缩物施加至合适的纯化基质,例如亲和基质或离子交换树脂。最后,可使用一个或多个反相HPLC步骤,以进一步纯化重组多肽。

[0475] 使用本领域普通技术人员熟知的技术,也可通过合成方法产生少于约100个氨基酸、通常少于约50个氨基酸的本发明多肽、其免疫原性片段和其它变体。例如,可用任何市售固相技术合成这样的多肽,所述技术为例如Merrifield固相合成法,其中将氨基酸序贯

添加到延长中的氨基酸链上。参见Merrifield, J. Am. Chem. Soc. 85:2149-2146(1963)。用于自动合成多肽的仪器市售自供应商例如Perkin Elmer/Applied BioSystems Division (Foster City, CA), 并可按照生产商的说明书进行操作。

[0476] 在某些具体的实施方案中, 多肽可以是包含如本文所述的多种多肽的融合蛋白, 或者是包含如本文所述的至少一种多肽和无关序列的融合蛋白, 所述蛋白的实例包括破伤风蛋白、结核病蛋白和肝炎蛋白(参见, 例如Stoute等., New Engl. J. Med. 336:86-91(1997))。融合配偶体可例如有助于提供T辅助细胞表位(免疫融合配偶体)、优选由人识别的T辅助细胞表位, 或者可有助于以高于天然重组蛋白的收率表达蛋白(表达增强子)。某些优选的融合配偶体既是免疫融合配偶体, 也是表达增强型融合配偶体。可选择其它融合配偶体, 以增加蛋白溶解度或者使蛋白能够靶向所需胞内区室。再一些融合配偶体包括便于蛋白纯化的亲和标记。

[0477] 通常可采用包括化学缀合在内的标准技术制备融合蛋白。优选在表达系统中以重组蛋白的形式表达融合蛋白, 使得相对于非融合蛋白制备水平提高。简而言之, 编码多肽组分的DNA序列可分别装配, 然后连接到合适表达载体中。将编码一种多肽组分的DNA序列的3'端, 通过或不通过肽接头, 与编码第二种多肽组分的DNA序列的5'端连接, 使得该序列的读框相同(in phase)。这允许翻译成保留两种组成多肽的生物活性的单一融合蛋白。

[0478] 可使用肽接头序列以将第一和第二种多肽组分间隔至足以保证每种多肽折叠成其二级和三级结构的间距。使用本领域众所周知的标准技术将这类肽接头序列掺入到融合蛋白中。可根据以下因素选择合适的肽接头序列:(1)其能够采用柔性延伸构象;(2)其不能采用会与第一和第二种多肽功能性表位相互作用的二级结构;和(3)缺乏可与多肽功能性表位反应的疏水性或带电荷残基。优选的肽接头序列含有Gly、Asn和Ser残基。其它接近中性氨基酸(例如Thr和Ala)也可用于接头序列。可有用地用作接头的氨基酸序列包括以下文献中公开的序列:Maratea等, Gene 40:39-46(1985);Murphy等, Proc. Natl. Acad. Sci. USA 83:8258-8262(1986);美国专利号4,935,233和美国专利号4,751,180。接头序列的长度通常可为1个至约50个氨基酸。当第一和第二种多肽具有可用于间隔开功能性结构域和防止位阻的非必需N-端氨基酸区时, 则无需接头序列。

[0479] 在优选实施方案中, 免疫融合配偶体来源于蛋白D, 一种革兰氏阴性菌嗜血流感杆菌B(*Haemophilus influenza B*)的表面蛋白(WO 91/18926)。优选蛋白D衍生物包含该蛋白的约前三分之一(例如N-端的前100-110个氨基酸), 并且蛋白D衍生物可被脂质化。在某些优选实施方案中, 将脂蛋白D融合配偶体的前109个残基包含在N-端, 以给多肽提供额外的外源T细胞表位和提高在大肠杆菌中的表达水平(因此起表达增强子的作用)。脂质尾部确保将抗原最佳呈递给抗原呈递细胞。其它融合配偶体包括来自流感病毒的非结构蛋白NS1(血凝素)。通常使用N-端81个氨基酸, 但也可使用包含T辅助细胞表位的不同片段。

[0480] 在另一个实施方案中, 免疫融合配偶体是称为LYTA的蛋白或其部分(优选C端部分)。LYTA来自肺炎链球菌(*Streptococcus pneumoniae*), 其合成称为酰胺酶LYTA的N-乙酰基-L-丙氨酸酰胺酶(由LytA基因编码;Gene 43:265-292(1986))。LYTA是一种特异性降解肽聚糖主链中的某些键的自溶素。LYTA蛋白C端结构域负责对胆碱或某些胆碱类似物例如DEAE的亲和力。该特性已经用于开发用于表达融合蛋白的大肠杆菌C-LYTA表达质粒。已描述在氨基端含有C-LYTA片段的杂合蛋白的纯化(参见Biotechnology 10:795-798(1992))。

在一个优选的实施方案中,可将LYTA的重复部分掺入到融合蛋白中。在C端区自残基178开始发现重复部分。特别优选的重复部分包括残基188-305。

[0481] T细胞

[0482] 免疫治疗组合物可另外或备选地包含对分枝杆菌抗原具有特异性的T细胞。通常可用标准方法在体外或离体制备这类细胞。例如,可使用市售细胞分离系统(例如Isolex™ System,得自Nexell Therapeutics, Inc.(Irvine, CA; 另见美国专利号5,240,856; 美国专利号5,215,926; WO 89/06280; WO 91/16116和WO 92/07243),从患者的骨髓、外周血或者骨髓或外周血部分中分离出T细胞。或者, T细胞可来源于相关或无关的人、非人类哺乳动物、细胞系或培养物。

[0483] 可用本发明多肽、编码所述多肽的多核苷酸、和/或表达所述多肽的抗原呈递细胞(APC)刺激T细胞。在足以允许产生对多肽具有特异性的T细胞的条件下和时间内进行这样的刺激。优选多肽或多核苷酸存在于递送载体(例如微球体)中,以促进产生特异性T细胞。

[0484] 如果T细胞特异性增殖、分泌细胞因子或杀伤用本发明多肽包被或表达本发明多肽编码基因的靶细胞,则认为所述T细胞对所述多肽具有特异性。可采用任何标准技术来评价T细胞特异性。例如,在铬释放测定或增殖测定中,与阴性对照相比,在溶解和/或增殖上刺激指数增加超过两倍,表明T细胞特异性。可例如如以下文献中所述进行这些测定:Chen等, Cancer Res. 54:1065-1070(1994)。或者,可通过各种已知技术进行T细胞增殖的测定。例如,可通过测定DNA合成的速率增加来检测T细胞增殖(例如通过用氘化胸苷脉冲标记T细胞培养物并测定掺入DNA的氘化胸苷的量)。与本发明多肽(100ng/ml-100μg/ml、优选200ng/ml-25μg/ml)接触3-7天,应导致T细胞增殖增加至少两倍。如标准细胞因子测定法所测,如上所述接触2-3小时应导致T细胞活化,所述测定法中细胞因子(例如TNF或IFN-γ)释放水平提高两倍,指示T细胞活化(参见Coligan等, Current Protocols in Immunology, 第1卷(1998))。响应多肽、多核苷酸或表达多肽的APC而活化的T细胞可以是CD4⁺和/或CD8⁺。用标准技术可扩繁蛋白特异性T细胞。在优选的实施方案中, T细胞来源于患者、相关供体或无关供体,并且在刺激和扩繁后给予患者。

[0485] 为了治疗目的,可在体外或体内大量增殖响应多肽、多核苷酸或APC而增殖的CD4⁺或CD8⁺T细胞。可以不同方式进行这类T细胞的体外增殖。例如,可在添加或不添加T细胞生长因子(例如白介素-2)和/或合成多肽的刺激细胞下,使T细胞重新暴露至多肽或对应于该多肽的免疫原性部分的短肽。或者,可通过克隆,大量扩繁在蛋白存在时增殖的一种或多种T细胞。用于克隆细胞的方法为本领域所熟知并包括有限稀释。

[0486] 药物组合物

[0487] 在另外的实施方案中,将本文公开的多核苷酸、多肽、T细胞和/或抗体组合物配制成药学上可接受或药理学上可接受的溶液剂,用于单独或者与一种或多种其它治疗形式联合给予细胞或动物。

[0488] 还应理解的是,如果需要的话,也可将表达本文所公开多肽的核酸区段(例如RNA或DNA)与其它作用剂(agent)联合给予,所述其它作用剂为例如其它蛋白或多肽或各种药理学活性剂(包括有效抗结核分枝杆菌感染的化疗药)。事实上,对于也可包括在内的其它组分几乎没有限制,只要另外的作用剂在与靶细胞或宿主组织接触时不会引起明显不良作用即可。因此,在具体情况下,可视需要将组合物与各种其它作用剂一起递送。可将这类组合

物从宿主细胞或其它生物来源中纯化出来,或者可如本文所述化学合成这类组合物。同样,这类组合物还可包括取代或衍生的RNA或DNA组合物。

[0489] 药学上可接受的赋形剂和载体溶液剂的配制为本领域技术人员所熟知,其视合适的给药和治疗方案的开发而定,所述方案针对将本文所述的具体组合物用于各种治疗方案中,包括例如口服、胃肠外、静脉内、鼻内和肌内给药和制剂。其它给药途径包括经由粘膜表面。

[0490] 通常,包含治疗有效量的制剂每次给药递送约0.1 μ g至约1000 μ g多肽,更通常每次给药递送约2.5 μ g至约100 μ g多肽。对于多核苷酸组合物而言,这些制剂通常每次给药递送约10 μ g至约20mg本发明多核苷酸,更通常每次给药递送约0.1mg至约10mg本发明多核苷酸。

[0491] 无疑,各种治疗有用组合物中活性化合物的量可按以下方式制备:合适的剂量包含在任何既定单位剂量的化合物中。诸如溶解度、生物利用度、生物半寿期、给药途径、产品保存期以及其它药理学考虑等因素由制备这类药物制剂领域的技术人员考虑,因而多种剂量和治疗方案都可为理想的。

[0492] 1.口服递送

[0493] 在某些应用中,本文所公开的药物组合物可经口服递送至动物。因而,这些组合物可与惰性稀释剂或可同化的可食用载体配制在一起,或者可将它们包封在硬壳或软壳明胶胶囊中,或者可将它们压制成片剂,或者可将它们直接与食物结合。

[0494] 活性化合物还可与赋形剂结合,并以可吸收片剂、口含片剂、锭剂、胶囊剂、酏剂、混悬剂、糖浆剂、糯米纸囊剂等形式使用(Mathiowitz等,1997;Hwang等,1998;美国专利5,641,515;美国专利5,580,579和美国专利5,792,451,所述文献各自通过引用以其整体结合到本文中)。片剂、锭剂、丸剂、胶囊剂等也可含有以下成分:粘合剂,如西黄蓍胶、阿拉伯胶、玉米淀粉或明胶;赋形剂,例如磷酸二钙;崩解剂,例如玉米淀粉、马铃薯淀粉、海藻酸等;润滑剂,例如硬脂酸镁;和甜味剂(例如可添加蔗糖、乳糖或糖精)或者矫味剂(例如薄荷、冬青油或樱桃香料)。当剂量单位形式是胶囊剂时,除了上述类型的材料之外,其还可含有液体载体。可存在各种其它材料,作为包衣或以另外改变剂量单位的物理形式。例如,片剂、丸剂或胶囊剂可用虫胶、糖或两者同时来包衣。酏剂糖浆可含有活性组分、作为甜味剂的蔗糖、作为防腐剂的羟苯基苯甲酸甲酯和对羟基苯甲酸丙酯、染色剂和矫味剂,例如樱桃或橘子香料。当然用于制备任何剂量单位形式的任何材料都应该是药物纯的且在所采用的量下基本无毒。另外,活性组分可掺入缓释制品和制剂中。

[0495] 为了口服,本发明组合物备选地可与一种或多种赋形剂以以下形式结合:漱口剂、洁齿剂、口含片剂、口腔喷雾剂或舌下口服制剂。例如,可将所需量的活性成分掺入到合适溶剂(例如硼酸钠溶液(Dobell液))中,制备漱口剂。或者,可将活性成分掺入到口服溶液剂(例如含硼酸钠、甘油和碳酸氢钾的溶液剂)或分散到洁齿剂中,或者以治疗有效量添加到以下组合物中:所述组合物可包含水、粘合剂、研磨剂、矫味剂、发泡剂和保湿剂。或者,可将组合物制成可置于舌下或在口腔内溶解的片剂或溶液剂形式。

[0496] 2.注射递送

[0497] 在某些情况下,理想的是经胃肠外、静脉内、肌内、皮内或甚至腹膜内递送本文所公开的药物组合物,如美国专利5,543,158、美国专利5,641,515和美国专利5,399,363中所述(所述各种文献通过引用以其整体结合到本文中)。可在适当混有表面活性剂(例如羟丙

基纤维素)的水中制备活性化合物作为游离碱或药理学上可接受的盐的溶液剂。也可在甘油、液态聚乙二醇及其混合物和油中制备分散剂。在普通贮藏和使用条件下,这些制剂含有防腐剂以防止微生物生长。

[0498] 适于注射用的药物形式包括无菌水溶液剂或分散剂以及用于临用时配制无菌注射溶液剂或分散剂的无菌粉末(美国专利5,466,468,通过引用以其整体具体结合到本文中)。在所有情况下,所述形式都必须无菌且必须在流动性上达到容易注射(syringability)的程度。其在制备和贮藏条件下必须稳定,而且必须以防止诸如细菌和真菌等微生物的污染作用的方式贮存。载体可以是溶剂或分散介质,包括例如水、乙醇、多元醇(例如甘油、丙二醇和液态聚乙二醇等)、其适当的混合物、和/或植物油。可例如通过使用包衣(例如卵磷脂)、通过保持所需粒度(在分散剂的情况下)以及通过使用表面活性剂,维持适当流动性。可通过各种抗菌药和抗真菌剂促进对微生物作用的预防,例如对羟基苯甲酸酯类、氯丁醇、苯酚、山梨酸、硫柳汞等。在许多情况下,优选包含等渗剂,例如糖或氯化钠。可通过在组合物中使用延迟吸收剂(例如单硬脂酸铝和明胶)延迟对注射用组合物的吸收。

[0499] 对于在水溶液剂中胃肠外给予而言,例如必要时溶液剂应适当缓冲,并且液体稀释剂先用足够盐水或葡萄糖赋予等渗性。这些特定水溶液剂尤其适用于静脉内、肌内、皮下和腹膜内给予。就此而论,根据本公开内容,可使用的无菌含水介质为本领域技术人员所知。例如,可将一个剂量溶于1ml等渗NaCl溶液中,然后加入到1000ml皮下灌注液中或者注射在预定输注部位(参见例如Remington's Pharmaceutical Sciences,第15版,第1035-1038和1570-1580页)。根据所治疗患者的状况,需要对剂量作出某些改动。负责给药的人员会在任何情况下确定个体受试者的合适剂量。此外,对于人类给药而言,制剂应当满足FDA生物制品部门标准所要求的无菌、热源性以及总体安全性和纯度标准。

[0500] 可按照需要,将所需量的活性化合物掺入具有上文列举的各种其它成分的合适溶剂中,接着过滤除菌,制备无菌注射用溶液剂。通常,通过将各种灭过菌的活性成分掺入到含有基础分散介质和上述其它所需成分的无菌溶媒中,制备分散剂。在用于制备无菌注射用溶液剂的无菌粉末的情况下,优选的制备方法是真空干燥和冷冻干燥技术,所述技术从其先前的无菌过滤溶液中得到活性成分和任何其它所需成分的粉末。

[0501] 本文所公开的组合物可配制成中性或盐形式。药学上可接受的盐,包括酸加成盐(与蛋白的游离氨基生成的)和与无机酸(例如盐酸或磷酸)生成的盐,或者与有机酸(例如乙酸、草酸、酒石酸、扁桃酸等)生成的盐。与游离羧基生成的盐也可衍生自无机碱(例如氢氧化钠、氢氧化钾、氢氧化铵、氢氧化钙或氢氧化铁)和有机碱(例如异丙胺、三甲胺、组氨酸、普鲁卡因等)。在配制后,就以剂量制剂相容的方式和以治疗有效量给予溶液剂。容易以各种剂型给予制剂,所述剂型例如注射用溶液剂、药物释放胶囊剂等。

[0502] 本文所用的“载体”包括任何和所有溶剂、分散介质、赋形剂(溶媒)、包衣剂、稀释剂、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂、缓冲剂、载体溶液剂、混悬剂、胶体等。这些介质和作用剂在药物活性物质中的应用为本领域所熟知。除非任何常规介质或作用剂与活性成分不相容,否则考虑其在治疗性组合物中的应用。辅助活性成分也可掺入组合物中。

[0503] 措词“药学上可接受的”是指当给予人时不会产生变态反应或类似不利反应的分子实体和组合物。本领域充分了解含作为活性成分的蛋白的含水组合物的制备。通常将这

类组合物制备成注射用液体溶液剂或混悬剂;也可制备成在注射前适于溶解或悬浮于液体中的固体形式。也可使制剂乳化。

[0504] 3.鼻腔和口腔递送

[0505] 在某些实施方案中,可通过鼻内喷雾、口腔喷雾、吸入和/或其它气雾剂递送载体而给予药物组合物。用于将基因、核酸和肽组合物直接递送至肺部的方法(例如通过鼻腔和口腔喷雾剂)已描述于例如美国专利5,756,353和美国专利5,804,212(所述各文献通过引用以其整体结合到本文中)。同样,利用鼻内微粒树脂(Takenaga等,1998)和溶血磷脂酰-甘油化合物(美国专利5,725,871,通过引用以其整体明确结合到本文中)递送药物,也为药学领域所熟知。同样,以聚四氟乙烯(polytetrafluoroethylene)支持基料形式的跨粘膜给药描述于美国专利5,780,045(通过引用以其整体明确结合到本文中)。

[0506] 4.脂质体、纳米囊和微粒介导的递送

[0507] 在某些实施方案中,本发明人考虑了脂质体、纳米囊、微粒、微球体、脂质粒、囊泡等在用于将本发明组合物导入合适宿主细胞中的用途。具体地讲,可将本发明组合物配制用于包囊于脂质粒、脂质体、囊泡、纳米球或纳米粒等中递送。

[0508] 可优选这样的制剂用于引入本文所公开的核酸或构建体的药学上可接受制剂。脂质体的形成和使用是本领域技术人员公知的(参见例如Couvreur等,1977;Couvreur,1988;Lasic,1998;所述文献描述了在针对胞内细菌感染和疾病的靶向抗生素治疗中脂质体和纳米囊的用途)。近来,开发出具有改善的血清稳定性和循环半衰期的脂质体(Gabizon和Papahadjopoulos,1988;Allen和Choun,1987;美国专利5,741,516,所述文献通过引用以其整体明确结合到本文中)。此外,有关脂质体和脂质体样制剂作为潜在药物载体的各种方法已有综述(Takakura,1998;Chandran等,1997;Margalit,1995;美国专利5,567,434;美国专利5,552,157;美国专利5,565,213;美国专利5,738,868和美国专利5,795,587,所述各文献通过引用以其整体明确结合到本文中)。

[0509] 脂质体已成功用于通常难以通过其它方法进行转染的大量细胞类型,包括T细胞悬物、原代肝细胞培养物和PC 12细胞(Renneisen等,1990;Muller等,1990)。另外,脂质体无DNA长度限制,而这通常在基于病毒的递送系统中会受到限制。脂质体已有效用于将基因、药物(Heath和Martin,1986;Heath等,1986;Balazsovits等,1989;Fresta和Puglisi,1996)、放射性治疗剂(Pikul等,1987)、酶(Imaizumi等,1990a;Imaizumi等,1990b)、病毒(Faller和Baltimore,1984)、转录因子和别构剂(Nicolau和Gersonde,1979)导入各种培养细胞系和动物体内。另外,已经完成了检查脂质体介导药物递送的有效性的若干成功临床试验(Lopez-Berestein等,1985a;1985b;Coune,1988;Sculier等,1988)。此外,一些研究表明,使用脂质体与系统给药后的自身免疫应答、毒性或生殖腺局限化(gonadal localization)无关(Mori和Fukatsu,1992)。

[0510] 脂质体由磷脂制成,磷脂分散在含水介质中并自发形成多层同心双层囊泡(亦称多层囊泡(MLV))。MLV的直径通常为25nm至4 μ m。超声处理MLV导致形成直径范围为**200-500 Å**的小单层囊泡(SUV),其核心含有水溶液。

[0511] 脂质体与细胞膜有类似之处,并且考虑了将脂质体在本发明中用作肽组合物的载体。它们广泛适用,因为水溶性和脂溶性物质都可被捕获,即分别在含水空间和双层本身之内。有可能的是,通过选择性修饰脂质体制剂,载药脂质体甚至可用于活性剂的位点特异

性递送。

[0512] 除了Couvreur等(1977;1988)的教导之外,以下信息可用于制备脂质体制剂。当分散在水中,磷脂可形成除脂质体之外的多种结构,这取决于脂质与水的摩尔比。在低比例时,脂质体是优选结构。脂质体的物理特性取决于pH、离子强度和二价阳离子的存在。脂质体可表现出对离子和极性物质的低通透性,但是在高温下经历明显改变其通透性的相变。相变包括从密集的有序结构(称为凝胶态)转变为松散堆积的较无序结构(称为流态)。这在特征性相变温度下发生,导致对离子、糖和药物的通透性增加。

[0513] 除了温度之外,暴露至蛋白也可改变脂质体通透性。某些可溶性蛋白例如细胞色素c与双层结合,使之变形并穿透它,从而改变通透性。胆固醇抑制蛋白的这种穿透,显然是通过更紧密地堆积磷脂。考虑到用于抗生素和抑制剂递送的最有用脂质体构成物将包含胆固醇。

[0514] 不同类型的脂质体包封溶质的能力不同。例如,MLV在包封溶质方面是中等有效的,但SUV非常低效。然而,SUV具有尺寸分布的均一性和再现性的优点,而大单层囊泡(LUV)提供了大小与包封效率间的折衷。通过醚蒸发制备它们,它们在溶质包封方面的效率是MLV的3-4倍。

[0515] 除了脂质体的特征之外,包封化合物的重要决定因素是所述化合物本身的理化性质。极性化合物包封于含水空间中,而非极性化合物与囊泡脂质双层结合。极性化合物通过渗透而释放或当双层破裂时释放,但非极性化合物仍然与双层结合,除非双层因温度或暴露至脂蛋白而瓦解。这两类都在相变温度下显示最大流出率。

[0516] 脂质体通过四种不同机制与细胞相互作用:通过网状内皮系统的吞噬细胞(例如巨噬细胞和嗜中性粒细胞)的胞吞作用;通过非特异性弱疏水力或静电力,或者通过与细胞表面组分的特异性相互作用,而吸附到细胞表面;通过将脂质体脂质双层插入到质膜中与细胞质膜融合,同时将脂质体内含物释放到细胞质内;和通过将脂质体脂质转移给细胞膜或亚细胞膜(或反之亦然),而与脂质体内含物无任何关系。通常难以确定是何种机制起作用,而且可能不止一种机制同时起作用。

[0517] 静脉注射的脂质体的命运和分布取决于其物理性质,例如大小、流动性和表面电荷。它们可在组织中保持数小时或数天(取决于其组成),而血液半寿期范围为数分钟至数小时。较大的脂质体例如MLV和LUV由网状内皮系统吞噬细胞快速接纳,但循环系统生理学在大多数部位都限制这类大物质的进出(exit)。它们仅可在毛细管内皮中存在大开口或孔隙的位置(例如肝或脾血窦)上进出。因此,这些器官是主要摄取部位。另一方面,SUV显示更广泛的组织分布,但仍高度隔离在肝和脾中。一般而言,这样的体内行为将脂质体的潜在靶向作用限制于仅其大尺寸可接近的器官和组织。这些器官和组织包括血、肝、脾、骨髓和淋巴样器官。

[0518] 靶向通常不是本发明的限制。然而,需要特异性靶向的话,可获得用于达到该目的的方法。抗体可用于与脂质体表面结合并将抗体及其药物内含物导向位于特定细胞型表面的特异性抗原性受体。碳水化合物决定簇(在细胞-细胞识别、相互作用和粘附方面起作用的糖蛋白或糖脂细胞表面组分)也可用作识别位点,因为它们具有将脂质体导向特定细胞型的潜力。通常,考虑采用静脉内注射脂质体制剂,但是也可想到其它给药途径。

[0519] 或者,本发明提供本发明组合物的药学上可接受的纳米囊制剂。纳米囊通常可以

稳定且可重现的方式包封化合物(Henry-Michelland等,1987;Quintanar-Guerrero等,1998;Douglas等,1987)。为了避免胞内聚合物过载的副作用,应当用体内能够降解的聚合物设计这类超细颗粒(大小约0.1 μ m)。考虑了将满足这些要求的生物可降解聚烷基-氰基丙烯酸酯纳米粒用于本发明中。可容易制备这类颗粒,如(Couvreur等,1980;1988;zur Muhlen等,1998;Zambaux等1998;Pinto-Alphandry等,1995和美国专利5,145,684,所述文献通过引用以其整体明确结合到本文中)所述。

[0520] 皮肤贴剂也可用于经皮递送。

[0521] 免疫原性组合物

[0522] 在本发明的某些优选实施方案中,提供免疫原性组合物。免疫原性组合物通常包含与免疫刺激剂联用的一种或多种多肽或多核苷酸,例如上文所论述的多肽或多核苷酸。免疫刺激剂可为提高或增强对外源抗原的免疫应答(抗体和/或细胞介导的)的任何物质。免疫刺激剂的实例包括佐剂、生物可降解微球体(例如丙交酯乙交酯共聚物(poly(lactic galactide)))和脂质体(其中掺入了化合物;参见,例如Fullerton,美国专利号4,235,877)。

[0523] 制备免疫原性组合物的一般性描述参见例如Powell和Newman,编著,Vaccine Design(the subunit and adjuvant approach(亚单位和佐剂方法))(1995)。本发明范围内的药物组合物和免疫原性组合物也可包含其它化合物,其可为生物活性的或非活性的。例如,在药物组合物或免疫原性组合物之内,可存在其它结核分枝杆菌抗原的一个或多个免疫原性部分(掺入到融合多肽中或作为单独化合物)。

[0524] 阐述性的免疫原性组合物可包含编码一种或多种上述多肽的多核苷酸(例如DNA),使得所述多肽原位产生(因而引起免疫应答)。如上所述,DNA可存在于任何本领域普通技术人员已知的各种递送系统中,所述递送系统包括核酸表达系统、细菌和病毒表达系统。许多基因递送技术为本领域所熟知,例如Rolland,Crit.Rev.Therap.Drug Carrier Systems 15:143-198(1998)及其中所引参考文献所述的技术。合适的核酸表达系统包含在患者中表达所需的DNA序列(例如合适的启动子和终止信号)。细菌递送系统包括给予表达多肽(例如在其细胞表面或分泌多肽)的细菌宿主细胞(例如分枝杆菌、芽胞杆菌或乳酸杆菌(Lactobacillus)菌株,包括卡介苗或乳酸乳球菌(Lactococcus lactis))(参见,例如Ferreira,等.,An Acad Bras Cienc(2005)77:113-124;和Raha,等,Appl Microbiol Biotechnol(2005)PubMedID 15635459)。在一个优选的实施方案中,可使用病毒表达系统(例如痘苗病毒或其它痘病毒、反转录病毒或腺病毒)将DNA引入,所述系统包括使用非病原性(缺陷型)、能复制的病毒。合适的系统公开于例如Fisher-Hoch等,Proc.Natl.Acad.Sci.USA 86:317-321(1989);Flexner.,Ann.N.Y.Acad.Sci.569:86-103(1989);Flexner等,Vaccine 8:17-21(1990);美国专利号4,603,112、4,769,330和5,017,487;WO 89/01973;美国专利号4,777,127;GB 2,200,651;EP 0,345,242;WO 91/02805;Berkner,Biotechniques 6:616-627(1988);Rosenfeld等,Science 252:431-434(1991);Kolls等,Proc.Natl.Acad.Sci.USA 91:215-219(1994);Kass-Eisler等,Proc.Natl.Acad.Sci.USA 90:11498-11502(1993);Guzman等,Circulation 88:2838-2848(1993);和Guzman等,Cir.Res.73:1202-1207(1993)。用于将DNA掺入这类表达系统的技术为本领域普通技术人员所熟知。DNA也可“裸的”,如例如Ulmer等,Science 259:1745-1749(1993)所述和由Cohen,Science 259:1691-1692(1993)综述。可通过将DNA包被至生物

可降解的珠粒上来增加裸DNA的摄取,所述珠粒被有效地转移至细胞内。显而易见的是,免疫原性组合物可包含多核苷酸和多肽组分两者。这类免疫原性组合物可提供增强的免疫应答。

[0525] 显而易见的是,免疫原性组合物可包含本文所提供多核苷酸和多肽的药学上可接受的盐。这类盐可制备自药学上可接受的无毒性的碱,包括有机碱(例如伯、仲和叔胺以及碱性氨基酸的盐)和无机碱(例如钠盐、钾盐、锂盐、铵盐、钙盐和镁盐)。

[0526] 尽管本领域普通技术人员已知的任何合适载体都可用于本发明的免疫原性组合物中,但是载体类型因给药模式而异。可将本发明组合物配制用于任何合适的给药方式,包括例如局部、口服、鼻内、静脉内、颅内、腹膜内、皮下或肌内给药。对于胃肠外给药例如皮下注射,载体优选包括水、盐水、醇、脂肪、蜡或缓冲剂。对于口服给药,可使用任何上述载体或固体载体,例如甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、滑石、纤维素、葡萄糖、蔗糖和碳酸镁。生物可降解微球体(例如聚乳酸酯(poly lactate)聚乙醇酸酯(polyglycolate))也可用作本发明药物组合物的载体。合适的生物可降解微球体公开于例如美国专利号4,897,268、5,075,109、5,928,647、5,811,128、5,820,883、5,853,763、5,814,344和5,942,252。也可以使用包含如美国专利号5,928,647所述的微粒-蛋白复合物的载体,该复合物能够在宿主中诱导I类-限制性细胞毒T淋巴细胞应答。

[0527] 这类组合物也可包含缓冲液(例如中性缓冲盐水或磷酸缓冲盐溶液)、碳水化合物(例如葡萄糖、甘露糖、蔗糖或葡聚糖)、甘露醇、蛋白、多肽或氨基酸(例如甘氨酸)、抗氧化剂、抑菌剂、螯合剂(例如EDTA或谷胱甘肽)、佐剂(例如氢氧化铝)、赋予制剂与受体血液等渗、低渗或弱高渗的溶质、助悬剂、增稠剂和/或防腐剂。或者,本发明组合物可配制成冻干物。也可以使用众所周知的技术,将化合物包封在脂质体中。

[0528] 任何各种免疫刺激剂都可用于本发明免疫原性组合物中。例如,可包含佐剂。大多数佐剂含有设计用于保护抗原不被快速代谢的物质(例如氢氧化铝或矿物油)以及免疫应答刺激剂(例如脂质A、百日咳博德特氏菌(*Bordetella pertussis*)蛋白或分枝杆菌蛋白或分枝杆菌衍生蛋白)。例如,可使用脱脂质化、脱糖脂化母牛分枝杆菌("pVac")。合适佐剂是市售的,例如弗氏不完全佐剂和弗氏完全佐剂(Difco Laboratories, Detroit, MI); Merck Adjuvant 65(Merck and Company, Inc., Rahway, NJ); AS01B、AS02A、AS15、AS-2及其衍生物(GlaxoSmithKline, Philadelphia, PA); CWS(来自结核杆菌的细胞壁骨架)、TDM(海藻糖二霉菌酸酯(trehalose dicorynomycolate))、Leif(利什曼虫延伸起始因子)、铝盐例如氢氧化铝凝胶(明矾)或磷酸铝;钙盐、铁盐或锌盐;酰化酪氨酸的不溶性混悬剂;酰化糖;阳离子或阴离子衍生的多糖;聚磷腈;生物可降解微球体;单磷酸脂质A(MPL®);和quil A(例如QS-21)。细胞因子(例如GM-CSF或白介素-2、白介素-7或白介素-12)也都可用作佐剂。

[0529] 佐剂是指在疫苗或治疗性组合物中增加对抗原的特异性免疫应答的组分(参见例如Edelman, AIDS Res. Hum Retroviruses 8:1409-1411(1992))。佐剂诱导Th1型和Th2型反应的免疫应答。Th1型细胞因子(例如IFN- γ 、IL-2和IL-12)倾向于诱导细胞介导的针对所给予抗原的免疫应答,而Th-2型细胞因子(例如IL-4、IL-5、IL-6、IL-10)倾向于诱导体液免疫应答。能够优先刺激Th-1细胞介导免疫应答的佐剂描述于WO 94/00153和WO 95/17209。

[0530] 在本文所提供的免疫原性组合物中,优选设计佐剂组合物,以诱导以Th1型为主的免疫应答。在施加本文所提供的免疫原性组合物之后,患者通常会产生包括Th1型和Th2型

反应在内的免疫应答。在反应主要为Th1型的一个优选实施方案中, Th1型细胞因子的水平将增加到比Th2型细胞因子水平更高的程度。使用标准测定, 可容易评价这些细胞因子的水平。有关细胞因子家族的综述参见Janeway等, Immunobiology, 第5版, 2001。

[0531] Rv3616c组合物通常包含一种或多种佐剂, 例如AS01B(含3-脱-O-酰化单磷酸脂质A(3D-MPL®)和QS21的脂质体制剂; 参见美国专利公开号2003/0143240); AS02A(3D-MPL®和QS21和水包油乳剂; 参见Bojang等, Lancet(2001)358:1927); ENHANZYN®(Detox); 3D-MPL®; 皂苷(包括Quil A及其组分例如QS21)和皂苷模拟物; CWS(来自结核杆菌的细胞壁骨架); TDM(海藻糖二霉菌酸酯); 氨基烷基氨基葡萄糖苷4-磷酸酯(AGP); 免疫刺激性的寡核苷酸例如CPG; Leif(利什曼虫延伸起始因子); 及其衍生物。在一个优选的实施方案中, 将Rv3616c多肽与一种或多种选自以下的佐剂一起给予: 含3D-MPL®和QS21的脂质体制剂(例如AS01B)以及3D-MPL®和QS21和水包油乳剂(例如AS02A)。佐剂系统AS01B和AS02A进一步描述于Pichyangkul, 等, Vaccine(2004)22:3831-40。

[0532] 当递送作为核酸的Rv3616c抗原时, 其可在例如病毒载体(即腺病毒载体)或突变细菌宿主细胞(即突变的无毒分枝杆菌、乳酸杆菌或芽孢杆菌宿主细胞, 包括卡介苗(BCG)和乳酸乳球菌)中递送。

[0533] 用于诱导占优势的Th1型反应的优选佐剂包括例如单磷酸脂质A(MPL®)(优选3-O-脱酰化单磷酸脂质A(3D-MPL®))任选与铝盐的组合(参见例如Ribi等, 1986, Immunology and Immunopharmacology of Bacterial Endotoxins(细菌内毒素的免疫学和免疫药理学), Plenum Publ. Corp., NY, 第407-419页; GB 2122204B; GB 2220211; 和US 4,912,094)。3D-MPL®的优选形式是含有直径小于0.2mm的小粒度乳剂形式, 其制备方法公开于W0 94/21292。含有单磷酸脂质A和表面活性剂的含水制剂已描述于W0 98/43670。例示性的优选佐剂包括AS01B(含MPL®和QS21的脂质体制剂)、含3D-MPL®和QS21的脂质体制剂、AS02A(MPL®和QS21和水包油乳剂)、3D-MPL®和QS21和水包油乳剂、以及AS15(得自GlaxoSmithKline)。MPL®佐剂得自GlaxoSmithKline(参见美国专利号4,436,727; 4,877,611; 4,866,034和4,912,094)。

[0534] 含有CpG的寡核苷酸(其中CpG二核苷酸未甲基化)也诱导占优势的Th1反应。CpG是DNA中存在的胞嘧啶-鸟苷二核苷酸基序的缩写。这样的寡核苷酸是众所周知的, 描述于例如W0 96/02555、W0 99/33488和美国专利号6,008,200和5,856,462。免疫刺激性DNA序列也描述于例如Sato等, Science 273:352(1996)。当CpG配制在免疫原性组合物中时, CpG通常以游离溶液与游离抗原一起给予(W0 96/02555; McCluskie和Davis, 见上文)或者与抗原共价缀合(W0 98/16247), 或者与载体例如氢氧化铝配制在一起((肝炎表面抗原)Davis等, 见上文; Brazolot-Millan等, Proc. Natl. Acad. Sci, USA, 1998, 95(26), 15553-8)。本领域已知CpG作为佐剂, 既可经系统途径给予也可经粘膜途径给予(W0 96/02555、EP 468520、Davis等, J. Immunol, 1998, 160(2):870-876; McCluskie和Davis, J. Immunol., 1998, 161(9):4463-6)。

[0535] 另一种优选的佐剂是皂苷或皂苷模拟物或衍生物, 例如Quil A、优选QS21(Aquila Biopharmaceuticals Inc., Framingham, MA), 其可单独使用或者与其它佐剂联用。例如, 一个增强的系统包括单磷酸脂质A(MPL®)与皂苷衍生物的组合, 例如如W0 94/00153所述的QS21与3D-MPL®的组合, 或者如W0 96/33739所述的其中QS21被胆固醇猝灭的较低反应原

性组合物。其它优选制剂包括水包油乳剂和生育酚。包含含QS21、3D-MPL®和生育酚的水包油乳剂的特别有效的佐剂制剂描述于W0 95/17210。其它用于本发明的皂苷佐剂包括QS7(描述于W0 96/33739和W0 96/11711)和QS17(描述于美国专利号5,057,540和EP 0 362 279B1)。

[0536] 或者皂苷制剂可与以下物质组合:由壳聚糖或其它聚阳离子聚合物构成的疫苗载体、聚丙交酯和丙交酯-乙交酯共聚物颗粒、基于聚-N-乙酰基葡萄糖胺的聚合物基质、由多糖或化学改性多糖构成的颗粒、脂质体和基于脂质的颗粒、由单甘油酯组成的颗粒等。也可在胆固醇存在下配制皂苷,以形成粒状结构例如脂质体或ISCOM®。此外,皂苷可与聚氧乙烯醚或聚氧乙烯酯一起配制成非粒状溶液或悬浮液,或者配制成粒状结构例如寡层薄膜(paucilamellar)脂质体或ISCOM®。皂苷也可与赋形剂例如CARBOPOL®一起配制以增加粘度,或者可用粉状赋形剂例如乳糖配制成干粉形式。

[0537] 在一个实施方案中,佐剂系统包含单磷酸脂质A与皂苷衍生物的组合,例如如W0 94/00153所述的QS21与3D-MPL®佐剂的组合,或者如W0 96/33739所述的其中QS21被含胆固醇的脂质体猝灭的较低反应原性组合物。其它合适的制剂包括水包油乳剂和生育酚。采用含QS21、3D-MPL®佐剂和生育酚的水包油乳剂的另一种合适佐剂制剂描述于W0 95/17210。

[0538] 另一种增强的佐剂系统包括含CpG的寡核苷酸与皂苷衍生物的组合、尤其是CpG与QS21的组合,参见W0 00/09159。合适的是,所述制剂另外包含水包油乳剂和生育酚。

[0539] 其它合适佐剂包括MONTANIDE® ISA 720(Seppic, France)、SAF(Chiron, California, United States)、ISCOMS®(CSL)、MF-59(Chiron)、SBAS系列佐剂(SmithKline Beecham, Rixensart, Belgium)、Detox(Corixa)、RC-529(Corixa)和其它氨基烷基氨基葡萄糖苷4-磷酸酯(AGP),例如描述于在审的美国专利申请顺序号08/853,826和09/074,720的佐剂(所述文献的公开内容都通过引用以其整体结合到本文中),以及聚氧乙烯醚佐剂例如W0 99/52549A1中所述的佐剂。SmithKline Beecham和Corixa Corporation现为GlaxoSmithKline的一部分。

[0540] 其它合适的佐剂包括以下通式(I)的佐剂分子:

[0541] $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{-A-R}$

[0542] 其中n为1-50,A为化学键或-C(O)-,R为C₁₋₅₀烷基或苯基C₁₋₅₀烷基。

[0543] 另一种引人关注的佐剂为例如如W02005/112991所述的志贺毒素b链。

[0544] 本发明的一个实施方案由包含通式(I)的聚氧乙烯醚的免疫原性组合物构成,其中n为1-50、优选4-24、最优选9;R组分为C₁₋₅₀、优选C_{4-C20}烷基和最优选C₁₂烷基,而A为键。聚氧乙烯醚的浓度应在0.1-20%、优选0.1-10%和最优选0.1-1%的范围内。优选的聚氧乙烯醚选自:聚氧乙烯-9-月桂醚、聚氧乙烯-9-硬脂酰(stearyl)醚、聚氧乙烯-8-硬脂酰(stearyl)醚、聚氧乙烯-4-月桂醚、聚氧乙烯-35-月桂醚和聚氧乙烯-23-月桂醚。聚氧乙烯醚例如聚氧乙烯月桂醚描述于Merck索引(第12版;条目7717)。这些佐剂分子描述于W0 99/52549。

[0545] 本文提供的任何免疫原性组合物可使用以下熟知方法制备,所述方法导致形成抗原、免疫应答增强剂和合适载体或赋形剂的组合。可作为缓释制剂(即在给药后实现化合物的缓慢释放的诸如胶囊剂、海绵剂或凝胶剂(例如由多糖组成)等制剂)的部分给予本文所述的组合物。这样的制剂通常可用众所周知的技术来制备(参见例如Coombes等,Vaccine

14:1429-1438(1996))并通过例如口服、直肠或皮下植入而给予,或通过在其所需靶部位植入而给予。缓释制剂可含有分散在载体基质中和/或包含在被速率控制膜包裹的贮库之内的多肽、多核苷酸或抗体。

[0546] 用于这类制剂的载体是生物相容的并且还可是生物可降解的;优选所述制剂提供相对恒定的活性组分释放水平。这类载体包括丙交酯-乙交酯共聚物、聚丙烯酸酯、乳胶、淀粉、纤维素、葡聚糖等的微粒。其它延迟释放载体包括超分子生物载体,其包括非液态的亲水核(例如交联多糖或寡糖)以及任选的包含两亲化合物例如磷脂的外层(参见例如美国专利号5,151,254和PCT申请WO 94/20078、WO 94/23701和WO 96/06638)。缓释制剂中活性化合物含量取决于植入部位,释放速率和预期持续时间。

[0547] 任何递送载体都可用于药物组合物和免疫原性组合物,以促进产生抗原特异性免疫应答。递送载体包括抗原呈递细胞(APC),例如树突细胞、巨噬细胞、B细胞、单核细胞和可改造成有效APC的其它细胞。这些细胞可(但非必需)经遗传修饰,以增加呈递抗原的容量、改善T细胞应答的激活和/或维持、和/或与受体免疫学上相容(即匹配的HLA单元型)。APC通常可从任何生物液体和器官中分离出,并且可为自体细胞、同种异体细胞、同源细胞或异种细胞。

[0548] 本发明的某些优选实施方案使用树突细胞或其祖细胞作为抗原呈递细胞。树突细胞是高度有效的APC(Banchereau和Steinman, *Nature* 392:245-251(1998)),并且已表明有效作为用于引起预防性或治疗性免疫的生理佐剂(参见Timmerman和Levy, *Ann. Rev. Med.* 50:507-529(1999))。一般而言,可根据树突细胞的典型形状(在原位星形,在体外可见明显胞质突(树突)),其高效摄取、加工和呈递抗原的能力以及其激活幼稚T细胞应答的能力,来鉴定树突细胞。当然,树突细胞可经改造以表达在体内或离体的树突细胞上不常见的特异性细胞表面受体或配体,而且本发明也考虑了这种经修饰的树突细胞。作为树突细胞的替代物,装载分泌型囊泡抗原的树突细胞(称为外来体)可用于免疫原性组合物中(参见Zitvogel等, *Nature Med.* 4:594-600(1998))。

[0549] 树突细胞及其祖细胞可得自外周血、骨髓、淋巴结、脾、皮肤、脐带血或任何其它合适组织或液体。例如,可通过将诸如GM-CSF、IL-4、IL-13和/或TNF α 等细胞因子的组合添加到从外周血中收获的单核细胞培养物中,使树突细胞离体分化。或者,可通过将GM-CSF、IL-3、TNF α 、CD40配体、LPS、flt3配体和/或诱导树突细胞分化、成熟和增殖的其它化合物的组合添加到培养基中,使从外周血、脐带血或骨髓中收获的CD34阳性细胞分化成树突细胞。

[0550] 树突细胞容易分类为“未成熟”和“成熟”细胞,这允许用简单方法区分两种充分表征的表型。然而,这样的命名不应理解为排除所有可能的分化中间阶段。未成熟树突细胞表征为具有高抗原摄取和加工能力的APC,这与Fc γ 受体和甘露糖受体的高度表达相关。成熟表型的典型特征在于这些标记的较低表达,但高度表达负责T细胞激活的细胞表面分子,例如I类和II类MHC、粘附分子(例如CD54和CD11)和共刺激分子(例如CD40、CD80、CD86和4-1BB)。

[0551] 通常可用编码蛋白(或其部分或其它变体)的多核苷酸转染APC,使得多肽在细胞表面上表达。这类转染可离体发生,然后可如本文所述使用包含这类转染细胞的药物组合物或免疫原性组合物。或者,可将靶向树突细胞或其它抗原呈递细胞的基因递送载体给予患者,导致在体内发生转染。例如,通常可使用本领域已知的任何方法,进行树突细胞的体

内和离体转染,所述方法为例如WO 97/24447所述的方法或基因枪方法,参见Mahvi等, *Immunology and Cell Biology* 75:456-460(1997)。可通过将树突细胞或祖细胞与多肽、DNA(裸的或在质粒载体内)或RNA一起孵育;或者与表达抗原的重组细菌或病毒(例如痘苗病毒、禽痘病毒、腺病毒或慢病毒载体)一起孵育,来给树突细胞加载抗原。加载之前,可将多肽与提供T细胞辅助的免疫配偶体(例如载体分子)共价缀合。或者,可用未缀合的免疫配偶体,单独、或在多肽存在下脉冲刺激树突细胞。

[0552] 免疫原性组合物和药物组合物可存在于单位剂量或多剂量容器(例如密封安瓿或小瓶)中。这样的容器优选是气密性的,以保持制剂的无菌性,直至使用。一般而言,制剂可在油性或水性溶媒中以混悬剂、溶液剂或乳剂贮存。或者,免疫原性组合物或药物组合物可贮存在临用前仅需加入无菌液体载体的冻干条件下。

[0553] 在某些实施方案中,在Rv3616c多肽(包括变体、免疫原性片段或融合蛋白)或编码所述多肽的多核苷酸的“初免”或首次给药之后,接着Rv3616c多肽(包括变体、免疫原性片段或融合蛋白)或编码所述多肽的多核苷酸的一次或多次“加强”或后续给药(“初免和加强”方法)。例如,首次给予Rv3616c多肽(包括变体、免疫原性片段或融合蛋白)或编码所述多肽的多核苷酸之后,接着进行一次或多次后续给予Rv3616c多肽(包括变体、免疫原性片段或融合蛋白)或编码所述多肽的多核苷酸。

[0554] 在一个实施方案中,首次给予Rv3616c多肽或多核苷酸之后,接着一次或多次后续给予Rv3616c多肽。在一个实施方案中,首次给予Rv3616c多肽或多核苷酸之后,接着一次或多次后续给予Rv3616c多核苷酸。通常首次或“初免”给药与第二次或“加强”给药之间间隔约2-12周,或多达4-6个月。后续“加强”给药间隔约6个月或长达1、2、3、4或5年。常规加强治疗(例如在蛋白初免给药之后接着蛋白加强给药)也可用于预防或治疗结核病(例如预防或治疗潜伏性结核病,尤其是预防或延迟结核病再活化)。

[0555] 抗体

[0556] “抗体”是指特异性结合并识别抗原的包含免疫球蛋白基因框架区的多肽或其片段。公知的免疫球蛋白基因包括 κ 、 λ 、 α 、 γ 、 δ 、 ϵ 和 μ 恒定区基因,以及众多的免疫球蛋白可变区基因。轻链分为 κ 或 λ 。重链分为 γ 、 μ 、 α 、 δ 或 ϵ ,它们继而又限定免疫球蛋白类别分别为IgG、IgM、IgA、IgD和IgE。

[0557] 例示性的免疫球蛋白(抗体)结构单元包含四聚体。每个四聚体由两对相同的多肽链组成,每对都具有有一条“轻”链(约25kDa)和一条“重”链(约50-70kDa)。各链的N-端限定约100-110个或更多个氨基酸的可变区,主要担负抗原识别。术语可变轻链(V_L)和可变重链(V_H)分别是指这些轻链和重链。

[0558] 抗体例如以完整免疫球蛋白形式或以用不同肽酶消化而产生的许多充分表征的片段形式存在。因此,例如胃蛋白酶在铰链区二硫键下消化抗体,产生 $F(ab)'_2$, $F(ab)'_2$ 是Fab的二聚体,Fab本身是通过二硫键与 V_H - C_H1 连接在一起的轻链。在温和条件下可将 $F(ab)'_2$ 还原,以打断铰链区的二硫键,从而将 $F(ab)'_2$ 二聚体转化成Fab'单体。Fab'单体本质上是带有铰链区部分的Fab(参见Fundamental Immunology(Paul编著,第3版,1993)。尽管根据完整抗体的消化来定义不同抗体片段,但是技术人员应认识到这些片段可通过化学法或通过使用重组DNA方法来从头合成。因此,本文所用的术语抗体也包括通过修饰完整抗体而产生的抗体片段,或用重组DNA方法从头合成的抗体片段(例如单链Fv)或用噬菌体展示

文库鉴定的抗体片段(参见例如McCafferty等,Nature 348:552-554(1990))。

[0559] 为了制备单克隆或多克隆抗体,可采用本领域已知的任何技术(参见例如Kohler和Milstein,Nature 256:495-497(1975);Kozbor等,Immunology Today 4:72(1983);Cole等,载于Monoclonal Antibodies and Cancer Therapy,第77-96页(1985))。用于产生单链抗体的技术(美国专利4,946,778)可适用于产生针对本发明多肽的抗体。另外,转基因小鼠或其它生物体例如其它哺乳动物也可用于表达人源化抗体。或者,噬菌体展示技术可用于鉴定特异性结合所选抗原的抗体和异聚Fab片段(参见例如McCafferty等,Nature 348:552-554(1990);Marks等,Biotechnology 10:779-783(1992))。

[0560] 措词“特异性(或选择性)结合”抗体或“与...特异性(或选择性)免疫反应”当提及蛋白或肽时,是指确定在蛋白和其它生物制品的异质群体中蛋白存在情况的结合反应。因此,在指定免疫测定条件下,特异抗体以背景的至少两倍与特定蛋白结合,而且基本上不与样品中存在的其它蛋白大量结合。在这类条件下与抗体的特异性结合可能需要根据其对特定蛋白的特异性而选择的抗体。例如,可选择针对融合蛋白产生的多克隆抗体,以获取仅与融合蛋白发生特异性免疫反应、而不与融合蛋白各个组分发生特异性免疫反应的多克隆抗体。可通过去掉与各个抗原交叉反应的抗体,来完成这样的选择。各种免疫测定形式都可用于选择与特定蛋白发生特异性免疫反应的抗体。例如,惯常将固相ELISA免疫测定用于选择与蛋白发生特异性免疫反应的抗体(参见例如Harlow和Lane,A Laboratory Manual(1988)和Using Antibodies:A Laboratory Manual(1998),关于对可用于测定特异性免疫反应的免疫测定形式和条件的描述)。通常,特异性或选择性反应应为背景信号或噪声的至少2倍,更通常高于背景的10、20或100倍(例如结合至其它分枝杆菌蛋白,例如其它结核分枝杆菌蛋白)。

[0561] 诊断

[0562] 在另一个方面,本发明提供将上述一种或多种多肽用于诊断潜伏性结核病的方法(例如使用常规形式的基于T细胞应答的测定或基于抗体的测定)。

[0563] 例如,本文提供用于在个体中测定潜伏性结核分枝杆菌感染的方法,其包括:

[0564] (a)从所述个体获得样品;

[0565] (b)使所述样品与分离多肽接触,所述分离多肽包含:

[0566] (i)Rv3616c蛋白序列;

[0567] (ii)Rv3616c蛋白序列的变体;或者

[0568] (iii)Rv3616c蛋白序列的免疫原性片段;

[0569] (c)定量所述样品反应。

[0570] 样品可例如为全血或纯化细胞。合适的是,样品包括外周血单核细胞。在本发明的一个实施方案中,个体为血清反应阳性的。在本发明的第二个实施方案中,个体为血清反应阴性的。

[0571] 合适的是,个体先前未针对结核分枝杆菌感染进行接种(例如合适的是,个体先前未用BCG接种)。

[0572] 样品反应可通过各种本领域技术人员已知的方法进行定量,所述方法包括监测淋巴细胞增殖或特定细胞因子或抗体的产生。例如T细胞ELISPOT可用于监测细胞因子例如干扰素 γ (IFN γ)、白介素2(IL2)和白介素5(IL5)。B细胞ELLISPOT可用于监测结核分枝杆菌

特定抗原的刺激。细胞应答也可通过使用胞内和胞外染色并通过流式细胞仪进行分析来表征。

[0573] 定量样品增殖反应的方法包括：

[0574] (i)用放射标记(例如氘化胸苷)脉冲培养细胞并监测氘摄取(例如气体闪烁)；

[0575] (ii)进行羧基荧光素二乙酸琥珀酰亚胺酯(CFSE)标记和使用流式细胞术荧光监测细胞分裂。

[0576] 定量样品细胞因子反应尤其包括监测干扰素 γ 产生。

[0577] 当使用这类定量方法时，抗原的阳性反应可通过至少2:1(例如至少3:1或至少5:1)的信噪比(S/N比)界定。

[0578] 在本发明的另一个方面提供利用皮试诊断潜伏性结核分枝杆菌感染的方法。本文所用“皮试”为直接在患者上进行的任何测定，其中在皮内注射如上述的Rv153c多肽(或其变体、免疫原性片段或者其编码核苷酸)之后，测得迟发型超敏(DTH)反应(例如膨胀、变红或皮炎)。可使用任何合适的足以使抗原组合与患者真皮细胞接触的装置实现这类注射，所述装置例如结核菌素注射器或1mL注射器。在一段时期后(例如注射后至少48小时、尤其48-72小时)测量所述反应。

[0579] DTH反应为细胞介导的免疫应答，其在先前已暴露于测试抗原的患者中更强烈。可使用尺目测所述应答。通常，直径大于约0.5cm、尤其大于约1.0cm的应答为阳性应答，指示先前感染结核分枝杆菌，该感染可表现或不表现为活动性疾病。

[0580] 为了用于皮试，Rv3616c组分适宜配制为含药理学上可接受的载体的药物组合物。合适的是，这类药物组合物中所用的载体为含合适防腐剂(例如苯酚和/或Tween 80™)的盐水溶液。

[0581] 本发明还提供用于任何上述诊断方法的试剂盒。这类试剂盒通常含两种或更多种对进行诊断测定所需的组分。组分可为化合物、试剂、容器和/或装置。

[0582] 例如，试剂盒内的一个容器可包含特异性结合至蛋白的单克隆抗体或其片段。如上所述，可将这类抗体或片段提供为连接至支持材料。一个或多个其它容器可装有待用于测定中的诸如试剂或缓冲剂等成分。或者，这类试剂盒也可包含如上所述的含适于直接或间接检测抗体结合的报道基团的检测试剂。

[0583] 或者，试剂盒可经设计以检测生物样品中编码蛋白的mRNA水平。这类试剂盒通常包含如上所述的至少一种与蛋白编码多核苷酸杂交的寡核苷酸探针或引物。这类寡核苷酸可用于例如PCR测定或杂交测定。可在这类试剂盒中存在的其它组分包括第二寡核苷酸和/或诊断试剂或容器，以有助于编码本发明蛋白的多核苷酸的检测。

[0584] 其它诊断试剂盒包括设计用于检测细胞介导的反应的试剂盒(其可例如用于本发明诊断方法中)。这类试剂盒通常包含：

[0585] (i)用于从受试者中获取合适细胞样品的装置；

[0586] (ii)用于用Rv3616c多肽(或其变体、其免疫原性片段或编码所述多肽的DNA)刺激所述细胞样品的工具；

[0587] (iii)用于检测或定量对刺激的细胞应答的工具。

[0588] 用于定量细胞应答的合适工具包括B细胞ELISPOT试剂盒或T细胞ELISPOT试剂盒，其为本领域技术人员已知。

- [0589] 一种可能的试剂盒包含：
- [0590] (a)本发明多肽；和
- [0591] (b)适于直接或间接检测抗体结合的检测试剂。
- [0592] 尤其引人关注的是专门用于定量T细胞应答的诊断试剂盒：
- [0593] 一种诊断试剂盒，其包含：
- [0594] (a)本发明多肽；和
- [0595] (b)足以使所述多肽与个体真皮细胞接触的装置。
- [0596] 一种诊断试剂盒，其包含：
- [0597] (a)本发明多肽；
- [0598] (b)足以使所述多肽与来自个体的样品（例如全血或更适合的PBMC）接触的装置；和
- [0599] (c)定量T细胞应答（例如增殖或IFN- γ 产生）的工具。

实施例

[0600] 仅以阐述的方式而非限制性方式提供以下实施例。本领域技术人员容易认识到，可改动或修改各种非关键性参数以得到基本类似的结果。

[0601] 实施例1-Rv3616c作为潜伏性TB疫苗靶标的鉴定

[0602] 基因Rv3616c编码保守的假定富含丙氨酸和甘氨酸的蛋白。

[0603] 如Murphy和Brown BMC.Infect.Dis.2007 7:84-99所述，根据对休眠阶段维持和侵染性相关结核分枝杆菌基因的泛基因组分析，选择Rv3616c。通过在模拟休眠条件下的细菌基因表达的已公布泛基因组DNA微阵列数据集的生物信息学元分析(meta-analysis)，对结核分枝杆菌中的潜在休眠阶段基因靶标进行优先排序。随后在整个基因组中对由基因编码的结核分枝杆菌蛋白进行亚细胞定位以鉴定疫苗靶标。

[0604] 简言之，在休眠模型中的实验条件差异很大，因而根据以下两个标准形成0-5的评分系统以标准化这些数据：1)实验条件与休眠状态的相关性和2)表达的等级次序。根据与临床出现休眠阶段结核分枝杆菌感染的潜在相关性，来调整具体实验数据集的最高分。表1显示步骤1所收集的数据集以及各个数据集的调整最高分。其它关于对生长必需的基因的数据集获得自己公布的使用基于转座子的敲除实验(TraSH)的研究。对生长无作用的基因被赋予零分。

[0605] 表1-结核分枝杆菌DNA微阵列基因表达和泛基因组基因敲除的来源、实验模型和评分标准(生长阶段必需性)。

[0606]

参考文献	实验模型	时间点:最高分 ^a
Betts JC 等 <i>Mol. Microbiol.</i> 2002 43:717-731	在受控 O ₂ 下的饥饿	96h: 3 24h: 2 4h: 1
Hampshire T 等 <i>Tuberculosis.(Edinb.)</i> 2004 84:228-238	在受控 O ₂ 下的营养耗竭	62 和 75d: 5 49d: 4 18d: 2
Muttucumaru DG 等. <i>Tuberculosis.(Edinb.)</i> 2004 84:239-246	Wayne 缺氧模型 [#]	14d (NRP-2): 4 7d (NRP-1): 2

[0607]

参考文献	实验模型	时间点:最高分 ^a
Voskuil MI 等 <i>Tuberculosis.(Edinb.)</i> 2004 84:218-227	Wayne 缺氧模型 [#]	30 和 80d: 5 14 和 20d: 4 10 和 12d: 3 6 和 8d: 2
Schnappinger D 等 <i>J. Exp. Med.</i> 2003 198:693-704	小鼠巨噬细胞的感染, +/- γ -INF	24 和 48h: 5
Karakousis PC 等 <i>J. Exp. Med.</i> 2004 200:647-657	中空纤维皮下植入小鼠中	10d: 3
Talaat AM 等 <i>Proc. Natl. Acad. Sci. U.S.A</i> 2004, 101:4602-4607	小鼠感染。收获自肺的 MTB ^b	28d: 3
Sasseti CM 等 <i>Mol. Microbiol.</i> 2003 48:77-84	在固体培养基上生长的 TraSH 突变文库	14d:5
Rengarajan J 等 <i>Proc.Natl.Acad.Sci.U.S.A</i> 2005, 102:8327-8332	用结核分枝杆菌 TraSH 突变文库感染小鼠巨噬细胞, +/- γ -INF	7d:5
Sasseti CM 等 <i>Proc.Natl.Acad.Sci.U.S.A</i> 2003 100:12989-12994	用结核分枝杆菌 TraSH 突变文库感染的 C57BL/6J 小鼠	7, 14, 28 和 56d:5

[0608] ^a根据休眠模型的相关性的最高分;h=小时;d=天。[0609] ^b来自 Balb/c 肺的结核分枝杆菌与通风培养 28d 的 MTB 的比。[0610] [#]Wayne LG 和 Hayes LG *Infect.Immun.*1996 64:2062-2069

[0611] 步骤2-在应用第二个标准:基因表达的等级次序时,将来自各个数据集的基因分数根据表达比(在实验条件下的倍数表达与比对数阶段液态培养物中的细胞)从高到低排列。将最高得分基因赋予该特定数据集的最高分(列于表1的第3栏(例如5、4...、1分)。将分数按顺序对每个基因递减0.005分直至零或者达到该数据集的末端。因而,当最高分为4分时,将排在第100位的基因赋予3.500的分数。对于最高分为5分而言,对1000个基因或者25%的结核分枝杆菌基因组赋予分数。对于数据采集自多个时间点的实验而言,将所有时间点的最高分用作最终分数。

[0612] 在步骤3中,将各个基因在各个实验条件下的分数收集至Microsoft Access数据库中。加上参考段(reference field)以促进优先化,例如Refseq ID、Genbank功能、Genbank注释、Tuberculist分类和KEGG与Sanger Center链接。通过组合来自不同研究和来源的数据,得到关于对休眠状态的生存最为关键的特定基因和途径的一致观点。

[0613] 在步骤4中,利用前400得分基因(约基因组的10%)补充以对生物化学途径、酶学、药物易处理性、与人类基因的同源性和其它现有知识进行经验计算(expert computational)和人工分析,得到治疗靶标的优先列表。大多数高分基因来自两个或三个组别相交的子集。

[0614] 在步骤5中,在整个基因组中进行由基因编码的结核分枝杆菌蛋白的亚细胞定位的鉴定。用于膜蛋白预测的探索性方法描述于Chalker等J.Bacteriol.2001 183:1259-1268。利用用梯形窗加权的GES疏水值(hydrophathy value)(Engelman DM等Annu.Rev.Biophys.Biophys.Chem.1986 15:321-353)产生平均疏水概况(H)(von Heijne G J.Mol.Biol.1992 225:487-494)。使用类似于TopPred II算法(Claros MG等.Comput.Appl.Biosci.1994 10:685-686)的初始步骤的程序,通过选择以最高H值(MaxH)为中心的19个氨基酸、出于进一步考虑掩蔽这些氨基酸和重复所述程序直至未留下H>0.5的峰,来对各个肽序列预测螺旋跨膜区段(TMS)。根据峰MaxH值、H>1.0的区段数以及推定TMS的分布和峰H值来指定亚细胞位置。选择MaxH截止值为1.15以使在分别含跨膜蛋白和胞质蛋白的两个SwissProtein释放34测试数据集之间的区分最大化(Boyd D等Protein Sci.19987:201-205)。将MaxH<1.15的蛋白分类为胞质蛋白,而MaxH>1.15且具有至少三个可能的TMS的蛋白分类为膜蛋白。将锚定蛋白定义为具有刚好两个TMS,一个起始于氨基酸(aa)35之前并且一个具有>1.15的H而另一个具有不低于0.5的H。将具有革兰氏阳性设置的SignalP特定用于分枝杆菌(M.bacterium)以鉴定在探索性分析中分类为胞质蛋白或“未知”蛋白中的分泌蛋白(Nielsen H等Protein Eng.1997 10:1-6)。

[0615] 按照以下数个标准,Rv3616c作为疫苗抗原排位很高:

[0616] (i)Rv3616c在所有休眠模型中一致为增量调节的。在整组于元分析中得分的3999个基因之中,Rv3616c在所有休眠模型中排在过表达基因的前四分位。Rv3616c的增量调节分数为6.52,其与最高基因分数22.28相比为有利的。

[0617] (ii)Rv3616c在小鼠脾感染模型中排列为对生存高度必需的(得分4.945,在可能得分为5的范围内)。

[0618] (iii)亚细胞定位预测Rv3616c蛋白为膜结合蛋白,因而具有明显的胞外暴露,这表明适合作为疫苗靶标。

[0619] (iv)Rv3616c可引发针对初始结核病攻击的保护反应。

[0620] (v)Rv3616c被公认为抗原。

[0621] 实施例2-Rv3616c表位预测

[0622] 方法

[0623] T细胞表位预测基于以下方法:

[0624]

预测	名称	URL/参考文献
CD4 和 CD8	Multipred	网址: antigen.i2r.a-star.edu.sg/multipred/ Zhang,G.L., Khan,A.M., Srinivasan,K.N., August,J.T. 和 Brusic,V. (2005) "MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides" <i>Nucleic Acids Res.</i> 33, W172 - W179.
	SVMHC	网址: www-bs.informatik.uni-tuebingen.de/SVMHC "Prediction of MHC class I binding peptides, using SVMHC." Pierre Dönnes 和 Arne Elofsson 载于: <i>BMC Bioinformatics</i> 2002 3: 25
CD4	ProPred	网址: www.imtech.res.in/raghava/propred/ Singh,H. 和 Raghava,G.P.S.(2001) "ProPred: Prediction of HLA-DR binding sites." <i>Bioinformatics</i> ,17(12), 1236-37.
	Tepitope2	载于基于以下的公众程序(house program): H. Bian, J. Hammer (2004) "Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE." <i>Methods</i> 34 : 468-75

[0625]

预测	名称	URL/参考文献
CD8	nHLA	网址: www.imtech.res.in/raghava/nhlapred/ Bhasin M.和 Raghava G P S (2006) "A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes"; J. Biosci. 32:31-42
	NetCTL	网址: www.cbs.dtu.dk/services/NetCTL/ "An integrative approach to CTL epitope prediction. A combined algorithm integrating MHC-I binding, TAP transport efficiency, and proteasomal cleavage predictions." Larsen M.V., Lundegaard C., Kasper Lamberth, Buus S., Brunak S., Lund O.,和Nielsen M. European Journal of Immunology. 35(8): 2295-303, 2005
	EpiJen	网址: www.jenner.ac.uk/EpiJen/ Doytchinova, I. A., P. Guan, D. R. Flower. "EpiJen: a server for multi-step T cell epitope prediction." BMC Bioinformatics, 2006, 7, 131.
	Syfpeithi	网址: www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm Hans-Georg Rammensee, Jutta Bachmann, Niels Nikolaus Emmerich, Oskar Alexander Bachor, Stefan Stevanovic: "SYFPEITHI: database for MHC ligands and peptide motifs." Immunogenetics (1999) 50: 213-219
	PredTAP	网址: antigen.i2r.a-star.edu.sg/predTAP/ Zhang,G.L., Petrovsky,N., Kwoh,C.K., August,J.T.和 Brusic,V. (2006) "PRED ^{TAP} : a system for prediction of peptide binding to the human transporter associated with antigen processing." Immunome Res. 2(1), 3.
	PAPROC	http://www.paproc2.de/paproc1/paproc1.html C. Kuttler, A.K. Nussbaum, T.P. Dick, H.-G. Rammensee, H. Schild, K.P. Haderler, "An algorithm for the prediction of proteasomal cleavages", J. Mol. Biol. 298 (2000), 417-429 A.K. Nussbaum, C. Kuttler, K.P. Haderler, H.-G. Rammensee, H. Schild, "PAProC: A Prediction Algorithm for Proteasomal Cleavages available on the WWW", Immunogenetics 53 (2001), 87-94

[0626] 结果

[0627] 表2-推定的Rv3616c人CD4+T细胞表位

[0628]

推定的 CD4 表位 号	氨基酸 位置	表位序列	SEQ ID No:	HLA 等位基因
1	5	FIIDPTISA	SEQ ID No: 29	DRB1_0301、DRB1_0401、 DRB1_1101
2	31	ILYSSLEYF	SEQ ID No: 30	DRB1_0301
3	36	LEYFEKALE	SEQ ID No: 31	DRB1_1301
4	63	YAGKNRNV	SEQ ID No: 32	DRB1_0801
5	87	LIHDQANAV	SEQ ID No: 33	DRB1_0301、DRB1_0401
6	111	FVRPVAVDL	SEQ ID No: 34	DRB1_0101
7	119	LTYPVVGHH	SEQ ID No: 35	DRB1_0401
8	121	YIPVVGHAL	SEQ ID No: 36	DRB1_0101
9	151	YLVVKTLIN	SEQ ID No: 37	DRB1_0401
10	152	LVVVKTLINA	SEQ ID No: 38	DRB1_1301
11	154	VKTLINATQ	SEQ ID No: 39	DRB1_0401
12	164	LKLLAKLAE	SEQ ID No: 40	DRB1_0301、DRB1_0801、 DRB1_1101、DRB1_1301
13	173	LVAAAIADI	SEQ ID No: 41	DRB1_0301、DRB1_1101、 DRB1_1301
14	181	IISDVADII	SEQ ID No: 42	DRB1_0301
15	197	WEFITNALN	SEQ ID No: 43	DRB1_0401
16	252	LFGAAGLSA	SEQ ID No: 44	DRB1_1501
17	264	LAHADSLAS	SEQ ID No: 45	DRB1_0401
18	270	LASSASLPA	SEQ ID No: 46	DRB1_0401
19	288	FGGLPSLAQ	SEQ ID No: 47	DRB1_0401

[0629] 表3-推定的Rv3616c人CD8+T细胞表位

[0630]

推定的 CD8表 位号	氨基酸 位置	表位序列	SEQ ID No:	HLA等位基因
1	5	FIIDPTISA	SEQ ID No: 48	A2
2	6	IIDPTISAI	SEQ ID No: 49	A_0101、A2
3	9	PTISAIDGL	SEQ ID No: 50	A2、A_0201、B7、B8
4	10	TISAIDGLY	SEQ ID No: 51	A1、A_0101、A3、A_0301
5	12	SAIDGLYDL	SEQ ID No: 52	A2、B_3501
6	13	AIDGLYDLL	SEQ ID No: 53	A_0101、A_0201、B44

[0631]

推定的 CD8表 位号	氨基酸 位置	表位序列	SEQ ID No:	HLA等位基因
7	17	LYDLLGIGI	SEQ ID No: 54	A24
8	25	IPNQGGILY	SEQ ID No: 55	B7、A_0101、B_3501、B51
9	30	GILYSSLEY	SEQ ID No: 56	A1、A_0101、A3、A_0301
10	33	YSSLEYFEK	SEQ ID No: 57	A1、A_0301
11	35	SLEYFEKAL	SEQ ID No: 58	A_0201、B7、Cw_0401、Cw_0602
12	38	YFEKALEEL	SEQ ID No: 59	A24、A_2402、B8、Cw_0401、Cw_0602
13	39	FEKALEELA	SEQ ID No: 60	B44、B_4403
14	69	NHVNFFQEL	SEQ ID No: 61	A24、Cw_0602
15	76	ELADLDRQL	SEQ ID No: 62	A_0201
16	77	LADLDRQLI	SEQ ID No: 63	A_0101、B51
17	79	DLDRQLISL	SEQ ID No: 64	A_0101、A_0201
18	80	LDRQLISLI	SEQ ID No: 65	A24、B7、B51
19	94	AVQTTRDIL	SEQ ID No: 66	B7
20	103	EGAKKGLEF	SEQ ID No: 67	A24、B7
21	107	KGLEFVRPV	SEQ ID No: 68	A_0201、B51
22	108	GLEFVRPVA	SEQ ID No: 69	A_0101、A_0301
23	109	LEFVRPVAV	SEQ ID No: 70	B44
24	111	FVRPVAVDL	SEQ ID No: 71	B7、B8、B_3501
25	113	RPVAVDLTY	SEQ ID No: 72	B7、A_0101、B_3501、B51
26	116	AVDLTYIPV	SEQ ID No: 73	A2、A_0201
27	120	TYIPVVGHA	SEQ ID No: 74	A24
28	121	YIPVVGHAL	SEQ ID No: 75	A_0101、A2、A_0201、B7、B8
29	129	LSAAFQAPF	SEQ ID No: 76	A1、B7、B_3501
30	130	SAAFQAPFC	SEQ ID No: 77	A_0201
31	131	AAFQAPFCA	SEQ ID No: 78	A_0301、B_3501
32	133	FQAPFCAGA	SEQ ID No: 79	A2、A_0201
33	135	APFCAGAMA	SEQ ID No: 80	B7、B_3501
34	136	PFCAGAMAV	SEQ ID No: 81	A3
35	141	AMAVVGGAL	SEQ ID No: 82	A2、A_0201、A24、B7
36	143	AVVGGALAY	SEQ ID No: 83	A1、A3、A_0301、B7
37	147	GALAYLVVK	SEQ ID No: 84	A3、A_0301
38	149	LAYLVVKTL	SEQ ID No: 85	B8、B44、B51
39	150	AYLVVKTLI	SEQ ID No: 86	A24
40	155	KTLINATQL	SEQ ID No: 87	A_0201、A2、A_0301、A24

[0632]

推定的 CD8表 位号	氨基酸 位置	表位序列	SEQ ID No:	HLA等位基因
41	156	TLINATQLL	SEQ ID No: 88	A2、A_0201、A3、A_0101、Cw_0401
42	158	INATQLLKL	SEQ ID No: 89	B7、B8、Cw_0602
43	159	NATQLLKLL	SEQ ID No: 90	A_2402、B7、B_3501、B44、Cw_0401、 Cw_0602
44	162	QLLKLLAKL	SEQ ID No: 91	A2、A_0201、A_0301、A_2402、B8、 Cw_0401、Cw_0602
45	165	KLLAKLAEL	SEQ ID No: 92	A2、A_0201、A_0301、B7、B8、Cw_0602
46	166	LLAKLAELV	SEQ ID No: 93	A2、A_0201、A_0101、B8
47	169	KLAELVAAA	SEQ ID No: 94	A2
48	170	LAELVAAAI	SEQ ID No: 95	A1、A24、B51
49	173	LVAAAIADI	SEQ ID No: 96	B7、B51
50	177	AIADIISDV	SEQ ID No: 97	A2、A_0201、Cw_0602
51	178	IADIISDVA	SEQ ID No: 98	A_0101、B_3501
52	182	ISDVADIHK	SEQ ID No: 99	A1、A_0301
53	192	TLGEVWEFI	SEQ ID No: 100	A2、A_0201
54	199	FITNALNGL	SEQ ID No: 101	A2
55	202	NALNGLKEL	SEQ ID No: 102	B51、A_2402、B_3501、Cw_0602
56	213	KLTGWVTGL	SEQ ID No: 103	A2、A_0201
57	214	LTGWVTGLF	SEQ ID No: 104	A1、A_0101、A24
58	225	GWSNLESFF	SEQ ID No: 105	A24
59	228	NLESFFAGV	SEQ ID No: 106	A2、A_0201
60	231	SFFAGVPGL	SEQ ID No: 107	A2、A_0201、A24、Cw_0401
61	238	GLTGATSGL	SEQ ID No: 108	A2、A_0201
62	246	LSQVTGLFG	SEQ ID No: 109	A1、B8
63	247	SQVTGLFGA	SEQ ID No: 110	A2
64	258	LSASSGLAH	SEQ ID No: 111	A1、A3、B7、B8
65	260	ASSGLAHAD	SEQ ID No: 112	A1、A3、A_0301
66	262	SGLAHADSL	SEQ ID No: 113	A_0201
67	263	GLAHADSLA	SEQ ID No: 114	A_0101、A_0201、A_0301
68	269	SLASSASLP	SEQ ID No: 115	A_0201、A_0301
69	271	ASSASLPAL	SEQ ID No: 116	B7
70	286	SGFGGLPSL	SEQ ID No: 117	A2、A_0201、B51
71	291	LPSLAQVHA	SEQ ID No: 118	B7、B_3501、B51
72	298	HAASTRQAL	SEQ ID No: 119	B7、B8、B_3501

[0633]

推定的 CD8表 位号	氨基酸 位置	表位序列	SEQ ID No:	HLA等位基因
73	301	STRQALRPR	SEQ ID No: 120	A3、A_0301
74	307	RPRADGPVG	SEQ ID No: 121	B7、B_0702、B8、B51
75	319	EQVGGQSQL	SEQ ID No: 122	B7、B44
76	350	GASKGTTTK	SEQ ID No: 123	A3、A_0301
77	351	ASKGTTTKK	SEQ ID No: 124	A3、A_0301
78	353	KGTTTKKYS	SEQ ID No: 125	A_0301、B8
79	368	TEDAERAPV	SEQ ID No: 126	B44

[0634] 从表2和3可知,Rv3616c包含许多预测的CD4+和CD8 T细胞表位。此外,该信息表明所述蛋白携带可由在世界范围内存在的HLA(即来自高加索、非洲、亚洲或拉美个体的HLA-参见网址www.allelefrequencies.net)所识别的表位。

[0635] 实施例3-Rv3616c表位鉴定

[0636] 制备一系列覆盖全长Rv3616c的30个重叠肽(关于详细资料参见图1和SEQ ID No: 127-156),并测试其刺激来自4个PPD+供体的PBMC的能力。

[0637] 示于图2的数据表明肽1-7和17-30对这些个体为免疫原性的。

[0638] 应指出的是,肽8-16(氨基酸残基92-215)在不同HLA型的其它个体中可为免疫原性的。

[0639] 实施例4-H37Rv同源物

[0640] 使用BLASTP GenBank搜索(H37Rv参比序列登记号NP_218133.1)鉴定了来自许多结核分枝杆菌菌株和BCG的Rv3616c序列:

[0641]

菌株	登记号	%同一性
CDC1551	NP_338263.1	99
F11	YP_001289574.1	99
Haarlem	ZP_02248979.1	99

[0642]

C	ZP_00877472.1	99
BCG	YP_979759.1	99

[0643] 同源序列的比对表明高水平的同一性。

[0644] 生物学测定

[0645] 针对Rv3616c的T细胞应答的定量

[0646] 可针对其在来自感染(例如潜伏性感染)个体的外周血单核细胞(PBMC)或全血制品中激活T细胞(诱导增殖和/或产生细胞因子)的能力筛选多肽。

[0647] 通常通过具有直径大于10mm的皮试且无症状、无Mtb阳性培养物、阴性的痰阴性以及无损伤(通过胸部X射线检测),来鉴定潜伏性感染个体。

[0648] 可基于PBMC样品或全血利用一系列体外测定:在抗原(或其变体/免疫原性片段(当适当时))存在下再刺激之后,可测定细胞增殖(通过CFSE/流式细胞仪测定)或定量细胞

因子的产生(存在于培养细胞上清液中,并通过ELISA测定,或者在CD4和CD8 T细胞胞内染色后,通过流式细胞仪分析)。

[0649] 例如,PBMC样品可通过按照标准程序进行Ficoll-Hypaque密度梯度离心获得自肝素化全血。然后可洗涤所述细胞并冷冻保存于液氮中直至测试(关于进一步资料参见Lalvani A等J.Infect.Dis.1999 180:1656-1664)。

[0650] T细胞增殖

[0651] 特异性免疫应答可通过使用氘化胸苷进行淋巴细胞增殖分析来表征。该技术评价在体外抗原刺激时的细胞扩繁。实际上,通过估计氘化胸苷掺入DNA中(与细胞数量的潜在变化紧密相关的过程)来确定细胞增殖。

[0652] 更适合的是,可使用羧基荧光素二乙酸琥珀酰亚胺酯(CFSE)来进行淋巴细胞增殖。CFSE通过与赖氨酸侧链和其它可用的胺基反应而自发地且不可逆地偶联至细胞内蛋白和细胞表面蛋白两者。当淋巴细胞分裂时,CFSE标记在子细胞之间等分,因而所述子细胞为亲代荧光性的一半。因此,在增殖细胞群体中减半的细胞荧光强度标记各个连续世代,并且容易通过流式细胞仪追踪(关于进一步细节参见Hodgkins,PD等J.Exp.Med.1996 184:277-281)。

[0653] 实际上,在解冻后,可将PBMC洗涤并用CFSE染色,然后再用含10ug/ml抗原的培养基(补充有谷氨酰胺、非必需氨基酸、丙酮酸盐和热失活的人AB血清的RPMI-1640)培养(2×10^5 个细胞)72hr。然后可收获细胞并使用表面染色表征其表型,以鉴定记忆CD8和CD4+T细胞。随后,可将流式细胞仪分析用于显示应答于各种抗原的淋巴细胞增殖程度(在体外刺激时CFSE强度降低的细胞的比例)。

[0654] 细胞因子产生

[0655] IFN- γ 产生(或其它细胞因子例如IL2、TNF- α 、IL5、IL12等的产生)可使用酶联免疫吸附测定(ELISA)进行测量。可用针对人IFN- γ 的小鼠单克隆抗体(PharMingen, San Diego, CA)在室温下于PBS中达4小时来包被ELISA板。然后将各个孔用含5%(W/V)脱脂奶粉的PBS于室温下封闭1小时。然后将板例如在PBS/0.2%TWEEN-20中洗涤6次,并将在培养基中稀释为1:2的样品在ELISA板中于室温下温育过夜。将板再次洗涤,然后可将例如在PBS/10%正常山羊血清中稀释为1:3000的多克隆兔抗-人IFN- γ 血清加至各个孔中。然后将板在室温下保温2小时,洗涤,然后可例如在PBS/5%脱脂奶粉中以1:2000稀释度加入辣根过氧化物酶偶联的抗-兔IgG(Sigma Chemical So., St. Louis, MO)。在室温下再次保温2小时后,将板洗涤并加入TMB底物。反应可在20分钟后用1N硫酸终止。然后可在450nm下使用570nm作为参比波长测定光密度。通常,导致两个重复都产生OD为培养于仅培养基的细胞平均OD的2倍的部分(fraction),可认为是阳性的。

[0656] 实施例5-在CB6F1小鼠中的免疫原性

[0657] 抗原的免疫原性在CB6F1小鼠(BALB/c与C57BL/6小鼠杂交的第一代)中评价。

[0658] 将CB6F1小鼠用0.5ug蛋白抗原联合Adjuvant System AS01E(含3D-MPL和QS21的脂质体佐剂)肌内免疫接种三次(在第0天、第14天和第28天)。

[0659] 实验设计如下:

[0660]

组别	第0天	第14天	第28天
----	-----	------	------

1	0.5ug Rv3616c/AS01E	0.5ug Rv3616c/AS01E	0.5ug Rv3616c/AS01E
---	---------------------	---------------------	---------------------

[0661] 总共24只小鼠用于方案组别中。

[0662] 在第21天(即第二次免疫接种后第7天)和第35天(即第三次免疫接种后第7天)采集并合并外周血淋巴细胞(PBL),然后在用覆盖目标序列的15聚物肽的合并物(pool)体外再刺激过夜之后,用流式细胞术测量抗原特异性CD4和CD8 T细胞应答(通过产生IL-2和/或IFN- γ 和/或TNF- α 的CD4或CD8 T细胞来测定)。

[0663] 通过使用短期抗原驱动表达的细胞因子表达的体外扩增来检测表达IL-2和/或IFN- γ 和/或TNF- α 的小鼠T细胞。

[0664] 简言之,将PharmLyse溶液(BD-Pharmingen)加至肝素化小鼠外周血以便溶解红细胞。将所得的PBL(外周血淋巴细胞)洗涤,随后在覆盖目标抗原序列的15-聚物肽(重叠11个氨基酸)的合并物以及1ug/ml CD28抗体和CD49d抗体(BD-Pharmingen)存在下温育。以终浓度为1ug/ml使用各个15-聚物肽。培养基对照也用CD28抗体和CD49d抗体刺激。

[0665] 在37°C、5%CO₂下开始培养2h后,将细胞因子分泌封闭化合物布雷菲德菌素-A(BD-Pharmingen)加入,并将细胞保持在37°C、5%CO₂下达另外4小时,然后于+4°C下温育过夜。

[0666] 然后收获细胞并用Pacific Blue偶联的抗-CD4(BD-克隆RM4-5,BD-Pharmingen)和多甲藻素叶绿素A蛋白(PerCp)花青素昔5.5(Cy5.5)偶联的抗-CD8 α (克隆53-6.7,BD-Pharmingen)抗体染色。

[0667] 然后将细胞洗涤、固定、透化(Cytofix-cytoperm试剂盒,BD-Pharmingen)以及用别藻蓝蛋白偶联的抗-IFN- γ 抗体(克隆XMG1.2,BDPharmingen)、异硫氰酸荧光素(FITC)偶联的抗-IL-2抗体(克隆JES6-5H4,Beckman Coulter)和藻红蛋白(PE)偶联的抗-TNF α 抗体(克隆MP6-XT22,BDPharmingen)染色。在最后的洗涤之后,在LSR II流式细胞仪(Beckton-Dickinson)上分析染色细胞。在CD8⁺子集中得到最少10,000个细胞。

[0668] 关于进一步的背景,参见Walzer T等Cell Immunol.2000 206(1):16-25和Maecker HT等J.Immunol.Methods 2001 255(1-2):27-40。

[0669] 作为阴性对照,某些细胞也在培养基中体外培养过夜(未刺激)。通过从肽刺激细胞产生的平均细胞因子反应中减去未刺激细胞产生的平均细胞因子反应来计算抗原特异性反应。

[0670] 在各个时间点和对于各个组别,从每组6只小鼠的4个合并物中收集数据。以下数据显示为产生IL-2和/或IFN- γ 和/或TNF- α 的CD4或CD8 T细胞%。对每个单独小鼠合并物(三角形)以及组别的平均值(条)作图。

[0671] 图3显示在第21天(即第二次免疫接种7天后),在用0.5ugRv3616c/AS01E免疫接种的小鼠中检测到Rv3616c特异性CD4和CD8T细胞应答。

[0672] 图4显示在第21天(即第二次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD4 T细胞应答的细胞因子概况。

[0673] 图5显示第21天(即第二次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD8 T细胞应答的细胞因子概况。

[0674] 图6显示在第35天(即第三次免疫接种7天后),在用0.5ug Rv3616c/AS01E免疫接种的小鼠中检测到Rv3616c特异性CD4和CD8 T细胞应答。第三剂增加CD4 T细胞应答但不增

加CD8 T细胞应答。由于技术问题,仅单个合并物获得数据。

[0675] 图7显示在第35天(即第三次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD4 T细胞应答的细胞因子概况。由于技术问题,仅单个合并物获得数据。

[0676] 图8显示在第35天(即第三次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD8 T细胞应答的细胞因子概况。由于技术问题,仅单个合并物获得数据。

[0677] 实施例6-在C57BL/6小鼠中的免疫原性

[0678] 抗原的免疫原性也在C57BL/6小鼠中评价。

[0679] 将C57BL/6小鼠用1 μ g蛋白抗原联合Adjuvant System AS01E(含3D-MPL和QS21的脂质体佐剂)肌内免疫接种三次(在第0天、第14天和第28天)。

[0680] 实验设计如下:

[0681]

组别	第0天	第14天	第28天
1	1 μ g Rv3616c/AS01E	1 μ g Rv3616c/AS01E	1 μ g Rv3616c/AS01E

[0682] 在第21天(即第二次免疫接种后第7天)和第35天(即第三次免疫接种后第7天)采集并合并外周血淋巴细胞(PBL),然后在用覆盖目标序列的15聚物肽的合并物体外再刺激过夜之后,用流式细胞术测量抗原特异性CD4和CD8 T细胞应答(通过产生IL-2和/或IFN- γ 和/或TNF- α 的CD4或CD8 T细胞来测定)。随后的程序如先前所述。

[0683] 作为阴性对照,某些细胞也在培养基中体外培养过夜(未刺激)。通过从肽刺激细胞产生的平均细胞因子反应中减去未刺激细胞产生的平均细胞因子反应来计算抗原特异性反应。

[0684] 在各个时间点和对于各个组别,从每组6只小鼠的4个合并物中收集数据。以下数据显示为产生IL-2和/或IFN- γ 和/或TNF- α 的CD4或CD8 T细胞%。对每个单独小鼠合并物(三角形)以及组别的平均值(条)作图。

[0685] 图9显示在第21天(即第二次免疫接种7天后),在用1 μ g Rv3616c/AS01E免疫接种的小鼠中检测到Rv3616c特异性CD4和CD8 T细胞应答,但抗原特异性CD8 T细胞应答非常低(因而未显示细胞因子概况数据)。

[0686] 图10显示第21天(即第二次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD4 T细胞应答的细胞因子概况。

[0687] 图11显示在第35天(即第三次免疫接种7天后),在用1 μ gRv3616c/AS01E免疫接种的小鼠中检测到Rv3616c特异性CD4和CD8 T细胞应答。第三剂免疫接种增加CD4 T细胞应答但仅轻微增加CD8 T细胞应答。

[0688] 图12显示第35天(即第三次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD4 T细胞应答的细胞因子概况。

[0689] 图13显示第35天(即第三次免疫接种7天后),来自Rv3616c肽合并物刺激的PBL(未除去培养基)的CD8 T细胞应答的细胞因子概况。

[0690] 实施例7-来自患有潜伏性TB的人的PBMC的体外识别

[0691] 在来自南非的4名无TB的健康成年人(PPD皮试=0mm)和8名潜伏性TB感染的健康

成年人(PPD皮试=15mm或以上)中,进行实验以评价特异于本发明抗原的外周T细胞应答。

[0692] PPD皮试数据

	个体 ID 号	硬块直径(mm)
	4	0
	5	0
[0693]	33	0
	38	0
	36	15
	46	15
	13	15
	7	16
	58	25
[0694]	74	26
	8	53
	60	55

[0695] 通过在分离的外周血单核细胞(PBMC)上经胞内细胞因子染色(ICS)测定测量细胞因子,来评价细胞介导的免疫(CMI)应答。

[0696] 所进行的ICS为先前所述方法(参见Von Eschen等,Hum.Vaccin.2009 5(7))的修改版。通过覆盖目标抗原整个序列的15聚物肽(重叠11个氨基酸)的一个合并物体外刺激PBMC。将细胞用肽刺激2小时,然后在布雷菲德菌素A存在下培养过夜,针对ICS进行处理并使用流式细胞术分析。测量表达IFN- γ 和/或TNF- α 和/或IL-17的抗原特异性CD3+CD4+或CD3+CD8+T细胞的频数。从肽合并物刺激的细胞所得的反应中减去培养基刺激的细胞应答。

[0697] ICS:抗体

[0698] 抗-CD3 PO(Invitrogen-目录号CD0330)

[0699] 抗-CD4 PB(BD-目录号558116)

[0700] 抗-CD8 APC-H7(BD-目录号641400)

[0701] 抗-IFN γ AF700(BD-Pharmingen-目录号557995)

[0702] 抗-TNF PE-Cy7(BD-Pharmingen-目录号557647)

[0703] 抗-IL17 AF647(BD-Pharmingen-目录号51-7178-71)

[0704] 将结果表示为每百万个CD3+CD4+T细胞中表达TNF- α 和IFN- γ 的抗原特异性CD3+CD4+T细胞数,因为这些细胞代表抗原特异性CD4 T细胞的主群体(去除了培养基所致的背景应答水平)。未检测到抗原特异性CD3+CD8+T细胞。图14显示:与在无感染个体中测得的非特异性CD4 T细胞应答相比,在8名潜伏性感染个体中检测到有6名存在抗原特异性CD4 T细胞应答(在个体号7和74中未测得)。

[0705] 总之,可指出Rv3616c抗原能够在CB6F1和C57BL/6小鼠两者中引发免疫应答。此外,细胞因子产生概况表明大部分抗原特异性T细胞表达多种Th1相关细胞因子(即引发多功能性T细胞应答)。重要的是,在免疫接种后CD4和CD8抗原特异性T细胞两者都出现,CD8细胞在潜伏性TB情况中可能尤其重要。根据来自南非的潜伏性感染个体中的高水平识别和无感染的受试者缺乏反应,确证Rv3616c与人类感染的相关性。因此,可预期Rv3616c在预防、治疗和诊断潜伏性结核病感染中有实质价值。

[0706] 尽管已通过用于明确理解的目的的阐述和实施例相当详细地描述前述发明,但是

对本领域普通技术人员显而易见的是,根据本发明的教导,可作出对本发明的某些改变和修改而不会背离所附权利要求的精神或范围。

[0707] 本申请提及的所有参考文献(包括专利和专利申请)通过引用以最大可能程度结合于本文中,如同各种出版物或专利申请具体且单独地指出通过引用结合。

[0708] 除非上下文另外需要,否则在整个说明书和随附的权利要求中,措辞“包括”及变体例如“包含”和“含有”应理解为意指包含所陈述的整数、步骤、整数组、步骤组,但并不排除任何其它整数、步骤、整数组或步骤组。

<110> 葛兰素史密丝克莱恩生物有限公司

<120> 新型组合物和方法

<130> VB63090PCT

<150> US61/083720

<151> 2008-07-25

<160> 159

<170> PatentIn version 3.5

<210> 1

<211> 392

<212> PRT

<213> 结核分枝杆菌

<220>

[0001]

<221> 其它特征

<223> 菌株 H37Rv

<400> 1

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1 5 10 15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
35 40 45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
50 55 60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65 70 75 80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
100 105 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
115 120 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
130 135 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145 150 155 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
165 170 175

[0002] Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Thr
180 185 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
195 200 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
210 215 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225 230 235 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
245 250 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
260 265 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe

	275	280	285	
	Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln			
	290	295	300	
	Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln			
	305	310	315	320
	Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met			
		325	330	335
	Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser			
		340	345	350
	Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr			
		355	360	365
[0003]	Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln			
	370	375	380	
	Lys Val Leu Val Arg Asn Val Val			
	385	390		
	<210> 2			
	<211> 1179			
	<212> DNA			
	<213> 结核分枝杆菌			
	<220>			
	<221> 其它特征			
	<223> 菌株 H37Rv			
	<400> 2			
	atgagcagag cgttcacatc cgatccaacg atcagtgcca ttgacggcgtt gtacgacctt			60
	ctggggattg gaatacccaa ccaagggggt atcctttact cctcactaga gtacttcgaa			120
	aaagccctgg aggagctggc agcagcgttt cggggtgatg gctggttagg ttcggccgog			180
	gacaaatacg ccggcaaaaa ccgcaaccac gtgaattttt tccaggaact ggcagacctc			240
[0004]				

```

gatcgtcagc tcatcagcct gatccacgac caggccaacg cgggccagac gaccgcgac      300
atcctggagg ggcgaagaa aggtctcgag ttctgctgcc cgggtgctgt ggacctgacc      360
tacatcccg tctctgggca cgcctatctg gccgccttcc aggcgccttt ttgcgcgggc      420
gcgatggcgc tagtggggcg cgccttgcc tacttggtcg tgaaaacgct gatcaacgcg      480
actcaactcc tcaaattgct tgccaaattg gcggagtgg tcgcggccgc cattgoggac      540
atcatttcgg atgtggcgga catcatcaag ggcacctctg gagaagtgtg ggagttcatc      600
acaaacgcgc tcaacggcct gaaagagctt tgggacaagc tcacggggtg ggtgaccgga      660
ctgtttcttc gaggggtgct gaacctggag tccttctttg cgggcgtccc cggettgcac      720
ggcgcgacca gggcttgct gcaagtgaact ggcttgctcg gtgcggccgg tctgtccgca      780
tcgtcgggct tggtcacgc ggatagcctg gcgagctcag ccagcttgcc cgccttgcc      840
ggcattgggg ggggttcgg ttttgggggc ttgccgagcc tggtcaggt ccatgcgcc      900
tcaactcggc aggcgctacg gccccgagct gatggcccg tcggcgcgc tgccgagcag      960
gtcggcgggc agtcgcagct ggtctccgc cagggttccc aaggtatggg cggaccgta      1020
ggcatgggc gcatgcacc ctcttcgggg gcgtcgaaag ggacgacgac gaagaagtac      1080
tcggaaggcg cggcgcggg cactgaagac gccgagcgc gccagtcga agctgacgcg      1140
ggcgggtggc aaaaggtgct ggtacgaaac gtctctaa      1179

```

<210> 3

<211> 392

<212> PRT

<213> 结核分枝杆菌

<220>

<221> 其它特征

<223> 菌株 CDC1551

<400> 3

```

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1           5           10          15

```

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
35 40 45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
50 55 60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65 70 75 80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
100 105 110

[0005]

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
115 120 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
130 135 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145 150 155 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
165 170 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
180 185 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
195 200 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
210 215 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225 230 235 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
245 250 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
260 265 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
275 280 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
290 295 300

[0006]

Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln
305 310 315 320

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
325 330 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
340 345 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
355 360 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
370 375 380

Lys Val Leu Val Arg Asn Val Val
385 390

<210> 4

<211> 392
 <212> PRT
 <213> 结核分枝杆菌

<220>
 <221> 其它特征
 <223> 菌株 F11

<400> 4

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
 1 5 10 15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
 20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
 35 40 45

[0007] Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
 50 55 60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
 65 70 75 80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
 85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
 100 105 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
 115 120 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
 130 135 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
 145 150 155 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
165 170 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
180 185 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
195 200 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
210 215 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225 230 235 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
245 250 255

[0008]

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
260 265 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
275 280 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
290 295 300

Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln
305 310 315 320

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
325 330 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
340 345 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
355 360 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
370 375 380

Lys Val Leu Val Arg Asn Val Val
385 390

<210> 5
<211> 392
<212> PRT
<213> 结核分枝杆菌

<220>
<221> 其它特征
<223> 菌株 Haarlem A

[0009] <400> 5

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1 5 10 15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
35 40 45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
50 55 60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65 70 75 80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
100 105 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
115 120 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
130 135 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145 150 155 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
165 170 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
180 185 190

[0010]

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
195 200 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
210 215 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225 230 235 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
245 250 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
260 265 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
275 280 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln

290

295

300

Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln
 305 310 315 320

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
 325 330 335

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
 340 345 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
 355 360 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
 370 375 380

[0011]

Lys Val Leu Val Arg Asn Val Val
 385 390

<210> 6
 <211> 392
 <212> PRT
 <213> 结核分枝杆菌

<220>
 <221> 其它特征
 <223> 菌株 C

<400> 6

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
 1 5 10 15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
 20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
 35 40 45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
50 55 60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65 70 75 80

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
100 105 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
115 120 125

[0012] Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
130 135 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145 150 155 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
165 170 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile
180 185 190

Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys
195 200 205

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
210 215 220

Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr
225 230 235 240

Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala
245 250 255

Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser
260 265 270

Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe
275 280 285

Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln
290 295 300

Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln
305 310 315 320

Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met
325 330 335

[0013]

Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser
340 345 350

Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr
355 360 365

Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln
370 375 380

Lys Val Leu Val Arg Asn Val Val
385 390

<210> 7
<211> 392
<212> PRT
<213> 牛分枝杆菌

<220>
<221> 其它特征

<223> 菌株 BCG

<400> 7

Met Ser Arg Val Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1 5 10 15

Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly Gly Ile Leu
20 25 30

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala Leu Glu Glu Leu Ala Ala
35 40 45

Ala Phe Pro Gly Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala
50 55 60

Gly Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu
65 70 75 80

[0014]

Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala Val Gln
85 90 95

Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys Gly Leu Glu Phe Val
100 105 110

Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly His Ala
115 120 125

Leu Ser Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val
130 135 140

Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn Ala
145 150 155 160

Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu Leu Val Ala Ala
165 170 175

Ala Ile Ala Asp Ile Ile Ser Asp Val Ala Asp Ile Ile Lys Gly Ile

	180	185	190
	Leu Gly Glu Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys 195	200	205
	Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg 210	215	220
	Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro Gly Leu Thr 225	230	235 240
	Gly Ala Thr Ser Gly Leu Ser Gln Val Thr Gly Leu Phe Gly Ala Ala 245	250	255
	Gly Leu Ser Ala Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser 260	265	270
[0015]	Ser Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe 275	280	285
	Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr Arg Gln 290	295	300
	Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly Ala Ala Ala Glu Gln 305	310	315 320
	Val Gly Gly Gln Ser Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met 325	330	335
	Gly Gly Pro Val Gly Met Gly Gly Met His Pro Ser Ser Gly Ala Ser 340	345	350
	Lys Gly Thr Thr Thr Lys Lys Tyr Ser Glu Gly Ala Ala Ala Gly Thr 355	360	365
	Glu Asp Ala Glu Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln 370	375	380

Lys Val Leu Val Arg Asn Val Val
385 390

<210> 8
<211> 110
<212> PRT
<213> 结核分枝杆菌

<220>
<221> 成熟肽
<222> (29)..(110)

<400> 8

Met Arg Leu Ser Leu Thr Ala Leu Ser Ala Gly Val Gly Ala Val Ala
-25 -20 -15

Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro Val Asp
-10 -5 -1 1

[0016]

Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu
5 10 15 20

Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser Pro Val
25 30 35

Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro Gln Arg
40 45 50

Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala Gln Tyr
55 60 65

Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr
70 75 80

<210> 9
<211> 97
<212> PRT

<213> 结核分枝杆菌

<400> 9

Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val Ala Ser Gln Ser
1 5 10 15

Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His Thr Ile Gly Gln Ala
20 25 30

Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His Gln Gly Glu Ser Ser
35 40 45

Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala Ala Ala Lys
50 55 60

Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu Gly Glu Ala Ala
65 70 75 80

[0017] Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ser Thr Tyr Thr Gly
85 90 95

Phe

<210> 10

<211> 94

<212> PRT

<213> 结核分枝杆菌

<400> 10

Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala His Gly Ala Met
1 5 10 15

Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu His Gln Ala Ile Val
20 25 30

Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly Gly Ala Gly Ser Val
35 40 45

Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg Asn Phe Gln Val Ile
50 55 60

Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln Ala Ala Gly Asn
65 70 75 80

Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser Trp Ala
85 90

<210> 11

<211> 132

<212> PRT

<213> 结核分枝杆菌

<400> 11

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1 5 10 15

[0018]

Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
20 25 30

Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
35 40 45

Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
50 55 60

Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
65 70 75 80

Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
85 90 95

Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp
100 105 110

Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
115 120 125

Gly Pro Pro Ala
130

<210> 12
<211> 195
<212> PRT
<213> 结核分枝杆菌

<400> 12

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
1 5 10 15

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
20 25 30

[0019] Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
35 40 45

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
50 55 60

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
65 70 75 80

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
85 90 95

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
100 105 110

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
115 120 125

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
130 135 140

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

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
165 170 175

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
180 185 190

Ala Ala Ser
195

<210> 13

<211> 391

<212> PRT

<213> 结核分枝杆菌

[0020] <400> 13

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

Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp
20 25 30

Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
35 40 45

Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
50 55 60

Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
65 70 75 80

Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
85 90 95

Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala
100 105 110

Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly
115 120 125

Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met
130 135 140

Trp Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala
145 150 155 160

Thr Ala Thr Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr
165 170 175

Ser Ala Gly Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser
180 185 190

[0021]

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
195 200 205

Gln Gln Leu Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu
210 215 220

Gly Gly Leu Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn
225 230 235 240

Met Val Ser Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val
245 250 255

Ser Met Thr Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala
260 265 270

Ala Ala Ala Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala
275 280 285

Met Ser Ser Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly

290 295 300

Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val
305 310 315 320

Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg
325 330 335

Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly
340 345 350

Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly
355 360 365

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met
370 375 380

[0022] Pro His Ser Pro Ala Ala Gly
385 390

<210> 14

<211> 423

<212> PRT

<213> 结核分枝杆菌

<400> 14

Met Asp Phe Gly Leu Leu Pro Pro Glu Val Asn Ser Ser Arg Met Tyr
1 5 10 15

Ser Gly Pro Gly Pro Glu Ser Met Leu Ala Ala Ala Ala Ala Trp Asp
20 25 30

Gly Val Ala Ala Glu Leu Thr Ser Ala Ala Val Ser Tyr Gly Ser Val
35 40 45

Val Ser Thr Leu Ile Val Glu Pro Trp Met Gly Pro Ala Ala Ala Ala
50 55 60

Met Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp Leu Ala Ala Thr Ala
65 70 75 80

Ala Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg Ala Ala Ala Glu Ala
85 90 95

Phe Gly Thr Ala Phe Ala Met Thr Val Pro Pro Ser Leu Val Ala Ala
100 105 110

Asn Arg Ser Arg Leu Met Ser Leu Val Ala Ala Asn Ile Leu Gly Gln
115 120 125

Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu Tyr Ala Glu Met Trp
130 135 140

Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu Gly Ala Ser Ala Ala
145 150 155 160

[0023]

Ala Ser Ala Leu Pro Pro Phe Thr Pro Pro Val Gln Gly Thr Gly Pro
165 170 175

Ala Gly Pro Ala Ala Ala Ala Ala Thr Gln Ala Ala Gly Ala Gly
180 185 190

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

Leu Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn Ala Asp Pro Leu Thr
210 215 220

Ser Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn Pro Gln Val Gly Ser
225 230 235 240

Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly Glu Leu Asp Val Ile
245 250 255

Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser Ile Ala Leu Ala Ile
260 265 270

Thr Asn Thr Ala Arg Pro Trp His Ile Gly Leu Tyr Gly Asn Ala Gly
275 280 285

Gly Leu Gly Pro Thr Gln Gly His Pro Leu Ser Ser Ala Thr Asp Glu
290 295 300

Pro Glu Pro His Trp Gly Pro Phe Gly Gly Ala Ala Pro Val Ser Ala
305 310 315 320

Gly Val Gly His Ala Ala Leu Val Gly Ala Leu Ser Val Pro His Ser
325 330 335

Trp Thr Thr Ala Ala Pro Glu Ile Gln Leu Ala Val Gln Ala Thr Pro
340 345 350

[0024]

Thr Phe Ser Ser Ser Ala Gly Ala Asp Pro Thr Ala Leu Asn Gly Met
355 360 365

Pro Ala Gly Leu Leu Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg
370 375 380

Gly Thr Thr Gly Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly
385 390 395 400

Gln Glu Asp Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro
405 410 415

Pro Pro Gly Asn Pro Pro Arg
420

<210> 15

<211> 95

<212> PRT

<213> 结核分枝杆菌

<220>

<221> 起始 MET

<222> (1).. (1)

<220>

<221> 成熟肽

<222> (2).. (95)

<400> 15

Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser
-1 1 5 10 15

Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly
20 25 30

Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser
35 40 45

[0025] Glu Ala Tyr Gln Gly Val Gln Gln Lys Trp Asp Ala Thr Ala Thr Glu
50 55 60

Leu Asn Asn Ala Leu Gln Asn Leu Ala Arg Thr Ile Ser Glu Ala Gly
65 70 75

Gln Ala Met Ala Ser Thr Glu Gly Asn Val Thr Gly Met Phe Ala
80 85 90

<210> 16

<211> 338

<212> PRT

<213> 结核分枝杆菌

<220>

<221> 成熟肽

<222> (43).. (338)

<400> 16

Met Gln Leu Val Asp Arg Val Arg Gly Ala Val Thr Gly Met Ser Arg
-40 -35 -30

Arg Leu Val Val Gly Ala Val Gly Ala Ala Leu Val Ser Gly Leu Val
-25 -20 -15

Gly Ala Val Gly Gly Thr Ala Thr Ala Gly Ala Phe Ser Arg Pro Gly
-10 -5 -1 1 5

Leu Pro Val Glu Tyr Leu Gln Val Pro Ser Pro Ser Met Gly Arg Asp
10 15 20

Ile Lys Val Gln Phe Gln Ser Gly Gly Ala Asn Ser Pro Ala Leu Tyr
25 30 35

Leu Leu Asp Gly Leu Arg Ala Gln Asp Asp Phe Ser Gly Trp Asp Ile
40 45 50

Asn Thr Pro Ala Phe Glu Trp Tyr Asp Gln Ser Gly Leu Ser Val Val
55 60 65 70

[0026]

Met Pro Val Gly Gly Gln Ser Ser Phe Tyr Ser Asp Trp Tyr Gln Pro
75 80 85

Ala Cys Gly Lys Ala Gly Cys Gln Thr Tyr Lys Trp Glu Thr Phe Leu
90 95 100

Thr Ser Glu Leu Pro Gly Trp Leu Gln Ala Asn Arg His Val Lys Pro
105 110 115

Thr Gly Ser Ala Val Val Gly Leu Ser Met Ala Ala Ser Ser Ala Leu
120 125 130

Thr Leu Ala Ile Tyr His Pro Gln Gln Phe Val Tyr Ala Gly Ala Met
135 140 145 150

Ser Gly Leu Leu Asp Pro Ser Gln Ala Met Gly Pro Thr Leu Ile Gly
155 160 165

Leu Ala Met Gly Asp Ala Gly Gly Tyr Lys Ala Ser Asp Met Trp Gly
170 175 180

Pro Lys Glu Asp Pro Ala Trp Gln Arg Asn Asp Pro Leu Leu Asn Val
185 190 195

Gly Lys Leu Ile Ala Asn Asn Thr Arg Val Trp Val Tyr Cys Gly Asn
200 205 210

Gly Lys Pro Ser Asp Leu Gly Gly Asn Asn Leu Pro Ala Lys Phe Leu
215 220 225 230

Glu Gly Phe Val Arg Thr Ser Asn Ile Lys Phe Gln Asp Ala Tyr Asn
235 240 245

Ala Gly Gly Gly His Asn Gly Val Phe Asp Phe Pro Asp Ser Gly Thr
250 255 260

[0027]

His Ser Trp Glu Tyr Trp Gly Ala Gln Leu Asn Ala Met Lys Pro Asp
265 270 275

Leu Gln Arg Ala Leu Gly Ala Thr Pro Asn Thr Gly Pro Ala Pro Gln
280 285 290

Gly Ala
295

<210> 17
<211> 325
<212> PRT
<213> 结核分枝杆菌

<220>
<221> 成熟肽
<222> (41)..(325)

<400> 17

Met Thr Asp Val Ser Arg Lys Ile Arg Ala Trp Gly Arg Arg Leu Met			
-40	-35	-30	-25
Ile Gly Thr Ala Ala Ala Val Val Leu Pro Gly Leu Val Gly Leu Ala			
	-20	-15	-10
Gly Gly Ala Ala Thr Ala Gly Ala Phe Ser Arg Pro Gly Leu Pro Val			
	-5	-1 1	5
Glu Tyr Leu Gln Val Pro Ser Pro Ser Met Gly Arg Asp Ile Lys Val			
10	15	20	
Gln Phe Gln Ser Gly Gly Asn Asn Ser Pro Ala Val Tyr Leu Leu Asp			
25	30	35	40
Gly Leu Arg Ala Gln Asp Asp Tyr Asn Gly Trp Asp Ile Asn Thr Pro			
	45	50	55
[0028] Ala Phe Glu Trp Tyr Tyr Gln Ser Gly Leu Ser Ile Val Met Pro Val			
	60	65	70
Gly Gly Gln Ser Ser Phe Tyr Ser Asp Trp Tyr Ser Pro Ala Cys Gly			
	75	80	85
Lys Ala Gly Cys Gln Thr Tyr Lys Trp Glu Thr Phe Leu Thr Ser Glu			
90	95	100	
Leu Pro Gln Trp Leu Ser Ala Asn Arg Ala Val Lys Pro Thr Gly Ser			
105	110	115	120
Ala Ala Ile Gly Leu Ser Met Ala Gly Ser Ser Ala Met Ile Leu Ala			
	125	130	135
Ala Tyr His Pro Gln Gln Phe Ile Tyr Ala Gly Ser Leu Ser Ala Leu			
	140	145	150
Leu Asp Pro Ser Gln Gly Met Gly Pro Ser Leu Ile Gly Leu Ala Met			

	155	160	165
	Gly Asp Ala Gly Gly Tyr Lys Ala Ala Asp Met Trp Gly Pro Ser Ser 170	175	180
	Asp Pro Ala Trp Glu Arg Asn Asp Pro Thr Gln Gln Ile Pro Lys Leu 185	190	195 200
	Val Ala Asn Asn Thr Arg Leu Trp Val Tyr Cys Gly Asn Gly Thr Pro 205	210	215
	Asn Glu Leu Gly Gly Ala Asn Ile Pro Ala Glu Phe Leu Glu Asn Phe 220	225	230
	Val Arg Ser Ser Asn Leu Lys Phe Gln Asp Ala Tyr Asn Ala Ala Gly 235	240	245
[0029]	Gly His Asn Ala Val Phe Asn Phe Pro Pro Asn Gly Thr His Ser Trp 250	255	260
	Glu Tyr Trp Gly Ala Gln Leu Asn Ala Met Lys Gly Asp Leu Gln Ser 265	270	275 280
	Ser Leu Gly Ala Gly 285		
	<210> 18		
	<211> 144		
	<212> PRT		
	<213> 结核分枝杆菌		
	<220>		
	<221> 起始 MET		
	<222> (1).. (1)		
	<220>		
	<221> 成熟肽		
	<222> (2).. (144)		

<400> 18

Met Ala Thr Thr Leu Pro Val Gln Arg His Pro Arg Ser Leu Phe Pro
 -1 1 5 10 15

Glu Phe Ser Glu Leu Phe Ala Ala Phe Pro Ser Phe Ala Gly Leu Arg
 20 25 30

Pro Thr Phe Asp Thr Arg Leu Met Arg Leu Glu Asp Glu Met Lys Glu
 35 40 45

Gly Arg Tyr Glu Val Arg Ala Glu Leu Pro Gly Val Asp Pro Asp Lys
 50 55 60

Asp Val Asp Ile Met Val Arg Asp Gly Gln Leu Thr Ile Lys Ala Glu
 65 70 75

Arg Thr Glu Gln Lys Asp Phe Asp Gly Arg Ser Glu Phe Ala Tyr Gly
 80 85 90 95

[0030]

Ser Phe Val Arg Thr Val Ser Leu Pro Val Gly Ala Asp Glu Asp Asp
 100 105 110

Ile Lys Ala Thr Tyr Asp Lys Gly Ile Leu Thr Val Ser Val Ala Val
 115 120 125

Ser Glu Gly Lys Pro Thr Glu Lys His Ile Gln Ile Arg Ser Thr Asn
 130 135 140

<210> 19

<211> 228

<212> PRT

<213> 结核分枝杆菌

<220>

<221> 成熟肽

<222> (24)..(228)

<400> 19

Met Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu Leu Cys
-20 -15 -10

Cys Ser Gly Val Ala Thr Ala Ala Pro Lys Thr Tyr Cys Glu Glu Leu
-5 -1 1 5

Lys Gly Thr Asp Thr Gly Gln Ala Cys Gln Ile Gln Met Ser Asp Pro
10 15 20 25

Ala Tyr Asn Ile Asn Ile Ser Leu Pro Ser Tyr Tyr Pro Asp Gln Lys
30 35 40

Ser Leu Glu Asn Tyr Ile Ala Gln Thr Arg Asp Lys Phe Leu Ser Ala
45 50 55

Ala Thr Ser Ser Thr Pro Arg Glu Ala Pro Tyr Glu Leu Asn Ile Thr
60 65 70

[0031]

Ser Ala Thr Tyr Gln Ser Ala Ile Pro Pro Arg Gly Thr Gln Ala Val
75 80 85

Val Leu Lys Val Tyr Gln Asn Ala Gly Gly Thr His Pro Thr Thr Thr
90 95 100 105

Tyr Lys Ala Phe Asp Trp Asp Gln Ala Tyr Arg Lys Pro Ile Thr Tyr
110 115 120

Asp Thr Leu Trp Gln Ala Asp Thr Asp Pro Leu Pro Val Val Phe Pro
125 130 135

Ile Val Gln Gly Glu Leu Ser Lys Gln Thr Gly Gln Gln Val Ser Ile
140 145 150

Ala Pro Asn Ala Gly Leu Asp Pro Val Asn Tyr Gln Asn Phe Ala Val
155 160 165

Thr Asn Asp Gly Val Ile Phe Phe Phe Asn Pro Gly Glu Leu Leu Pro
170 175 180 185

Glu Ala Ala Gly Pro Thr Gln Val Leu Val Pro Arg Ser Ala Ile Asp
190 195 200

Ser Met Leu Ala
205

<210> 20
<211> 355
<212> PRT
<213> 结核分枝杆菌

<220>
<221> 成熟肽
<222> (33).. (355)

<400> 20

[0032]

Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
-30 -25 -20

Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
-15 -10 -5 -1

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
1 5 10 15

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
20 25 30

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
35 40 45

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
50 55 60

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln

65	70	75	80
Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala	85	90	95
Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly	100	105	110
Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly	115	120	125
Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu	130	135	140
Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr	145	150	155
[0033] Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser	165	170	175
Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr	180	185	190
Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala	195	200	205
Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly	210	215	220
Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu	225	230	235
Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val	245	250	255
Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile	260	265	270

Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
275 280 285

Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp Gln
290 295 300

Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
305 310 315 320

Pro Pro Ala

<210> 21
<211> 323
<212> PRT
<213> 人工序列

[0034] <220>
<223> 成熟 Mtb32A 的 Ser/Ala 突变体

<400> 21

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
1 5 10 15

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
20 25 30

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
35 40 45

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
50 55 60

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
65 70 75 80

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala

	85	90	95
	Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly		
	100	105	110
	Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly		
	115	120	125
	Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu		
	130	135	140
	Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr		
	145	150	155
	Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala		
	165	170	175
[0035]	Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr		
	180	185	190
	Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala		
	195	200	205
	Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly		
	210	215	220
	Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu		
	225	230	235
	Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val		
	245	250	255
	Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile		
	260	265	270
	Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp		
	275	280	285

Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr Trp Gln
290 295 300

Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
305 310 315 320

Pro Pro Ala

<210> 22

<211> 96

<212> PRT

<213> 结核分枝杆菌

<400> 22

Met Ser Gln Ile Met Tyr Asn Tyr Pro Ala Met Leu Gly His Ala Gly
1 5 10 15

[0036]

Asp Met Ala Gly Tyr Ala Gly Thr Leu Gln Ser Leu Gly Ala Glu Ile
20 25 30

Ala Val Glu Gln Ala Ala Leu Gln Ser Ala Trp Gln Gly Asp Thr Gly
35 40 45

Ile Thr Tyr Gln Ala Trp Gln Ala Gln Trp Asn Gln Ala Met Glu Asp
50 55 60

Leu Val Arg Ala Tyr His Ala Met Ser Ser Thr His Glu Ala Asn Thr
65 70 75 80

Met Ala Met Met Ala Arg Asp Thr Ala Glu Ala Ala Lys Trp Gly Gly
85 90 95

<210> 23

<211> 723

<212> PRT

<213> 人工序列

<220>

<223> Mtb72f

<400> 23

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1 5 10 15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
20 25 30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
35 40 45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
50 55 60

[0037] Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65 70 75 80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
85 90 95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
100 105 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
115 120 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
130 135 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145 150 155 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
165 170 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
180 185 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
195 200 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
210 215 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225 230 235 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
245 250 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
260 265 270

[0038]

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
275 280 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
290 295 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305 310 315 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
325 330 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
340 345 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
355 360 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
370 375 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Gln Ala Val Gln Thr Ala
405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
450 455 460

[0039] Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
530 535 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545 550 555 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr

	565	570	575
	Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val		
	580	585	590
	Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln		
	595	600	605
	Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala		
	610	615	620
	Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly		
	625	630	635 640
	Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly		
	645	650	655
[0040]	Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu		
	660	665	670
	Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr		
	675	680	685
	Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser		
	690	695	700
	Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr		
	705	710	715 720
	Ala Ala Ser		
	<210> 24		
	<211> 723		
	<212> PRT		
	<213> 人工序列		
	<220>		

<223> M72

<400> 24

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1 5 10 15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
20 25 30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
35 40 45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
50 55 60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65 70 75 80

[0041]

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
85 90 95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
100 105 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
115 120 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
130 135 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145 150 155 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
165 170 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser

	180	185	190
	Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr		
	195	200	205
	Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala		
	210	215	220
	Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr		
	225	230	235 240
	Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu		
	245	250	255
	Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn		
	260	265	270
[0042]	Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe		
	275	280	285
	Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe		
	290	295	300
	Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala		
	305	310	315 320
	Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met		
	325	330	335
	Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly		
	340	345	350
	Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro		
	355	360	365
	His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met		
	370	375	380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Gln Ala Val Gln Thr Ala
405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465 470 475 480

[0043]

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
530 535 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545 550 555 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
565 570 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
580 585 590

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
595 600 605

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
610 615 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625 630 635 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
645 650 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
660 665 670

[0044]

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
675 680 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
690 695 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705 710 715 720

Ala Ala Ser

<210> 25
<211> 702
<212> PRT
<213> 人工序列

<220>
<223> Mtb71f

<400> 25

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val
1 5 10 15

Val Ala Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn
20 25 30

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

Pro Pro Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly
50 55 60

Ala Ala Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn
65 70 75 80

[0045] Asn Tyr Glu Leu Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp Ala
85 90 95

His Gly Ala Met Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu His
100 105 110

Gln Ala Ile Val Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly Gly
115 120 125

Ala Gly Ser Val Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg Asn
130 135 140

Phe Gln Val Ile Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val Gln
145 150 155 160

Ala Ala Gly Asn Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser Ser
165 170 175

Trp Ala Thr Ser Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu Val
180 185 190

Ala Ser Gln Ser Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His Thr
195 200 205

Ile Gly Gln Ala Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His Gln
210 215 220

Gly Glu Ser Ser Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val Ala
225 230 235 240

Ala Ala Ala Lys Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn Leu
245 250 255

Gly Glu Ala Ala Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ala Ser
260 265 270

Thr Tyr Thr Gly Phe Asp Ile Met Asp Phe Gly Leu Leu Pro Pro Glu
275 280 285

[0046]

Val Asn Ser Ser Arg Met Tyr Ser Gly Pro Gly Pro Glu Ser Met Leu
290 295 300

Ala Ala Ala Ala Ala Trp Asp Gly Val Ala Ala Glu Leu Thr Ser Ala
305 310 315 320

Ala Val Ser Tyr Gly Ser Val Val Ser Thr Leu Ile Val Glu Pro Trp
325 330 335

Met Gly Pro Ala Ala Ala Ala Met Ala Ala Ala Ala Thr Pro Tyr Val
340 345 350

Gly Trp Leu Ala Ala Thr Ala Ala Leu Ala Lys Glu Thr Ala Thr Gln
355 360 365

Ala Arg Ala Ala Ala Glu Ala Phe Gly Thr Ala Phe Ala Met Thr Val
370 375 380

Pro Pro Ser Leu Val Ala Ala Asn Arg Ser Arg Leu Met Ser Leu Val
385 390 395 400

Ala Ala Asn Ile Leu Gly Gln Asn Ser Ala Ala Ile Ala Ala Thr Gln
405 410 415

Ala Glu Tyr Ala Glu Met Trp Ala Gln Asp Ala Ala Val Met Tyr Ser
420 425 430

Tyr Glu Gly Ala Ser Ala Ala Ala Ser Ala Leu Pro Pro Phe Thr Pro
435 440 445

Pro Val Gln Gly Thr Gly Pro Ala Gly Pro Ala Ala Ala Ala Ala
450 455 460

Thr Gln Ala Ala Gly Ala Gly Ala Val Ala Asp Ala Gln Ala Thr Leu
465 470 475 480

[0047] Ala Gln Leu Pro Pro Gly Ile Leu Ser Asp Ile Leu Ser Ala Leu Ala
485 490 495

Ala Asn Ala Asp Pro Leu Thr Ser Gly Leu Leu Gly Ile Ala Ser Thr
500 505 510

Leu Asn Pro Gln Val Gly Ser Ala Gln Pro Ile Val Ile Pro Thr Pro
515 520 525

Ile Gly Glu Leu Asp Val Ile Ala Leu Tyr Ile Ala Ser Ile Ala Thr
530 535 540

Gly Ser Ile Ala Leu Ala Ile Thr Asn Thr Ala Arg Pro Trp His Ile
545 550 555 560

Gly Leu Tyr Gly Asn Ala Gly Gly Leu Gly Pro Thr Gln Gly His Pro
565 570 575

Leu Ser Ser Ala Thr Asp Glu Pro Glu Pro His Trp Gly Pro Phe Gly

	580	585	590
	Gly Ala Ala Pro Val Ser Ala Gly Val Gly His Ala Ala Leu Val Gly		
	595	600	605
	Ala Leu Ser Val Pro His Ser Trp Thr Thr Ala Ala Pro Glu Ile Gln		
	610	615	620
	Leu Ala Val Gln Ala Thr Pro Thr Phe Ser Ser Ser Ala Gly Ala Asp		
	625	630	635 640
	Pro Thr Ala Leu Asn Gly Met Pro Ala Gly Leu Leu Ser Gly Met Ala		
	645	650	655
	Leu Ala Ser Leu Ala Ala Arg Gly Thr Thr Gly Gly Gly Gly Thr Arg		
	660	665	670
[0048]	Ser Gly Thr Ser Thr Asp Gly Gln Glu Asp Gly Arg Lys Pro Pro Val		
	675	680	685
	Val Val Ile Arg Glu Gln Pro Pro Pro Gly Asn Pro Pro Arg		
	690	695	700
	<210> 26		
	<211> 920		
	<212> PRT		
	<213> 人工序列		
	<220>		
	<223> M72-Mtb9.9-Mtb9.8		
	<400> 26		
	Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly		
	1	5	10 15
	Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg		
	20	25	30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
35 40 45
Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
50 55 60
Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65 70 75 80
Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
85 90 95
Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
100 105 110
Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
115 120 125
[0049] Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
130 135 140
Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145 150 155 160
Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
165 170 175
Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
180 185 190
Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
195 200 205
Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
210 215 220
Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr

225	230	235	240
Val Pro Pro Pro	Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu		
	245	250	255
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn			
	260	265	270
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe			
	275	280	285
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe			
	290	295	300
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala			
	305	310	315
[0050] Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met			
	325	330	335
Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly			
	340	345	350
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro			
	355	360	365
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met			
	370	375	380
Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met			
	385	390	395
Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala			
	405	410	415
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly			
	420	425	430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
500 505 510

[0051] Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
530 535 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
545 550 555 560

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
565 570 575

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
580 585 590

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
595 600 605

Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
610 615 620

Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
625 630 635 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
645 650 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
660 665 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
675 680 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
690 695 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705 710 715 720

[0052]

Ala Ala Ser Ser Thr Met Thr Ile Asn Tyr Gln Phe Gly Asp Val Asp
725 730 735

Ala His Gly Ala Met Ile Arg Ala Gln Ala Ala Ser Leu Glu Ala Glu
740 745 750

His Gln Ala Ile Val Arg Asp Val Leu Ala Ala Gly Asp Phe Trp Gly
755 760 765

Gly Ala Gly Ser Val Ala Cys Gln Glu Phe Ile Thr Gln Leu Gly Arg
770 775 780

Asn Phe Gln Val Ile Tyr Glu Gln Ala Asn Ala His Gly Gln Lys Val
785 790 795 800

Gln Ala Ala Gly Asn Asn Met Ala Gln Thr Asp Ser Ala Val Gly Ser
805 810 815

Ser Trp Ala Thr Ser Met Ser Leu Leu Asp Ala His Ile Pro Gln Leu
820 825 830

Val Ala Ser Gln Ser Ala Phe Ala Ala Lys Ala Gly Leu Met Arg His
835 840 845

Thr Ile Gly Gln Ala Glu Gln Ala Ala Met Ser Ala Gln Ala Phe His
850 855 860

Gln Gly Glu Ser Ser Ala Ala Phe Gln Ala Ala His Ala Arg Phe Val
865 870 875 880

Ala Ala Ala Ala Lys Val Asn Thr Leu Leu Asp Val Ala Gln Ala Asn
885 890 895

Leu Gly Glu Ala Ala Gly Thr Tyr Val Ala Ala Asp Ala Ala Ala Ala
900 905 910

[0053]

Ser Thr Tyr Thr Gly Phe Pro Trp
915 920

<210> 27
<211> 1010
<212> PRT
<213> 人工序列

<220>
<223> M103

<400> 27

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1 5 10 15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
20 25 30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
35 40 45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
50 55 60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65 70 75 80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
85 90 95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
100 105 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
115 120 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
130 135 140

[0054]

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145 150 155 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
165 170 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
180 185 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
195 200 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
210 215 220

Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225 230 235 240

Val Pro Pro Pro	Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
245	250 255
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn	
260	265 270
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe	
275	280 285
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe	
290	295 300
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala	
305	310 315 320
Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met	
325	330 335
[0055] Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly	
340	345 350
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro	
355	360 365
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met	
370	375 380
Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met	
385	390 395 400
Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala	
405	410 415
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly	
420	425 430
Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala	

	435	440	445
	Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln		
	450	455	460
	Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser		
	465	470	475 480
	Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly		
		485	490 495
	Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val		
		500	505 510
	Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile		
		515	520 525
[0056]	Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu		
	530	535	540
	Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val		
	545	550	555 560
	Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr		
		565	570 575
	Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val		
		580	585 590
	Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln		
		595	600 605
	Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala		
		610	615 620
	Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly		
	625	630	635 640

Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
645 650 655

Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
660 665 670

Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
675 680 685

Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala
690 695 700

Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
705 710 715 720

Ala Ala Ser Ser Gly Phe Ser Arg Pro Gly Leu Pro Val Glu Tyr Leu
725 730 735

[0057]

Gln Val Pro Ser Pro Ser Met Gly Arg Asp Ile Lys Val Gln Phe Gln
740 745 750

Ser Gly Gly Asn Asn Ser Pro Ala Val Tyr Leu Leu Asp Gly Leu Arg
755 760 765

Ala Gln Asp Asp Tyr Asn Gly Trp Asp Ile Asn Thr Pro Ala Phe Glu
770 775 780

Trp Tyr Tyr Gln Ser Gly Leu Ser Ile Val Met Pro Val Gly Gly Gln
785 790 795 800

Ser Ser Phe Tyr Ser Asp Trp Tyr Ser Pro Ala Cys Gly Lys Ala Gly
805 810 815

Cys Gln Thr Tyr Lys Trp Glu Thr Phe Leu Thr Ser Glu Leu Pro Gln
820 825 830

Trp Leu Ser Ala Asn Arg Ala Val Lys Pro Thr Gly Ser Ala Ala Ile
835 840 845

Gly Leu Ser Met Ala Gly Ser Ser Ala Met Ile Leu Ala Ala Tyr His
850 855 860

Pro Gln Gln Phe Ile Tyr Ala Gly Ser Leu Ser Ala Leu Leu Asp Pro
865 870 875 880

Ser Gln Gly Met Gly Pro Ser Leu Ile Gly Leu Ala Met Gly Asp Ala
885 890 895

Gly Gly Tyr Lys Ala Ala Asp Met Trp Gly Pro Ser Ser Asp Pro Ala
900 905 910

Trp Glu Arg Asn Asp Pro Thr Gln Gln Ile Pro Lys Leu Val Ala Asn
915 920 925

[0058]

Asn Thr Arg Leu Trp Val Tyr Cys Gly Asn Gly Thr Pro Asn Glu Leu
930 935 940

Gly Gly Ala Asn Ile Pro Ala Glu Phe Leu Glu Asn Phe Val Arg Ser
945 950 955 960

Ser Asn Leu Lys Phe Gln Asp Ala Tyr Asn Ala Ala Gly Gly His Asn
965 970 975

Ala Val Phe Asn Phe Pro Pro Asn Gly Thr His Ser Trp Glu Tyr Trp
980 985 990

Gly Ala Gln Leu Asn Ala Met Lys Gly Asp Leu Gln Ser Ser Leu Gly
995 1000 1005

Ala Gly
1010

<210> 28
<211> 1148
<212> PRT
<213> 人工序列

<220>
<223> M114

<400> 28

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1 5 10 15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
20 25 30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
35 40 45

[0059] Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
50 55 60

Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp
65 70 75 80

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
85 90 95

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
100 105 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
115 120 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
130 135 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145 150 155 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
165 170 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
180 185 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr
195 200 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
210 215 220

Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225 230 235 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
245 250 255

[0060]

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
260 265 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
275 280 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
290 295 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305 310 315 320

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
325 330 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
340 345 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
355 360 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
370 375 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Gln Ala Val Gln Thr Ala
405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
435 440 445

[0061]

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
530 535 540

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val

[0062]

545	550	555	560
Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr			
565	570	575	
Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val			
580	585	590	
Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln			
595	600	605	
Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala			
610	615	620	
Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly			
625	630	635	640
Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly			
645	650	655	
Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu			
660	665	670	
Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr			
675	680	685	
Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ala			
690	695	700	
Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr			
705	710	715	720
Ala Ala Ser Ser Thr Met Asp Phe Gly Leu Leu Pro Pro Glu Val Asn			
725	730	735	
Ser Ser Arg Met Tyr Ser Gly Pro Gly Pro Glu Ser Met Leu Ala Ala			
740	745	750	

Ala Ala Ala Trp Asp Gly Val Ala Ala Glu Leu Thr Ser Ala Ala Val
755 760 765

Ser Tyr Gly Ser Val Val Ser Thr Leu Ile Val Glu Pro Trp Met Gly
770 775 780

Pro Ala Ala Ala Ala Met Ala Ala Ala Ala Thr Pro Tyr Val Gly Trp
785 790 795 800

Leu Ala Ala Thr Ala Ala Leu Ala Lys Glu Thr Ala Thr Gln Ala Arg
805 810 815

Ala Ala Ala Glu Ala Phe Gly Thr Ala Phe Ala Met Thr Val Pro Pro
820 825 830

[0063] Ser Leu Val Ala Ala Asn Arg Ser Arg Leu Met Ser Leu Val Ala Ala
835 840 845

Asn Ile Leu Gly Gln Asn Ser Ala Ala Ile Ala Ala Thr Gln Ala Glu
850 855 860

Tyr Ala Glu Met Trp Ala Gln Asp Ala Ala Val Met Tyr Ser Tyr Glu
865 870 875 880

Gly Ala Ser Ala Ala Ala Ser Ala Leu Pro Pro Phe Thr Pro Pro Val
885 890 895

Gln Gly Thr Gly Pro Ala Gly Pro Ala Ala Ala Ala Ala Thr Gln
900 905 910

Ala Ala Gly Ala Gly Ala Val Ala Asp Ala Gln Ala Thr Leu Ala Gln
915 920 925

Leu Pro Pro Gly Ile Leu Ser Asp Ile Leu Ser Ala Leu Ala Ala Asn
930 935 940

Ala Asp Pro Leu Thr Ser Gly Leu Leu Gly Ile Ala Ser Thr Leu Asn
945 950 955 960

Pro Gln Val Gly Ser Ala Gln Pro Ile Val Ile Pro Thr Pro Ile Gly
965 970 975

Glu Leu Asp Val Ile Ala Leu Tyr Ile Ala Ser Ile Ala Thr Gly Ser
980 985 990

Ile Ala Leu Ala Ile Thr Asn Thr Ala Arg Pro Trp His Ile Gly Leu
995 1000 1005

Tyr Gly Asn Ala Gly Gly Leu Gly Pro Thr Gln Gly His Pro Leu
1010 1015 1020

Ser Ser Ala Thr Asp Glu Pro Glu Pro His Trp Gly Pro Phe Gly
1025 1030 1035

[0064]

Gly Ala Ala Pro Val Ser Ala Gly Val Gly His Ala Ala Leu Val
1040 1045 1050

Gly Ala Leu Ser Val Pro His Ser Trp Thr Thr Ala Ala Pro Glu
1055 1060 1065

Ile Gln Leu Ala Val Gln Ala Thr Pro Thr Phe Ser Ser Ser Ala
1070 1075 1080

Gly Ala Asp Pro Thr Ala Leu Asn Gly Met Pro Ala Gly Leu Leu
1085 1090 1095

Ser Gly Met Ala Leu Ala Ser Leu Ala Ala Arg Gly Thr Thr Gly
1100 1105 1110

Gly Gly Gly Thr Arg Ser Gly Thr Ser Thr Asp Gly Gln Glu Asp
1115 1120 1125

Gly Arg Lys Pro Pro Val Val Val Ile Arg Glu Gln Pro Pro Pro
1130 1135 1140

Gly Asn Pro Pro Arg
1145

<210> 29
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 29

Phe Ile Ile Asp Pro Thr Ile Ser Ala
1 5

[0065]

<210> 30
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 30

Ile Leu Tyr Ser Ser Leu Glu Tyr Phe
1 5

<210> 31
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 31

Leu Glu Tyr Phe Glu Lys Ala Leu Glu
1 5

<210> 32
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 32

Tyr Ala Gly Lys Asn Arg Asn His Val

1 5

<210> 33
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 33

Leu Ile His Asp Gln Ala Asn Ala Val
1 5

<210> 34
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 34

Phe Val Arg Pro Val Ala Val Asp Leu
1 5

[0066]

<210> 35
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 35

Leu Thr Tyr Ile Pro Val Val Gly His
1 5

<210> 36
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 36

Tyr Ile Pro Val Val Gly His Ala Leu
1 5

<210> 37
<211> 9

<212> PRT
<213> 结核分枝杆菌

<400> 37

Tyr Leu Val Val Lys Thr Leu Ile Asn
1 5

<210> 38
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 38

Leu Val Val Lys Thr Leu Ile Asn Ala
1 5

[0067]

<210> 39
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 39

Val Lys Thr Leu Ile Asn Ala Thr Gln
1 5

<210> 40
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 40

Leu Lys Leu Leu Ala Lys Leu Ala Glu
1 5

<210> 41
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 41

Leu Val Ala Ala Ala Ile Ala Asp Ile
1 5

<210> 42

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 42

Ile Ile Ser Asp Val Ala Asp Ile Ile
1 5

<210> 43

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 43

Trp Glu Phe Ile Thr Asn Ala Leu Asn
1 5

[0068]

<210> 44

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 44

Leu Phe Gly Ala Ala Gly Leu Ser Ala
1 5

<210> 45

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 45

Leu Ala His Ala Asp Ser Leu Ala Ser
1 5

<210> 46

<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 46

Leu Ala Ser Ser Ala Ser Leu Pro Ala
1 5

<210> 47
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 47

Phe Gly Gly Leu Pro Ser Leu Ala Gln
1 5

[0069]

<210> 48
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 48

Phe Ile Ile Asp Pro Thr Ile Ser Ala
1 5

<210> 49
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 49

Ile Ile Asp Pro Thr Ile Ser Ala Ile
1 5

<210> 50
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 50

Pro Thr Ile Ser Ala Ile Asp Gly Leu
1 5

<210> 51
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 51

Thr Ile Ser Ala Ile Asp Gly Leu Tyr
1 5

<210> 52
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 52

[0070] Ser Ala Ile Asp Gly Leu Tyr Asp Leu
1 5

<210> 53
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 53

Ala Ile Asp Gly Leu Tyr Asp Leu Leu
1 5

<210> 54
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 54

Leu Tyr Asp Leu Leu Gly Ile Gly Ile
1 5

<210> 55
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 55

Ile Pro Asn Gln Gly Gly Ile Leu Tyr
1 5

<210> 56
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 56

Gly Ile Leu Tyr Ser Ser Leu Glu Tyr
1 5

[0071]

<210> 57
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 57

Tyr Ser Ser Leu Glu Tyr Phe Glu Lys
1 5

<210> 58
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 58

Ser Leu Glu Tyr Phe Glu Lys Ala Leu
1 5

<210> 59
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 59

Tyr Phe Glu Lys Ala Leu Glu Glu Leu
1 5

<210> 60

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 60

Phe Glu Lys Ala Leu Glu Glu Leu Ala
1 5

<210> 61

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 61

[0072]

Asn His Val Asn Phe Phe Gln Glu Leu
1 5

<210> 62

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 62

Glu Leu Ala Asp Leu Asp Arg Gln Leu
1 5

<210> 63

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 63

Leu Ala Asp Leu Asp Arg Gln Leu Ile
1 5

<210> 64
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 64

Asp Leu Asp Arg Gln Leu Ile Ser Leu
1 5

<210> 65
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 65

Leu Asp Arg Gln Leu Ile Ser Leu Ile
1 5

[0073] <210> 66
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 66

Ala Val Gln Thr Thr Arg Asp Ile Leu
1 5

<210> 67
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 67

Glu Gly Ala Lys Lys Gly Leu Glu Phe
1 5

<210> 68
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 68

Lys Gly Leu Glu Phe Val Arg Pro Val
1 5

<210> 69

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 69

Gly Leu Glu Phe Val Arg Pro Val Ala
1 5

<210> 70

<211> 9

<212> PRT

<213> 结核分枝杆菌

[0074]

<400> 70

Leu Glu Phe Val Arg Pro Val Ala Val
1 5

<210> 71

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 71

Phe Val Arg Pro Val Ala Val Asp Leu
1 5

<210> 72

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 72

Arg Pro Val Ala Val Asp Leu Thr Tyr
1 5

<210> 73
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 73

Ala Val Asp Leu Thr Tyr Ile Pro Val
1 5

<210> 74
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 74

Thr Tyr Ile Pro Val Val Gly His Ala
1 5

[0075]

<210> 75
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 75

Tyr Ile Pro Val Val Gly His Ala Leu
1 5

<210> 76
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 76

Leu Ser Ala Ala Phe Gln Ala Pro Phe
1 5

<210> 77
<211> 9
<212> PRT

<213> 结核分枝杆菌

<400> 77

Ser Ala Ala Phe Gln Ala Pro Phe Cys
1 5

<210> 78

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 78

Ala Ala Phe Gln Ala Pro Phe Cys Ala
1 5

<210> 79

<211> 9

<212> PRT

<213> 结核分枝杆菌

[0076] <400> 79

Phe Gln Ala Pro Phe Cys Ala Gly Ala
1 5

<210> 80

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 80

Ala Pro Phe Cys Ala Gly Ala Met Ala
1 5

<210> 81

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 81

Pro Phe Cys Ala Gly Ala Met Ala Val

1 5

<210> 82
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 82

Ala Met Ala Val Val Gly Gly Ala Leu
1 5

<210> 83
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 83

Ala Val Val Gly Gly Ala Leu Ala Tyr
1 5

[0077]

<210> 84
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 84

Gly Ala Leu Ala Tyr Leu Val Val Lys
1 5

<210> 85
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 85

Leu Ala Tyr Leu Val Val Lys Thr Leu
1 5

<210> 86
<211> 9

<212> PRT
<213> 结核分枝杆菌

<400> 86

Ala Tyr Leu Val Val Lys Thr Leu Ile
1 5

<210> 87
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 87

Lys Thr Leu Ile Asn Ala Thr Gln Leu
1 5

[0078] <210> 88
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 88

Thr Leu Ile Asn Ala Thr Gln Leu Leu
1 5

<210> 89
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 89

Ile Asn Ala Thr Gln Leu Leu Lys Leu
1 5

<210> 90
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 90

Asn Ala Thr Gln Leu Leu Lys Leu Leu
1 5

<210> 91
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 91

Gln Leu Leu Lys Leu Leu Ala Lys Leu
1 5

<210> 92
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 92

Lys Leu Leu Ala Lys Leu Ala Glu Leu
1 5

[0079]

<210> 93
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 93

Leu Leu Ala Lys Leu Ala Glu Leu Val
1 5

<210> 94
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 94

Lys Leu Ala Glu Leu Val Ala Ala Ala
1 5

<210> 95

<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 95

Leu Ala Glu Leu Val Ala Ala Ala Ile
1 5

<210> 96
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 96

Leu Val Ala Ala Ala Ile Ala Asp Ile
1 5

[0080] <210> 97
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 97

Ala Ile Ala Asp Ile Ile Ser Asp Val
1 5

<210> 98
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 98

Ile Ala Asp Ile Ile Ser Asp Val Ala
1 5

<210> 99
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 99

Ile Ser Asp Val Ala Asp Ile Ile Lys
1 5

<210> 100
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 100

Thr Leu Gly Glu Val Trp Glu Phe Ile
1 5

<210> 101
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 101

[0081] Phe Ile Thr Asn Ala Leu Asn Gly Leu
1 5

<210> 102
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 102

Asn Ala Leu Asn Gly Leu Lys Glu Leu
1 5

<210> 103
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 103

Lys Leu Thr Gly Trp Val Thr Gly Leu
1 5

<210> 104
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 104

Leu Thr Gly Trp Val Thr Gly Leu Phe
1 5

<210> 105
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 105

Gly Trp Ser Asn Leu Glu Ser Phe Phe
1 5

[0082]

<210> 106
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 106

Asn Leu Glu Ser Phe Phe Ala Gly Val
1 5

<210> 107
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 107

Ser Phe Phe Ala Gly Val Pro Gly Leu
1 5

<210> 108
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 108

Gly Leu Thr Gly Ala Thr Ser Gly Leu
1 5

<210> 109

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 109

Leu Ser Gln Val Thr Gly Leu Phe Gly
1 5

<210> 110

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 110

[0083]

Ser Gln Val Thr Gly Leu Phe Gly Ala
1 5

<210> 111

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 111

Leu Ser Ala Ser Ser Gly Leu Ala His
1 5

<210> 112

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 112

Ala Ser Ser Gly Leu Ala His Ala Asp
1 5

<210> 113
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 113

Ser Gly Leu Ala His Ala Asp Ser Leu
1 5

<210> 114
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 114

Gly Leu Ala His Ala Asp Ser Leu Ala
1 5

[0084]

<210> 115
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 115

Ser Leu Ala Ser Ser Ala Ser Leu Pro
1 5

<210> 116
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 116

Ala Ser Ser Ala Ser Leu Pro Ala Leu
1 5

<210> 117
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 117

Ser Gly Phe Gly Gly Leu Pro Ser Leu
1 5

<210> 118

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 118

Leu Pro Ser Leu Ala Gln Val His Ala
1 5

<210> 119

<211> 9

<212> PRT

<213> 结核分枝杆菌

[0085]

<400> 119

His Ala Ala Ser Thr Arg Gln Ala Leu
1 5

<210> 120

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 120

Ser Thr Arg Gln Ala Leu Arg Pro Arg
1 5

<210> 121

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 121

Arg Pro Arg Ala Asp Gly Pro Val Gly
1 5

<210> 122
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 122

Glu Gln Val Gly Gly Gln Ser Gln Leu
1 5

<210> 123
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 123

Gly Ala Ser Lys Gly Thr Thr Thr Lys
1 5

[0086] <210> 124
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 124

Ala Ser Lys Gly Thr Thr Thr Lys Lys
1 5

<210> 125
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 125

Lys Gly Thr Thr Thr Lys Lys Tyr Ser
1 5

<210> 126
<211> 9
<212> PRT

<213> 结核分枝杆菌

<400> 126

Thr Glu Asp Ala Glu Arg Ala Pro Val
1 5

<210> 127

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 127

Met Ser Arg Ala Phe Ile Ile Asp Pro Thr Ile Ser Ala Ile Asp Gly
1 5 10 15

Leu Tyr Asp Leu
20

[0087]

<210> 128

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 128

Ile Asp Gly Leu Tyr Asp Leu Leu Gly Ile Gly Ile Pro Asn Gln Gly
1 5 10 15

Gly Ile Leu Tyr
20

<210> 129

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 129

Asn Gln Gly Gly Ile Leu Tyr Ser Ser Leu Glu Tyr Phe Glu Lys Ala
1 5 10 15

Leu Glu Glu Leu
20

<210> 130
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 130

Glu Lys Ala Leu Glu Glu Leu Ala Ala Ala Phe Pro Gly Asp Gly Trp
1 5 10 15

Leu Gly Ser Ala
20

<210> 131
<211> 20
<212> PRT
<213> 结核分枝杆菌

[0088]

<400> 131

Asp Gly Trp Leu Gly Ser Ala Ala Asp Lys Tyr Ala Gly Lys Asn Arg
1 5 10 15

Asn His Val Asn
20

<210> 132
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 132

Lys Asn Arg Asn His Val Asn Phe Phe Gln Glu Leu Ala Asp Leu Asp
1 5 10 15

Arg Gln Leu Ile
20

<210> 133
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 133

Asp Leu Asp Arg Gln Leu Ile Ser Leu Ile His Asp Gln Ala Asn Ala
1 5 10 15

Val Gln Thr Thr
20

<210> 134
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 134

[0089] Ala Asn Ala Val Gln Thr Thr Arg Asp Ile Leu Glu Gly Ala Lys Lys
1 5 10 15

Gly Leu Glu Phe
20

<210> 135
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 135

Ala Lys Lys Gly Leu Glu Phe Val Arg Pro Val Ala Val Asp Leu Thr
1 5 10 15

Tyr Ile Pro Val
20

<210> 136
<211> 27
<212> PRT
<213> 结核分枝杆菌

<400> 136

Phe Val Arg Pro Val Ala Val Asp Leu Thr Tyr Ile Pro Val Val Gly
1 5 10 15

His Ala Leu Ser Ala Ala Phe Gln Ala Pro Phe
20 25

<210> 137

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 137

Ala Ala Phe Gln Ala Pro Phe Cys Ala Gly Ala Met Ala Val Val Gly
1 5 10 15

Gly Ala Leu Ala
20

[0090]

<210> 138

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 138

Val Val Gly Gly Ala Leu Ala Tyr Leu Val Val Lys Thr Leu Ile Asn
1 5 10 15

Ala Thr Gln Leu
20

<210> 139

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 139

Leu Ile Asn Ala Thr Gln Leu Leu Lys Leu Leu Ala Lys Leu Ala Glu

1 5 10 15

Leu Val Ala Ala
20

<210> 140
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 140

Leu Ala Glu Leu Val Ala Ala Ala Ile Ala Asp Ile Ile Ser Asp Val
1 5 10 15

Ala Asp Ile Ile
20

[0091] <210> 141
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 141

Ser Asp Val Ala Asp Ile Ile Lys Gly Thr Leu Gly Glu Val Trp Glu
1 5 10 15

Phe Ile Thr Asn
20

<210> 142
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 142

Val Trp Glu Phe Ile Thr Asn Ala Leu Asn Gly Leu Lys Glu Leu Trp
1 5 10 15

Asp Lys Leu Thr

20

<210> 143
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 143

Glu Leu Trp Asp Lys Leu Thr Gly Trp Val Thr Gly Leu Phe Ser Arg
1 5 10 15

Gly Trp Ser Asn
20

<210> 144
<211> 20
<212> PRT
<213> 结核分枝杆菌

[0092] <400> 144

Phe Ser Arg Gly Trp Ser Asn Leu Glu Ser Phe Phe Ala Gly Val Pro
1 5 10 15

Gly Leu Thr Gly
20

<210> 145
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 145

Gly Val Pro Gly Leu Thr Gly Ala Thr Ser Gly Leu Ser Gln Val Thr
1 5 10 15

Gly Leu Phe Gly
20

<210> 146

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 146

Gln Val Thr Gly Leu Phe Gly Ala Ala Gly Leu Ser Ala Ser Ser Gly
1 5 10 15

Leu Ala His Ala
20

<210> 147

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 147

Ser Ser Gly Leu Ala His Ala Asp Ser Leu Ala Ser Ser Ala Ser Leu
1 5 10 15

[0093]

Pro Ala Leu Ala
20

<210> 148

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 148

Ala Ser Leu Pro Ala Leu Ala Gly Ile Gly Gly Gly Ser Gly Phe Gly
1 5 10 15

Gly Leu Pro Ser
20

<210> 149

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 149

Gly Phe Gly Gly Leu Pro Ser Leu Ala Gln Val His Ala Ala Ser Thr
1 5 10 15

Arg Gln Ala Leu
20

<210> 150

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 150

Ala Ser Thr Arg Gln Ala Leu Arg Pro Arg Ala Asp Gly Pro Val Gly
1 5 10 15

Ala Ala Ala Glu
20

[0094]

<210> 151

<211> 20

<212> PRT

<213> 结核分枝杆菌

<400> 151

Pro Val Gly Ala Ala Ala Glu Gln Val Gly Gly Gln Ser Gln Leu Val
1 5 10 15

Ser Ala Gln Gly
20

<210> 152

<211> 19

<212> PRT

<213> 结核分枝杆菌

<400> 152

Gln Leu Val Ser Ala Gln Gly Ser Gln Gly Met Gly Gly Pro Val Gly
1 5 10 15

Met Gly Gly

<210> 153
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 153

Pro	Val	Gly	Met	Gly	Gly	Met	His	Pro	Ser	Ser	Gly	Ala	Ser	Lys	Gly
1				5					10					15	

Thr Thr Thr Lys
20

[0095] <210> 154
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 154

Ser	Lys	Gly	Thr	Thr	Thr	Lys	Lys	Tyr	Ser	Glu	Gly	Ala	Ala	Ala	Gly
1				5					10					15	

Thr Glu Asp Ala
20

<210> 155
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 155

Ala	Ala	Gly	Thr	Glu	Asp	Ala	Glu	Arg	Ala	Pro	Val	Glu	Ala	Asp	Ala
1				5					10					15	

Gly Gly Gly Gln
20

<210> 156
<211> 20
<212> PRT
<213> 结核分枝杆菌

<400> 156

Arg Ala Pro Val Glu Ala Asp Ala Gly Gly Gly Gln Lys Val Leu Val
1 5 10 15

Arg Asn Val Val
20

<210> 157
<211> 1053
<212> PRT
<213> 结核分枝杆菌

<400> 157

[0096] Met Asn Phe Ser Val Leu Pro Pro Glu Ile Asn Ser Ala Leu Ile Phe
1 5 10 15

Ala Gly Ala Gly Pro Glu Pro Met Ala Ala Ala Ala Thr Ala Trp Asp
20 25 30

Gly Leu Ala Met Glu Leu Ala Ser Ala Ala Ala Ser Phe Gly Ser Val
35 40 45

Thr Ser Gly Leu Val Gly Gly Ala Trp Gln Gly Ala Ser Ser Ser Ala
50 55 60

Met Ala Ala Ala Ala Ala Pro Tyr Ala Ala Trp Leu Ala Ala Ala Ala
65 70 75 80

Val Gln Ala Glu Gln Thr Ala Ala Gln Ala Ala Ala Met Ile Ala Glu
85 90 95

Phe Glu Ala Val Lys Thr Ala Val Val Gln Pro Met Leu Val Ala Ala

	100	105	110
	Asn Arg Ala Asp Leu Val Ser Leu Val Met Ser Asn Leu Phe Gly Gln		
	115	120	125
	Asn Ala Pro Ala Ile Ala Ala Ile Glu Ala Thr Tyr Glu Gln Met Trp		
	130	135	140
	Ala Ala Asp Val Ser Ala Met Ser Ala Tyr His Ala Gly Ala Ser Ala		
	145	150	155 160
	Ile Ala Ser Ala Leu Ser Pro Phe Ser Lys Pro Leu Gln Asn Leu Ala		
	165	170	175
	Gly Leu Pro Ala Trp Leu Ala Ser Gly Ala Pro Ala Ala Ala Met Thr		
	180	185	190
[0097]	Ala Ala Ala Gly Ile Pro Ala Leu Ala Gly Gly Pro Thr Ala Ile Asn		
	195	200	205
	Leu Gly Ile Ala Asn Val Gly Gly Gly Asn Val Gly Asn Ala Asn Asn		
	210	215	220
	Gly Leu Ala Asn Ile Gly Asn Ala Asn Leu Gly Asn Tyr Asn Phe Gly		
	225	230	235 240
	Ser Gly Asn Phe Gly Asn Ser Asn Ile Gly Ser Ala Ser Leu Gly Asn		
	245	250	255
	Asn Asn Ile Gly Phe Gly Asn Leu Gly Ser Asn Asn Val Gly Val Gly		
	260	265	270
	Asn Leu Gly Asn Leu Asn Thr Gly Phe Ala Asn Thr Gly Leu Gly Asn		
	275	280	285
	Phe Gly Phe Gly Asn Thr Gly Asn Asn Asn Ile Gly Ile Gly Leu Thr		
	290	295	300

Gly Asn Asn Gln Ile Gly Ile Gly Gly Leu Asn Ser Gly Thr Gly Asn
305 310 315 320

Phe Gly Leu Phe Asn Ser Gly Ser Gly Asn Val Gly Phe Phe Asn Ser
325 330 335

Gly Asn Gly Asn Phe Gly Ile Gly Asn Ser Gly Asn Phe Asn Thr Gly
340 345 350

Gly Trp Asn Ser Gly His Gly Asn Thr Gly Phe Phe Asn Ala Gly Ser
355 360 365

Phe Asn Thr Gly Met Leu Asp Val Gly Asn Ala Asn Thr Gly Ser Leu
370 375 380

[0098] Asn Thr Gly Ser Tyr Asn Met Gly Asp Phe Asn Pro Gly Ser Ser Asn
385 390 395 400

Thr Gly Thr Phe Asn Thr Gly Asn Ala Asn Thr Gly Phe Leu Asn Ala
405 410 415

Gly Asn Ile Asn Thr Gly Val Phe Asn Ile Gly His Met Asn Asn Gly
420 425 430

Leu Phe Asn Thr Gly Asp Met Asn Asn Gly Val Phe Tyr Arg Gly Val
435 440 445

Gly Gln Gly Ser Leu Gln Phe Ser Ile Thr Thr Pro Asp Leu Thr Leu
450 455 460

Pro Pro Leu Gln Ile Pro Gly Ile Ser Val Pro Ala Phe Ser Leu Pro
465 470 475 480

Ala Ile Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala
485 490 495

Asn Ile Thr Val Gly Ala Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser
500 505 510

Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala Asn Ile Thr Val Gly Ala
515 520 525

Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala
530 535 540

Thr Thr Pro Ala Asn Ile Thr Val Gly Ala Phe Ser Leu Pro Gly Leu
545 550 555 560

Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr Pro Ala Asn Ile
565 570 575

Thr Val Gly Ala Phe Ser Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn
580 585 590

[0099]

Ile Pro Ala Ala Thr Thr Pro Ala Asn Ile Thr Val Gly Ala Phe Ser
595 600 605

Leu Pro Gly Leu Thr Leu Pro Ser Leu Asn Ile Pro Ala Ala Thr Thr
610 615 620

Pro Ala Asn Ile Thr Val Ser Gly Phe Gln Leu Pro Pro Leu Ser Ile
625 630 635 640

Pro Ser Val Ala Ile Pro Pro Val Thr Val Pro Pro Ile Thr Val Gly
645 650 655

Ala Phe Asn Leu Pro Pro Leu Gln Ile Pro Glu Val Thr Ile Pro Gln
660 665 670

Leu Thr Ile Pro Ala Gly Ile Thr Ile Gly Gly Phe Ser Leu Pro Ala
675 680 685

	Ile His Thr Gln Pro Ile Thr Val Gly Gln Ile Gly Val Gly Gln Phe	
	690	700
	Gly Leu Pro Ser Ile Gly Trp Asp Val Phe Leu Ser Thr Pro Arg Ile	
	705	720
	Thr Val Pro Ala Phe Gly Ile Pro Phe Thr Leu Gln Phe Gln Thr Asn	
	725	735
	Val Pro Ala Leu Gln Pro Pro Gly Gly Gly Leu Ser Thr Phe Thr Asn	
	740	750
	Gly Ala Leu Ile Phe Gly Glu Phe Asp Leu Pro Gln Leu Val Val His	
	755	765
	Pro Tyr Thr Leu Thr Gly Pro Ile Val Ile Gly Ser Phe Phe Leu Pro	
	770	780
[0100]	Ala Phe Asn Ile Pro Gly Ile Asp Val Pro Ala Ile Asn Val Asp Gly	
	785	800
	Phe Thr Leu Pro Gln Ile Thr Thr Pro Ala Ile Thr Thr Pro Glu Phe	
	805	815
	Ala Ile Pro Pro Ile Gly Val Gly Gly Phe Thr Leu Pro Gln Ile Thr	
	820	830
	Thr Gln Glu Ile Ile Thr Pro Glu Leu Thr Ile Asn Ser Ile Gly Val	
	835	845
	Gly Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr Thr Pro	
	850	860
	Pro Leu Thr Ile Asp Pro Ile Asn Leu Thr Gly Phe Thr Leu Pro Gln	
	865	880
	Ile Thr Thr Pro Pro Ile Thr Thr Pro Pro Leu Thr Ile Asp Pro Ile	

	885	890	895
	Asn Leu Thr Gly Phe Thr Leu Pro Gln Ile Thr Thr Pro Pro Ile Thr		
	900	905	910
	Thr Pro Pro Leu Thr Ile Glu Pro Ile Gly Val Gly Gly Phe Thr Thr		
	915	920	925
	Pro Pro Leu Thr Val Pro Gly Ile His Leu Pro Ser Thr Thr Ile Gly		
	930	935	940
	Ala Phe Ala Ile Pro Gly Gly Pro Gly Tyr Phe Asn Ser Ser Thr Ala		
	945	950	955 960
	Pro Ser Ser Gly Phe Phe Asn Ser Gly Ala Gly Gly Asn Ser Gly Phe		
	965	970	975
[0101]	Gly Asn Asn Gly Ser Gly Leu Ser Gly Trp Phe Asn Thr Asn Pro Ala		
	980	985	990
	Gly Leu Leu Gly Gly Ser Gly Tyr Gln Asn Phe Gly Gly Leu Ser Ser		
	995	1000	1005
	Gly Phe Ser Asn Leu Gly Ser Gly Val Ser Gly Phe Ala Asn Arg		
	1010	1015	1020
	Gly Ile Leu Pro Phe Ser Val Ala Ser Val Val Ser Gly Phe Ala		
	1025	1030	1035
	Asn Ile Gly Thr Asn Leu Ala Gly Phe Phe Gln Gly Thr Thr Ser		
	1040	1045	1050
	<210> 158		
	<211> 450		
	<212> PRT		
	<213> 结核分枝杆菌		
	<400> 158		

Met Ser Glu Leu Ser Val Ala Thr Gly Ala Val Ser Thr Ala Ser Ser
1 5 10 15

Ser Ile Pro Met Pro Ala Gly Val Asn Pro Ala Asp Leu Ala Ala Glu
20 25 30

Leu Ala Ala Val Val Thr Glu Ser Val Asp Glu Asp Tyr Leu Leu Tyr
35 40 45

Glu Cys Asp Gly Gln Trp Val Leu Ala Ala Gly Val Gln Ala Met Val
50 55 60

Glu Leu Asp Ser Asp Glu Leu Arg Val Ile Arg Asp Gly Val Thr Arg
65 70 75 80

Arg Gln Gln Trp Ser Gly Arg Pro Gly Ala Ala Leu Gly Glu Ala Val
85 90 95

[0102]

Asp Arg Leu Leu Leu Glu Thr Asp Gln Ala Phe Gly Trp Val Ala Phe
100 105 110

Glu Phe Gly Val His Arg Tyr Gly Leu Gln Gln Arg Leu Ala Pro His
115 120 125

Thr Pro Leu Ala Arg Val Phe Ser Pro Arg Thr Arg Ile Met Val Ser
130 135 140

Glu Lys Glu Ile Arg Leu Phe Asp Ala Gly Ile Arg His Arg Glu Ala
145 150 155 160

Ile Asp Arg Leu Leu Ala Thr Gly Val Arg Glu Val Pro Gln Ser Arg
165 170 175

Ser Val Asp Val Ser Asp Asp Pro Ser Gly Phe Arg Arg Arg Val Ala
180 185 190

Asp Ile Thr Asn Asn Ala Arg Gly Glu Val Ala Ser Leu Gly Phe Leu
100 105 110

Ile Ser Leu Trp Ala Gly Ser Ser Ala Ile Ser Ala Phe Val Asp Ala
115 120 125

Val Val Glu Ala His Asp Gln Thr Pro Leu Arg His Pro Val Arg Gln
130 135 140

Arg Phe Phe Ala Leu Phe Leu Tyr Val Val Met Leu Val Phe Leu Val
145 150 155 160

Ala Thr Ala Pro Val Met Val Val Gly Pro Arg Lys Val Ser Glu His
165 170 175

Ile Pro Glu Ser Leu Ala Asn Leu Leu Arg Tyr Gly Tyr Tyr Pro Ala
180 185 190

[0105]

Leu Ile Leu Gly Leu Thr Val Gly Val Ile Leu Leu Tyr Arg Val Ala
195 200 205

Leu Pro Val Pro Leu Pro Thr His Arg Leu Val Leu Gly Ala Val Leu
210 215 220

Ala Ile Ala Val Phe Leu Ile Ala Thr Leu Gly Leu Arg Val Tyr Leu
225 230 235 240

Ala Trp Ile Thr Arg Thr Gly Tyr Thr Tyr Gly Ala Leu Ala Thr Pro
245 250 255

Ile Ala Phe Leu Leu Phe Ala Phe Phe Gly Gly Phe Ala Ile Met Leu
260 265 270

Gly Ala Glu Leu Asn Ala Ala Val Gln Glu Glu Trp Pro Ala Pro Ala
275 280 285

Thr His Ala His Arg Leu Gly Asn Trp Leu Lys Ala Arg Ile Gly Val
290 295 300

[0106] Gly Thr Thr Thr Tyr Ser Ser Thr Ala Gln His Ser Ala Val Ala Ala
305 310 315 320

Glu Pro Pro Ser

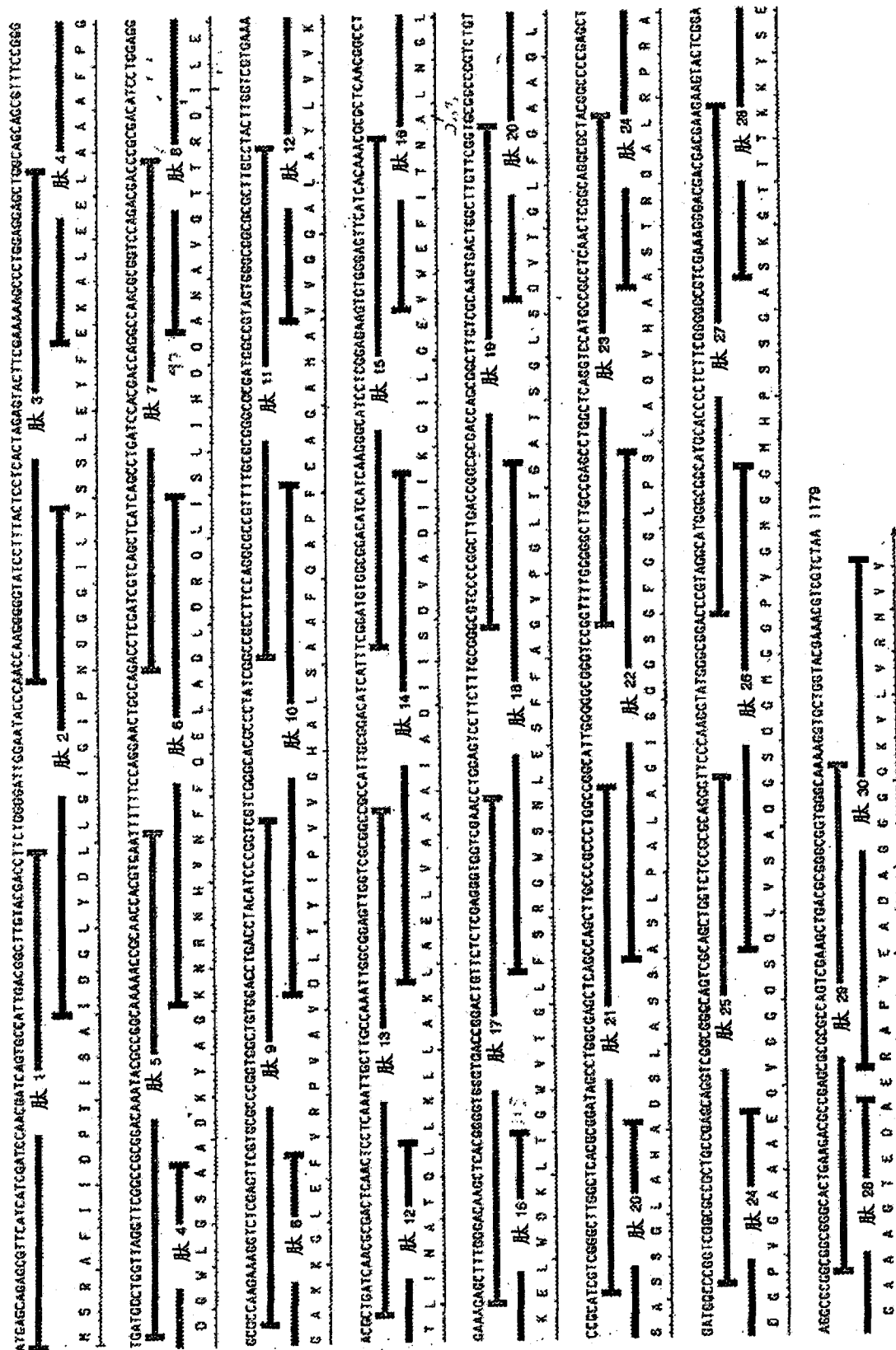


图 1

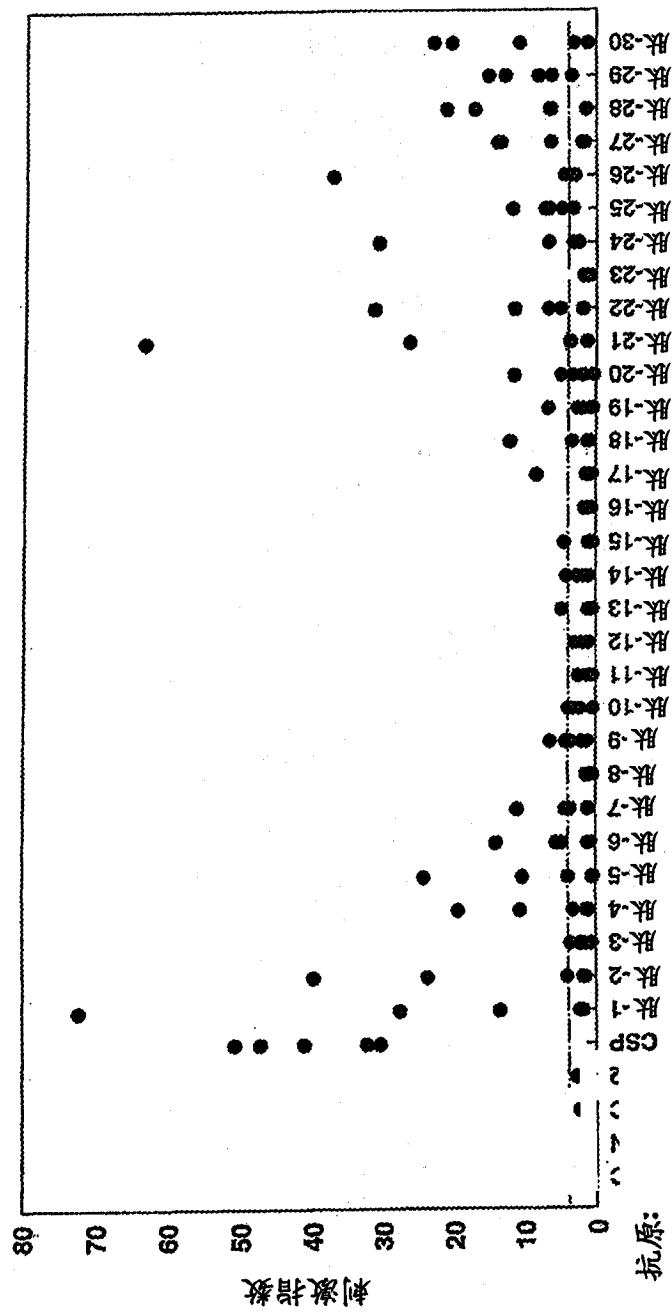


图2

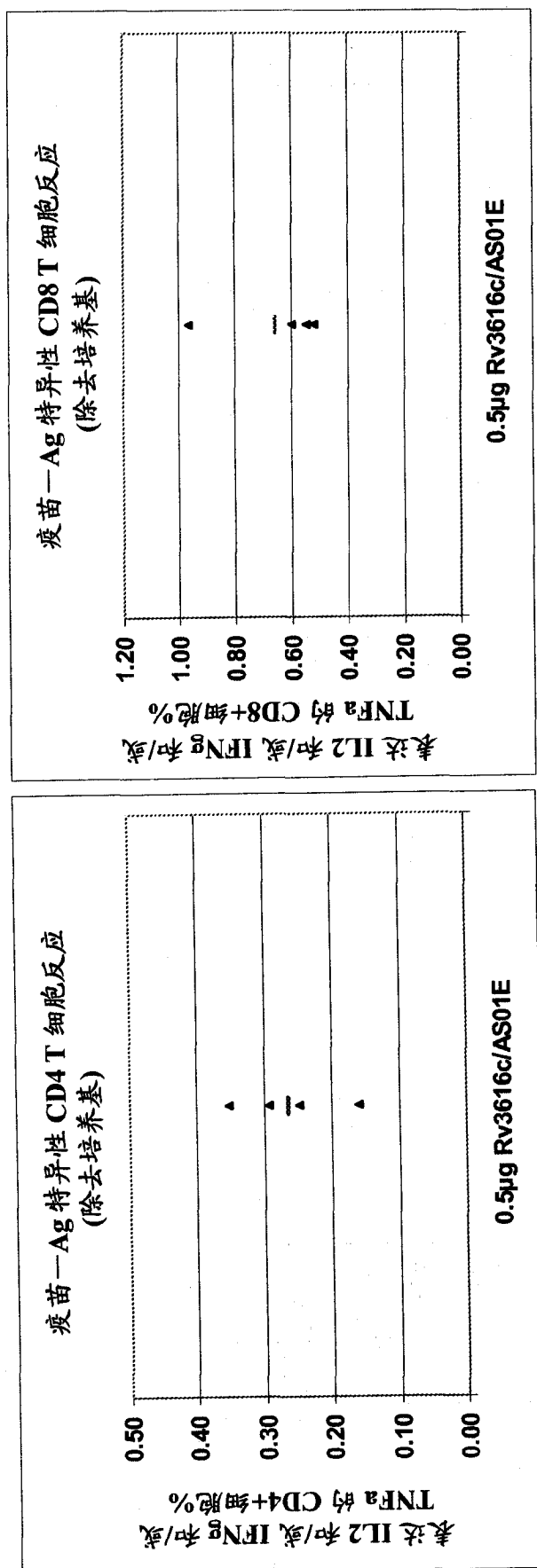


图3

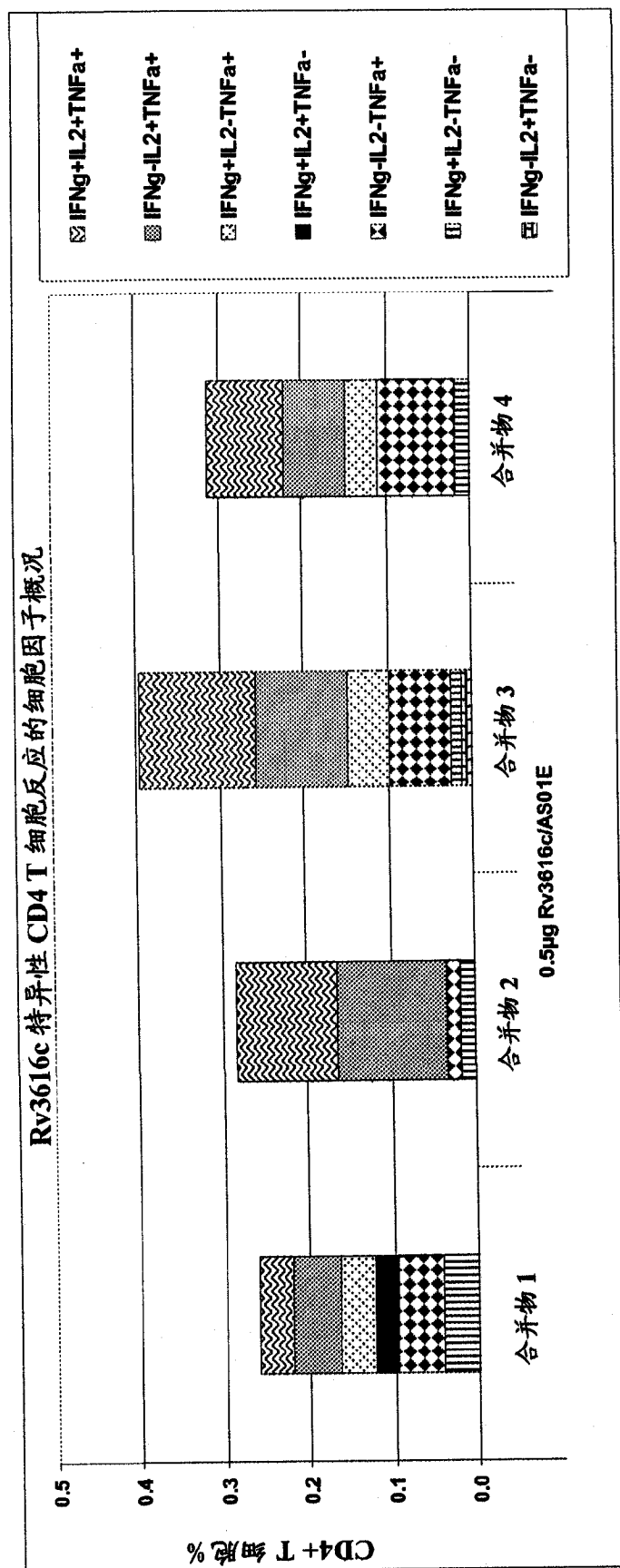


图4

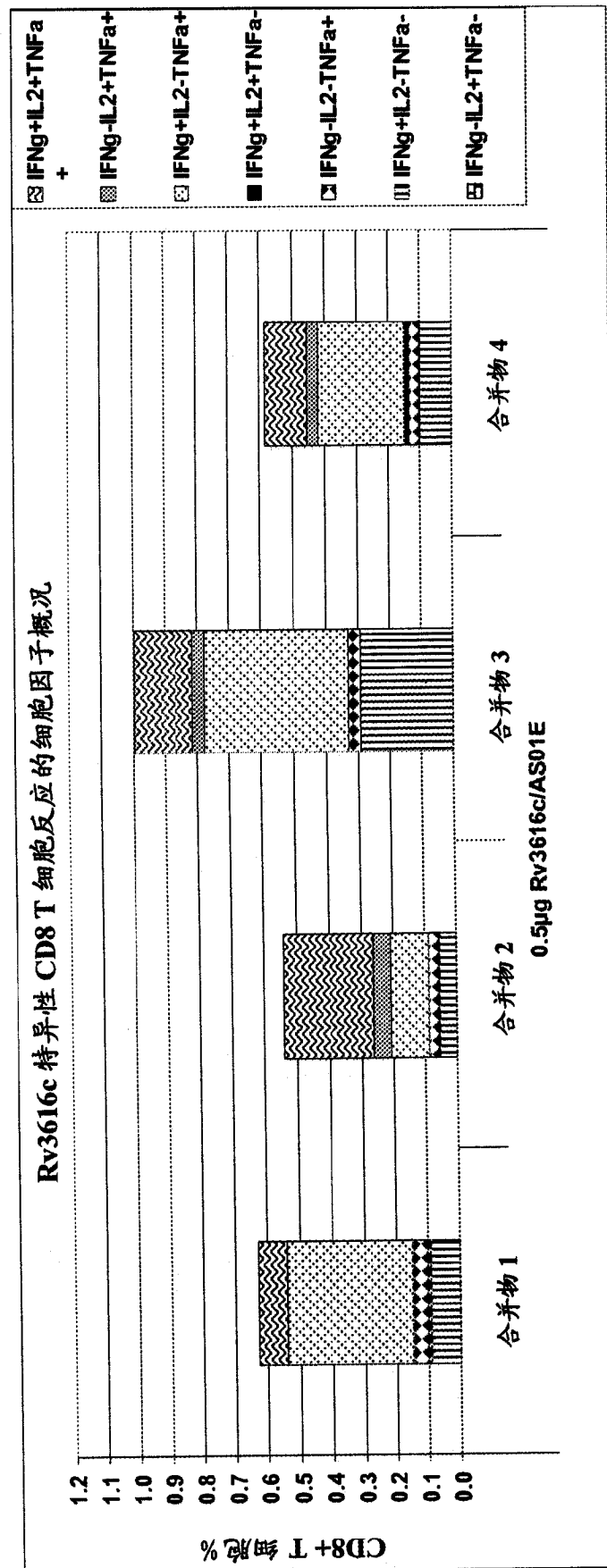


图5

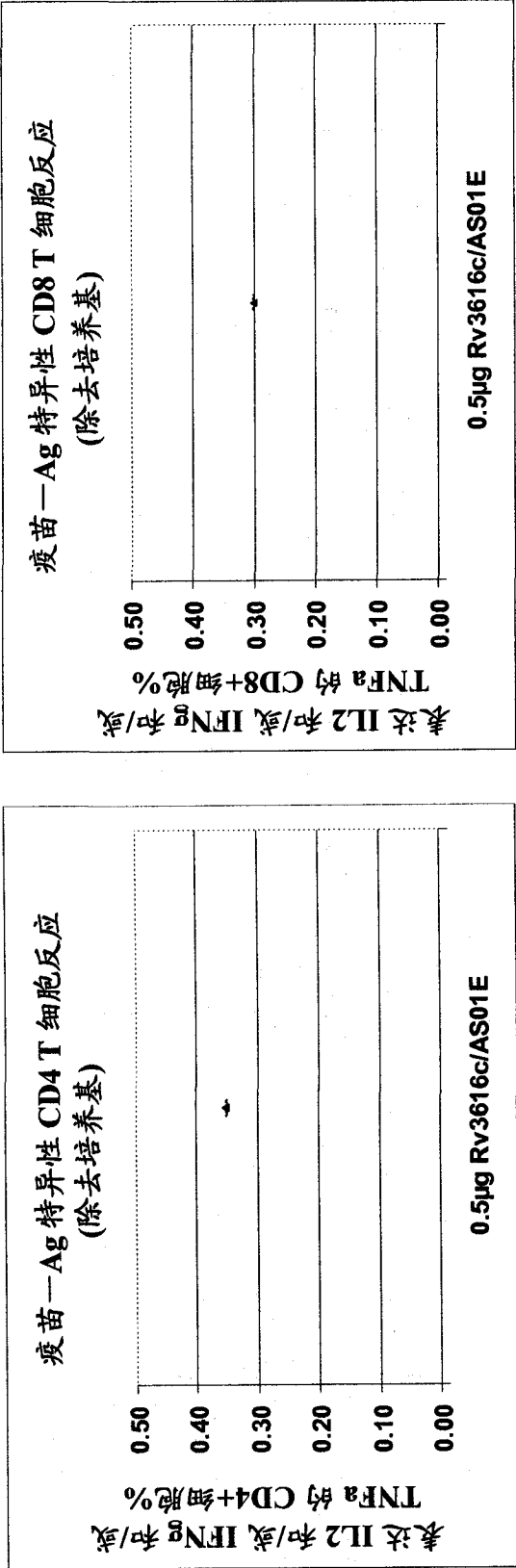


图6

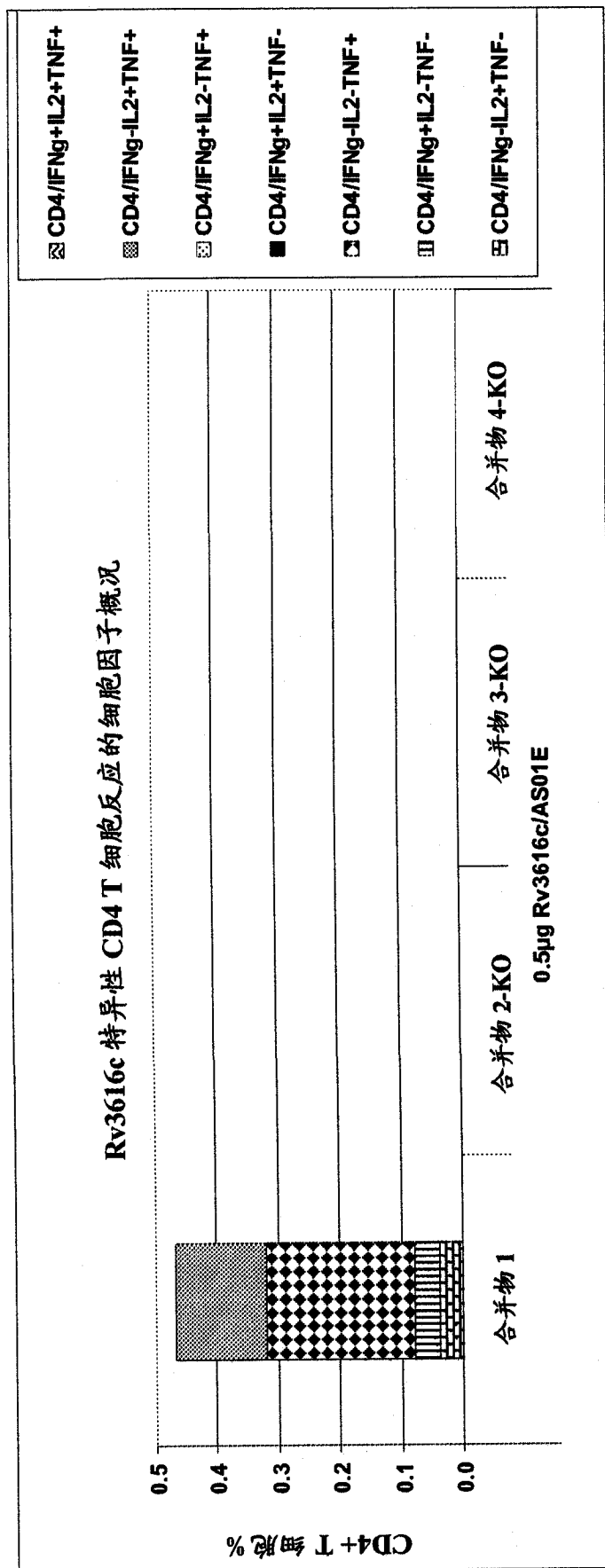


图7

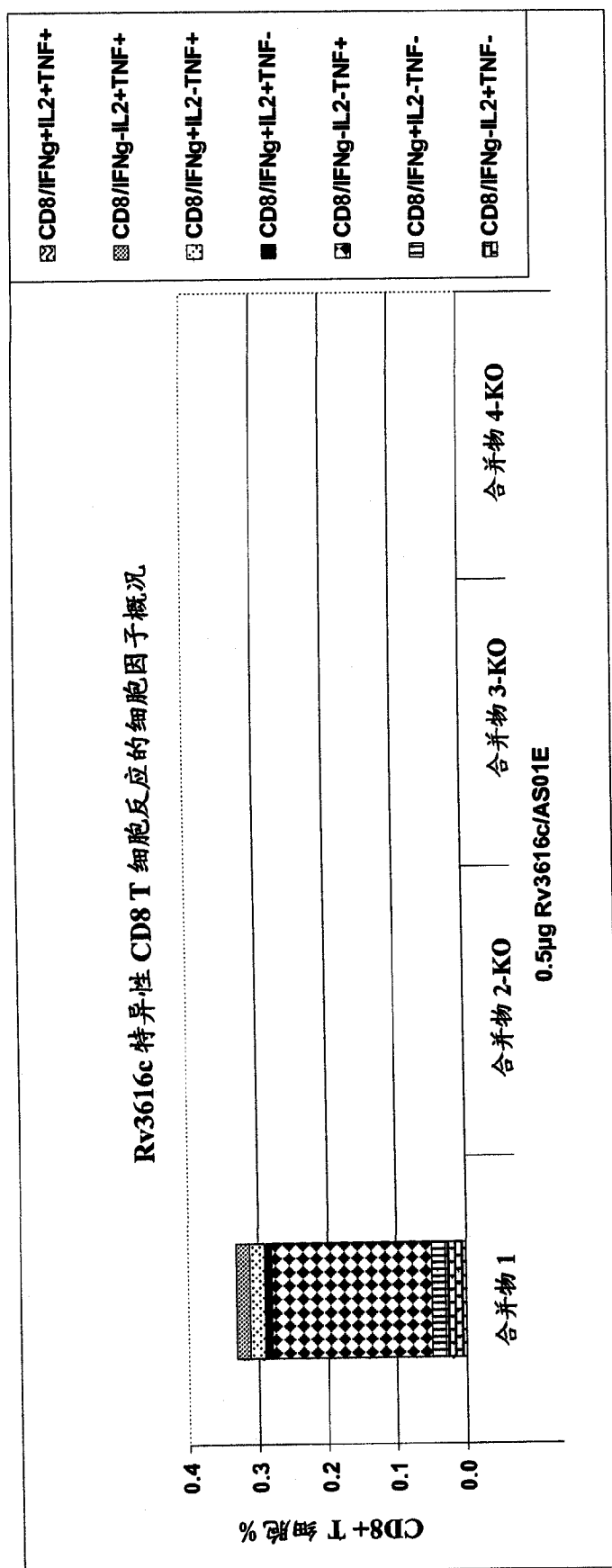


图8

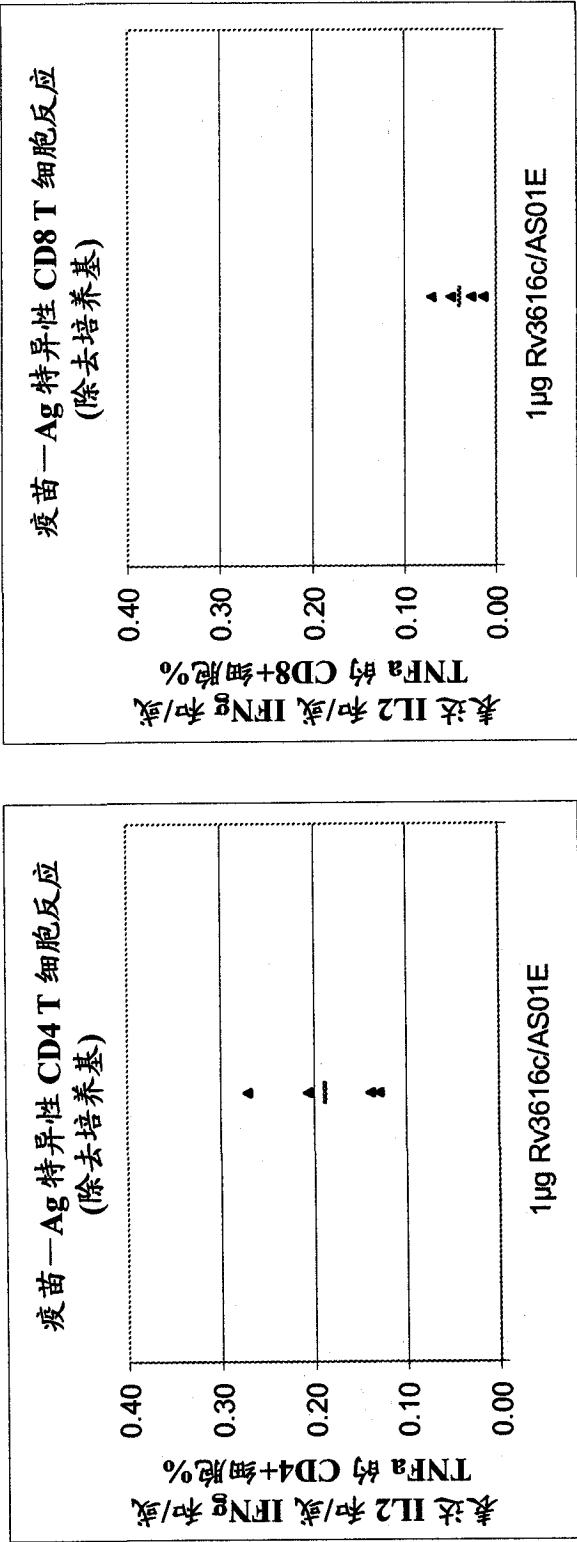


图6

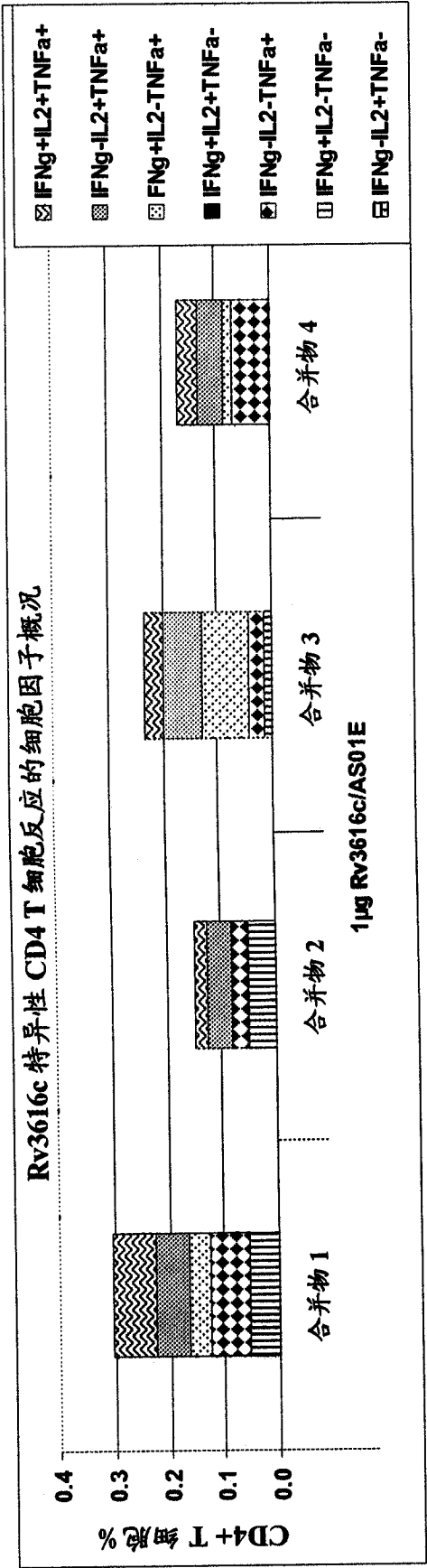


图10

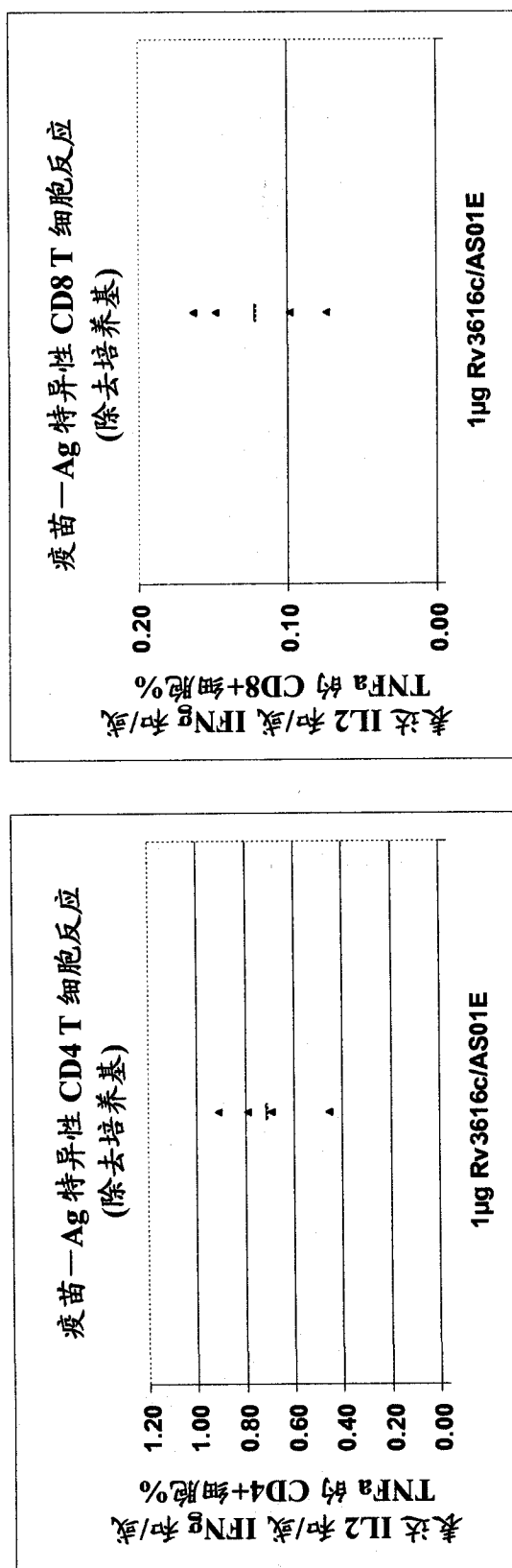


图11

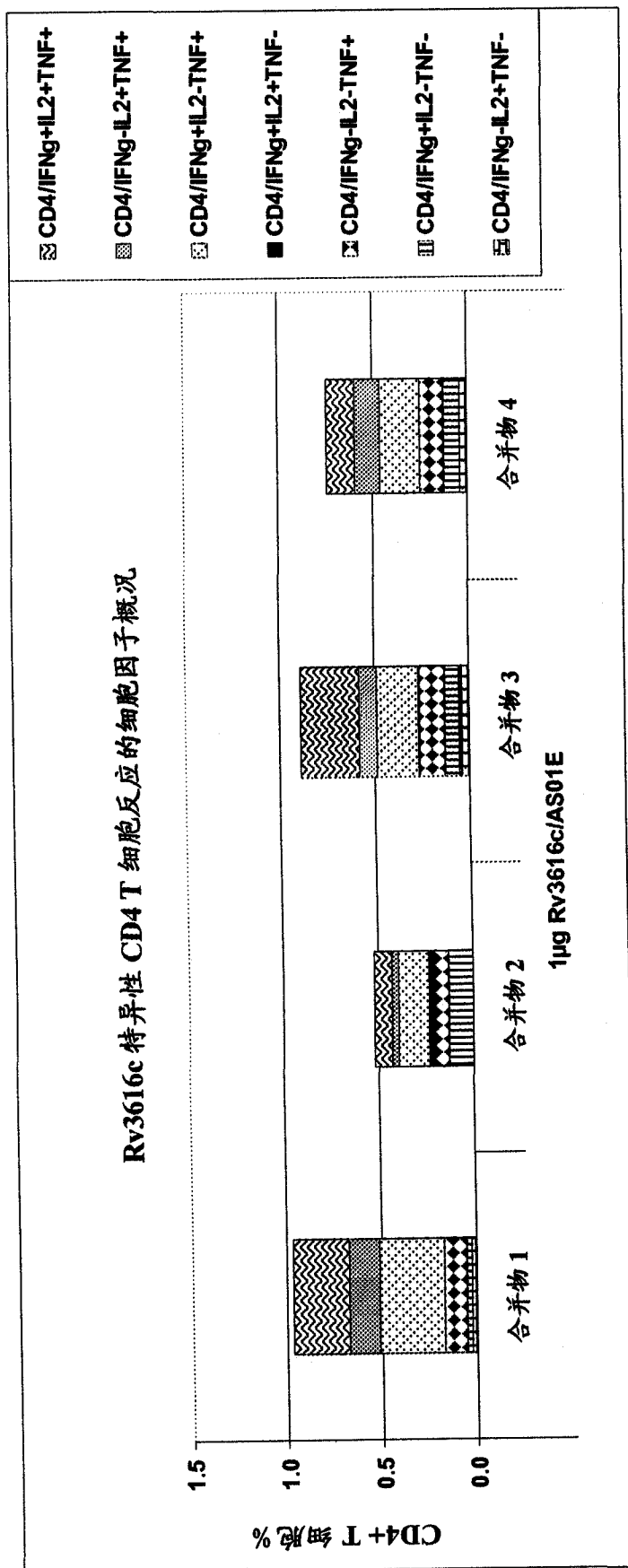


图12

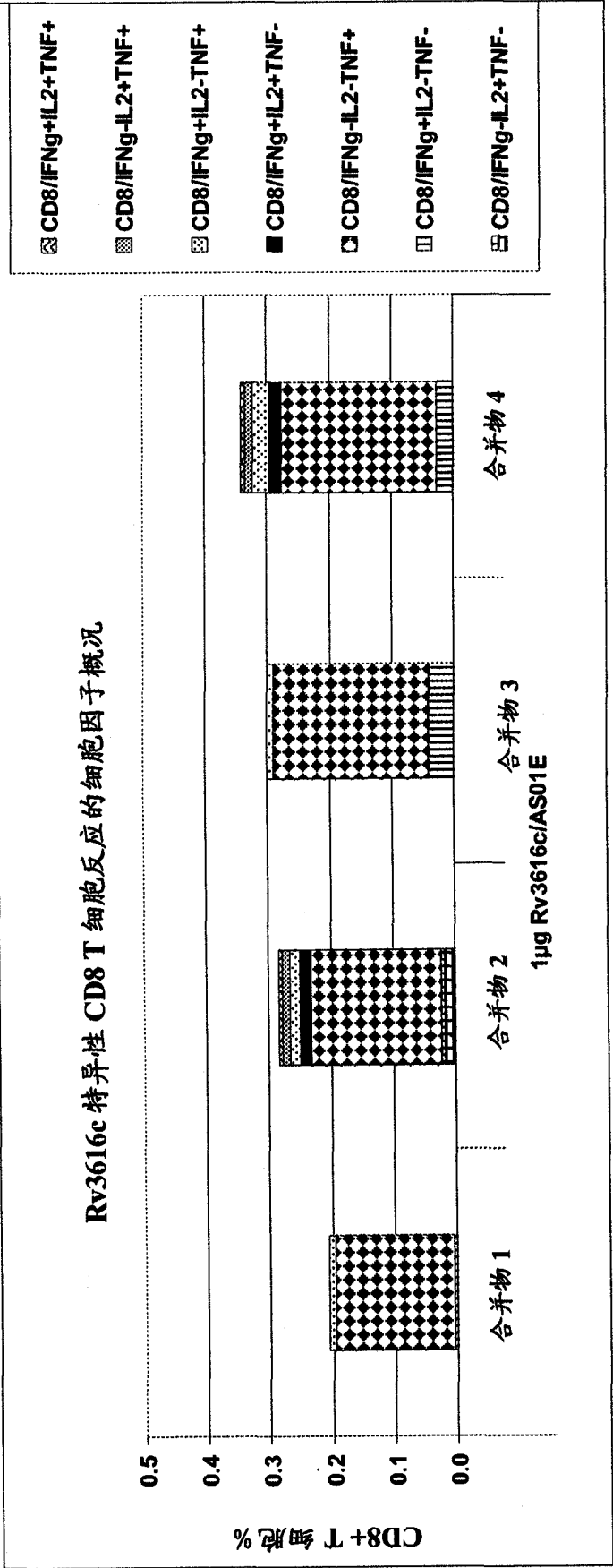


图13

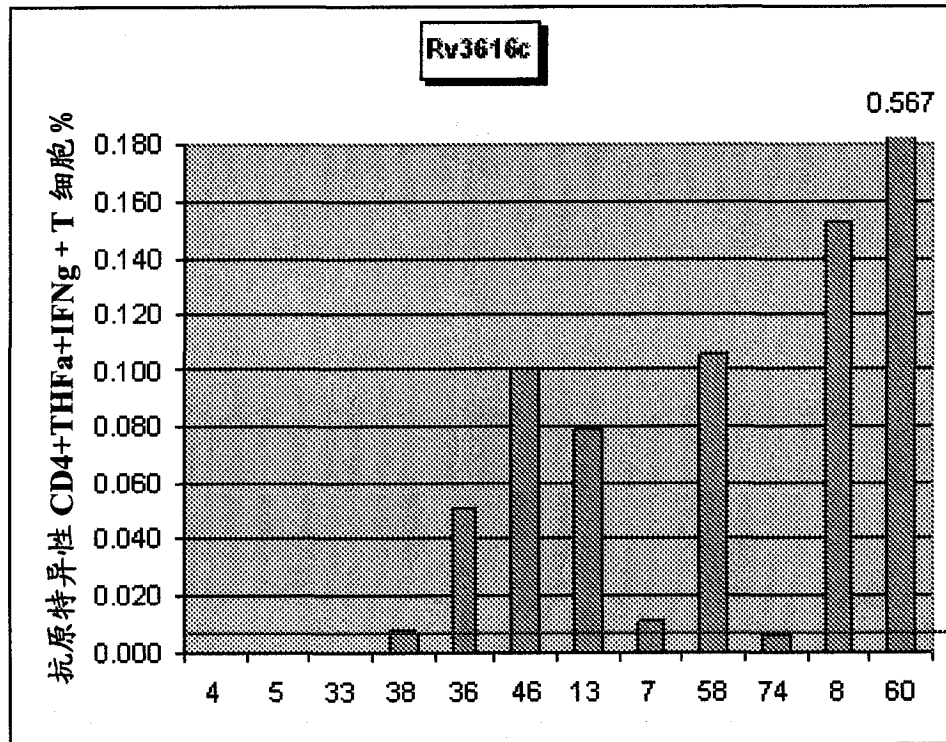


图14