



(12)发明专利申请

(10)申请公布号 CN 106866801 A

(43)申请公布日 2017. 06. 20

(21)申请号 201710068231.X

(22)申请日 2009.07.24

(30)优先权数据

61/083699 2008.07.25 US

(62)分案原申请数据

200980138405.9 2009.07.24

(71)申请人 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

申请人 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 翟建伟 鲁炜

(51)Int. Cl.

C07K 14/35(2006.01)

C07K 16/12(2006.01)

C07K 19/00(2006.01)

C12N 15/31(2006.01)

C12N 15/62(2006.01)

C12P 21/02(2006.01)

A61K 39/04(2006.01)

A61P 31/06(2006.01)

G01N 33/569(2006.01)

权利要求书7页 说明书62页

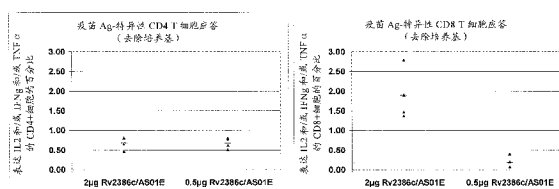
序列表72页 附图9页

(54)发明名称

结核病RV2386C蛋白、组合物及其用途

(57)摘要

本发明涉及结核病RV2386C蛋白、组合物及其用途,还涉及多肽,其包含(i)Rv2386c蛋白序列;(ii)Rv2386c蛋白序列的变体;或者(iii)Rv2386c蛋白序列的免疫原性片段。本发明的其它方面涉及相关的多核苷酸、融合蛋白以及治疗或预防结核病的方法。



1. 分离多肽,其包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段。
2. 如权利要求1所述的多肽,其用作药剂。
3. 如权利要求1或2所述的多肽,其包含Rv2386c蛋白序列。
4. 如权利要求3所述的多肽,其由Rv2386c蛋白序列组成。
5. 如权利要求1或2所述的多肽,其包含Rv2386c蛋白序列的变体。
6. 如权利要求5所述的多肽,其包含与Rv2386c蛋白序列具有至少70%的一致性的Rv2386c蛋白序列的变体。
7. 如权利要求6所述的多肽,其包含与Rv2386c蛋白序列具有至少90%的一致性的Rv2386c蛋白序列的变体。
8. 如权利要求7所述的多肽,其包含与Rv2386c蛋白序列具有至少95%的一致性的Rv2386c蛋白序列的变体。
9. 如权利要求5所述的多肽,其由Rv2386c蛋白序列的变体组成。
10. 如权利要求9所述的多肽,其由与Rv2386c蛋白序列具有至少70%的一致性的Rv2386c蛋白序列的变体组成。
11. 如权利要求10所述的多肽,其由与Rv2386c蛋白序列具有至少90%的一致性的Rv2386c蛋白序列的变体组成。
12. 如权利要求11所述的多肽,其由与Rv2386c蛋白序列具有至少95%的一致性的Rv2386c蛋白序列的变体组成。
13. 如权利要求1或2所述的多肽,其包含Rv2386c蛋白序列的免疫原性片段。
14. 如权利要求13所述的多肽,其包含来自Rv2386c蛋白序列的最少10个氨基酸残基的免疫原性片段。
15. 如权利要求14所述的多肽,其包含来自Rv2386c蛋白序列的最少20个氨基酸残基的免疫原性片段。
16. 如权利要求15所述的多肽,其包含来自Rv2386c蛋白序列的最少50个氨基酸残基的免疫原性片段。
17. 如权利要求13所述的多肽,其由Rv2386c蛋白序列的免疫原性片段组成。
18. 如权利要求17所述的多肽,其由来自Rv2386c蛋白序列的最少10个氨基酸残基的免疫原性片段组成。
19. 如权利要求18所述的多肽,其由来自Rv2386c蛋白序列的最少20个氨基酸残基的免疫原性片段组成。
20. 如权利要求19所述的多肽,其由来自Rv2386c蛋白序列的最少50个氨基酸残基的免疫原性片段组成。
21. 如权利要求1-20中任意一项所述的多肽,其中该Rv2386c蛋白序列来自结核分枝杆菌。
22. 如权利要求21所述的多肽,其中该Rv2386c蛋白具有SEQ ID NO: 1的序列。
23. 如权利要求1-22中任意一项所述的多肽,其用于治疗或预防结核病。

24. 如权利要求23所述的多肽,其用于治疗结核病。
25. 如权利要求23所述的多肽,其用于预防结核病。
26. 如权利要求23所述的多肽,其用于治疗潜伏结核病。
27. 如权利要求23所述的多肽,其用于预防潜伏结核病。
28. 如权利要求23所述的多肽,其用于预防结核病的再活化。
29. 如权利要求23所述的多肽,其用于延缓结核病的再活化。
30. 如权利要求23-29中任意一项所述的多肽,其中该结核病与结核分枝杆菌感染相关。
31. 包含编码多肽的核酸序列的分离多核苷酸,该多肽包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段。
32. 如权利要求31所述的多核苷酸,其用作药剂。
33. 如权利要求31或32所述的多核苷酸,其包含编码多肽的核酸序列,该多肽包含Rv2386c蛋白序列。
34. 如权利要求31或32所述的多核苷酸,其包含编码多肽的核酸序列,该多肽包含Rv2386c蛋白序列的变体。
35. 如权利要求34所述的多核苷酸,其包含编码多肽的核酸序列,该多肽包含与Rv2386c蛋白序列具有至少70%的一致性的Rv2386c蛋白序列的变体。
36. 如权利要求31或32所述的多核苷酸,其包含编码多肽的核酸序列,该多肽包含Rv2386c蛋白序列的免疫原性片段。
37. 如权利要求36所述的多核苷酸,其包含编码多肽的核酸序列,该多肽包含来自Rv2386c蛋白序列的至少10个氨基酸残基的免疫原性片段。
38. 如权利要求31-37中任意一项所述的多核苷酸,其包含编码多肽的核酸序列,该多肽由以下组成:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段。
39. 如权利要求38所述的多核苷酸,其由编码多肽的核酸序列组成,该多肽由以下组成:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段。
40. 如权利要求31-39中任意一项所述的多核苷酸,其中该Rv2386c蛋白序列来自结核分枝杆菌。
41. 如权利要求40所述的多核苷酸,其中该Rv2386c蛋白具有SEQ ID NO: 1的序列。
42. 如权利要求41所述的多核苷酸,其包含SEQ ID NO: 2的序列或其片段。
43. 如权利要求31-42中任意一项所述的多核苷酸,其用于治疗或预防结核病。
44. 如权利要求43所述的多核苷酸,其用于治疗结核病。

45. 如权利要求43所述的多核苷酸,其用于预防结核病。
46. 如权利要求43所述的多核苷酸,其用于治疗潜伏结核病。
47. 如权利要求43所述的多核苷酸,其用于预防潜伏结核病。
48. 如权利要求43所述的多核苷酸,其用于预防结核病的再活化。
49. 如权利要求43所述的多核苷酸,其用于延缓结核病的再活化。
50. 如权利要求43-49中任意一项所述的多核苷酸,其中该结核病与结核分枝杆菌感染相关。
51. 药物组合物,其包含:
 - (a) 多肽,其包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段;或者
 - (b) 包含编码(a)的多肽的核酸序列的多核苷酸;以及
 - (c) 药学上可接受的载体或赋形剂。
52. 免疫原性组合物,其包含:
 - (a) 多肽,其包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段;或者
 - (b) 包含编码(a)的多肽的核酸序列的多核苷酸;以及
 - (c) 非特异性免疫应答增强剂。
53. 如权利要求52所述的免疫原性组合物,其中该非特异性免疫应答增强剂为佐剂。
54. 包含编码多肽的核酸序列的表达载体,该多肽包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段。
55. 以权利要求54所述的表达载体转化的分离宿主细胞。
56. 重组表达多肽的分离宿主细胞,该多肽包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段。
57. 如权利要求56所述的分离宿主细胞,其选自大肠杆菌、酵母和哺乳动物细胞。
58. 如权利要求56所述的分离宿主细胞,其为卡介杆菌。
59. 生产多肽的方法,该多肽包含:
 - (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段;

所述方法包括在宿主细胞内重组表达所述多肽的步骤。

60. 治疗或预防结核病的方法,其包括向有此需要的对象施用安全和有效量的多肽,该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体;或者
- (iii) Rv2386c蛋白序列的免疫原性片段;

其中所述多肽诱导免疫应答。

61. 治疗或预防结核病的方法,其包括向有此需要的对象施用安全和有效量的多核苷酸,该多核苷酸包含编码多肽的核酸序列,该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体;或者
- (iii) Rv2386c蛋白序列的免疫原性片段;

其中所述多核苷酸诱导免疫应答。

62. 如权利要求60或61所述的方法,其中该多肽或多核苷酸作为药物组合物施用。

63. 如权利要求60或61所述的方法,其中该多肽或多核苷酸作为免疫原性组合物施用。

64. 如权利要求60-63中任意一项所述的方法,其中该对象为哺乳动物。

65. 如权利要求64所述的方法,其中该对象为人。

66. 如权利要求60-65中任意一项所述的方法,其中该对象患有活动性结核病。

67. 如权利要求60-65中任意一项所述的方法,其中该对象患有潜伏结核病。

68. 如权利要求60-65中任意一项所述的方法,其中该对象未患有结核病。

69. 如权利要求60-68中任意一项所述的方法,其中该对象之前用卡介杆菌 (BCG) 免疫。

70. 如权利要求60-69中任意一项所述的方法,其中该Rv2386c蛋白序列来自结核分枝杆菌。

71. 如权利要求70所述的方法,其中该Rv2386c蛋白具有SEQ ID NO: 1的序列。

72. 如权利要求60-71中任意一项所述的方法,其进一步包括施用一种或多种可有效治疗结核病的化疗剂。

73. 如权利要求72所述的方法,其中该一种或多种化疗剂选自异烟肼和利福平。

74. 如权利要求60-73中任意一项所述的方法,进一步包括施用至少一种另外的结核分枝杆菌抗原。

75. 如权利要求74所述的方法,其中该另外的结核分枝杆菌抗原以多肽形式提供。

76. 如权利要求75所述的方法,其中该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体;或者
- (iii) Rv2386c蛋白序列的免疫原性片段;

且该另外的结核分枝杆菌抗原以融合蛋白的形式提供。

77. 如权利要求74所述的方法,其中该另外的结核分枝杆菌抗原以多核苷酸形式提供。

78. 如权利要求77所述的方法,其中该多核苷酸包含编码多肽的核酸序列,该多肽包含:

- (i) Rv2386c蛋白序列;
 - (ii) Rv2386c蛋白序列的变体;或者
 - (iii) Rv2386c蛋白序列的免疫原性片段;
- 且该另外的结核分枝杆菌抗原以单个多核苷酸序列的形式提供。

79. 如权利要求78所述的方法,其中所述单个多核苷酸序列编码融合蛋白,该融合蛋白包含该Rv2386c抗原和该另外的结核分枝杆菌抗原。

80. 如权利要求74所述的方法,其中该另外的结核分枝杆菌抗原以卡介杆菌 (BCG) 形式提供。

81. 如权利要求60-80中任意一项所述的方法,其中结核病的治疗或预防指结核病的治疗。

82. 如权利要求60-80中任意一项所述的方法,其中结核病的治疗或预防指结核病的预防。

83. 如权利要求60-80中任意一项所述的方法,其中结核病的治疗或预防指潜伏结核病的治疗。

84. 如权利要求60-80中任意一项所述的方法,其中结核病的治疗或预防指潜伏结核病的预防。

85. 如权利要求60-80中任意一项所述的方法,其中结核病的治疗或预防指结核病再活化的预防。

86. 如权利要求60-80中任意一项所述的方法,其中结核病的治疗或预防指结核病再活化的延缓。

87. 如权利要求60-86中任意一项所述的方法,其中该结核病与结核分枝杆菌感染相关。

88. 如权利要求87所述的方法,其中该结核分枝杆菌感染与结核分枝杆菌的多重耐药菌株有关。

89. 如权利要求60-88中任意一项所述的方法,其诱导针对结核分枝杆菌感染的免疫应答。

90. 多肽在用于治疗或预防结核病的药剂生产中的用途,该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体;或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

91. 包含编码多肽的核酸序列的多核苷酸在用于治疗或预防结核病的药剂生产中的用途,该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体;或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

92. 包含多肽的融合蛋白,该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体;或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

93. 包含编码含有多肽的融合蛋白的核苷酸序列的多核苷酸, 该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体; 或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

94. 包含第一核苷酸序列的多核苷酸, 该第一核苷酸序列在适度严格条件下与编码多肽的第二核苷酸序列的补体选择性杂交, 该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体; 或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

95. 如权利要求94所述的多核苷酸, 其包含在高度严格条件下与编码多肽的第二核苷酸序列的补体选择性杂交的第一核苷酸序列, 该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体; 或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

96. 如权利要求94或95所述的多核苷酸, 其包含与SEQ ID NO: 2的补体选择性杂交的第一核苷酸序列。

97. 与多肽特异性结合的抗体或其片段, 该多肽包含:

- (i) Rv2386c蛋白序列;
- (ii) Rv2386c蛋白序列的变体; 或者
- (iii) Rv2386c蛋白序列的免疫原性片段。

98. 诊断试剂盒, 其包含:

- (a) 如权利要求1所述的多肽;
- (b) 使所述多肽与来自个体的样本(例如, 全血或更合适地为PBMC)充分接触的设备;

和

- (c) 对样本的T细胞应答进行定量的装置。

99. 如权利要求98所述的诊断试剂盒, 其包含:

- (a) 如权利要求1所述的多肽;
- (b) 使所述多肽与来自个体的样本(例如, 全血或更合适地为PBMC)充分接触的设备;

和

- (c) 对样本的T细胞应答进行定量和表征的装置。

100. 诊断试剂盒, 其包含:

- (a) 如权利要求1所述的多肽;
- (b) 使所述多肽与来自个体的样本(例如, 全血或更合适地为PBMC)充分接触的设备;

和

- (c) 对样本的B细胞应答进行定量的装置。

101. 诊断试剂盒, 其包含:

- (a) 如权利要求1所述的多肽; 和
- (b) 使所述多肽与患者的真皮细胞充分接触的设备。

102. 诊断试剂盒, 其包含:

- (a) 如权利要求1所述的多肽;和
- (c) 对来自个体样本(例如,全血或更合适地为PBMC)的T细胞应答进行定量的装置。

103. 诊断试剂盒,其包含:

- (a) 如权利要求1所述的多肽;和
- (c) 对来自个体样本(例如,全血或更合适地为PBMC)的B细胞应答进行定量的装置。

结核病RV2386C蛋白、组合物及其用途

[0001] 本申请是申请日为2009年7月24日的中国专利申请200980138405.9“结核病RV2386C蛋白、组合物及其用途”的分案申请。

技术领域

[0002] 本发明涉及用于治疗或预防结核病(特别是用于治疗或预防潜伏结核病以及预防或延缓结核病的再活化)的多肽和多核苷酸(以及涉及相关方法)。本发明进一步涉及包含所述多肽和多核苷酸的药物和免疫原性组合物,以及涉及结核病(特别是潜伏结核病)的诊断方法。

背景技术

[0003] 结核病(TB)是一种由结核分枝杆菌(*Mycobacterium tuberculosis*)和其它分枝杆菌种类的感染引起的慢性传染疾病。它是世界上发展中国家的一种主要疾病,并成为发达区域日益严重的问题。超过20亿人被认为感染了TB杆菌,每年新增约920万TB病例和170万死亡病例。其中10%感染TB杆菌的人会发展为活动性TB,每位患有活动性TB的人平均每年感染10-15其他人。尽管在全球的年均发病率已经达到顶峰,由于人口增长,死亡数和病例数依然在上升(World Health Organisation Tuberculosis Facts 2008)。

[0004] 结核分枝杆菌通过呼吸道途径感染个体。肺泡巨噬细胞吸入该细菌,但它能够通过抑制具有酸性溶酶体的吞噬体融合存活和增殖。随后发生涉及CD4⁺和CD8⁺T细胞的复杂免疫应答,最终导致肉芽肿的形成。结核分枝杆菌成为病原体的关键在于该分离但未根除的细菌可长期存在,使得个体容易在随后发展为活动性TB。

[0005] 在感染后的第一年,少于5%的感染个体会发展为活动性TB。该肉芽肿可存在数十年,并被认为在缺乏氧和营养的休眠状态下包含活结核分枝杆菌。然而,最近表明,大部分处于休眠状态的细菌位于遍布体内的非巨噬细胞的细胞类型中(Locht等人,Expert Opin.Biol.Ther.2007 7(11):1665-1677)。当宿主的天然免疫和病原体之间的平衡发生变化时(例如由于免疫抑制事件)则发展为活动性TB(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TB和活动性TB之间的平衡的动力学假设(Cardana P-J Inflammation&Allergy-Drug Targets 2006 6:27-39;Cardana P-J Infection 2009 37(2):80-86)。

[0007] 尽管感染在相当长时间内可以是无症状的,活动性疾病最常显现为肺的急性炎症,从而导致疲劳、体重下降、发烧和持续咳嗽。如果未经治疗,通常可导致严重的并发症和死亡。

[0008] 结核病通常可采用长期抗生素疗法进行控制,尽管该种治疗并不足以防止该疾病的扩散。感染的个体可以是无症状的,但在一定时间具有传染性。此外,尽管对治疗方案的顺应性是关键,但患者的行为很难监测。一些患者并未完成治疗周期,这可能导致无效治疗和耐药性的形成。

[0009] 多重耐药性TB (MDR-TB) 是一种对一线药物没有应答的形式。所有TB病例中有5%是MDR-TB, 每年估计有490,000新MDR-TB病例。当耐受针对MDR-TB开发的二线药物时, 产生广泛耐药TB (XDR-TB)。据估计每年产生40,000实际无法治疗的XDR-TB新病例 (World Health Organisation 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可能无法从感染个体根除, 并可保留为能够再活化的潜伏感染。

[0011] 为了控制结核病扩散, 对疾病的有效免疫和准确早期诊断是最为重要的。

[0012] 潜伏TB感染的诊断通常是采用结核菌素皮试来完成的, 该皮试包括皮内暴露至结核菌素蛋白-纯化的衍生物 (PPD)。抗原特异性T细胞应答导致注射后在注射部位48-72小时的可测量硬结, 其表示暴露至分枝菌抗原。然而, 该测试的灵敏度和特异性是一个问题, 且接种了BCG的个体并不能始终容易地与感染个体相区别 (这对于BCG无法对潜伏感染形成保护的事实特别重要)。一般而言, 已经接受BCG但未被结核分枝杆菌感染的个体显示了直径小于10mm的PPD反应, 而具有直径在10mm以上的PPD反应的人被认为已经被结核分枝杆菌感染。然而, 这一原则不适用于由于HIV感染而免疫抑制的个体, 后者可导致直径小于10mm的PPD反应; 或者在发病地区, 被非结核分枝杆菌感染的人能显示直径在10mm以上的PPD反应。

[0013] 近些年在体外基于T细胞的测定上已经取得了进展, 该测定基于干扰素-gamma释放并采用比PPD对结核分枝杆菌更具特异性的抗原, 称为ESAT-6和CFP-10。这些高特异性测试看来至少具有结核菌素皮试的灵敏度, 且由于BCG接种而显示更少的交叉反应性。有关潜伏TB诊断的最近综述参见Pai M等人Expert Rev.Mol.Diagn.20066 (3):413-422。然而, 由于ESAT-6/CFP-10是初期抗原, 基于ESAT-6/CFP-10的测定仅能在新近感染的人员中最佳实施。因此, 与潜伏结核病特异性相关的新抗原的鉴定有助于能够检测更长期潜伏感染的更灵敏测定的开发。

[0014] 对于治疗和预防结核病 (特别是治疗和预防潜伏TB以及预防TB的再活化) 的有效策略存在着需求。

发明内容

[0015] 本发明总体涉及作为TB抗原 (特别是与潜伏TB相关的抗原) 的Rv2386c的鉴定, 并涉及预防和治疗TB (尤其是预防和/或治疗潜伏TB以及预防或延缓TB再活化) 的相关方法和用途。

[0016] 本发明提供了分离的多肽, 其包含:

[0017] (i) Rv2386c蛋白序列;

[0018] (ii) Rv2386c蛋白序列的变体; 或者

[0019] (iii) Rv2386c蛋白序列的免疫原性片段。

[0020] 本发明还提供了多肽, 其包含:

[0021] (i) Rv2386c蛋白序列;

[0022] (ii) Rv2386c蛋白序列的变体; 或者

[0023] (iii) Rv2386c蛋白序列的免疫原性片段;

[0024] 该多肽用作药剂。

[0025] 本发明另一方面涉及治疗或预防TB的方法, 其包括向有此需要的对象施用安全和

有效量的多肽,该多肽包含:

[0026] (i) Rv2386c蛋白序列;

[0027] (ii) Rv2386c蛋白序列的变体;或者

[0028] (iii) Rv2386c蛋白序列的免疫原性片段;

[0029] 其中所述多肽诱导免疫应答,特别是抗结核分枝杆菌的免疫应答。

[0030] 多肽在用于制备治疗或预防TB的药物中的用途,该多肽包含:

[0031] (i) Rv2386c蛋白序列;

[0032] (ii) Rv2386c蛋白序列的变体;或者

[0033] (iii) Rv2386c蛋白序列的免疫原性片段;

[0034] 该用途代表了本发明的另一方面。

[0035] 本发明提供了包含编码多肽的核酸序列的分离多核苷酸,该多肽包含:

[0036] (i) Rv2386c蛋白序列;

[0037] (ii) Rv2386c蛋白序列的变体;或者

[0038] (iii) Rv2386c蛋白序列的免疫原性片段。

[0039] 还提供了包含编码多肽的核酸序列的分离多核苷酸,该多肽包含:

[0040] (i) Rv2386c蛋白序列;

[0041] (ii) Rv2386c蛋白序列的变体;或者

[0042] (iii) Rv2386c蛋白序列的免疫原性片段;

[0043] 该多核苷酸用作药剂。

[0044] 本发明另一方面涉及治疗或预防TB的方法,其包括向有此需要的对象施用安全和有效量的包含编码多肽的核酸序列的多核苷酸,该多肽包含:

[0045] (i) Rv2386c蛋白序列;

[0046] (ii) Rv2386c蛋白序列的变体;或者

[0047] (iii) Rv2386c蛋白序列的免疫原性片段;

[0048] 其中所述多核苷酸诱导免疫应答,特别是抗结核分枝杆菌的免疫应答。

[0049] 包含编码多肽的核酸序列的多核苷酸在用于制备治疗或预防TB的药物中的用途,该多肽包含:

[0050] (i) Rv2386c蛋白序列;

[0051] (ii) Rv2386c蛋白序列的变体;或者

[0052] (iii) Rv2386c蛋白序列的免疫原性片段;

[0053] 该用途代表了本发明的另一方面。

[0054] 另外,提供了药物组合物,其包含:

[0055] (a) 多肽,其包含:

[0056] (i) Rv2386c蛋白序列;

[0057] (ii) Rv2386c蛋白序列的变体;或者

[0058] (iii) Rv2386c蛋白序列的免疫原性片段;或者

[0059] (b) 包含编码(a)的多肽的核酸序列的多核苷酸;

[0060] 以及

[0061] (c) 药学上可接受的载体或赋形剂。

- [0062] 进一步地,提供了免疫原性组合物,其包含:
- [0063] (a) 多肽,其包含:
- [0064] (i) Rv2386c蛋白序列;
- [0065] (ii) Rv2386c蛋白序列的变体;或者
- [0066] (iii) Rv2386c蛋白序列的免疫原性片段;或者
- [0067] (b) 包含编码(a)的多肽的核酸序列的多核苷酸;
- [0068] 以及
- [0069] (c) 非特异性免疫应答增强剂。
- [0070] 适当地,该药物和免疫原性组合物包含了Rv2386c和第二异种结核病抗原成分的组合。
- [0071] 还提供了包含编码多肽的核酸序列的表达载体,该多肽包含:
- [0072] (i) Rv2386c蛋白序列;
- [0073] (ii) Rv2386c蛋白序列的变体;或者
- [0074] (iii) Rv2386c蛋白序列的免疫原性片段。
- [0075] 用所述表达载体转化的宿主细胞构成了本发明进一步的方面。另外提供了重组表达多肽的宿主细胞,该多肽包含:
- [0076] (i) Rv2386c蛋白序列;
- [0077] (ii) Rv2386c蛋白序列的变体;或者
- [0078] (iii) Rv2386c蛋白序列的免疫原性片段。
- [0079] 进一步提供了生产多肽的方法,该多肽包含:
- [0080] (i) Rv2386c蛋白序列;
- [0081] (ii) Rv2386c蛋白序列的变体;或者
- [0082] (iii) Rv2386c蛋白序列的免疫原性片段;
- [0083] 所述方法包括在宿主细胞内重组表达所述多肽的步骤。
- [0084] 另外提供了与多肽特异性结合的抗体或其片段,该多肽包含:
- [0085] (i) Rv2386c蛋白序列;
- [0086] (ii) Rv2386c蛋白序列的变体;或者
- [0087] (iii) Rv2386c蛋白序列的免疫原性片段。
- [0088] 还提供了所述抗体在诊断中的用途(例如诊断结核病的方法,其包括在来自测试对象的生物样本中确定与本发明多肽特异性结合的抗体或其片段的存在)。
- [0089] 还提供了诊断试剂盒,其包含:
- [0090] (a) 本发明的多肽;
- [0091] (b) 使所述多肽与来自个体的样本(例如,全血或更合适地为PBMC)充分接触的设备;和
- [0092] (c) 对样本的T细胞应答进行定量的装置。
- [0093] 本发明的另一方面涉及诊断试剂盒,其包含:
- [0094] (a) 本发明的多肽;和
- [0095] (b) 使所述多肽与患者的真皮细胞充分接触的设备。
- [0096] 在一个实施方式中,接受本发明的多肽、多核苷酸或组合物的对象可患有活动性

结核病(例如被结核分枝杆菌活动性感染)。在第二实施方式中,该对象可具有潜伏结核病(例如,被结核分枝杆菌潜伏感染)。在第三实施方式中,该对象可未患结核病(例如,未被结核分枝杆菌感染)。

[0097] 接受本发明的多肽、多核苷酸或组合物的对象可能之前已经进行结核接种(例如,针对结核分枝杆菌的感染接种),例如已用卡介杆菌(BCG)接种。替代性地,接受本发明的多肽、多核苷酸或组合物的对象可未曾进行结核接种(例如,未针对结核分枝杆菌的感染接种),例如未曾用卡介杆菌(BCG)接种。

附图说明

[0098] 图1:在第21天(即,第二次接种后7天)来自免疫CB6F1小鼠的表达IFN-gamma和/或IL-2和/或TNF- α 细胞因子的CD4和CD8细胞的百分比。

[0099] 图2:免疫CB6F1小鼠中抗原特异性CD4应答在第21天(即,第二次接种后7天)的细胞因子分布(profile)。

[0100] 图3:免疫CB6F1小鼠中抗原特异性CD8应答在第21天(即,第二次接种后7天)的细胞因子分布。

[0101] 图4:在第35天(即,第三次接种后7天)来自免疫CB6F1小鼠的表达IFN-gamma和/或IL-2和/或TNF- α 细胞因子的CD4和CD8细胞的百分比。

[0102] 图5:免疫CB6F1小鼠中抗原特异性CD4应答在第35天(即,第三次接种后7天)的细胞因子分布。

[0103] 图6:免疫CB6F1小鼠中抗原特异性CD8应答在第35天(即,第三次接种后7天)的细胞因子分布。

[0104] 图7:在第21天(即,第二次接种后7天)来自免疫C57BL/6小鼠的表达IFN-gamma和/或IL-2和/或TNF- α 细胞因子的CD4和CD8细胞的百分比。

[0105] 图8:免疫C57BL/6小鼠中抗原特异性CD4应答在第21天(即,第二次接种后7天)的细胞因子分布。

[0106] 图9:免疫C57BL/6小鼠中抗原特异性CD8应答在第21天(即,第二次接种后7天)的细胞因子分布。

[0107] 列出的序列描述

[0108] SEQ ID No:1:来自结核分枝杆菌H37Rv菌株的Rv2386c的多肽序列。

[0109] SEQ ID No:2:来自结核分枝杆菌H37Rv菌株的Rv2386c的多核苷酸序列。

[0110] SEQ ID No:3:来自结核分枝杆菌CDC1551菌株的Rv2386c的多肽序列。

[0111] SEQ ID No:4:来自结核分枝杆菌F11菌株的Rv2386c的多肽序列。

[0112] SEQ ID No:5:来自结核分枝杆菌Haarlem A菌株的Rv2386c的多肽序列。

[0113] SEQ ID No:6:来自结核分枝杆菌C菌株的Rv2386c的多肽序列。

[0114] SEQ ID No:7:来自BCG的Rv2386c的多肽序列。

[0115] SEQ ID No:8:Mtb8.4的多肽序列。

[0116] SEQ ID No:9:Mtb9.8的多肽序列。

[0117] SEQ ID No:10:Mtb9.9的多肽序列。

[0118] SEQ ID No:11:Ra12的多肽序列。

- [0119] SEQ ID No:12:Ra35的多肽序列。
- [0120] SEQ ID No:13:TbH9的多肽序列。
- [0121] SEQ ID No:14:Mtb40的多肽序列。
- [0122] SEQ ID No:15:Mtb41的多肽序列。
- [0123] SEQ ID No:16:ESAT-6的多肽序列。
- [0124] SEQ ID No:17:Ag85A的多肽序列。
- [0125] SEQ ID No:18:Ag85B的多肽序列。
- [0126] SEQ ID No:19:alpha-晶体蛋白的多肽序列。
- [0127] SEQ ID No:20:MPT64的多肽序列。
- [0128] SEQ ID No:21:Mtb32A的多肽序列。
- [0129] SEQ ID No:22:Ser/Ala突变成熟Mtb32A的多肽序列。
- [0130] SEQ ID No:23:TB10.4的多肽序列。
- [0131] SEQ ID No:24:Mtb72f的多肽序列。
- [0132] SEQ ID No:25:M72的多肽序列。
- [0133] SEQ ID No:26:Mtb71f的多肽序列。
- [0134] SEQ ID No:27:M92融合体的多肽序列。
- [0135] SEQ ID No:28:M103融合体的多肽序列。
- [0136] SEQ ID No:29:M114融合体的多肽序列。
- [0137] SEQ ID No:30:推定人CD4细胞表位1。
- [0138] SEQ ID No:31:推定人CD4细胞表位2。
- [0139] SEQ ID No:32:推定人CD4细胞表位3。
- [0140] SEQ ID No:33:推定人CD4细胞表位4。
- [0141] SEQ ID No:34:推定人CD4细胞表位5。
- [0142] SEQ ID No:35:推定人CD4细胞表位6。
- [0143] SEQ ID No:36:推定人CD4细胞表位7。
- [0144] SEQ ID No:37:推定人CD4细胞表位8。
- [0145] SEQ ID No:38:推定人CD4细胞表位9。
- [0146] SEQ ID No:39:推定人CD4细胞表位10。
- [0147] SEQ ID No:40:推定人CD4细胞表位11。
- [0148] SEQ ID No:41:推定人CD4细胞表位12。
- [0149] SEQ ID No:42:推定人CD4细胞表位13。
- [0150] SEQ ID No:43:推定人CD4细胞表位14。
- [0151] SEQ ID No:44:推定人CD4细胞表位15。
- [0152] SEQ ID No:45:推定人CD4细胞表位16。
- [0153] SEQ ID No:46:推定人CD4细胞表位17。
- [0154] SEQ ID No:47:推定人CD4细胞表位18。
- [0155] SEQ ID No:48:推定人CD4细胞表位19。
- [0156] SEQ ID No:49:推定人CD4细胞表位20。
- [0157] SEQ ID No:50:推定人CD4细胞表位21。

- [0158] SEQ ID No:51:推定人CD4细胞表位22。
- [0159] SEQ ID No:52:推定人CD4细胞表位23。
- [0160] SEQ ID No:53:推定人CD8细胞表位1。
- [0161] SEQ ID No:54:推定人CD8细胞表位2。
- [0162] SEQ ID No:55:推定人CD8细胞表位3。
- [0163] SEQ ID No:56:推定人CD8细胞表位4。
- [0164] SEQ ID No:57:推定人CD8细胞表位5。
- [0165] SEQ ID No:58:推定人CD8细胞表位6。
- [0166] SEQ ID No:59:推定人CD8细胞表位7。
- [0167] SEQ ID No:60:推定人CD8细胞表位8。
- [0168] SEQ ID No:61:推定人CD8细胞表位9。
- [0169] SEQ ID No:62:推定人CD8细胞表位10。
- [0170] SEQ ID No:63:推定人CD8细胞表位11。
- [0171] SEQ ID No:64:推定人CD8细胞表位12。
- [0172] SEQ ID No:65:推定人CD8细胞表位13。
- [0173] SEQ ID No:66:推定人CD8细胞表位14。
- [0174] SEQ ID No:67:推定人CD8细胞表位15。
- [0175] SEQ ID No:68:推定人CD8细胞表位16。
- [0176] SEQ ID No:69:推定人CD8细胞表位17。
- [0177] SEQ ID No:70:推定人CD8细胞表位18。
- [0178] SEQ ID No:71:推定人CD8细胞表位19。
- [0179] SEQ ID No:72:推定人CD8细胞表位20。
- [0180] SEQ ID No:73:推定人CD8细胞表位21。
- [0181] SEQ ID No:74:推定人CD8细胞表位22。
- [0182] SEQ ID No:75:推定人CD8细胞表位23。
- [0183] SEQ ID No:76:推定人CD8细胞表位24。
- [0184] SEQ ID No:77:推定人CD8细胞表位25。
- [0185] SEQ ID No:78:推定人CD8细胞表位26。
- [0186] SEQ ID No:79:推定人CD8细胞表位27。
- [0187] SEQ ID No:80:推定人CD8细胞表位28。
- [0188] SEQ ID No:81:推定人CD8细胞表位29。
- [0189] SEQ ID No:82:推定人CD8细胞表位30。
- [0190] SEQ ID No:83:推定人CD8细胞表位31。
- [0191] SEQ ID No:84:推定人CD8细胞表位32。
- [0192] SEQ ID No:85:推定人CD8细胞表位33。
- [0193] SEQ ID No:86:推定人CD8细胞表位34。
- [0194] SEQ ID No:87:推定人CD8细胞表位35。
- [0195] SEQ ID No:88:推定人CD8细胞表位36。
- [0196] SEQ ID No:89:推定人CD8细胞表位37。

- [0197] SEQ ID No:90:推定人CD8细胞表位38。
- [0198] SEQ ID No:91:推定人CD8细胞表位39。
- [0199] SEQ ID No:92:推定人CD8细胞表位40。
- [0200] SEQ ID No:93:推定人CD8细胞表位41。
- [0201] SEQ ID No:94:推定人CD8细胞表位42。
- [0202] SEQ ID No:95:推定人CD8细胞表位43。
- [0203] SEQ ID No:96:推定人CD8细胞表位44。
- [0204] SEQ ID No:97:推定人CD8细胞表位45。
- [0205] SEQ ID No:98:推定人CD8细胞表位46。
- [0206] SEQ ID No:99:推定人CD8细胞表位47。
- [0207] SEQ ID No:100:推定人CD8细胞表位48。
- [0208] SEQ ID No:101:推定人CD8细胞表位49。
- [0209] SEQ ID No:102:推定人CD8细胞表位50。
- [0210] SEQ ID No:103:推定人CD8细胞表位51。
- [0211] SEQ ID No:104:推定人CD8细胞表位52。
- [0212] SEQ ID No:105:推定人CD8细胞表位53。
- [0213] SEQ ID No:106:推定人CD8细胞表位54。
- [0214] SEQ ID No:107:推定人CD8细胞表位55。
- [0215] SEQ ID No:108:推定人CD8细胞表位56。
- [0216] SEQ ID No:109:推定人CD8细胞表位57。
- [0217] SEQ ID No:110:推定人CD8细胞表位58。
- [0218] SEQ ID No:111:推定人CD8细胞表位59。
- [0219] SEQ ID No:112:推定人CD8细胞表位60。
- [0220] SEQ ID No:113:推定人CD8细胞表位61。
- [0221] SEQ ID No:114:推定人CD8细胞表位62。
- [0222] SEQ ID No:115:推定人CD8细胞表位63。
- [0223] SEQ ID No:116:推定人CD8细胞表位64。
- [0224] SEQ ID No:117:推定人CD8细胞表位65。
- [0225] SEQ ID No:118:推定人CD8细胞表位66。
- [0226] SEQ ID No:119:推定人CD8细胞表位67。
- [0227] SEQ ID No:120:推定人CD8细胞表位68。
- [0228] SEQ ID No:121:推定人CD8细胞表位69。
- [0229] SEQ ID No:122:推定人CD8细胞表位70。
- [0230] SEQ ID No:123:推定人CD8细胞表位71。
- [0231] SEQ ID No:124:推定人CD8细胞表位72。
- [0232] SEQ ID No:125:推定人CD8细胞表位73。
- [0233] SEQ ID No:126:推定人CD8细胞表位74。
- [0234] SEQ ID No:127:推定人CD8细胞表位75。
- [0235] SEQ ID No:128:推定人CD8细胞表位76。

- [0236] SEQ ID No:129:推定人CD8细胞表位77。
- [0237] SEQ ID No:130:推定人CD8细胞表位78。
- [0238] SEQ ID No:131:推定人CD8细胞表位79。
- [0239] SEQ ID No:132:推定人CD8细胞表位80。
- [0240] SEQ ID No:133:推定人CD8细胞表位81。
- [0241] SEQ ID No:134:推定人CD8细胞表位82。
- [0242] SEQ ID No:135:推定人CD8细胞表位83。
- [0243] SEQ ID No:136:推定人CD8细胞表位84。
- [0244] SEQ ID No:137:推定人CD8细胞表位85。
- [0245] SEQ ID No:138:推定人CD8细胞表位86。
- [0246] SEQ ID No:139:推定人CD8细胞表位87。
- [0247] SEQ ID No:140:推定人CD8细胞表位88。
- [0248] SEQ ID No:141:推定人CD8细胞表位89。
- [0249] SEQ ID No:142:推定人CD8细胞表位90。
- [0250] SEQ ID No:143:推定人CD8细胞表位91。
- [0251] SEQ ID No:144:推定人CD8细胞表位92。
- [0252] SEQ ID No:145:推定人CD8细胞表位93。
- [0253] SEQ ID No:146:推定人CD8细胞表位94。
- [0254] SEQ ID No:147:推定人CD8细胞表位95。
- [0255] SEQ ID No:148:推定人CD8细胞表位96。
- [0256] SEQ ID No:149:推定人CD8细胞表位97。
- [0257] SEQ ID No:150:推定人CD8细胞表位98。
- [0258] SEQ ID No:151:推定人CD8细胞表位99。
- [0259] SEQ ID No:152:推定人CD8细胞表位100。
- [0260] SEQ ID No:153:推定人CD8细胞表位101。
- [0261] SEQ ID No:154:推定人CD8细胞表位102。
- [0262] SEQ ID No:155:来自结核分枝杆菌H37Rv菌株的Rv1753c的多肽序列。
- [0263] SEQ ID No:156:来自结核分枝杆菌H37Rv菌株的Rv2707c的多肽序列。

具体实施方式

[0264] 目前,用活菌接种是诱导保护性免疫最有效的方法。用于该目的的最常用的分枝杆菌是卡介杆菌 (BCG),这是60多年前开发的牛分枝杆菌的无毒性菌株。然而,BCG的安全性和有效性成为争论的来源-尽管在儿童中显示了对严重疾病的保护,BCG并不能预防成年人中潜伏TB的形成或者肺部疾病的再活化。此外,在一些国家,例如美国,并未用这种药剂来给普通人群接种。

[0265] 几乎所有目前在临床开发中的新一代TB疫苗均被设计为暴露前 (pre-exposure) 疫苗。其中包括了亚基疫苗,其在加强由先前的BCG接种诱导的免疫中特别有效,还包括高级活分枝杆菌疫苗,其用于将BCG替换为更有效和/或更安全的菌株。尽管这些疫苗的目的在于改进对感染的耐受性,它们在潜伏TB病例中作为暴露后或治疗性疫苗的效力可能相对

较低 (Lin MY等人Endocrine, Metabolic & Immune Disorders-Drug Targets 2008;15-29)。

[0266] 这些蛋白中有多种在分枝杆菌感染早期被强烈表达, 据显示它们可在动物接种模型中提供很强的保护效力。然而, 用感染早期高度表达的抗原接种可能无法提供处理感染后期的最佳免疫应答。对感染后期的充分控制可能需要T细胞, 其特异性针对在该时期表达的特定抗原。

[0267] 直接靶向持续休眠菌的暴露后疫苗可能有助于针对TB再活化的保护, 从而强化TB控制, 或者甚至能清除感染。因此, 靶向潜伏TB的疫苗可以显著和经济地降低全球TB感染率。

[0268] 基于后期抗原的亚基疫苗还可与早期抗原联合使用以提供多阶段疫苗。替代性地, 后期抗原可用于补充和改善BCG接种 (通过促进BCG应答或者通过形成高级重组BCG菌株)。

[0269] Rv2386c, 也称为MbtI, 可通过将分支酸盐转化为水杨酸盐催化分枝杆菌素生物合成中的初始转化 (Zwahlen等人Biochemistry 2007 46:954-964)。

[0270] Rv2386之前已被鉴定为与应激状态下的表达相关, 尽管在豚鼠模型中作为DNA疫苗施用时尚未形成保护 (Vipond等人Vaccine 2006 24:6340-6350)。

[0271] 最近, 基于对结核分枝杆菌基因组的全基因组生物信息学分析 (Zvi等人BMC Medical Genetics 2008 1:18) 和对活动性和潜伏感染个体中差异表达蛋白的测试 (Schuck SD等人PLoS ONE 2009 4(5):e5590) 已经提出了一些结核分枝杆菌疫苗候选物。

[0272] 尽管已显示巨噬细胞是分枝杆菌免疫性的首要效应物, T细胞是该种免疫性的主要诱导物。T细胞在针对结核病的保护中的关键作用解释为人免疫缺陷病毒感染个体中 (由于CD4+T细胞的相关耗竭) 增加的TB再活化速率。此外, 在针对结核分枝杆菌初次免疫应答的顶点进行的CD4+T细胞的转移已显示赋予T细胞缺陷型小鼠中针对结核分枝杆菌的保护 (Orme等人J. Exp. Med. 1983 158:74-83)。

[0273] 分枝杆菌活性CD4+T细胞据显示是 γ -干扰素 (IFN- γ) 的有效生产者, 这进而显示触发在小鼠中巨噬细胞的抗分枝杆菌作用 (Flynn等人J. Exp. Med. 1993 178:2249-2254)。尽管对IFN- γ 在人体内的作用所知更少, 研究已显示, 1, 25-二羟基-维生素D3单独或者与IFN- γ 或肿瘤坏死因子- α 组合激活人巨噬细胞以抑制结核分枝杆菌感染。此外, 已知IFN- γ 刺激人巨噬细胞来制备1, 25-二羟基-维生素D3。类似地, 已显示白介素-12 (IL-12) 在刺激针对结核分枝杆菌感染的耐受性上起作用。有关结核分枝杆菌感染免疫学的综述可参见Chan & Kaufmann, tuberculosis: Pathogenesis, Protection and Control (Bloom编辑, 1994), tuberculosis (第2版, Rom和Garay, 编辑2003), 和Harrison's Principles of Internal Medicine, 第150章, pp 953-966 (第16版, Braunwald, 等人, 编辑2005)。

[0274] 本发明总体涉及作为TB抗原 (特别是与潜伏TB相关的抗原) 的Rv2386c的鉴定, 并涉及预防和治疗TB (尤其是预防和/或治疗潜伏TB以及预防或延缓TB再活化) 的相关方法和用途。

[0275] 因此, 本发明提供了Rv2386c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, 以用于治疗或预防TB。适当地, 该用途可特别针对潜伏TB的预防和治疗 (尤其是潜伏TB的治疗)。替代性地, 该用途可针对TB再活化的预防或延缓 (特别是TB再

活化的延缓,例如延缓数月、数年或甚至无限期)。

[0276] 术语“结核复合物(complex)的分枝杆菌种类”包括通常被认为引起结核病疾病的种类,以及引起免疫缺陷患者(例如患有AIDS的患者)的结核病和肺部疾病的分枝杆菌环境和机会种,例如结核分枝杆菌(*M.tuberculosis*)、牛分枝杆菌(*M.bovis*)或非洲分枝杆菌(*M.africanum*)、BCG、鸟结核分枝杆菌(*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*)、牝牛分枝杆菌(*M.vaccae*)、偶发分枝杆菌(*M.fortuitum*)以及瘰癧分枝杆菌(*M.scrofulaceum*) (参见,例如,Harrison's Principles of Internal Medicine,第150章,pp 953-966 (第16版,Braunwald,等人,编辑,2005)。本发明特别涉及结核分枝杆菌的感染。

[0277] 术语“活动性感染”指具有已显现的疾病症状和/或损伤(合适地具有显现的疾病症状)的感染(例如,结核分枝杆菌感染)。

[0278] 术语“非活动性感染”、“潜伏感染”或“潜伏性感染”指未显示疾病症状和/或损伤(合适地未显示疾病症状)的感染(例如,结核分枝杆菌感染)。

[0279] 术语“原发结核病”指在感染(例如,结核分枝杆菌感染)后直接形成的临床疾病(疾病症状的显现)。参见Harrison's Principles of Internal Medicine,第150章,pp 953-966 (第16版,Braunwald,等人,编辑,2005)。

[0280] 术语“继发性结核病”或“原发后结核病”指潜伏、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's Principles of Internal Medicine,第150章,pp 953-966 (第16版,Braunwald,等人,编辑,2005)。

[0281] 术语“结核病再活化”指感染测试呈阳性(例如,在结核菌素皮试中呈阳性,合适地在体外基于T细胞的测定中呈阳性)但没有明显的疾病症状的个体随后显现疾病症状。该阳性诊断测试表明该个体被感染,然而,该个体可能曾显示或未曾显示已进行过充分治疗以将结核病带入非活动性或潜伏状态的活动的疾病的症状。应当认识到,预防、延缓或治疗结核病再活化的方法可在显现活动性疾病症状的个体内启动。

[0282] 术语“耐药性”结核病指感染(例如,被结核分枝杆菌感染),其中该感染菌株不被一种或多种能够有效治疗结核病的所谓“一线”化疗剂(例如,异烟肼、利福平、乙胺丁醇、链霉素和吡嗪酰胺)所压制或杀死(即,耐受)。

[0283] 术语“多重耐药性”结核病指耐受两种或更多种能够有效治疗结核病的“一线”化疗剂的感染菌株的感染(例如,被结核分枝杆菌感染)。

[0284] “化疗剂”指本领域已知并用于治疗结核病(例如,被结核分枝杆菌感染)的药剂。用于治疗结核病的示范性的药剂包括,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、利福霉素类(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、阿齐霉素和氟喹诺酮类。用于治疗非耐药性结核病的“一线”或“前线”化疗剂包括异烟肼、利福平、乙胺丁醇、链霉素和吡嗪酰胺。用于治疗对一种或多种“一线”药物显示了耐药性的结核病的“二线”化疗剂包括氧氟沙星、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。参见Goodman and Gilman's The Pharmacological Basis of Therapeutics,Hardman and Limbird编辑,2001第48章中综述的这些药剂。

[0285] 术语“多肽”、“肽”和“蛋白”在此可互换使用以表示氨基酸残基的聚合物。该术语还可用于其中一个或多个氨基酸残基为相应天然存在氨基酸的人工化学模拟物的氨基酸聚合物,还可用于天然存在氨基酸聚合物和非天然存在氨基酸聚合物。合适地,根据本发明的多肽将仅由天然存在氨基酸残基(特别是由遗传密码编码的那些氨基酸)组成。

[0286] 术语“氨基酸”指天然存在和合成的氨基酸,以及以类似于天然存在氨基酸的方式发挥功能的氨基酸类似物和氨基酸模拟物。天然存在的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。氨基酸类似物指与天然存在的氨基酸具有相同基本化学结构的化合物,即,与氢连接的 α 碳,羧基基团,氨基基团以及R基团,例如,高丝氨酸、正亮氨酸、蛋氨酸亚砷、蛋氨酸甲基砷。该种类类似物具有修饰的R基团(例如,正亮氨酸)或修饰的肽骨架,但保留了与天然存在的氨基酸相同的基本化学结构。氨基酸模拟物指具有与氨基酸的一般化学结构不同的结构,但功能上与天然存在的氨基酸类似的化合物。合适地,氨基酸为天然存在的氨基酸或氨基酸类似物,特别是天然存在的氨基酸,尤其是由遗传密码编码的氨基酸。

[0287] “核酸”指脱氧核糖核苷酸或核糖核苷酸及其单链或双链形式的聚合物。该术语包括含有合成、天然存在的和非天然存在的、与参考核酸具有类似结合特性的且与参考核苷酸以类似方式代谢的已知核苷酸类似物或修饰的骨架残基或连锁的核酸。该种类类似物的范例包括,但不限于,硫代磷酸酯、氨基磷酸酯、膦酸甲酯、手性-膦酸甲酯、2-O-核糖核苷酸甲酯、肽核酸(PNAs)。合适地,术语“核酸”指天然存在的脱氧核糖核苷酸或核糖核苷酸及其聚合物。

[0288] 除非另行指明,特定的核酸序列还意指包含其保守修饰的变体(例如,简并密码子置换)和互补序列,以及明确指出的序列(合适地,它指明示指出的序列)。特别地,简并密码子置换可通过生成其中一个或多个选定的(或者所有的)密码子的第三位置被混合碱基和/或脱氧肌苷残基所置换的序列来实现(Batzer等人,Nucleic Acid Res.19:5081(1991); Ohtsuka等人,J.Biol.Chem.260:2605-2608(1985);Rossolini等人,Mol.Cell.Probes 8:91-98(1994))。术语核酸可与基因、cDNA、mRNA、寡核苷酸和多核苷酸互换使用。

[0289] 本文的氨基酸可用常见的三个字母的符号表示或者可通过IUPAC-IUB生物化学命名委员会推荐的单字母符号表示。同样地,核苷酸也可采用其通常接受的单字母密码来表示。

[0290] 本文所用的术语“Rv2386c蛋白序列”指SEQ ID No:1中提供的多肽序列或其同源物,其来自结核复合物的分枝杆菌种类,例如,结核分枝杆菌、牛分枝杆菌或非洲分枝杆菌等种类,或者环境性和机会性的并引起免疫缺陷宿主(例如,患有AIDS的患者)的机会性感染如肺部感染的分枝杆菌种类,例如,BCG,鸟结核分枝杆菌、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、牝牛分枝杆菌、偶发分枝杆菌和瘰癧分枝杆菌(参见,例如,Harrison's Principles of Internal Medicine,第150章,pp 953-966,第16版,Braunwald,等人,编辑,2005)。

[0291] 为了确保在接种宿主中的高效力比例,疫苗的成分应当在具有临床意义的菌株中良好保藏。合适地,Rv2386c蛋白来自于结核分枝杆菌H37Rv(即,SEQ ID No:1中提供的多肽序列)或者是来自其它结核分枝杆菌菌株(例如CDC1551、F11、Haarlem A和C菌株)的同源物。与耐药性相关的结核分枝杆菌菌株(例如,MDR或特别是XDR)是Rv2386c蛋白序列的特别

有价值的基础。感兴趣的菌株包括：

[0292] CDC1551-可传递且强毒性菌株

[0293] Haarlem家族(例如Haarlem A)-在拥挤人群中发现的耐药性菌株。结核分枝杆菌菌株中的Haarlem家族成员已在世界各地被发现。该家族的第一个代表成员发现于荷兰的Haarlem。

[0294] KZN4207-从南非KwaZulu-Natal的患者得到的药物敏感分离物。

[0295] KZN1435-从南非KwaZulu-Natal的患者得到的多重耐药(MDR)分离物。

[0296] KZN605-从南非KwaZulu-Natal的患者得到的广泛耐药(XDR)分离物。

[0297] C-在纽约市高度传递。在一项研究中发现该菌株在注射药物使用者中更为常见，且耐受活性氮中间体(Friedman等人J.Infect.Dis.1997 176(2):478-84)。

[0298] 94M4241A-1994年在旧金山从一名出生于中国的患者分离得到。该菌株曾通过基因缺失分析进行分析(Gagneux等人,PNAS 2006 103(8):2869-2873)。

[0299] 02_1987-2002年在旧金山从一名出生于南韩的患者分离得到。该菌株曾通过基因缺失分析进行分析(Gagneux等人,PNAS 2006 103(8):2869-2873)。

[0300] T92-1999年在旧金山从一名出生于菲律宾的患者分离得到。该菌株发表于Hirsh等人PNAS 2004 101:4871-4876。

[0301] T85-1998年在旧金山从一名出生于中国的患者分离得到。该菌株发表于Hirsh等人PNAS 2004 101:4871-4876。

[0302] EAS054-1993年在旧金山从一名出生于印度的患者分离得到。该菌株曾通过基因缺失分析进行分析(Gagneux等人,PNAS 2006 103(8):2869-2873)。

[0303] Gagneux等人,PNAS 2006 103(8):2869-2873和Herbert等人Infect.Immun.2007 75(12):5798-5805为已知存在的结核分枝杆菌菌株范围提供了有价值的背景知识。

[0304] 最适宜地,Rv2386c蛋白选自SEQ ID No:1和3-7、特别是SEQ ID No:1和3-6(例如SEQ ID No:1)中提供的多肽序列。

[0305] 特别感兴趣的多核苷酸是包含(例如由其组成)编码如下的序列的那些多核苷酸：

[0306] (i) Rv2386c蛋白序列；

[0307] (ii) Rv2386c蛋白序列的变体；或者

[0308] (iii) Rv2386c蛋白序列的免疫原性片段。

[0309] 多核苷酸将合适地包含(例如由其组成)编码Rv2386c蛋白的免疫原性片段的SEQ ID NO:2的变体或SEQ ID NO:2的片段。

[0310] 组合

[0311] 本发明的Rv2386c相关的多肽可进一步包含设计用于增强其免疫原性或在其它方面改善这些抗原的其它成分。例如，可通过在抗原的一端添加一段组氨酸残基(通常称为his-tag)来促进多肽抗原的改善分离。

[0312] 术语“his-tag”指参考序列中插入的一串组氨酸残基，通常是六个残基。为了最小化对与参考序列相关的活性的破坏，his-tag一般被插在N-末端，通常紧邻起始蛋氨酸残基，或者插在C-末端。它们对于天然序列通常是异源的，但由于它们可通过改善与固定化金属亲和层析树脂(IMAC)的蛋白结合促进分离而被整合入天然序列。通常而言，从引发针对参考蛋白的所需免疫应答而言his-tag的存在与否并不重要。然而，为了避免针对his-tag

本身的不良反应的风险,据认为最好最小化his-tag的长度,例如,降低至四个或更少的残基,特别是两个残基(或者完全不用his-tag)。

[0313] 为了提高引发的免疫应答的强度和/或幅度,本发明的组合物、多肽和核酸可包含来自分枝杆菌种类(特别是结核分枝杆菌)的本发明抗原和/或另外的异源多肽(或编码它们的多核苷酸)的多个拷贝。

[0314] 本领域技术人员将认识到,当一些成分组合使用时,精确的存在可发生变化。例如,Rv2386c成分和本发明的抗原或另外的异源抗原成分的另外拷贝可以如下方式存在:

[0315] (1) 两种单独的多肽成分;

[0316] (2) 包含两种多肽成分的融合蛋白;

[0317] (3) 一种多肽和一种多核苷酸成分;

[0318] (4) 两种单独的多核苷酸成分;

[0319] (5) 编码两种单独多肽成分的单个多核苷酸;或者

[0320] (6) 编码包含两种多肽成分的融合蛋白的单个多核苷酸。

[0321] 这种灵活性可同样应用于三种或更多种成分组合使用的情况。然而,为方便起见,通常期望当存在多种成分时,它们包含于单个融合蛋白中或编码单个融合蛋白的多核苷酸中。在本发明的一个实施方式中,所有的抗原成分以多肽提供(例如,在单个融合蛋白中)。在本发明的一个替代性的实施方式中,所有抗原成分以多核苷酸提供(例如,单个多核苷酸,例如编码单个融合蛋白的多核苷酸)。

[0322] 在用于表示核酸的部分时,术语“异源的”表示该核酸包含两种或更多种被发现性质上关系彼此不相同的子序列。例如,该核酸通常重组生成,具有两个或更多个来自无关基因的序列并排列为生成新的功能核酸,例如,来自一个来源的启动子和来自另一来源的编码区。类似地,异源蛋白表示该蛋白包含两种或更多种被发现性质上关系彼此不相同的子序列(例如,融合蛋白)。

[0323] “融合多肽”或“融合蛋白”指具有直接或通过氨基酸接头共价连接的至少两种异源多肽(例如,至少两种分枝杆菌属多肽)的蛋白。形成融合蛋白的多肽通常将C-末端连接至N-末端,尽管它们也可将C-末端连接至C-末端,N-末端连接至N-末端,或者将N-末端连接至C-末端。融合蛋白的多肽可以呈任意顺序。该术语还指构成该融合蛋白的抗原的保守修饰的变体、多形态变体、等位基因、突变体、免疫原性片段和种间同源物。结合分枝杆菌抗原描述于Cole等人,Nature 393:537 (1998),其披露了整个结核分枝杆菌基因组。对应于结核分枝杆菌抗原的来自其它分枝杆菌种类的抗原可通过例如采用本文所述的序列比较算法或者本领域技术人员已知的其它方法(例如,杂交测定和抗体结合测定)进行鉴定。

[0324] 术语“融合”指融合蛋白中两个多肽之间的共价键。多肽通常通过肽键彼此直接连接或经氨基酸接头连接。任选地,肽可通过本领域技术人员已知的非肽共价键连接。

[0325] 可与Rv2386c结合的示范性结核分枝杆菌抗原包括如下(例如(i)至(xii)的一种或多种)的一种或多种(例如,1-5种,如1-3种,特别是1种):

[0326] (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以SEQ ID No:101显示)和Coler等人Journal of Immunology 1998161:2356-2364中。特别感兴趣的是成熟Mtb8.4序列,其缺失前导信号肽(即,W097/09428的SEQ ID No:102的氨基酸残基15-96)。Mtb8.4的全长多肽序列以SEQ ID No:8显示;

[0327] (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的片段披露于W098/53075的SEQ ID No: 110-124, 特别感兴趣的是SEQ ID No: 119和120) 以及Coler等人Vaccine 200927:223-233 (特别是其中图2所示的活性片段) 中。Mtb9.8的全长多肽序列以SEQ ID No: 9显示;

[0328] (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等人Journal of Experimental Medicine 2000 7:551-559 (MTI的片段披露于W098/53075的SEQ ID No: 17和51-66, 特别感兴趣的是SEQ ID No: 17、51、52、53、56和62-65)。一些MTI多肽变体描述于W098/53075的SEQ ID No: 21、23、25、27、29和31以及Alderson等人Journal of Experimental Medicine 2000 7:551-559中。Mtb9.9的全长多肽序列以SEQ ID No: 10显示;

[0329] (iv) Ra12 (也称为Mtb32A C-末端抗原), 其多肽序列描述于W001/98460的SEQ ID No: 10和Skeiky等人Journal of Immunology 2004 172:7618-7682中。Ra12的全长多肽序列以SEQ ID No: 11显示;

[0330] (v) Ra35 (也称为Mtb32A N-末端抗原), 其多肽序列描述于W001/98460的SEQ ID No: 8和Skeiky等人Journal of Immunology 2004 172:7618-7682中。Ra35的全长多肽序列以SEQ ID No: 12显示;

[0331] (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等人Infection and Immunity 1999 67 (6):2941-2950和Skeiky等人Journal of Immunology 2004 172:7618-7682中。TbH9的全长多肽序列以SEQ ID No: 13显示;

[0332] (vii) Mtb40 (也称为HTCC1和Rv3616c), 其多肽序列描述于W098/53075的SEQ ID No: 138 (cDNA描述于SEQ ID No: 137)。Mtb40的全长多肽序列以SEQ ID No: 14显示;

[0333] (viii) 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) 和Skeiky等人Journal of Immunology 2000 165:7140-7149中。Mtb41的全长多肽序列以SEQ ID No: 15显示;

[0334] (ix) ESAT-6 (也称为esxA和Rv3875), 其多肽序列描述于W097/09428的SEQ ID No: 103 (cDNA以SEQ ID No: 104描述) 和Sorensen等人Infection and Immunity 1995 63 (5):1710-1717中。ESAT-6的全长多肽序列以SEQ ID No: 16显示;

[0335] (x) Ag85复合抗原 (例如Ag85A, 也称为fbpA和Rv3804c; 或者Ag85B, 也称为fbpB和Rv1886c), 其在例如, Content等人Infection and Immunity 1991 59:3205-3212和Huygen等人Nature Medicine 1996 2 (8):893-898中讨论。Ag85A的全长多肽序列以SEQ ID No: 17显示 (特别感兴趣的是残基43-338的成熟蛋白, 即, 缺失信号肽)。Ag85B的全长多肽序列以SEQ ID No: 18显示 (特别感兴趣的是残基41-325的成熟蛋白, 即, 缺失信号肽);

[0336] (xi) Alpha-晶体蛋白 (也称为hspX和Rv2031c), 其描述于Verbon等人Journal of Bacteriology 1992 174:1352-1359和Frischia等人Clinical and Experimental Immunology 1995 102:53-57 (特别感兴趣的是对应于残基71-91、21-40、91-110和111-130的片段)。alpha-晶体蛋白的全长多肽序列以SEQ ID No: 19显示;

[0337] (xii) Mpt64 (也称为Rv1980c), 其描述于Roche等人Scandinavian Journal of Immunology 1996 43:662-670。MPT64的全长多肽序列以SEQ ID No: 20显示 (特别感兴趣的

是残基24-228的成熟蛋白,即,缺失信号肽);

[0338] (xiii) Mtb32A,其多肽序列描述于W001/98460的SEQ ID No:2(全长)和SEQ ID No:4的残基8-330(成熟),特别是具有催化三联体突变的至少一个的变体(例如,催化丝氨酸残基,其可例如突变为丙氨酸)。Mtb32A的全长多肽序列以SEQ ID No:21显示。具有Ser/Ala突变的Mtb32A的成熟形式以SEQ ID No:22显示;

[0339] (xiv) TB10.4, TB10.4的全长多肽序列以SEQ ID No:23显示;

[0340] (xv) Rv1753c,来自于结核分枝杆菌H37Rv的Rv1753c的全长多肽序列以SEQ ID No:155显示;和/或

[0341] (xv) Rv2707c,来自于结核分枝杆菌H37Rv的Rv2707c的全长多肽序列以SEQ ID No:156显示。

[0342] 或其组合,例如(如(a)至(g)的组合):

[0343] (a) Ra12、TbH9和Ra35成分的组合,例如呈融合蛋白形式,如Mtb72f。Mtb72f的多肽序列描述于W02006/117240的SEQ ID No:6(cDNA以SEQ ID No:5描述)以及Skeiky等人Journal of Immunology 2004 172:7618-7682(其中结合了任选的His-tag来帮助纯化,当应用于本发明中时,合适地,Mtb72f不具有该任选的组氨酸残基)。Mtb72f的多肽序列以SEQ ID No:24显示;

[0344] (b) Ra12、TbH9和Ser/Ala突变的Ra35(即,催化丝氨酸残基被丙氨酸替换)成分的组合,例如呈融合蛋白形式,如M72。M72的多肽序列描述于W02006/117240的SEQ ID No:4(cDNA以SEQ ID No:3描述),其中结合了任选的双组氨酸来帮助生产,当应用于本发明中时,M72也可结合双组氨酸,但合适地,M72不含该任选的双组氨酸(即,特别感兴趣的是W02006/117240的SEQ ID No:4的残基4-725)。M72的多肽序列以SEQ ID No:25显示;

[0345] (c) Mtb8.4、Mtb9.8、Mtb9.9和Mtb41成分的组合,例如呈融合蛋白形式,如Mtb71f。Mtb71f的多肽序列描述于W099/051748的SEQ ID No:16(cDNA以SEQ ID No:15描述),其中结合了任选的His-tag来帮助纯化,当应用于本发明中时,合适地,Mtb71f对应于来自W099/051748的SEQ ID No:16的氨基酸残基9-710。Mtb71f的多肽序列以SEQ ID No:26显示;

[0346] (d) Mtb72f或M72(合适地不具有任选的帮助表达的组氨酸残基)与Mtb9.8和Mtb9.9的组合,例如呈融合蛋白形式。M72-Mtb9.9-Mtb9.8融合体的多肽序列以SEQ ID No:27(M92融合)显示,在用于本发明时,该M72-Mtb9.9-Mtb9.8融合体可任选地在起始蛋氨酸残基后结合双组氨酸以帮助生产。

[0347] (e) Mtb72f或M72(合适地不具有任选的帮助表达的组氨酸残基)与Ag85B的组合,例如呈融合蛋白形式,如Mtb103f。Mtb103f的多肽序列描述于W003/070187的SEQ ID No:18(cDNA以SEQ ID No:10描述),其结合了任选的His-tag以帮助纯化,当用于本发明时,合适地,Mtb103f对应于W003/070187的SEQ ID No:18的氨基酸残基8-1016。还特别感兴趣的是M103,即Mtb103f结合了Ra35成分中的Ser/Ala突变,当用于本发明时,合适地,M103对应于W003/070187的SEQ ID No:18的氨基酸残基8-1016,其中在第710位的Ser残基被替换为Ala。M103的多肽序列以SEQ ID No:28显示,当用于本发明时,该M72-Mtb9.9-Mtb9.8融合体可任选地在起始蛋氨酸残基后结合双组氨酸以帮助生产;

[0348] (f) Mtb72f或M72(合适地不具有任选的帮助表达的组氨酸残基)与Mtb41的组合,例如呈融合蛋白形式,如Mtb114f。Mtb114f的多肽序列描述于W003/070187的SEQ ID No:16

(cDNA描述于SEQ ID No:9),其结合了任选的His-tag以帮助纯化,当用于本发明时,合适地,Mtb114f对应于W003/070187的SEQ ID No:16的氨基酸残基8-1154。还特别感兴趣的是M114,即Mtb114f结合了Ra35成分中的Ser/Ala突变,当用于本发明时,合适地,M114对应于W003/070187的SEQ ID No:16的氨基酸残基8-1154,其中在第710位的Ser残基被替换为Ala。M114的多肽序列以SEQ ID No:29显示,当用于本发明时,该M72-Mtb9.9-Mtb9.8融合体可任选地在起始蛋氨酸残基后结合双组氨酸以帮助生产;

[0349] (g) Ag85B和ESAT-6成分的组合,例如在Doherty等人Journal of Infectious Diseases 2004 190:2146-2153所述的融合体中;和/或

[0350] (h) Ag85B和TB10.4成分的组合,例如在Dietrich等人Journal of Immunology 2005 174(10):6332-6339;190:2146-2153所述的融合体中。

[0351] 特别感兴趣的是Rv2386c成分和Mtb40成分的组合。显然,该种组合可任选地包含其它附加的抗原成分(例如,M72成分)。

[0352] 另一感兴趣的组合包含Rv2386c成分和M72成分。

[0353] 另一感兴趣的组合包含Rv2386c成分和Rv1753c成分。

[0354] 其它感兴趣的组合包括那些包含Rv2386c成分和Rv2707c成分的组合。

[0355] 另外的感兴趣的组合包含Rv2386c成分和alpha-晶体蛋白成分。

[0356] 本领域技术人员将认识到,这些组合并不依赖于上文(i)-(xvi)和(a)-(h)中所述的特定序列,且所述序列的保守修饰的变体(例如,具有至少70%一致性,例如至少80%一致性,特别是至少90%一致性,尤其至少95%一致性)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%,特别是至少70%,尤其是至少80%)可用于实现相同的实际效果。

[0357] 上述单独抗原序列的每一个还披露于Cole等人Nature 1998393:537-544和Camus Microbiology 2002 148:2967-2973。公众已可获得结核分枝杆菌H37Rv的基因组,例如可见Wellcome Trust Sanger Institute的网站(www.sanger.ac.uk/Projects/M_tuberculosis/)以及其它途径。

[0358] 上述抗原中的多个还披露于美国专利申请号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以及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,其每一个在此引入作为参考。

[0359] 本发明的组合物、多肽和核酸还可包含来自其它来源的另外的多肽。例如,本发明的组合物和融合蛋白可包含多肽或编码多肽的核酸,其中该多肽增强抗原的表达,例如,NS1,流感病毒蛋白(参见例如W099/40188和W093/04175)。本发明的核酸可根据所选种类,例如人(当进行体内表达时)或特定细菌(当进行多肽生产时)的密码子偏好进行工程改造。

[0360] Rv2386c成分还可与一种或多种抗结核病(例如,结核分枝杆菌感染)有效的化疗剂施用。这些化疗剂的范例包括,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、利福霉素类(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、阿齐霉素和氟喹诺酮类。该种化疗可由主

治医生采用优选药物组合的判断来确定。用于治疗非耐药性结核病(例如,结核分枝杆菌感染)的“一线”化疗剂包括异烟肼、利福平、乙胺丁醇、链霉素和吡嗪酰胺。用于治疗已证明对一种或多种“一线”药物有耐药性的结核病(例如,结核分枝杆菌感染)的“二线”化疗剂包括氧氟沙星、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。

[0361] 传统的化疗剂通常在相对较长的周期内给药(约9个月)。传统化疗剂与根据本发明的Rv2386c成分给药的组合可缩短化疗治疗周期(例如,缩短至8个月、7个月、6个月、5个月、4个月、3个月或更短)而不降低效力。

[0362] 特别感兴趣的是Rv2386c成分与卡介杆菌(BCG)的联用。例如,以重组表达Rv2386c(或其如本文所述的变体或片段)的修饰BCG的形式。替代性地,该Rv2386c成分可用于通过同时给药或通过促进之前的BCG免疫来增强对象对于BCG免疫的应答。当用于增强对象对BCG接种的应答时,该Rv2386c成分可显然地以多肽或多核苷酸(任选地结合上述另外的抗原成分)的形式提供。

[0363] 本领域技术人员将认识到,成分的组合不必同时给药,并可按如下方式应用:单独或组合;同时、相继或在短时间内;通过相同或不同的途径。然而,为方便起见,通常期望将成分的组合作为单种组合物给药(当给药方案相容时)。

[0364] 本发明的多肽、多核苷酸和组合物通常向人给药,但对包括家养哺乳动物(例如,狗、猫、兔、大鼠、小鼠、豚鼠、仓鼠、南美栗鼠)和农用哺乳动物(例如,牛、猪、绵羊、山羊、马)等其它哺乳动物也有效。

[0365] 免疫原性片段

[0366] T细胞表位是被T细胞(例如,CD4+或CD8+T细胞)识别的氨基酸的短连续片段。T细胞表位的鉴定可通过本领域技术人员公知的表位定位实验实现(参见,例如Paul, *Fundamental Immunology*, 第3版, 243-247 (1993); Beißbarth等人 *Bioinformatics* 2005 21 (Suppl.1): i29-i37)。

[0367] 替代性地,可采用实施例所述的方法预测表位。

[0368] 由于结核病中T细胞应答的关键参与,显然包含至少一个T细胞表位的全长Rv2386c多肽的片段将具有免疫原性并有助于免疫保护。该种片段在此称为免疫原性片段。

[0369] 根据本发明的免疫原性片段将通常包含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或至少20个连续氨基酸),特别是至少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)。合适地,该免疫原性片段将是该全长多肽序列长度的至少20%,例如至少50%、至少70%或至少80%。

[0370] 应当理解,在不同的远系交配群体(例如人)中,不同的HLA类型表示该特定表位可能不被该群体所有成员所识别。因此,为了最大化识别的水平以及对多肽的免疫应答的强度,通常期望该免疫原性片段包含来自该全长序列的多个表位(适当时为全部表位)。

[0371] 可以使用的Rv2386c蛋白的具体片段包括包含至少一个CD4+表位、适当地包含至少两个CD4+表位、特别是包含所有CD4+表位(例如实施例和SEQ ID No.: 30-52中所述的那些表位,特别是与多个HLA等位基因相关的表位,例如与2、3、4、5或更多个等位基因相关的表位)的那些Rv2386c蛋白。

[0372] 可以使用的Rv2386c蛋白的其它片段包括包含至少一个CD8表位、适当地包含至少两个CD8表位、特别是包含所有CD8表位(例如实施例和SEQ ID No.: 53-154中所述的那些表

位,特别是与多个HLA等位基因相关的表位,例如与2、3、4、5或更多个等位基因相关的表位)的那些Rv2386c蛋白。

[0373] 当使用全长多肽的单独片段时,该种片段在其引发的应答达到参考序列在PBMC或全血对特定抗原的体外再刺激测定(例如,在数小时至长达两周之间,例如长达一天、1天至1周或者1至2周的时间内再刺激)中活性的至少20%、合适地至少50%且特别是至少75%(例如至少90%)时被认为是免疫原性的,其中该测定通过培养悬浮液中淋巴组织增殖、细胞因子生成(通过ELISA、CBA等测量)或者通过胞内和胞外染色(例如,采用特异性针对免疫标记物的抗体,例如CD3、CD4、CD8、IL2、TNF α 、IFN γ 、CD40L、CD69等)然后以流式细胞仪分析T和B细胞应答特征来测量细胞的活化。合适地,当片段引发参考序列在T细胞增殖和/或IFN- γ 生成测定中的活性的至少20%、合适地至少50%且特别是至少75%(例如至少90%)的应答时被认为具有免疫原性。

[0374] 在一些情况下,全长多肽的多个片段(其可与全长序列重叠或不重叠并且覆盖或不覆盖该全长序列的全部)可被用于获得与全长序列本身同等的生物应答。例如,至少两个(例如三个、四个或五个)上述免疫原性片段组合提供的活性是该参考序列在PBMC或全血的体外再刺激测定(例如,T细胞增殖和/或IFN- γ 生成测定)中活性的至少50%、合适地至少75%且尤其是至少90%。

[0375] 变体

[0376] “变体”或“保守修饰变体”可同时用于氨基酸和核酸序列。对于特定的核酸序列,保守修饰的变体指编码相同或基本相同的氨基酸序列的那些核酸,或者当该核酸不编码氨基酸序列时,指基本相同的序列。

[0377] 由于遗传密码的简并,大量功能相同的核酸编码任意给定蛋白。例如,密码子GCA、GCC、GCG和GCU均编码氨基酸丙氨酸。因此,在丙氨酸由密码子限定的每个位置,该密码子可被改为所述的任意相应的密码子而不改变编码的多肽。该种核酸变体可导致“沉默”或“简并”变体,其为保守修饰变异的一种。本文编码多肽的每个核酸序列也描述了该核酸每个可能的沉默变异。本领域技术人员将认识到核酸中的每个密码子(除了AUG,其通常是蛋氨酸的唯一密码子,以及TGG,其通常是色氨酸的唯一密码子)可被修饰以得到功能相同的分子。因此,每个所述的序列中暗示了编码多肽的核酸的每个沉默变异。

[0378] 本发明的多核苷酸与参考序列相比可包含多个沉默变异(例如,1-50个,例如1-25个,特别是1-5个,尤其是1个密码子被更改)。本发明的多核苷酸与参考序列相比可包含多个非沉默保守变异(例如,1-50个,例如1-25个,特别是1-5个,尤其是1个密码子被更改)。非沉默变异是那些可导致编码氨基酸序列的变化(通过氨基酸残基的置换、缺失或添加)的变异。本领域技术人员将认识到,特定的多核苷酸序列可同时包含沉默和非沉默保守变异。

[0379] 对于蛋白序列的变体,本领域技术人员将认识到改变、添加或缺失单个氨基酸或少量百分比的氨基酸的多肽的单独置换、缺失或添加是“保守修饰变体”,其中该改变导致以功能类似的氨基酸置换氨基酸或者残基的置换/缺失添加基本不影响该变体的生物功能。

[0380] 提供功能类似的氨基酸的保守置换表是本领域公知的。该种保守修饰的变体附加于且不排除本发明多形态变体、种间同源物以及等位基因。

[0381] 本发明的多肽与参考序列相比可包含多个保守置换(例如,1-50个,例如1-25个,

特别是1-10个,尤其是1个氨基酸残基被更改)。总体而言,该种保守置换将落入下文限定的氨基酸分组之一,尽管在一些情况下,也可能有其它置换而可能基本不影响该抗原的免疫原性。以下八组各包含了彼此可进行典型保守置换的氨基酸:

- [0382] 1) 丙氨酸(A),甘氨酸(G);
- [0383] 2) 天冬氨酸(D),谷氨酸(E);
- [0384] 3) 天冬酰胺(N),谷氨酰胺(Q);
- [0385] 4) 精氨酸(R),赖氨酸(K);
- [0386] 5) 异亮氨酸(I),亮氨酸(L),蛋氨酸(M),缬氨酸(V);
- [0387] 6) 苯丙氨酸(F),酪氨酸(Y),色氨酸(W);
- [0388] 7) 丝氨酸(S),苏氨酸(T);以及
- [0389] 8) 半胱氨酸(C),蛋氨酸(M)
- [0390] (参见,例如,Creighton,Proteins 1984)。

[0391] 合适地,该种置换不发生在表位区域,因此对该抗原的免疫原性不具有明显影响。

[0392] 蛋白变体还可包括相对于参考序列其中插入了另外氨基酸的那些蛋白变体,例如,该种插入可发生在1-10个位置(例如1-5个位置,合适地1或2个位置,特别是1个位置),并可以,例如,在每个位置添加50或更少的氨基酸(例如20个或更少,特别是10个或更少,尤其是5个或更少)。合适地,该种插入不发生在表位区域,因此对该抗原的免疫原性不具有明显影响。插入的一个实施例包括一段短的组氨酸残基(例如,2-6个残基)以帮助目的抗原的表达和/或纯化。

[0393] 蛋白变体包括相对于参考序列氨基酸已经缺失的那些蛋白变体,例如,该种缺失可发生在1-10个位置(例如1-5个位置,合适地1或2个位置,特别是1个位置),并可以,例如,涉及在每个位置缺失50或更少的氨基酸(例如20个或更少,特别是10个或更少,尤其是5个或更少)。合适地,该种缺失不发生在表位区域,因此对该抗原的免疫原性不具有明显影响。

[0394] 本领域技术人员将认识到特定的蛋白变体可包括置换、缺失和添加(或其任意组合)。

[0395] 确定抗原表位区域的方法在实施例中描述和示范。

[0396] 变体优选地显示相对于相关参考序列至少约70%一致性,更优选地至少约80%一致性和最优选地至少约90%一致性(例如至少约95%,至少约98%或至少约99%)。

[0397] 上下文中两个或更多核酸或多肽序列的“一致性”或百分比“一致性”指,当比较和比对得到比较窗口中的最大对应或指定区域采用下列序列比较算法之一测量或通过人工比对和视觉观察时,两个或更多个序列或子序列彼此相同或具有一定百分比的相同(即,相对指定区域70%一致性,任选地75%、80%、85%、90%、95%、98%或99%的一致性)的氨基酸残基或核苷酸。这样的序列随后可被称为“基本一致的”。该定义还指测试序列的结果(compliment)。任选地,该一致性存在于长度至少约25至约50个氨基酸或核苷酸的区域,或者任选地长度为75-100个氨基酸或核苷酸的区域。合适地,该比较在对应于参考序列全长的窗口进行。

[0398] 在序列比较中,通常将一个序列作为参考序列,将测试序列与之比较。当使用序列比较算法时,将测试和参考序列输入计算机,指定子序列坐标,并在需要时指定序列算法程序参数。可采用默认程序参数,或指定供选参数。随后序列比较算法将根据程序参数计算测

试序列相对于参考序列的序列一致性的百分比。

[0399] 本文所用的“比较窗口”指对两个序列进行最佳比对后,其中序列可与相同数量连续位置的参考序列比较的片段。通过比对序列进行比较的方法已为本领域所熟知。通过序列的最佳比对进行比较可通过下述方法实现,例如,通过Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981) 的局部同源性算法,通过Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970) 的同源性比对算法,通过Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988) 的相似性搜索法,通过这些算法的计算机执行 (GAP, BESTFIT, FASTA, 以及Wisconsin遗传学软件包中的TFASTA, Genetics Computer Group, 575 Science Dr., Madison, WI), 或者通过人工比对和视觉观察 (参见,例如, *Current Protocols in Molecular Biology* (Ausubel等人, 编辑, 1995 supplement))。

[0400] 一种有用算法的实施例为PILEUP。PILEUP采用渐进的,双序列比对建立了来自一组相关序列的多重序列比对,从而显示相关性和序列一致性百分比。它还绘制了可显示用于建立比对的聚类关系的树状图或系统树图。PILEUP采用了简化的Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987) 的渐进比对法。所采用的方法与Higgins & Sharp, *CABIOS* 5:151-153 (1989) 所描述的方法类似。该程序可对高达300个序列进行比对,每个序列的最大长度为5,000个核苷酸或氨基酸。该多重比对程序始于对两个最为相似的序列的成对比对,从而产生两个比对序列的聚类。然后将该聚类与下一个最相关的序列或比对序列的聚类进行比对。两个序列聚类通过将两个单独序列的成对比对的简单扩展进行比对。最终比对可通过一系列的渐进,双序列比对实现。该程序通过指定特定的序列及其氨基酸或核苷酸坐标为序列比较区域并指定程序参数来运行。使用PILEUP时采用下列参数通过比较参考序列和其它测试序列以确定序列一致性关系百分比:默认gap weight (3.00), 默认gap length weight (0.10), 以及weighted end gaps。PILEUP可以从GCG序列分析软件包的7.0版等版本中获取 (Devereaux 等人, *Nuc. Acids Res.* 12:387-395 (1984))。

[0401] 适于确定序列一致性和序列相似性百分比的算法的另一实施例为BLAST和BLAST2.0算法,其分别描述于Altschul等人, *Nuc. Acids Res.* 25:3389-3402 (1977) 和Altschul等人, *J. Mol. Biol.* 215:403-410 (1990)。可进行BLAST分析的软件可通过国家生物技术信息中心公开获取 (网址在www.ncbi.nlm.nih.gov)。该算法包括首先通过鉴定待测序列中长度为W的短字来鉴定高得分序列对 (HSPs), 其中该短字与数据库序列中相同长度的字比对时匹配或满足某些正数值阈值得分T。T指邻近字得分阈值 (Altschul等人, 同前)。这些初始邻近字采样数 (word hits) 作为种子启动对包含它们的更长HSPs的搜索。该字采样数沿着每个序列的两个方向伸展,直至累计比对得分增加。对核苷酸序列的累计分数采用参数M (对一对匹配残基的奖赏分数;总是>0) 和N (对不匹配残基的惩罚分数;总是<0) 进行计算。对于氨基酸序列,可使用得分矩阵来计算累计分数。以下情况下字采样数向各方向的伸展被停止:累计比对分数较其最大获得值小数量X;由于一个或多个负得分残基比对的累积,累计分数降至0或以下;或者达到了任一序列的末端。BLAST算法的参数W、T和X确定了比对的灵敏度和速度。BLASTN程序 (用于核苷酸序列) 用默认字长 (w) 为11, 预期 (E) 为5, M=5, N=-4, 并同时对比两条链进行比较。对于氨基酸序列, BLASTP程序用默认字长 (w) 为3, 和预期 (E) 为10, 以及该BLOSUM62得分矩阵 (参见Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) 的比对 (B) 为50, 预期 (E) 为10, M=5, N=-4, 并

对两条链进行比较。

[0402] BLAST算法还可进行两个序列间的相似度统计学分析(例如,参见Karlin & Altschul, Proc. Nat'l Acad. Sci. USA 90:5873-5787 (1993))。由BLAST算法所提供的一种相似度检测为最小总可能性(P(N)),它对两个核苷酸或氨基酸序列之间偶然产生匹配的可能性提供了指示。例如,当测试核酸与参考核酸之间比较得到的最小总可能性小于约0.2,更优选小于约0.01,最优选小于约0.001时,该核酸被认为与参考核酸相类似。

[0403] 本发明还涉及多核苷酸,其包含在适度严格条件(例如高度严格条件)下与编码多肽的第二核苷酸序列的补体选择性杂交的第一核苷酸序列,该多肽包含:

[0404] (i) Rv2386c蛋白序列;

[0405] (ii) Rv2386c蛋白序列的变体;或者

[0406] (iii) Rv2386c蛋白序列的免疫原性片段。

[0407] 短语“高度严格杂交条件”指探针与其目标子序列(通常在核酸的复杂混合物中)而不与其它序列进行杂交时的条件。高度严格条件为序列依赖性的且在不同的环境中将不同。更长序列特别在更高的温度下杂交。对于核酸杂交的全面指导可参见Tijssen, *Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Probes*, “Overview of principles of hybridisation and the strategy of nucleic acid assays” (1993)。一般而言,高度严格条件可选定为较限定离子强度pH下具体序列的热熔点(T_m)低大约5-10°C。 T_m 为(在限定离子强度、pH和核酸浓度下)平衡时50%的与目标互补的探针与目标序列进行杂交时的温度(由于目标序列为过量存在,因此在平衡时在 T_m 下50%的探针被占用)。高度严格条件将是在pH7.0至8.3下盐浓度小于约1.0M钠离子,通常为约0.01至1.0M钠离子浓度(或其它盐),且对于短探针(例如,10至50个核苷酸),温度为至少约30°C,以及对长探针(例如,大于50个核苷酸),温度为至少约60°C的条件。高度严格条件还可通过添加去稳定剂例如甲酰胺实现。对于选择性或特异性杂交,阳性信号是背景的至少两倍,任选地为背景杂交的10倍。

[0408] 示例性的高度严格杂交条件可以是如下条件:50%甲酰胺、5x SSC以及1%SDS,42°C下孵育,或者5x SSC,1%SDS,65°C下孵育,并在65°C下用0.2x SSC和0.1%SDS洗涤。

[0409] 当核酸编码的多肽基本一致时,在高度严格条件下不能彼此杂交的该核酸仍然功能相同。这发生于,例如,当核酸的拷贝是通过遗传密码允许的最大密码子简并所生成时。在该种情况下,该核酸通常在适度严格杂交条件下杂交。

[0410] 示例性的“适度严格杂交条件”包括在37°C下于40%甲酰胺,1M NaCl,1%SDS的缓冲液中杂交,并在45°C下于1x SSC中洗涤。阳性的杂交是背景的至少两倍。本领域普通技术人员将容易认识到,可采用替代性的杂交和洗涤条件以提供具有类似严格度的条件。

[0411] 短语“选择性(或特异性)杂交”指在严格杂交条件下,当特定核苷酸序列存在于复合混合物(例如,总细胞或文库DNA或RNA)中时,一种分子仅与该特定序列结合、双联或杂交。

[0412] 在任意情况下,多肽序列的变体将具有与参考序列基本相同的活性(对于多核苷酸,变体多核苷酸序列将编码与参考序列具有基本相同活性的多肽)。基本相同的活性指参考序列在PBMC或全血对特定抗原的体外再刺激测定(例如,再刺激数小时至长达两周之间,例如长达一天、1天至1周或者1至2周的时间)中活性的至少50%、合适地至少75%且特别是

至少90%，其中该测定通过培养悬浮液中淋巴组织增殖、细胞因子生成(通过ELISA、CBA等测量)或者胞内和胞外染色(例如，采用特异性针对免疫标记物的抗体，例如CD3、CD4、CD8、IL2、TNF α 、IFN γ 、CD40L、CD69等)然后以流式细胞仪分析T和B细胞应答特征来测量细胞的活化。合适地，基本相同的活性指参考序列在T细胞增殖和/或IFN-gamma生成测定中的活性的至少50%、合适地至少75%且尤其是至少90%。

[0413] 多核苷酸组合物

[0414] 本文所用的术语“多核苷酸”指经过分离不含特定种类的总基因组DNA的分子。因此，编码多肽的多核苷酸指包含一个或多个编码序列，且从获得该多核苷酸的种类的总基因组DNA基本分离或纯化的多核苷酸片段。

[0415] 本领域技术人员将能理解，本发明的多核苷酸可包括表达或可改造为表达蛋白、多肽、肽等的基因组序列、基因组外和质粒编码的序列以及更小的工程改造基因片段。该种片段可天然分离，或人工修饰合成。

[0416] 本文所用的“分离”指多核苷酸与其它编码序列基本分开，且该多核苷酸不包含大比例的无关编码DNA，例如大染色体片段或其它功能基因或多肽编码区。分离的核酸与位于该基因侧翼并编码与该基因不同的蛋白的其它开放阅读框架相隔离。当然，这是指原始分离的DNA片段，且不排除随后通过人工向该片段添加的基因或编码区。

[0417] 本领域技术人员将能理解，多核苷酸可以是单链(编码或反义)或双链的，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的方式与DNA分子相对应，还包括不包含内含子的mRNA分子。本发明的多核苷酸中可存在但不必须存在另外的编码或非编码序列，且多核苷酸可以(但不必须)与其它分子和/或支撑材料连接。

[0418] 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性序列)或可包含该种序列的变体、生物或功能等价物。多核苷酸变体可包含一个或多个置换、添加、缺失和/或插入，如下文进一步描述，优选地使得所编码多肽的免疫原性相对于参考蛋白未减小。对编码多肽的免疫原性的影响可总体上通过下文所述方法评估。

[0419] 在另外的实施方式中，本发明提供了包含与本文所述序列的一个或多个相同或互补的序列的各种长度的连续片段的多核苷酸和多肽。例如，本发明提供了包含本文所述的参考序列的至少约30、40、50、75、100、150、200、300、400、500或1000或者更多个连续核苷酸以及其中任意中间长度的连续核苷酸的多核苷酸。可以容易地理解，上下文中的“中间长度”指举例数值之间的任意长度，例如30、31、32等；50、51、52、53等；100、101、102、103等；150、151、152、153等；包括所有200-500、500-1000中的整数，等等。

[0420] 此外，本领域技术人员将能理解，由于遗传密码简并的结果，存在多种编码本文所述多肽的核苷酸序列。这些多核苷酸中一部分与任意天然基因的核苷酸序列的一致性相对低。然而，本发明特别预期了由于密码子使用的不同而变化的多核苷酸，例如，对人和/或灵长动物密码子选择优化的多核苷酸。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的范围内。等位基因是由于核苷酸的一个或多个突变(例如缺失、添加和/或置换)而被改变的内源性基因。所得的mRNA和蛋白可以(但非必须)具有不同的结构或功能。等位基因可采用标准技术鉴定(例如杂交、扩增和/或数据库序列比较)。

[0421] 多核苷酸鉴定和表征

[0422] 多核苷酸可通过各种成熟技术中的任意一种进行鉴定、制备和/或操作。例如,多核苷酸可通过筛选cDNA的微阵列(将在下文详细描述)进行鉴定。该种筛选可,例如,采用Synteni微阵列(Palo Alto,CA)参照生产商的说明实施(且基本如Schena等人,Proc.Natl.Acad.Sci.USA 93:10614-10619(1996)和Heller等人,Proc.Natl.Acad.Sci.USA 94:2150-2155(1997)所述)。替代性地,多核苷酸可由cDNA扩增得到,该cDNA从表达本文所述的蛋白的细胞(例如结核分枝杆菌细胞)制备。该种多核苷酸可通过聚合酶链式反应(PCR)扩增。对于该方法,序列特异性引物可基于本文提供的序列设计,以及可购买或合成得到。

[0423] 多核苷酸的扩增部分可用于通过公知技术从合适的文库(例如,结核分枝杆菌cDNA文库)分离全长基因。在该种技术中,可采用适于扩增的一种或多种多核苷酸探针或引物筛选文库(cDNA或基因组)。优选地,文库可经大小选择以包括更大的分子。随机引物文库也可优选用于鉴定基因的5'和上游区域。基因组文库优选用于获得内含子和延伸5'序列。

[0424] 对于杂交技术,可采用公知技术标记部分序列(例如,通过切口平移或以³²P进行末端标记)。然后通过包含具有标记探针的变性细菌集落(或包含噬菌斑的菌苔)的杂交过滤器(hybridising filters)总体筛选细菌或噬菌体文库(参见Sambrook等人,Molecular Cloning: A Laboratory Manual(2000))。选择和扩增杂交菌落或噬菌斑,且分离DNA以供进一步分析。通过例如采用来自该部分序列的引物和来自该载体的引物的PCR分析cDNA克隆以确定添加序列的量。可生成限制性图谱和部分序列以鉴定一种或多种重叠克隆。然后可采用标准技术(可包括生成一系列缺失克隆)确定完整的序列。然后所得的重叠序列可被装配入单个连续序列。可采用公知技术通过连接适合的片段生成全长cDNA分子。

[0425] 替代性地,有多种扩增技术以从部分cDNA序列获得全长编码序列。在这些技术中,扩增通常通过PCR进行。各种市售的试剂盒中的任意一种均可用于实施该扩增步骤。引物可采用例如本领域已知的软件进行设计。引物优选具有22-30个核苷酸的长度,具有至少50%的GC含量且在约68℃-72℃的温度下退火至目标序列。扩增区域可按上述方式测序,且重叠序列被装配进入连续序列。

[0426] 一种这样的扩增技术是反向PCR(参见Triglia等人,Nucl.Acids Res.16:8186(1988)),其采用限制性内切酶生成该基因已知区域的片段。该片段然后通过分子内连接环化,并用作PCR的模板,该PCR采用从已知区域衍生的趋异(divergent)引物。在替代性的方法中,邻近部分序列的序列可通过采用针对接头序列的引物和特异性针对已知区域的引物的扩增重新获得。该扩增的序列通常进行第二轮扩增,该第二轮扩增采用相同的接头引物和特异性针对已知区域的第二引物。WO 96/38591中描述了该步骤的一种变异,其中采用了从已知序列的相反方向启动延伸的两种引物。另一种该类技术被称为“cDNA末端快速扩增”或RACE。该技术包括采用内部引物和外部引物,其与polyA区域或载体序列杂交,以鉴定已知序列的5'和3'的序列。另外的技术包括捕捉PCR(Lagerstrom等人,PCR Methods Applic.1:111-19(1991))和步行PCR(Parker等人,Nucl.Acids.Res.19:3055-60(1991))。采用扩增的其它方法也被用于获得全长cDNA序列。

[0427] 在某些情况下,有可能通过在表达序列标记(EST)数据库(例如可从GenBank获得)中提供的序列的分析获得全长cDNA序列。对重叠ESTs的检索通常可采用公知程序(例如NCBI BLAST检索)进行,该ESTs可用于生成连续全长序列。全长DNA序列还可通过基因组片

段分析获得。

[0428] 多核苷酸在宿主细胞中的表达

[0429] 编码多肽、或融合蛋白或其功能等价物的多核苷酸序列或其片段可用于重组DNA分子以引导多肽在合适的宿主细胞中的表达。由于遗传密码的固有简并,可生成编码基本相同或功能等同的氨基酸序列的其它DNA序列,且这些序列可用于克隆或表达给定多肽。

[0430] 本领域技术人员将能理解,生成具有非天然存在的密码子的多肽编码核苷酸序列在某些情况下是有利的。例如,可以选择被特定的原核或真核宿主所优选的密码子以提高蛋白表达率或生成具有所需特性(例如具有比从天然存在的序列生成的转录物的半衰期更长的半衰期)的重组RNA转录物。

[0431] 此外,该多核苷酸序列可通过本领域通常已知的方法工程改造,从而为了各种理由改变多肽编码序列,该理由包括但不限于,调节基因产物的克隆、加工和/或表达的改变。例如,通过随机片段化的DNA改组和基因片段的PCR重组以及合成寡核苷酸可用于对核苷酸序列进行改造。此外,定点突变可用于插入新的限制性内切位点、改变糖基化模式、改变密码子偏好、生成剪切变体或引入突变等等。

[0432] 天然、修饰或重组核酸序列可被连接至异源性序列以编码融合蛋白。例如,为了在肽文库筛选多肽活性的抑制剂,编码能被市售的抗体识别的嵌合蛋白是有用的。融合蛋白还可通过工程改造以包含位于多肽编码序列和异源性蛋白序列之间的裂解位点,使得该多肽可被裂解并从异源性部分纯化。

[0433] 编码所需多肽的序列可采用本领域公知的方法完全或部分合成(参见Caruthers, M.H.等人, Nucl. Acids Res. Symp. Ser. pp 215-223 (1980), Horn等人, Nucl. Acids Res. Symp. Ser. pp 225-232 (1980))。替代性地,该蛋白本身可利用合成多肽的氨基酸序列或其部分的化学方法生产。例如,肽合成可利用各种固相技术(Roberge等人, Science 269: 202-204 (1995))实现,且可例如利用ABI 431A肽合成仪(Perkin Elmer, Palo Alto, CA)实现自动化合成。

[0434] 新合成的肽可通过制备型高压液相色谱(例如, Creighton, Proteins, Structures and Molecular Principles (1983))或本领域其它相当的技术基本上纯化。合成肽的组合物可通过氨基酸分析或测序(例如, Edman降解方法)确认。另外,多肽或其任一部分的氨基酸序列可在直接合成中改变和/或通过化学方法与来自其它蛋白或其任意部分的序列合并,以生成变体多肽。

[0435] 为了表达所需的多肽,编码该多肽或其功能等同物的核苷酸序列可被插入合适的表达载体,即,包含插入编码序列转录和翻译的必需元件的载体。可采用本领域技术人员公知的方法构建包含编码目标多肽的序列和适当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、合成技术以及体内遗传重组。这些技术描述于Sambrook等人, Molecular Cloning, A Laboratory Manual (2000), 以及Ausubel等人, Current Protocols in Molecular Biology (每年更新)。

[0436] 各种表达载体/宿主系统可用于包含和表达多核苷酸序列。其包括,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;以酵母表达载体转化的酵母;以病毒表达载体(如杆状病毒)感染的昆虫细胞系统;以病毒表达载体(例如,花椰菜花叶病毒, CaMV; 烟草花叶病毒, TMV)或以细菌表达载体(例如, Ti或pBR322质粒)转化的植物

细胞系统;或动物细胞系统。

[0437] 存在于表达载体中的“控制元件”或“调节序列”是载体的非翻译区域——增强子、启动子、5' 和3' 未翻译区域——其与宿主细胞蛋白相互作用以进行转录和翻译。该元件可在强度和特异性上变化。根据所用的载体系统和宿主,可采用任意数量的合适转录和翻译元件,包括组成型或诱导型启动子。例如,在细菌系统中克隆时,可以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)等的杂交lacZ启动子。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。如果有必要生成包含编码多肽的序列的多个拷贝的细胞系,基于SV40或EBV的载体可有利地与具有合适的可选择标记一起使用。

[0438] 在细菌系统中,根据被表达多肽的目标用途,可选择多种表达载体。例如,在需要较大的量时,例如对于抗体的诱导,可使用引导能被容易地纯化的融合蛋白高水平表达的载体。该载体包括,但不限于,多功能大肠杆菌克隆和表达载体如BLUESCRIPT(Stratagene),其中该编码目标多肽的序列可被连接进入载体,并与氨基末端Met和的 β -半乳糖苷酶的后续7个残基的序列同框,从而生成杂交蛋白;pIN载体(Van Heeke & Schuster, J. Biol. Chem. 264:5503-5509 (1989))等;pGEX载体(Promega, Madison, Wis.)也可用于表达作为具有谷胱甘肽S-转移酶(GST)的融合蛋白的外源多肽。一般而言,该融合蛋白为可溶的,并能通过吸附至谷胱甘肽-琼脂糖珠,然后在游离谷胱甘肽存在下洗脱从裂解的细胞容易地纯化。在该系统中制备的蛋白可被设计为包括肝素、凝血酶或XA因子蛋白酶裂解位点,从而可根据需要使得克隆的目标多肽从GST部分释放。

[0439] 在酿酒酵母(*Saccharomyces cerevisiae*)中,可使用一些包含组成型或诱导型启动子的载体,例如alpha因子、醇氧化酶和PGH。包含组成型或诱导型启动子的其它载体包括GAP、PGK、GAL和ADH。综述可参见Ausubel等人(同上)以及Grant等人,Methods Enzymol. 153:516-544 (1987)和Romas等人Yeast 8 423-88 (1992)。

[0440] 当采用植物表达载体时,编码多肽的序列的表达可通过多种启动子的任意一种驱动。例如,CaMV的35S和19S启动子等病毒启动子可单独使用或与来自TMV的omega前导序列合并使用(Takamatsu, EMBO J. 6:307-311 (1987))。替代性地,可使用作为RUBISCO的小亚基或热休克启动子等植物启动子(Coruzzi等人, EMBO J. 3:1671-1680 (1984); Broglie等人, Science 224:838-843 (1984); 和Winter等人, Results Probl. Cell Differ. 17:85-105 (1991))。这些构建体可通过直接DNA转化或病原体介导的转染被引入植物细胞。这些技术描述于一些通常可获得的综述中(参见,例如, Hobbs in McGraw Hill Yearbook of Science and Technology pp 191-196 (1992))。

[0441] 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,苜蓿银纹夜蛾(*Autographa californica*)核多角体病毒(AcNPV)被用作载体来表达草地夜蛾(*Spodoptera frugiperda*)细胞或粉纹夜蛾幼虫(*Trichoplusia larvae*)中的外源基因。编码该多肽的序列可被克隆进入病毒的非必需区,例如多角体蛋白基因,并置于多角体蛋白启动子的控制下。该多肽编码序列的成功插入将使该多角体蛋白基因失活,并生成缺乏外壳蛋白的重组病毒。该重组病毒可随后用于感染,例如,可表达目标多肽的草地夜蛾细胞或粉纹夜蛾幼虫(Engelhard等人, Proc. Natl. Acad. Sci. U.S.A. 91:3224-3227 (1994))。

[0442] 在哺乳动物宿主细胞中,通常可获得一些基于病毒的表达系统。例如,当采用腺病

毒作为表达载体时,可将编码感兴趣的多肽的序列连接进入由晚期启动子和三联前导序列组成的腺病毒转录/翻译复合物。在病毒基因组的非必需E1区和E3区中的插入可用于获得能在感染宿主细胞内表达多肽的活病毒(Logan&Shenk,Proc.Natl.Acad.Sci.U.S.A.81:3655-3659(1984))。此外,转录增强子例如劳斯肉瘤病毒(RSV)增强子,可用于增强哺乳动物宿主细胞中的表达。用腺病毒载体进行工作的方法和方案综述于Wold,Adenovirus Methods and Protocols,1998。有关腺病毒载体用途的其它参考可参见Adenovirus:A Medical Dictionary,Bibliography,and Annotated Research Guide to Internet References,2004。

[0443] 特定的起始信号也可用于实现编码目标多肽的序列的更有效翻译。这类信号包括ATG起始密码子和邻近序列。当序列编码该多肽时,其起始密码子和上游序列被插入合适的表达载体,可无需另外的转录或翻译控制信号。然而,当仅插入了编码序列或其部分时,应当提供包含ATG起始密码子的外源性翻译控制信号。此外,该起始密码子应当位于正确的阅读框内以确保整个插入物的翻译。外源性翻译元件和起始密码子可来自天然和合成的各种来源。可通过包含适用于所用的特定细胞系统的增强子来增强表达效力,该增强子如描述于文献中的那些(Scharf.等人,Results Probl.Cell Differ.20:125-162(1994))。

[0444] 此外,可选择宿主细胞菌株的调节插入序列表达或者以所需方式加工表达蛋白的能力。这些多肽的修饰包括,但不限于,乙酰化、羧基化、糖基化、磷酸化、酯化和酰化。裂解该蛋白的“prepro”形式的翻译后加工也可用于促进正确插入、折叠和/或功能。不同的宿主细胞(例如,CHO、HeLa、MDCK、HEK293和WI38)对该种翻译后活动具有特定的细胞结构和特征机制,可经选择以确保该外源蛋白的正确修饰和加工。

[0445] 通常优选稳定表达以进行重组蛋白的长期高产量生产。例如,稳定表达目标多核苷酸的细胞系可用表达载体转化,该表达载体可能包含病毒复制起点和/或外源性表达元件以及在相同或不同载体上的选择性标记基因。在导入载体后,可将细胞在富集培养基中生长1-2天,然后转入选择性培养基。该选择性标记的目的在于赋予对选择的耐受性,其存在允许成功表达该导入序列的细胞的生长和回收。稳定转化细胞的耐受性克隆可采用适于该细胞类型的组织培养技术增殖。

[0446] 任意数量的选择系统可用于回收转化细胞系。这些选择系统包括,但不限于,可分别应用于tk.sup.⁻或aprt.sup.⁻细胞的单纯疱疹病毒胸苷激酶(Wigler等人,Cell 11:223-32(1977))和腺嘌呤磷酸核糖转移酶(Lowy等人,Cell 22:817-23(1990))基因。此外,抗代谢产物、抗生素或除草剂抗性可用作选择的基础;例如,赋予对氨甲喋呤的抗性的dhfr(Wigler等人,Proc.Natl.Acad.Sci.U.S.A.77:3567-70(1980));赋予对氨基糖苷类、新霉素和G-418的抗性的npt(Colbere-Garapin等人,J.Mol.Biol.150:1-14(1981));以及分别赋予对氯磺隆(chlorsulfuron)和膦丝菌素乙酰基转移酶的抗性的als和pat(Murry,同上文)。其它的可选择基因已有描述,例如,trpB,其允许细胞利用吲哚来代替色氨酸,或者hisD,其允许细胞利用组氨醇来代替组氨酸(Hartman & Mulligan,Proc.Natl.Acad.Sci.U.S.A.85:8047-51(1988))。最近,可见标记的使用已经得到普及,其中这些标记物如花青素、β-葡萄糖醛酸酶及其底物GUS以及萤光素酶及其底物萤光素不仅被广泛用于识别转化体,还被广泛用于对由特定载体系统得到的瞬时或稳定蛋白表达进行定量(Rhodes等人,Methods Mol.Biol.55:121-131(1995))。

[0447] 尽管标记基因表达的存在/缺失表明目标基因也存在,它的存在和表达可能需要确认。例如,如果编码多肽的序列被插入标记基因序列内,包含序列的重组细胞可通过标记基因功能的缺失进行鉴定。替代性地,标记基因可与多肽编码序列在单个启动子的控制下串联。该标记基因响应诱导或选择的表达通常还指示了该串联基因的表达。

[0448] 替代性地,包含和表达所需多核苷酸序列的宿主细胞可通过本领域技术人员已知的各种方法进行鉴别。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,其包括用于核酸或蛋白的检测和/或定量的基于膜、溶液或芯片的技术。

[0449] 本领域已知采用特异性针对多核苷酸编码产物的多克隆或单克隆抗体检测和测量该产物表达的各种方案。实施例包括酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)和荧光激活细胞分类(FACS)。在一些应用中优选采用对在给定多肽上的两个非干扰表位具有活性的单克隆抗体的二位的、基于单克隆的免疫测定,但也可采用竞争性结合测定。这些和其它测定描述于,其中包括,Hampton等人,Serological Methods,a Laboratory Manual (1990)和Maddox等人,J.Exp.Med.158:1211-1216(1983)。

[0450] 各种标记和结合技术已为本领域技术人员所知,并可用于各种核酸和氨基酸测定。生成用于检测与多核苷酸相关的序列的标记的杂交或PCR探针的方法包括低聚标记、切口平移、末端标记或采用标记核苷酸的PCR扩增。替代性地,该序列或其任意部分可被克隆进入载体以生成mRNA探针。该载体在本领域已知,并且已有市售,其可用于通过添加合适的RNA聚合酶例如T7、T3或SP6以及标记的核苷酸体外合成RNA探针。这些方法可采用各种市售的试剂盒进行。可以使用的合适的报道分子或标记包括放射性核素、酶、荧光的、化学发光的或者生色的试剂以及底物、辅因子、抑制剂、磁性颗粒等。

[0451] 用目标多核苷酸序列转化的宿主细胞可在适于蛋白从细胞培养物中表达和回收的条件下进行培养。根据所用的序列和/或载体,通过重组细胞生成的蛋白可被分泌或包含在胞内。本领域技术人员将能理解,包含多核苷酸的表达载体可被设计为包含信号序列,该信号序列引导编码的多肽通过原核或真核细胞膜分泌。其它重组构建体可用于将编码目标多肽的序列连接至编码能促进可溶性蛋白纯化的多肽结构域的核苷酸序列。该种有助纯化的结构域包括,但不限于,金属螯合肽(例如允许在固定化金属上纯化的组氨酸-色氨酸模块,允许在固定化免疫球蛋白上纯化的蛋白A结构域,以及在FLAGS伸展/亲和纯化系统使用的结构域(Immunex Corp,Seattle Wash.))。在该纯化结构域和编码多肽之间包含的可裂解接头序列例如特异性针对XA因子或肠激酶(Invitrogen,San Diego,Calif.)可用于促进纯化。一种该类表达载体提供了包含目标多肽和在硫氧还蛋白或肠激酶裂解位点之前的核酸编码6个组氨酸残基的融合蛋白的表达。该组氨酸残基促进了如Porath等人,Prot.Exp.Purif.3:263-281(1992)所述的在IMAC(固定化金属亲和层析)上的纯化,该肠激酶裂解位点提供了从该融合蛋白纯化所需多肽的方法。Kroll等人,DNA Cell Biol.12:441-453(1993))提供了包含融合蛋白的载体的讨论。

[0452] 体内多核苷酸传递技术

[0453] 在另外的实施方式中,包含本发明的一种或多种多核苷酸的遗传构建体被体内导入细胞。这可通过本领域公知的任意多种技术实现,下文概述了其中一些技术以进行说明。

[0454] 1.腺病毒

[0455] 用于体内传递一种或多种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”意在包括如下构建体,该构建体包含了足以(a)支持构建体的包装和(b)以正义和反义方向表达已经克隆在构建体中的多核苷酸的腺病毒序列。当然,在反义构建体方面,表达不需要该基因产物被合成。

[0456] 该表达载体包含腺病毒的遗传工程改造的形式。有关腺病毒,一种36kb的线性双链DNA病毒的遗传结构的知识允许用高达7kb的外源序列置换大部分的腺病毒DNA (Grunhaus & Horwitz, 1992)。与逆转录病毒不同,宿主细胞的腺病毒感染并不导致染色体整合,因为腺病毒DNA可以游离方式复制而没有潜在的基因毒性。此外,腺病毒是结构稳定的,且在广泛扩增后未检测到基因组重排。腺病毒可实际感染所有上皮细胞,而与其细胞周期阶段无关。目前为止,腺病毒感染似乎仅与轻微的疾病有关(例如人的急性呼吸道疾病)。

[0457] 腺病毒特别适合用作基因转移载体,因为它具有中等大小的基因组、容易操作、高效价、广泛的靶细胞范围以及高传染性。病毒基因组的两端均含有100-200碱基对反向重复(ITRs),它们是病毒DNA复制和包装所必需的顺式元件。基因组的早期(E)和晚期(L)区包含了被病毒DNA复制的起始所分开的不同转录单元。E1区(E1A和E1B)编码负责调节病毒基因组和少量细胞基因的转录的蛋白。E2区(E2A和E2B)区的表达导致用于病毒DNA复制的蛋白的合成。这些蛋白涉及DNA复制、晚期基因表达和宿主细胞关闭(Renan, 1990)。晚期基因的产物(包括大部分病毒衣壳蛋白)仅在主要的晚期启动子(MLP)形成的单个主要转录物的重要加工后表达。MLP(位于16.8m.u.)在感染晚期特别有效,从该启动子形成的所有mRNA具有5'-三联前导(TPL)序列,这使得它们成为翻译的优选mRNA。

[0458] 在当前的系统中,重组腺病毒是从穿梭载体和前病毒载体之间的同源重组生成的。由于两个前病毒载体之间可能的重组,可从该过程生成野生型的腺病毒。因此,重要的是从单个噬菌斑分离单个的病毒克隆,并检验其基因组结构。

[0459] 目前复制缺陷性的腺病毒载体的生成和繁殖取决于代号293的独特的辅助细胞系,它是通过Ad5DNA片段和组成表达E1蛋白由人胚胎肾细胞转化得到(Graham等人,1977)。由于E3区在腺病毒基因组中是非必需的(Jones & Shenk, 1978),目前的腺病毒载体在293细胞的帮助下携带E1、D3或两个区中的外源DNA(Graham & Prevec, 1991)。腺病毒在性质上可以包装约105%的野生型基因组(Ghosh-Choudhury等人,1987),提供约额外2kB DNA的容量。结合E1和E3区中约5.5kB的可置换DNA,目前的腺病毒载体的最大容量在7.5kB以下,或者是载体总长的约15%。超过80%的腺病毒的病毒基因组保留在载体骨架中,并且是载体携带的细胞毒性的来源。此外,E1-缺失病毒的复制缺陷是不完整的。例如,在高多重性感染(MOI)下现有载体中观察到了病毒基因表达的遗漏(Mulligan, 1993)。

[0460] 辅助细胞系可来自于人胚胎肾细胞、肌肉细胞、造血细胞或其它人胚胎间充质或上皮细胞等人类细胞。替代性地,该辅助细胞可来自于人腺病毒允许的其它哺乳动物种类的细胞。这些细胞包括,例如,Vero细胞或其它猴胚胎间充质或上皮细胞。如上所述,目前优选的辅助细胞系是293。

[0461] Racher等人(1995)披露了培养293细胞和繁殖腺病毒的改良方法。在一种形式下,可通过将单个细胞接种进入含100-200ml培养基的1升硅化转瓶(Techno, Cambridge, UK)中生长天然细胞聚集物。在40rpm下搅拌后,用台盼蓝评估细胞活性。在另一形式中,Fibra-Cel微载体(Bibby Sterlin, Stone, UK) (5g/l)按如下应用。将重新悬浮于5ml培养基的细胞接种物添加至250ml Erlenmeyer烧瓶中的载体(50ml)并静置,偶尔搅拌,持续1至4小时。然

后用50ml新鲜培养基更换该培养基并开始振荡。为了生成病毒,将细胞生长至约80%汇合,然后更换培养基(至25%的终体积),并以0.05的MOI添加腺病毒。将培养物静置过夜,然后将体积增加至100%,并开始振荡另外72小时。

[0462] 除了该腺病毒载体是复制缺陷型或者至少条件性缺陷的要求之外,腺病毒载体的性质对于本发明成功实施并不认为是关键的。腺病毒可以是42种不同的已知血清型或亚组A-F的任意一种。亚组C的5型腺病毒是优选的起始材料,以获得用于本发明的条件性复制缺陷型腺病毒载体,因为5型腺病毒是人腺病毒,人们已知有关于它的大量生化和遗传信息,且已经被历史性地用于以腺病毒为载体的大多数构建体中。

[0463] 如上所述,根据本发明的典型载体是复制缺陷型的,且不具有腺病毒E1区。因此,可极为方便地将编码目标基因的多核苷酸导入该E1-编码序列被去除的位置。然而,该构建体在腺病毒序列内的插入位置对于本发明而言并非关键。编码目标基因的多核苷酸也可被插入Karlsson等人(1986)所述的E3替换载体中替代缺失的E3区或者插入辅助细胞系或辅助病毒与E4缺陷互补的E4区。

[0464] 腺病毒易于生长和操作,并显示了广泛的体外和体内宿主范围。该组病毒可以高效价获得,例如 10^9 - 10^{11} 噬菌斑形成单位/ml,且它们具有高感染性。腺病毒的生命周期不要求整合进入宿主细胞基因组。由腺病毒载体传递的外源基因是游离的,且因此对于宿主细胞具有低基因毒性。在用野生型腺病毒接种的研究中未有副作用的报导(Couch等人,1963; Top等人,1971),证明了它们作为体内基因传递载体的安全性和治疗潜力。

[0465] 腺病毒载体已用于真核基因表达(Levrero等人,1991;Gomez-Foix等人,1992)和疫苗开发(Grunhaus & Horwitz,1992;Graham & Prevec,1992)。最近,动物研究显示,重组腺病毒可用于基因疗法(Stratford-Perricaudet & Perricaudet,1991;Stratford-Perricaudet等人,1990;Rich等人,1993)。向不同组织施用重组腺病毒的研究包括气管滴注(Rosenfeld等人,1991;Rosenfeld等人,1992)、肌肉注射(Ragot等人,1993)、外周静脉注射(Herz & Gerard,1993)以及立体定向接种至脑中(Le Gal La Salle等人,1993)。

[0466] 腺病毒载体可来源于人腺病毒。替代性地,它们可来源于其它物种的腺病毒,例如,黑猩猩,其可具有一定优势,即该病毒载体不会被针对许多人类对象体内循环的人腺病毒的抗体所中和(参见,例如:Tatsis N等人Gene Therapy 2006 13:421-429)。

[0467] 35型腺病毒相对不常见,因此它们对于载体本身具有低水平的预存免疫,且已被用作正在开发中的某些结核病疫苗的传递系统(参见,例如,Radosevic等人Infection and Immunity 2007 75(8):4105-4115)。35型腺病毒在本发明中作为传递载体也具有特定的价值。

[0468] 2. 逆转录病毒

[0469] 逆转录病毒是一组单链RNA病毒,其特征在于通过逆转录过程在感染细胞内将它们的RNA转化为双链DNA的能力(Coffin,1990)。然后所得的DNA作为前病毒稳定地整合进入细胞染色体并引导病毒蛋白的合成。该整合导致病毒基因序列保留在接受者细胞及其后代中。该逆转录病毒基因组包含三个基因,gag、pol和env,其分别编码衣壳蛋白、聚合酶和包膜成分。在gag基因上游发现的序列包含了将基因组包装进入病毒体的信号。两个长末端重复(LTR)序列存在于病毒基因组的5'和3'末端。它们包含强启动子和增强子序列,且也为宿主细胞基因组的整合所需(Coffin,1990)。

[0470] 为了构建逆转录病毒载体,编码一种或多种目标寡核苷酸或多核苷酸序列的核酸被插入病毒基因组中某些病毒序列的位置,以生成具有复制缺陷的病毒。为了生成病毒体,构建了包含gag、pol和env基因但不包含LTR和包装成分的包装细胞系(Mann等人,1983)。当包含cDNA以及逆转录病毒LTR和包装序列的重组质粒被导入该细胞系时(例如通过磷酸钙沉淀导入),该包装序列允许该重组质粒的RNA转录物被包装进入病毒颗粒,后者可随后分泌进入培养基(Nicolas&Rubenstein,1988;Temin,1986;Mann等人,1983)。然后收集包含重组逆转录病毒的培养基,任选地进行浓缩,并用于基因转移。逆转录病毒载体能够感染各种各样的细胞类型。然而,整合和稳定表达需要宿主细胞的分裂(Paskind等人,1975)。

[0471] 最近已开发设计用于允许特异性靶向逆转录病毒载体的新型方法,该方法基于通过向该病毒包膜化学添加乳糖残基对逆转录病毒进行化学修饰。该修饰可允许通过唾液酸糖蛋白受体特异性感染肝细胞。

[0472] 还设计了靶向重组逆转录病毒的不同方法,其中使用了针对逆转录病毒包膜蛋白和针对特异性细胞受体的生物素化抗体。该抗体通过采用链霉亲和素与生物素成分偶联(Roux等人,1989)。采用针对主要组织相容性复合物I或II类抗原的抗体,它们证明了亲嗜性病毒对携带这些表面抗原的各种人细胞的体外感染(Roux等人,1989)。

[0473] 3.腺相关病毒

[0474] AAV(Ridgeway,1988;Hermonat&Muzyczka,1984)是一种细小病毒,其作为腺病毒储藏中的污染被发现。它是一种未与任何疾病关联的独特病毒(85%US人群中存在抗体)。它还被分类为依赖病毒,因为它的复制依赖于辅助病毒(例如腺病毒)的存在。已分离出五种血清型,其中AAV-2得到了最好的鉴定。

[0475] AAV具有单链线性DNA,其被衣壳化进入衣壳蛋白VP1、VP2和VP3,以形成直径为20-24nm的二十面体病毒体(Muzyczka&McLaughlin,1988)。

[0476] AAV DNA大约为4.7千碱基长。它包含两个阅读框,并且侧翼有两个ITR。在AAV基因组中有两个主要的基因: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。这些末端重复为染色体整合唯一必需的AAV顺式元件。因此,AAV可被用作载体,该载体去除了所有病毒编码序列并由用于传递的基因盒取代。已鉴定了三个病毒启动子,并根据它们的图谱位置命名为p5、p19和p40。p5和p19的转录导致了rep蛋白的生成,p40的转录生成衣壳蛋白(Hermonat&Muzyczka,1984)。

[0477] 有多个因素激发研究人员研究以rAAV作为表达载体的可能性。其中一个因素是传递基因以整合进入宿主染色体的要求令人吃惊的少。其必须有145-bp ITR,这仅是AAV基因组的6%。这为在载体中装配4.5-kb DNA插入物提供了空间。尽管这种携带能力可能阻止AAV传递大的基因,它十分适于传递反义构建体。

[0478] 由于其安全性,AAV也是传递载体的良好选择。存在一种相对复杂的援救机制:移动rAAV不仅需要野生型腺病毒,还需要AAV基因。同样地,AAV不具有病原性,并且不与任何疾病相关。病毒编码序列的去除使得对病毒基因表达的免疫反应最小化,因此,rAAV不会激发炎症反应。

[0479] 4.作为表达构建体的其它病毒载体

[0480] 其它的病毒载体可在本发明中用作表达构建体以将寡核苷酸或多核苷酸序列传递至宿主细胞。可采用从牛痘病毒(Ridgeway,1988;Coupar等人,1988)、慢病毒、脊髓灰质

炎病毒和疱疹病毒等病毒衍生的载体。其它的痘病毒衍生载体,例如鸡痘衍生载体,也预期可被使用。它们提供了对于各种哺乳动物细胞的多个具有吸引力的特征(Friedmann,1989; Ridgeway,1988; Coupar等人,1988; Horwich等人,1990)。

[0481] 随着最近对缺陷型肝炎B病毒的认识,对于不同病毒序列的结构-功能关系已经有了新的认识。体外研究显示,即使删除了基因组的多达80%,该病毒仍能保留辅助依赖性包装和反转录的能力(Horwich等人,1990)。这提示了该基因组的大部分可被外源性遗传材料替换。嗜肝性和持续性(整合)对于肝定向基因转移而言是特别有吸引力的特性。Chang等人(1991)将氯霉素乙酰转移酶(CAT)基因导入鸭肝炎B病毒基因组中聚合酶、表面和前表面(pre-surface)编码序列处。它与野生型病毒共转染进入鸟类肝癌细胞系。包含高效价重组病毒的培养基用于感染原始小鸭肝细胞。在转染后至少24天检测到了稳定的CAT基因表达(Chang等人,1991)。

[0482] 另外的“病毒”载体包括类病毒颗粒(VLPs)和噬菌体。

[0483] 5. 非病毒载体

[0484] 为了实现本发明的寡核苷酸或多核苷酸序列的表达,该表达构建体必须被传递进入细胞。该传递可如在转化细胞系的实验室步骤中在体外实现,或者如某些疾病阶段的治疗在体内或离体实现。如上所述,一种优选的用于传递的机制是通过病毒感染,其中该表达构建体被包装进入感染性病毒颗粒中。

[0485] 一旦该表达构建体被传递进入细胞,编码所需寡核苷酸或多核苷酸序列的核酸可在不同的位置定位和表达。在某些实施方式中,编码该构建体的核酸可被稳定整合进入该细胞的基因组中。这种整合可通过同源性重组(基因替换)特定地定位和定向或者可整合进入随机的非特异性位置(基因增强)。在进一步的实施方式中,该核酸可作为单独的附加型DNA片段被稳定维持在细胞中。这些核酸片段或“附加体”编码足以允许独立于宿主细胞周期或与宿主细胞周期同步的维持和复制的序列。表达构建体如何传递进入细胞和该核酸在细胞内的位置取决于所用的表达构建体的类型。

[0486] 在本发明的某些实施方式中,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由裸重组DNA或质粒简单组成。该构建体的传递可通过,例如,物理或化学透过细胞膜的任意方法实施。这特别适用于体外传递,但也可用于体内传递。Dubensky等人(1984)以磷酸钙沉淀的形式成功地将多瘤病毒DNA注入了成年和新生小鼠的肝和脾,显示了活性病毒复制和急性感染。Benvenisty & Reshef (1986)还证实了直接腹腔内注射磷酸钙沉淀质粒导致转染基因的表达。可以预期编码目标基因的DNA也可以类似方式体内传递并表达基因产物。

[0487] 用于将裸DNA表达构建体传递进入细胞的本发明的另一实施方式可包括粒子轰击。该方法取决于将DNA-包被的微粒加速至高速以允许它们穿透细胞膜并进入细胞而不杀伤细胞的能力(Klein等人,1987)。已经开发了多种用于加速小颗粒的装置。其中一种这样的装置依靠高压放电生成电流,并进而提供推动力(Yang等人,1990)。所用的微粒包含生物惰性物质如钨或金珠。

[0488] 所选的器官包括已经被体内轰击的大鼠和小鼠的肝、皮肤和肌肉组织(Yang等人,1990; Zelenin等人,1991)。这可能要求手术暴露该组织或细胞,以消除枪和靶器官之间的任何干扰组织,即,离体处理。又,编码特定基因的DNA可通过该方法被传递并仍然被引入。

[0489] 细菌也可用作传递方法(例如,李斯特菌,参见W02004/11048),特别是BCG。

[0490] 多肽组合物

[0491] 在其它方面,本发明提供了多肽组合物。

[0492] 一般地,本发明的多肽是分离多肽(即,与其通常发现天然伴随的那些成分相分离)。

[0493] 例如,当与天然系统中部分或所有共存物质相分离时,天然存在的蛋白是分离的。优选地,该种多肽的纯度为至少约90%,更优选至少约95%,最优选至少约99%。当例如多核苷酸被克隆进入并非天然环境一部分的载体时,该多核苷酸被认为是分离的。

[0494] 多肽可通过各种公知技术中的任意一种制备。由上述DNA序列编码的重组多肽可利用本领域普通技术人员已知的各种表达载体中的任意一种由该DNA序列容易地制备。表达可在以包含了编码重组多肽的DNA分子的表达载体转化或转染的任意合适宿主细胞中实现。合适的宿主细胞包括原核细胞、酵母和更高级的真核细胞,例如哺乳动物细胞和植物细胞。优选地,所用的宿主细胞为大肠杆菌、酵母或哺乳动物细胞系如COS或CHO。来自于将重组蛋白或多肽分泌进入培养基的合适宿主/载体系统的上清液可首先通过已有市售的过滤器浓缩。浓缩后,该浓缩物可被应用至合适的纯化基质如亲和基质或离子交换树脂。最后,可采用一个或多个反相HPLC步骤来进一步纯化重组多肽。

[0495] 本发明的多肽、其免疫原性片段以及其它的变体具有小于约100个氨基酸,且通常小于约50个氨基酸,并且还能利用本领域普通技术人员公知的技术采用合成手段生成。例如,这些多肽可采用市售的固相技术的任意一种合成,例如Merrifield固相合成法,其中氨基酸被顺序添加至生长中的氨基酸链。参见Merrifield, J. Am. Chem. Soc. 85:2149-2146 (1963)。自动化合成多肽的设备可从如Perkin Elmer/Applied BioSystems Division (Foster City, CA) 等供应商处购得,并可参照生产商的说明进行操作。

[0496] 在某些特定实施方式中,多肽可以是包含多个本文所述的多肽的融合蛋白,或者是包含至少一个本文所述的多肽和一个不相关序列的融合蛋白,该种蛋白的范例包括破伤风、结核病和肝炎蛋白(参见,例如,Stoute等人, New Engl. J. Med. 336:86-91 (1997))。融合伙伴可以,例如,协助提供T辅助表位(免疫原性融合伙伴),优选为被人识别的T辅助表位,或者可协助以比天然重组蛋白更高的产率表达该蛋白(表达增强子)。某些优选的融合伙伴既是免疫原性的也是表达增强的融合伙伴。可选择其它的融合伙伴,以增强该蛋白的溶解度或者使该蛋白靶向所需的胞内区室。其它进一步的融合伙伴包括亲和标记,其促进该蛋白的纯化。

[0497] 融合蛋白通常可采用包括化学缀合的标准技术制备。优选地,融合蛋白作为重组蛋白表达,允许在表达系统中相对于非融合蛋白以更高水平生成。简言之,编码该多肽成分的DNA序列可被单独装配,并连接进入合适的表达载体。编码一个多肽成分的DNA序列的3'端经过或不经肽接头被连接至编码第二多肽成分的DNA序列的5'端,使得该序列的阅读框架协调(in phase)。这允许翻译成为保留了两种成分多肽的生物活性的单个融合蛋白。

[0498] 肽接头序列可用于通过足以确保每个多肽能折叠成为其二级和三级结构的距离来分隔该第一和第二多肽成分。该种肽接头序列通过本领域公知的标准技术被结合进入融合蛋白。合适的肽接头序列可根据如下因素选择:(1) 它们采用灵活延伸构象的能力;(2) 不会形成与第一和第二多肽的功能性表位相互作用的二级结构;以及(3) 不具有可能与多肽

功能性表位相互反应的疏水或带电残基。优选的肽接头序列包含Gly、Asn和Ser残基。其它接近中性的氨基酸,例如Thr和Ala也可用于接头序列。有可能用作接头的氨基酸序列包括披露于Maratea等人, Gene 40:39-46 (1985); Murphy等人, Proc. Natl. Acad. Sci. USA 83: 8258-8262 (1986); 美国专利4,935,233和美国专利4,751,180中的那些。接头序列的长度通常可为1至约50个氨基酸。当第一和第二多肽具有可用于分隔功能结构域和防止空间干扰的非必需N末端氨基酸区时,不要求有接头序列。

[0499] 在优选的实施方式中,免疫原性融合伙伴来自于蛋白D,这是一种革兰氏阴性菌B型流感嗜血杆菌的表面蛋白(WO 91/18926)。优选地,蛋白D衍生物包含了该蛋白的约首三分之一(例如,N末端首100-110氨基酸),且蛋白D衍生物可被脂化。在某些优选的实施方式中,脂蛋白D融合伙伴的首109个残基被包括在N末端以提供具有另外的外源性T细胞表位的多肽并提高在大肠杆菌中的表达水平(因此被用作表达增强子)。脂质尾端确保了对抗原呈递细胞最好的抗原呈递。其它的融合伙伴包括来自流感病毒NS1(血凝素蛋白)的非结构性蛋白。典型地,可采用N末端81个氨基酸,尽管可采用包括了T辅助表位的不同片段。

[0500] 在另一实施方式中,该免疫原性融合伙伴是已知为LYTA或其部分(优选C末端部分)的蛋白。LYTA来自于肺炎链球菌,其合成已知为酰胺酶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骨架中的某些键的自溶素。LYTA蛋白的C-末端结构域负责对胆碱或对部分胆碱类似物如DEAE的亲合性。这种特性已被利用来开发用于表达融合蛋白的大肠杆菌C-LYTA表达质粒。在氨基末端包含C-LYTA片段的杂交蛋白的纯化已有描述(参见Biotechnology 10:795-798 (1992))。在优选的实施方式中,LYTA的重复部分可被结合进入融合蛋白。重复部分发现于C末端区,起始于残基178。特别优选的重复部分结合残基188-305。

[0501] T细胞

[0502] 免疫治疗性组合物还可包含,或者替代性地包含,特异性针对分枝杆菌抗原的T细胞。该种细胞通常可采用标准方法在体外或离体制备。例如,T细胞可采用已有市售的细胞分离系统,例如可由Nexell Therapeutics, Inc. (Irvine, CA) 获得的Isolex™系统,从患者骨髓、外周血或者骨髓或外周血的部分分离得到(还可参见美国专利5,240,856、美国专利5,215,926、WO 89/06280、WO 91/16116以及WO 92/07243)。替代性地,T细胞可由相关或不相关的人、非人哺乳动物、细胞系或培养物中获得。

[0503] T细胞可以本发明的多肽、编码该种多肽的多核苷酸和/或表达该种多肽的抗原呈递细胞(APC)进行刺激。在一定条件下进行该种刺激并进行一段足以允许生成特异性针对该多肽的T细胞的时间。优选地,该多肽或多核苷酸存在于传递载体中,例如微球体中,从而促进特异性T细胞的生成。

[0504] 当T细胞特异性增殖、分泌细胞因子或者杀死涂覆该多肽或者表达编码该多肽的基因的目标细胞时,该T细胞被认为特异性针对本发明的多肽。T细胞的特异性可利用多种标准技术中的任意一种进行评估。例如,在铬释放测定或增殖测定中,与阴性对照相比在裂解和/或增殖上超过两倍的刺激指数表示T细胞特异性。这些测定可例如,如Chen等人, Cancer Res. 54:1065-1070 (1994) 所述实施。替代性地,T细胞的增殖的检测可通过多种已知技术实现。例如,T细胞增殖可通过测量DNA合成的增加率进行检测(例如,通过氘化胸腺嘧啶对T细胞进行脉冲标记培养并测量结合进入DNA的氘化胸腺嘧啶的数量)。与本发明的

多肽(100ng/ml-100μg/ml,优选200ng/ml-25μg/ml)接触3-7天应导致T细胞增殖至少上升两倍。如上所述接触2-3小时应导致T细胞的激活,这可通过标准细胞因子测定测得,其中细胞因子(例如,TNF或IFN-γ)释放水平上升两倍是T细胞激活的指征(参见Coligan等人,Current Protocols in Immunology,vol.1(1998))。在响应多肽、多核苷酸或多肽表达APC中已被激活的T细胞可以是CD4⁺和/或CD8⁺。蛋白特异性T细胞可采用标准技术扩增。在优选的实施方式中,该T细胞来自于患者、相关供体或不相关供体,且在刺激和扩增后被施用至患者。

[0505] 为治疗目的,在响应多肽、多核苷酸或APC中增殖的CD4⁺或CD8⁺T细胞可在体外或体内大量扩增。该种T细胞的体外增殖可通过多种方式实现。例如,该种T细胞可被重新暴露至多肽,或暴露至对应于该多肽的免疫原性部分的短肽,暴露时添加或不添加T细胞生长因子,例如,白介素-2,和/或合成多肽的刺激细胞。替代性地,在该蛋白存在下增殖的一种或多种T细胞可通过克隆大量扩增。克隆细胞的方法已为本领域公知,且包括有限稀释。

[0506] 药物组合物

[0507] 在另外的实施方式中,本文所述的多核苷酸、多肽、T细胞和/或抗体组合物将在药学可接受或生理可接受的溶液中配制,以单独或联合一种或多种其它治疗形式向细胞或动物给药。

[0508] 还应当理解,如果需要,表达本文所述的多肽的核酸片段(例如,RNA或DNA)可与其它试剂以及例如,其它蛋白或多肽或各种药学活性试剂,包括有效抗结核分枝杆菌感染的化疗剂联合给药。事实上,对可以包含的其它成分实际上没有限制,只要该另外的药剂在与靶细胞或宿主组织接触时不引起明显的副作用。该组合物因此可根据具体实例中的需求与各种其它试剂一同传递。该种组合物可从宿主细胞或其它生物来源纯化,或者替代性地按本文所述进行化学合成。同样地,该种组合物可进一步包含取代的或衍生化的RNA或DNA组合物。

[0509] 正如将本文所述的特定组合物用于各种治疗方案的合适的给药和治疗方案的开发,药学可接受的赋形剂和载体溶液的制剂已为本领域技术人员公知,包括,例如,口服、肠胃外、静脉内、鼻内和肌肉内给药和制剂。其它的给药途径包括通过粘膜表面给药。

[0510] 典型地,包含治疗有效量的制剂每次给药传递约0.1μg至约1000μg的多肽,更典型地每次给药传递约2.5μg至约100μg多肽。对于多核苷酸组合物,这些制剂典型地每次给药传递约10μg至约20mg本发明的多核苷酸,更典型地每次给药传递约0.1mg至约10mg本发明的多核苷酸。

[0511] 自然地,在每种治疗有用的组合物中活性化合物的量可以如下方式制备:以化合物的任意给定单位剂量获得合适的剂量。制备该种药物制剂的本领域技术人员将预期例如溶解性、生物利用度、生物半衰期、给药途径、产品保质期以及其它药理方面等因素,这样,可期望各种剂量和治疗方案。

[0512] 1.口服传递

[0513] 在某些应用中,本文披露的药物组合物可通过口服给药传递给动物。这样,这些组合物可联合惰性稀释剂或联合可吸收食用载体配制,或者它们可被包封在硬壳或软壳凝胶胶囊中,或者它们可被压缩为片剂,或者它们可被直接混入日常食物中。

[0514] 该活性化合物甚至可混入赋形剂并以可摄取片剂、口腔片剂、糖锭、胶囊、酞剂、悬

浮剂、糖浆、圆片等形式使用(Mathiowitz等人,1997;Hwang等人,1998;美国专利5,641,515;美国专利5,580,579和美国专利5,792,451,每一个在此全文特别引入作为参考)。这些片剂、糖锭、丸剂、胶囊等还可包含如下成分:粘合剂,如黄芪树胶、阿拉伯胶、玉米淀粉或明胶;赋形剂,如磷酸氢钙;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;以及甜味剂,例如可加入蔗糖、乳糖或糖精,或者调味剂,例如薄荷、冬青树油或樱桃调味剂。当该剂量单位形式为胶囊时,它可在上述类型的物质外包含液体载体。各种其它的材料可作为涂层存在,或以调节该剂量单位的物理形式。例如,片剂、丸剂或胶囊可以虫胶、糖或两者涂覆。酏剂的糖浆可包含活性成分、作为甜味剂的蔗糖、作为防腐剂的甲基和丙基苯甲酸酯、染料和调味剂如樱桃或柑橘口味。当然,用于制备任意剂量单位形式的任意材料应当是制药纯的,且在所用的量下基本无毒。此外,该活性成分可被混合进入缓释制剂和配方。

[0515] 对于口服给药,本发明的组合物可替代性地以漱口水、洁牙剂、口腔片剂、口服喷雾或舌下口服给药制剂的形式混合一种或多种赋形剂。例如,漱口水可通过将所需量的活性成分混合进入合适的溶剂如硼酸钠溶液(Dobell溶液)中进行制备。替代性地,该活性成分可被混合进入口服溶液,例如包含硼酸钠、甘油和碳酸氢钾的口服溶液中,或者分散在洁牙剂中,或者以治疗有效量添加至可能包含水、粘合剂、研磨料、调味剂、起泡剂和润湿剂的组合物中。替代性地,该组合物可被制成可置于舌下或以其他方式溶解于口中的片剂或溶液形式。

[0516] 2.可注射传递

[0517] 在某些情况下,将需要肠胃外、静脉内、肌肉内、皮内或甚至腹腔传递本文所述的药物组合物,如美国专利5,543,158、美国专利5,641,515和美国专利5,399,363所述(每一个在此全文具体引入作为参考)。作为游离碱或药学上可接受的盐的活性化合物的溶液可在与表面活性剂(例如羟丙基纤维素)适当混合的水中制备。还可在甘油、液体聚乙二醇及其混合物以及在油中制备分散剂。在常规贮存和使用条件下,这些制剂包含防腐剂以防止微生物生长。

[0518] 适于注射用途的药物形式包括无菌水溶液或分散液以及用于可无菌注射溶液或分散液的临时制剂的无菌粉末(美国专利5,466,468,特别在此全文引入作为参考)。在所有情况下,该形式应当无菌且就容易注射而言应当是液体。它在生产和贮存条件下应保持稳定,并应当抗微生物(例如细菌和真菌)的污染。该载体可以是一种溶剂或分散介质,其包含,例如,水、乙醇、多羟基化合物(例如,甘油,丙二醇,以及液体聚乙二醇等)及其适当的混合物和/或植物油。适当的流动性可通过多种方式维持,例如,通过使用如卵磷脂等涂层,通过针对分散剂时对所需的粒度的维持,以及通过表面活性剂的使用。微生物作用的预防可通过各种抗菌和抗真菌剂(例如,对羟基苯甲酸酯,氯丁醇,苯酚,山梨酸,硫柳汞等)实现。在许多情况下,该组合物中优选包含等张剂,例如,糖或氯化钠。可注射组合物的延长吸收可通过在组合物中使用延缓吸收的试剂(例如,单硬脂酸铝和明胶)来实现。

[0519] 对于在水溶液中进行肠胃外给药,例如,该溶液应当在需要时适当缓冲,且液体稀释剂先用足量的盐水或葡萄糖成为等张液体。这些特定的水溶液特别适用于静脉内、肌肉内、皮下和腹膜内给药。就此而言,根据本披露,可以使用的无菌水性介质将为本领域技术人员公知。例如,一个剂量可溶解于1ml等张NaCl溶液中,并被添加至1000ml的皮下灌注液

体中或者在预期灌注部位注射(参见,例如,Remington's Pharmaceutical Sciences,第15版,pp 1035-1038和1570-1580)。根据待治疗对象的状况,在剂量上可能需要发生一些变化。在任意情况下,负责给药的人员将确定针对个体对象的合适剂量。此外,对于人类给药,制剂应满足FDA生物制品标准办公室所要求的无菌性、致热性和总体安全性和纯度标准。

[0520] 无菌可注射溶液可通过将含于合适溶剂的所需量的活性化合物混合进入各种上述列举的其它成分后根据要求过滤灭菌制备得到。一般而言,分散剂可通过将各种无菌活性成分与包含基础分散介质和上述列举的所需其它成分的无菌载体结合得到。对于用于制备无菌可注射溶液的无菌粉末,其优选制备方法为真空干燥和冷冻干燥技术,其得到的粉末包含活性成分和从其之前无菌过滤溶液得到的任何附加所需成分。

[0521] 本文披露的组合物可配制为中性或盐的形式。药学可接受的盐包括酸加成盐(与蛋白的游离氨基形成),其可与无机酸(例如,氢氯酸、磷酸)或有机酸(例如乙酸、草酸、酒石酸、扁桃酸等)形成。也可由无机碱(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、三乙胺、组氨酸、普鲁卡因等)衍生得到与游离羧酸基团形成的盐。在配制时,溶液将以与剂型相容的方式和治疗有效的量给药。该制剂可以各种剂型(例如可注射溶液、药物释放胶囊等)容易地给药。

[0522] 本文所用的“载体”包括任意和所有的溶剂、分散介质、溶媒、涂层、稀释剂、抗菌剂和抗真菌剂、等张和吸收延缓剂、缓冲剂、载体溶液、悬浮液、胶体等。用于药物活性物质的该种介质和试剂的使用为本领域公知。除了与该活性成分不相容的任意传统介质或试剂外,其在治疗性组合物中的用途在预期之内。补充性活性成分也可混合进入该组合物。

[0523] 短语“药学上可接受的”指在给药给人时不生成过敏性或类似不良反应的分子实体和组合物。包含作为活性成分的蛋白的水性组合物的制备已被本领域充分理解。典型地,该种组合物被制备为可注射的,既可作为液体溶液也可作为悬浮液;也可制备在注射前溶解于或悬浮于液体中的固体形式。该制剂也可被乳化。

[0524] 3. 鼻部和口腔传递

[0525] 在某些实施方式中,该药物组合物可通过鼻内喷雾、口腔喷雾、吸入和/或其它气溶胶传递载体传递。用于直接向肺传递基因、核酸和肽组合物的方法,例如通过鼻和口腔气溶胶喷雾传递已描述于,例如,美国专利5,756,353和美国专利5,804,212(每一个特别在此全文引入作为参考)。同样地,采用鼻内微粒树脂(Takenaga等人,1998)和溶血磷脂甘油化合物(美国专利5,725,871,特别在此全文引入作为参考)的药物传递也已为制药领域公知。同样地,以聚四氟乙烯支撑基质形式的穿粘膜药物传递描述于美国专利5,780,045(特别地在此全文引入作为参考)。

[0526] 4. 脂质体、纳米胶囊和微粒介导的传递

[0527] 在某些实施方式中,本发明人预期采用脂质体、纳米胶囊、微粒、微球体、脂质颗粒、囊泡等将本发明的组合物导入合适的宿主细胞。具体而言,本发明的组合物可被配制为包封于脂质颗粒、脂质体、囊泡、纳米球体或纳米颗粒等进行传递。

[0528] 该种制剂对本文披露的核酸或构建体的药学上可接受的制剂引入而言是优选的。脂质体的形成和使用通常已为本领域技术人员公知(参见,例如,Couvreux等人,1977; Couvreux,1988; Lasic,1998;其描述了用于胞内细菌感染和疾病的靶向抗生素疗法中脂质体和纳米胶囊的使用)。最近开发了具有改进的血清稳定性和循环半衰期的脂质体

(Gabizon & Papahadjopoulos, 1988; Allen 和 Choun, 1987; 美国专利 5,741,516, 特别地在此全文引入作为参考)。此外, 对于作为潜在药物载体的脂质体和类脂质体制剂的各种方法已有综述 (Takakura, 1998; Chandran 等人, 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, 每一篇特别地在此全文引入作为参考)。

[0529] 脂质体已成功地用于一系列通常耐受其它方法转染 (包括 T 细胞悬浮液、原代肝细胞培养物和 PC 12 细胞) 的细胞类型 (Renneisen 等人, 1990; Muller 等人, 1990)。此外, 脂质体不受到 DNA 长度限制, 而基于病毒的传递系统受到这种限制。脂质体已被用于有效地将基因、药物 (Heath & Martin, 1986; Heath 等人, 1986; Balazsovits 等人, 1989; Fresta & Puglisi, 1996)、放射性治疗剂 (Pikul 等人, 1987)、酶 (Imaizumi 等人, 1990a; Imaizumi 等人, 1990b)、病毒 (Fallar & Baltimore, 1984)、转录因子和别构剂 (Nicolau & Gersonde, 1979) 引入各种培养的细胞系和动物中。此外, 多个成功的临床试验已考察了脂质体介导的药物传递的效力 (Lopez-Berestein 等人, 1985a; 1985b; Coune, 1988; Sculier 等人, 1988)。进一步地, 多个研究显示脂质体的使用与全身传递后的自身免疫应答、毒性或性腺局部化不相关 (Mori & Fukatsu, 1992)。

[0530] 脂质体是由分散在水性介质中并自发地形成多层同心双层囊泡 (也称为多层囊泡 (MLVs)) 的磷脂形成的。MLVs 通常具有 25nm 至 4 μ m 的直径。对 MLVs 进行超声可形成在核内包含水溶液的直径在 200-500 Å 范围内的小单层囊泡 (SUVs)。

[0531] 脂质体与细胞膜具有相似性, 且预期在本发明中被用作肽组合物的载体。它们可被广泛适用, 因为水溶性和脂溶性物质均可被包埋在内, 即, 分别包埋在水性空间内和其自身双层内。该携带药物的脂质体甚至有可能通过选择性调整该脂质体制剂而被用于活性药剂的定点传递。

[0532] 除 Couvreur 等人 (1977; 1988) 的教导外, 如下信息可被用于生成脂质体制剂。根据脂质和水的摩尔比, 当分散在水中时, 磷脂可形成多种脂质体以外的结构。在较低的比例下, 脂质体是优选的结构。脂质体的物理特征取决于 pH、离子强度以及二价阳离子的存在。脂质体可对离子和极性物质显示较低的穿透性, 但在升高的温度下则会经历相变, 从而显著改变它们的穿透性。该相变包括由紧密堆积的有序的结构 (称为凝胶态) 转变为松散堆积的、较为无序的结构 (称为液态)。这发生在特征相变温度下, 并导致对离子、糖和药物的穿透性的提高。

[0533] 除温度外, 暴露至蛋白可改变脂质体的穿透性。一些可溶性蛋白, 例如细胞色素 c, 结合、变形和穿透该双层, 从而引起穿透性的变化。胆固醇显然通过使磷脂更为紧密堆积而抑制蛋白的穿透。预期用于抗生素和抑制剂传递的最有用的脂质体形式将包含胆固醇。

[0534] 不同类型的脂质体具有不同的包埋溶解物的能力。例如, MLVs 对于包埋溶解物较为有效, 而 SUVs 则极为无效。SUVs 可提供在粒度分布上的均匀性和重现性的优势, 然而, 大的单层囊泡 (LUVs) 提供了粒度和包埋效率之间的折中。这些大单层囊泡通过醚蒸发制备, 且比 MLVs 具有高三至四倍的溶解物包埋效率。

[0535] 除脂质体特征外, 在包埋化合物中的一个重要的决定因素在于该化合物本身的理化特性。极性化合物被包埋在水性空间内, 而非极性化合物结合至囊泡的脂质双层。极性化合物通过穿透或在双层破裂时被释放, 而非极性化合物则会保持与双层连接, 直至其被温

度或暴露至脂蛋白破坏。两种类型均在相变温度下显示最大流出速率。

[0536] 脂质体通过四种不同的机制与细胞相互作用:通过网状内皮组织系统的吞噬细胞如巨噬细胞和中性粒细胞的胞吞作用;通过非特异性弱疏水或静电力或通过与细胞表面成分的特异性相互作用吸附至细胞表面;通过脂质体的脂质双层插入质膜与浆细胞膜融合,同时将脂质体内容释放进入细胞质;以及通过将脂质体的脂质转移至细胞或亚细胞膜或者反向转移,而不与脂质体内容关联。通常难以确定哪种机制在运作,且可能有一种以上的机制同时运作。

[0537] 静脉注射的脂质体的去向和处理取决于它们的物理特性,例如大小、流动性和表面电荷。它们可在组织中存在数小时或数天,这取决于它们的组成,而在血液中的半衰期可有数分钟至数小时。更大的脂质体,例如MLVs和LUVs,迅速被网状内皮组织系统的吞噬细胞摄取,但循环系统的生理学在多数部位保留了该种大脂质体。它们仅存在于毛细血管内皮中具有大开口或孔隙处,例如肝或脾的窦状隙排出。因此,这些器官成为主要的摄取部位。另一方面,SUVs显示了更广的组织分布,但仍然在肝脏和脾脏高度集中。一般而言,这种体内表现将脂质体限制为仅潜在靶向它们的大尺寸可到达的器官和组织。其包括血液、肝、脾、骨髓和淋巴器官。

[0538] 就本发明而言,靶向通常不是限制。然而,当需要特异性靶向时,可存在实现这一的方法。抗体可用于结合至脂质体表面,并将抗体及其药物成分引导至位于特定细胞类型表面上的特异性抗原受体。碳水化合物决定簇(在细胞-细胞识别、相互作用和吸附中具有作用的糖蛋白或糖脂类细胞-表面成分)也可被用作识别位点,因为它们有潜力将脂质体引导至特定的细胞类型。通常预期采用脂质体制剂的静脉注射,但也可采用其它给药途径。

[0539] 替代性地,本发明提供了本发明的组合物的药学上可接受的纳米胶囊制剂。纳米胶囊通常可以稳定和可重现的方式包埋化合物(Henry-Michelland等人,1987;Quintanar-Guerrero等人,1998;Douglas等人,1987)。为了避免由于胞内聚合过载的副作用,应采用能够体内降解的聚合物设计超细颗粒(尺寸约为0.1 μ m)。满足这些需求的可生物降解的聚烷基-氰基丙烯酸酯纳米颗粒预期用于本发明。该种颗粒可根据描述容易地制备(Couvreux等人,1980;1988;zur Muhlen等人,1998;Zambaux等1998;Pinto-Alphandry等人,1995以及美国专利5,145,684,特别全文引入作为参考)。

[0540] 皮肤贴剂也可用于经皮传递。

[0541] 免疫原性组合物

[0542] 在本发明某些优选的实施方式中提供了免疫原性组合物。该免疫原性组合物通常包含一种或多种多肽或多核苷酸(例如上述多肽和多核苷酸),并联合免疫刺激剂。免疫刺激剂可以是增强或加强对外源性抗原的免疫应答(抗体和/或细胞介导)的任意物质。免疫刺激剂的实例包括佐剂、可生物降解的微球体(例如,polylactic galactide)和脂质体(化合物合并入其中;参见,例如,Fullerton,美国专利4,235,877)。

[0543] 免疫原性组合物的制备总体描述于,例如,Powell & Newman,编辑,Vaccine Design(the subunit and adjuvant approach)(1995)。在本发明范围内的药物组合物和免疫原性组合物还可包含其它化合物,其可具有生物活性或不具活性。例如,在药物或免疫原性组合物中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,其或者结合在融合多肽内,或者作为单独的化合物。

[0544] 示范性的免疫原性组合物可包含编码一种或多种本文所述多肽的多核苷酸(例如,DNA),从而原位生成该多肽(以诱发免疫应答)。如上文指出,DNA可存在于本领域普通技术人员已知的各种传递系统(包括核酸表达系统、细菌和病毒表达系统)的任意一种内。各种基因传递技术已在本领域公知,例如描述于Rolland,Crit.Rev.Therap. Drug Carrier Systems 15:143-198(1998)(在此引用)。合适的核酸表达系统包含了在患者体内表达的必需DNA序列(例如合适的启动子和终止信号)。细菌传递系统包括施用表达该多肽(例如,在细胞表面或分泌该多肽)的细菌宿主细胞(例如,分枝杆菌、杆菌或乳酸菌菌株,包括Calmette-Guerrin杆菌或乳酸乳球菌)(参见,例如,Ferreira,等人,An Acad Bras Cignc (2005) 77:113-124;和Raha,等人,Appl Microbiol Biotechnol (2005) PubMedID 15635459)。在优选的实施方式中,该DNA可利用病毒表达系统(例如,牛痘或其它水痘病毒、逆转录病毒或腺病毒)引入,其可包括采用非病原性(缺陷型)、可复制病毒。合适的系统披露于,例如,Fisher-Hoch等人,Proc.Natl.Acad.Sci.USA 86:317-321(1989);Flexner等人,Ann.N.Y.Acad.Sci.569:86-103(1989);Flexner等人,Vaccine 8:17-21(1990);美国专利4,603,112、4,769,330和5,017,487;WO 89/01973;美国专利4,777,127;GB 2,200,651;EP 0,345,242;WO 91/02805;Berkner,Biotechniques 6:616-627(1988);Rosenfeld等人,Science 252:431-434(1991);Kolls等人,Proc.Natl.Acad.Sci.USA 91:215-219(1994);Kass-Eisler等人,Proc.Natl.Acad.Sci.USA 90:11498-11502(1993);Guzman等人,Circulation 88:2838-2848(1993);以及Guzman等人,Cir.Res.73:1202-1207(1993)。将DNA结合进入该种表达系统的技术为本领域普通技术人员公知。该DNA可以是“裸”DNA,例如,如Ulmer等人,Science 259:1745-1749(1993)所述和Cohen,Science 259:1691-1692(1993)综述。可通过将该DNA涂覆在可被有效运输进入细胞的可生物降解的珠子上提高裸DNA的摄入。显而易见地,免疫原性组合物可同时包含多核苷酸和多肽成分。该种免疫原性组合物可提供增强的免疫应答。

[0545] 显而易见地,免疫原性组合物可包含本文提供的多核苷酸和多肽的药学上可接受的盐。该种盐可由药学上可接受的无毒性碱制备,该无毒性碱包括有机碱(例如,伯、仲和叔胺和碱性氨基酸的盐)和无机碱(例如,钠、钾、锂、铵、钙和镁盐)。

[0546] 尽管本领域普通技术人员已知的任意合适的载体可被用于本发明的免疫原性组合物,载体的类型将取决于给药的模式。本发明的组合物可配制为用于任意合适的给药方式,包括,例如,局部、口服、鼻内、静脉内、颅内、腹腔内、皮下或肌肉内给药。对于肠胃外给药,如皮下注射,该载体优选包含水、盐水、醇、脂肪、蜡或缓冲液。对于口服给药,可采用任意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、蔗糖以及碳酸镁。可生物降解的微球体(例如,聚乳酸聚乙二醇)也可被用作本发明的药物组合物的载体。合适的可生物降解微球体披露于,例如,美国专利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。还可采用包含美国专利5,928,647中所述的微粒-蛋白复合物的载体,其能在宿主中诱导I类限制性细胞毒性T淋巴细胞应答。

[0547] 该种组合物还可包含缓冲液(例如,中性缓冲盐水或磷酸盐缓冲盐水)、碳水化合物(例如,葡萄糖、甘露糖、蔗糖或右旋糖)、甘露醇、蛋白、多肽或氨基酸如甘氨酸、抗氧化剂、抑菌剂、螯合剂如EDTA或谷胱甘肽、佐剂(例如,氢氧化铝)、使得该制剂与接受者血液等

张、低渗或弱高渗的溶解物、悬浮剂、增稠剂和/或防腐剂。替代性地,本发明的组合物可配制为冻干制品。化合物还可利用公知技术被包封在脂质体内。

[0548] 各种免疫刺激剂的任意一种可用于本发明的免疫原性组合物。例如,可以包含佐剂。大部分佐剂包含经设计用于防止抗原快速分解的物质,例如氢氧化铝或矿物油,以及免疫应答的刺激剂,例如脂质A、百日咳杆菌或分枝杆菌种类或分枝杆菌衍生的蛋白。例如,可使用去脂化、去糖脂化的牦牛分枝杆菌(“pVac”)。合适的佐剂已有市售,例如,弗氏不完全佐剂和完全佐剂(Difco Laboratories, Detroit, MI); Merck佐剂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 (海藻糖二白喉菌酸(dicorynomycolate)), Leif (利什曼原虫延长起始因子), 铝盐例如氢氧化铝凝胶(明矾)或磷酸铝; 钙、铁或锌盐; 酰化酪氨酸的不溶性悬浮液; 酰化糖; 阳离子或阴离子衍生化的多糖; 聚磷腈; 可生物降解的微球体; 单磷酸脂质A(MPL®); 以及quil A (例如, QS-21)。细胞因子如GM-CSF或白介素-2、-7或-12也可用作佐剂。

[0549] 佐剂指疫苗或治疗性组合物中提高针对抗原的特异性免疫应答的成分(参见, 例如Edelman, AIDS Res. Hum Retroviruses 8:1409-1411 (1992))。佐剂诱导Th1-型和Th2-型免疫应答。Th1-型细胞因子(例如, IFN- γ 、IL-2和IL-12)易于诱导针对施用抗原的细胞介导的免疫应答, 而Th2-型细胞因子(例如, IL-4、IL-5、IL-6、IL-10)易于诱导体液免疫应答。优选能刺激Th-1细胞介导的免疫应答的佐剂描述于WO 94/00153和WO 95/17209。

[0550] 在本文提供的免疫原性组合物中, 该佐剂组合物优选设计为主要诱导Th1型免疫应答。在应用本文提供的免疫原性组合物后, 患者通常支持包括Th1-型和Th2-型应答的免疫应答。在应答主要是Th1-型的优选实施方式中, Th1-型细胞因子的水平将增加至远远大于Th2-型细胞因子的水平。这些细胞因子的水平可通过标准测定容易地评估。该族细胞因子的综述可参见Janeway, 等人, Immunobiology, 第5版, 2001。

[0551] 该Rv2386c组合物通常可在脂质体制剂中包含一种或多种佐剂, 例如, AS01B (3-去-O-酰化单磷酸脂质A (3D-MPL®) 和QS21; 参见, 美国专利公布2003/0143240); AS02A (3D-MPL®和QS21以及水包油乳液; 参见, Bo Jang, 等人, Lancet (2001) 358:1927); , ENHANZYN® (Detox); 3D-MPL®; 皂甙, 包括Quil A及其成分, 例如QS21和皂甙类似物; CWS (来自结核杆菌的细胞壁骨架); TDM (海藻糖二白喉菌酸); 氨基烷基葡萄糖苷4-磷酸酯 (AGPs); 免疫刺激性寡核肽, 例如CPG; Leif (利什曼原虫延长起始因子); 及其衍生物。在优选的实施方式中, Rv2386c多肽与一种或多种佐剂一起给药, 所述佐剂选自在脂质体制剂例如AS01B中的3D-MPL®和QS21以及3D-MPL®和QS21和水包油乳液(例如AS02A)。佐剂系统AS01B和AS02A进一步描述于Pichyangkul, 等人, Vaccine (2004) 22:3831-40。

[0552] 在作为核酸传递Rv2386c抗原时, 其可在病毒载体(即, 腺病毒载体)或突变细菌宿主细胞(即, 突变体, 无毒性分枝杆菌, 乳酸杆菌或杆菌宿主细胞, 包括卡介杆菌 (BCG) 和乳酸乳球菌)中传递。

[0553] 用于诱发主要的Th1-型应答的优选佐剂包括, 例如, 单磷酸脂质A(MPL®)、优选3-O-去酰化单磷酸脂质A (3D-MPL®) 任选地与铝盐的组合(参见, 例如Ribi, 等人, 1986, Immunology and Immunopharmacology of Bacterial Endotoxins, Plenum Publ. Corp., NY, pp 407-419; GB 2122204B; GB 2220211; 以及US 4,912,094)。优选的3D-MPL®形式为

具有直径小于0.2mm的小颗粒尺寸的乳液形式,其生产方法披露于W0 94/21292。包含单磷酸脂质A和表面活性剂的水性制剂披露于W0 98/43670。示范性的优选佐剂包括可从GlaxoSmithKline获得的AS01B(含于脂质体制剂的MPL®和QS21)、在脂质体制剂内的3D-MPL®和QS21、AS02A(MPL®和QS21和水包油乳液)、3D-MPL®和QS21和水包油乳液以及AS15。MPL®佐剂可从GlaxoSmithKline获得(参见美国专利4,436,727;4,877,611;4,866,034和4,912,094)。

[0554] 含CpG的寡核苷酸(其中CpG二核苷酸是未甲基化的)也诱导主要的Th1应答。CpG是存在于DNA中的胞苷-鸟苷二核苷酸基序的缩写。该种寡核苷酸众所周知,并描述于,例如,W0 96/02555、W0 99/33488和美国专利6,008,200和5,856,462。免疫刺激DNA序列还描述于,例如,Sato等人,Science 273:352(1996)。在配制进入免疫原性组合物中时,CpG通常在自由溶液中与游离抗原(W0 96/02555;McCluskie和Davis,见上文)一起给药,或者与抗原共价缀合(W0 98/16247),或者与氢氧化铝等载体配制((肝炎表面抗原)Davis等人见上;Brazolot-Millan等人,Proc.Natl.Acad.Sci.,USA,1998,95(26),15553-8)。在本领域中已知CpG是能通过全身和粘膜途径给药的佐剂(W0 96/02555,EP 468520,Davis等人,J.Immunol,1998,160(2):870-876;McCluskie和Davis,J.Immunol.,1998,161(9):4463-6)。

[0555] 另一优选的佐剂是皂苷或皂苷类似物或衍生物,例如,Quil A,优选QS21(Aquila Biopharmaceuticals Inc.,Framingham,MA),其可单独使用或与其它佐剂组合使用。例如,强化的系统包括了单磷酸脂质A(MPL®)和皂苷衍生物的组合,例如W0 94/00153所述的QS21和3D-MPL®的组合,或者反应原性更低的组合物,其中QS21用胆固醇淬火,如W0 96/33739所述。其它优选的制剂包含水包油乳液和生育酚。一种特别有效的佐剂制剂包含了含于水包油乳液的QS21、3D-MPL®和生育酚,其描述于W0 95/17210。本发明使用的其它皂苷佐剂包括QS7(描述于W0 96/33739和W0 96/11711)和QS17(描述于美国专利5,057,540和EP0 362 279 B1)。

[0556] 替代性地,该皂苷制剂可与疫苗载体合并,所述疫苗由壳聚糖或其它聚阳离子聚合物、聚交酯和聚交酯-共-乙交酯颗粒、聚-N-乙酰基葡萄糖胺基聚合物基质、由多糖或化学修饰的多糖构成的颗粒、基于脂质体和脂质的颗粒、由甘油单酯组成的颗粒等组成。该皂苷还可在胆固醇存在下配制形成微粒结构,例如脂质体或ISCOM®。此外,该皂苷可与聚氧乙烯醚或酯一起配制于非微粒溶液或悬浮液中,或配制在微粒结构如paucilamellar脂质体或ISCOM®中。皂苷还可与赋形剂如CARBOPOL®配制以提高粘度,或者与粉状赋形剂如乳糖配制为干粉形式。

[0557] 在一个实施方式中,该佐剂系统包括了单磷酸脂质A和皂苷衍生物的组合,例如W0 94/00153所述的QS21和3D-MPL®的组合,或者反应原性更低的组合物,其中QS21用含胆固醇的脂质体淬火,如W0 96/33739所述。其它合适的制剂包含水包油乳液和生育酚。另一合适的佐剂制剂采用W0 95/17210中所述的含于水包油乳液的QS21、3D-MPL®佐剂和生育酚。

[0558] 另一强化佐剂系统包含含有CpG寡核苷酸和皂苷衍生物的组合,特别是W0 00/09159中所述的CpG和QS21的组合。该制剂适当地另外包含水包油乳液和生育酚。

[0559] 其它合适的佐剂包括MONTANIDE® ISA 720(Seppic,France)、SAF(Chiron,

Califomia, United States), ISCOMS® (CSL)、MF-59 (Chiron)、SBAS系列佐剂 (SmithKline Beecham, Rixensart, Belgium)、Detox (Corixa)、RC-529 (Corixa) 以及其它氨基烷基葡萄糖苷4-磷酸酯类 (AGPs), 例如在待审美国专利申请08/853,826和09/074,720中披露的那些, 其披露在此全文引入作为参考, 以及如WO 99/52549 A1中所述的聚氧乙烯醚佐剂。SmithKline Beecham和Corixa Corporation现在属于GlaxoSmithKline的一部分。

[0560] 其它合适的佐剂包括通式 (I) 的佐剂分子:

[0561] $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{-A-R}$

[0562] 其中, n 为1-50, A 是键或 $-\text{C}(\text{O})-$, R 是 C_{1-50} 烷基或苯基 C_{1-50} 烷基。

[0563] 其它感兴趣的佐剂是志贺毒素b链, 其使用例如, 如WO2005/112991所述。

[0564] 本发明的一个实施方式由免疫原性组合物组成, 该免疫原性组合物包含通式 (I) 的聚氧乙烯醚, 其中 n 介于1和50之间, 优选4-24, 最优选9; R 成分为 C_{1-50} , 优选 $\text{C}_4\text{-C}_{20}$ 烷基, 最优选 C_{12} 烷基, 且 A 为键。聚氧乙烯醚的浓度应当在0.1-20%范围内, 优选在0.1-10%范围内, 最优选在0.1-1%范围内。优选的聚氧乙烯醚选自如下: 聚氧乙烯-9-月桂醚、聚氧乙烯-9-steoryl醚、聚氧乙烯-8-steoryl醚、聚氧乙烯-4-月桂醚、聚氧乙烯-35-月桂醚以及聚氧乙烯-23-月桂醚。聚氧乙烯醚例如聚氧乙烯月桂醚描述于Merck目录 (第12版: entry 7717)。这些佐剂分子描述于WO 99/52549。

[0565] 本文提供的任何免疫原性组合物可通过公知的获得抗原、免疫应答增强剂和合适载体或赋形剂的组合的方法进行制备。本文所述的组合物可作为缓释制剂 (即, 可使化合物在给药后缓慢释放的制剂, 例如胶囊、海绵体或凝胶 (例如, 由多糖组成)) 的一部分进行给药。该种制剂通常可采用公知技术 (参见, 例如, Coombes等人, Vaccine 14:1429-1438 (1996)) 制备, 并通过, 例如, 口服、直肠或皮下灌输或者通过在所需的靶点灌输进行给药。缓释制剂可包含分散在载体基质和/或包含在被控速膜围绕的储器内的多肽、多核苷酸或抗体。

[0566] 用于该种制剂内的载体是生物相容的, 并且也是可生物降解的; 优选地, 该制剂提供相对恒定的活性成分释放水平。该种载体包括聚 (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。其它的缓释载体包括超分子生物载体, 其包含非液体亲水核 (例如, 交联多糖或寡糖), 以及, 可选地, 包含两亲化合物如磷脂的外层 (参见, 例如, 美国专利5,151,254和PCT申请WO 94/20078、WO/94/23701和WO 96/06638)。在缓释制剂内包含的活性化合物的量取决于灌输的部位、释放的速率和预期时间。

[0567] 多种传递载体中的任意一种可用于药物组合物和免疫原性组合物中以促进生成抗原特异性免疫应答。传递载体包括抗原呈递细胞 (APCs), 例如树突细胞、巨噬细胞、B细胞、单核细胞以及可被工程改造为有效APCs的其它细胞。该种细胞可以 (但非必须) 被工程改造以提高呈递抗原的能力, 以改善T细胞应答的激活和/或维持, 和/或与接受者免疫原性相容 (即, 匹配的HLA单元型)。APCs通常可从各种生物液体和器官的任意一种分离, 并可以是自体的、同种异体的、同基因的或异种细胞。

[0568] 本发明的某些优选的实施方式采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APCs (Banchereau & Steinman, Nature 392:245-251 (1998)), 且显示为诱发预防性或治疗性免疫性的有效的生理学佐剂 (参见 Timmerman & Levy, Ann.Rev.Med.50:507-529 (1999))。总体而言, 树突细胞可基于其典型的形状 (原位星形、在

体外有可见的显著胞质突(树突))、它们高效摄取、处理和呈递抗原的能力以及激活幼稚T细胞应答的能力进行鉴定。树突细胞当然可以被工程改造以表达在体内或离体树突细胞上不常见的特异性细胞表面受体或配体,且该种改造的树突细胞在本发明的预期之中。作为树突细胞的替代,分泌囊泡抗原负载树突细胞(称为外染色体)可用于免疫原性组合物(参见Zitvogel等人,Nature Med.4:594-600(1998))。

[0569] 树突细胞和祖细胞可从外周血、骨髓、淋巴结、脾、皮肤、脐带血或任意其它合适的组织或液体获得。例如,树突细胞可通过将细胞因子如GM-CSF、IL-4、IL-13和/或TNF α 的组合添加至从外周血收获的单核细胞的培养物以离体分化。替代性地,从外周血、脐带血或骨髓收获的CD34阳性细胞可通过向培养基中添加GM-CSF、IL-3、TNF α 、CD40配体、LPS、f1t3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以分化成为树突细胞。

[0570] 树突细胞可方便地分类为“不成熟”和“成熟”细胞,这允许在两种充分鉴定的表型之间以简单方式进行区分。然而,这种命名法不应解释为排除所有可能的分化中间阶段。不成熟的树突细胞表征为具有高抗原摄取和处理能力的APC,其与Fc γ 受体和甘露糖受体的高表达相关。成熟表型通常通过这些标记物的低表达,负责T细胞激活的细胞表面分子如I类和II类MHC、吸附分子(例如,CD54和CD11)和共刺激分子(例如,CD40、CD80、CD86和4-1BB)的高表达进行表征。

[0571] APCs通常可用编码蛋白(或其部分或其它变体)的多核苷酸转染,从而使多肽在细胞表面表达。该种转染可离体进行,且包含该种转染细胞的药物组合物或免疫原性组合物可按照本文所述使用。替代性地,靶向树突或其它抗原呈递细胞的基因传递载体可被给药至患者,从而形成体内转染。例如,树突细胞的体内和离体转染通常可利用本领域已知的任何方法(例如描述于W0 97/24447中的那些)进行,或者通过Mahvi等人,Immunology and Cell Biology 75:456-460(1997)所述的基因枪方法进行。树突细胞的抗原负载可通过用多肽、DNA(裸DNA或在质粒载体内)或RNA孵育树突细胞或祖细胞实现,或者通过抗原表达重组细菌或病毒(例如,痘苗、禽痘、腺病毒或慢病毒载体)实现。在负载之前,多肽可被共价缀合至提供T细胞辅助的免疫原性伙伴(例如,载体分子)。替代性地,树突细胞可单独的或在多肽存在下用非缀合免疫原性伙伴进行脉冲。

[0572] 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内,例如密封的安瓿或小瓶。该种容器优选严密密封,以将制剂保持无菌至使用。总体而言,制剂可作为悬浮液、溶液或含于油性或水性溶媒中的乳液储存。替代性地,免疫原性组合物或药物组合物可储存在冷冻干燥条件下,仅需要在使用前添加无菌液体载体。

[0573] 在一些实施方式中,“引发”或第一次施用Rv2386c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,然后进行一次或多次“促进”或后续施用Rv2386c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(“引发和促进”法)。例如,第一次施用Rv2386c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,然后一次或多次后续施用Rv2386c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。

[0574] 在一个实施方式中,首次施用Rv2386c多肽或多核苷酸,然后一次或多次后续施用Rv2386c多肽。在一个实施方式中,首次施用Rv2386c多肽或多核苷酸,然后一次或多次后续施用Rv2386c多核苷酸。通常该首次或“引发”施用和该第二次或“促进”施用的间隔约2-12

周,或长达4-6月。后续的“促进”施用可相隔约6个月,或者相隔长达1、2、3、4或5年。传统的促进治疗(例如,蛋白引发施用,然后进行蛋白促进施用)也可用于预防或治疗结核病(例如,预防或治疗潜伏结核病,特别是预防或延缓结核病再活化)。

[0575] 抗体

[0576] “抗体”指包含来自免疫球蛋白基因或其片段的框架区的多肽,其特异性结合和识别抗原。该识别的免疫球蛋白基因包括kappa、lambda、alpha、gamma、delta、epsilon和mu恒定区基因,以及众多免疫球蛋白可变区基因。轻链被分类为kappa或lambda。重链被分类为gamma、mu、alpha、delta或epsilon,其分别依次定义了免疫球蛋白类别IgG、IgM、IgA、IgD和IgE。

[0577] 示范性的免疫球蛋白(抗体)结构单元包含四聚物。每个四聚物由两个相同的多肽链对组成,每个对具有一条“轻”链(约25kDa)和一条重”链(约50-70kDa)。每条链的N-末端定义了约100至110或更多氨基酸的可变区,其主要负责抗原识别。术语可变轻链(V_L)和可变重链(V_H)分别指轻链和重链。

[0578] 抗体作为例如完整免疫球蛋白或多个充分鉴定的以各种肽酶消化生成的片段存在。因此,例如,胃蛋白酶消化铰链区中的二硫键下的抗体以生成Fab的二聚物F(ab)'₂,其自身是通过二硫键连接至V_H-C_{H1}的轻链。该F(ab)'₂可在温和条件下还原以打断铰链区中的二硫键,从而将F(ab)'₂二聚物转化为Fab'单体。该Fab'单体基本是具有铰链区部分的Fab(参见Fundamental Immunology (Paul编辑,第3版,1993)。尽管各种抗体片段可从完整抗体的消化进行定义,本领域技术人员将理解,该种片段可通过化学方法或重组DNA方法从头合成。因此,本文所用的术语抗体还包括通过全抗体修饰、或采用重组DNA方法从头合成(例如,单链Fv)、或采用噬菌体展示库鉴定的抗体片段(参见,例如,McCafferty等人,Nature 348:552-554(1990))。

[0579] 为制备单克隆或多克隆抗体,可采用本领域已知的任意技术(参见,例如,Kohler & Milstein, Nature 256:495-497 (1975); Kozbor等人, Immunology Today 4:72 (1983); Cole等人, pp 77-96, Monoclonal Antibodies and Cancer Therapy (1985))。用于生成单链抗体的技术(美国专利4,946,778)可适用于生成针对本发明多肽的抗体。此外,转基因小鼠,或其它生物如其它哺乳动物,可用于表达人源化抗体。替代性地,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等人, Nature 348:552-554 (1990); Marks等人, Biotechnology 10:779-783 (1992))。

[0580] 对于蛋白或肽,短语与抗体的“特异性(或选择性)结合”或“特异性(或选择性)免疫反应”指在蛋白和其它生物物质的异种群体中确定该蛋白存在的结合反应。因此,在指定的免疫测定条件下,该特异性抗体与特定蛋白的结合是背景的至少两倍,且基本不会与样本中存在的其它蛋白大量结合。在该种条件下特异性结合至抗体可能需要选择对特定蛋白具有特异性的抗体。例如,可选择针对融合蛋白的多克隆抗体,从而仅获得与融合蛋白特异性免疫反应且不与该融合蛋白的单独成分特异性免疫反应的多克隆抗体。这种选择可通过减去与单个抗原交叉反应的抗体实现。多种免疫测定形式可用于选择能与特定蛋白进行特异性免疫反应的抗体。例如,固相ELISA免疫测定可常规用于选择与蛋白特异性免疫反应的抗体(可用于确定特异性免疫反应性的免疫测定形式和条件的描述可参见,例如,Harlow & Lane, Antibodies, A Laboratory Manual (1988) 以及 Using Antibodies: A Laboratory

Manual (1998))。典型地,特异性或选择性反应将是背景信号或噪声的至少两倍,更典型地,是背景(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)的10、20或100倍以上。

[0581] 诊断

[0582] 在另一方面,本发明提供了用一种或多种上述多肽诊断结核病的方法(例如采用常规形式的基于T细胞应答的测定或基于抗体的测定)。

[0583] 例如,提供了一种测定个体中先前的结核分枝杆菌感染的方法,该方法包括:

[0584] (a) 从该个体获得样本;

[0585] (b) 用分离多肽接触所述样本,该分离多肽包括:

[0586] (i) Rv2386c蛋白序列;

[0587] (ii) Rv2386c蛋白序列的变体;或者

[0588] (iii) Rv2386c蛋白序列的免疫原性片段;

[0589] (c) 对样本的应答进行定量。

[0590] 该样本可以是例如全血或纯化细胞。适当地,该样本将包含外周血单核细胞(PBMC)。在本发明的一个实施方式中,该个体血清反应呈阳性。在本发明的第二个实施方式中,该个体血清反应呈阴性。

[0591] 合适地,该个体之前未针对结核分枝杆菌感染进行接种(例如,合适地,该个体之前未用BCG接种)。

[0592] 该样本的应答可通过本领域技术人员已知的一系列手段进行定量,包括监测淋巴细胞增殖或特异性细胞因子或抗体的生成。例如,T-细胞ELISPOT可用于监测细胞因子例如干扰素gamma (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异性抗原的刺激。细胞应答还可通过使用胞内和胞外染色并通过流式细胞仪分析来进行表征。

[0593] 对样本增殖应答的定量方法包括:

[0594] (i) 用放射性标记(例如,氚化胸腺嘧啶)对培养细胞进行脉冲,并监测氚的摄取(例如,气体闪烁)。

[0595] (ii) 二醋酸羧基荧光素琥珀酰亚胺酯(CFSE)标记并用流式细胞仪对细胞分裂进行荧光监测。

[0596] 对样本细胞因子应答的定量具体包括对干扰素gamma生成的监测。

[0597] 在使用该种定量方法时,对抗原的阳性应答可通过至少2:1(例如,至少3:1或至少5:1)的信噪比(S/N比)进行定义。

[0598] 在本发明进一步的方面提供了采用皮试诊断结核分枝杆菌感染的方法。本文所用的“皮试”是直接对患者进行的任意测定,其中在皮内注射上述Rv153c多肽(或其变体、免疫原性片段或编码它们的核苷酸)后测量延缓型超敏(DTH)反应(例如肿胀、发红或皮肤炎)。该种注射可用足以将抗原组合与患者真皮细胞接触的任意装置实现,例如结核菌素注射器或1mL注射器。该反应在一定时间后测量,例如在注射后至少48小时,特别是48-72小时。

[0599] DTH反应是一种细胞介导的免疫应答,这种应答在之前曾暴露至受试抗原的患者中更大。该应答可采用尺子肉眼测量。总体而言,直径大于约0.5cm,特别是直径大于约1.0cm的应答为阳性应答,是之前有结核分枝杆菌感染的指征,其可能显示或不显示为活动

性疾病。

[0600] 为用于皮试,该Rv2386c成分合适地被配制为包含生理可接受载体的药物组合物。合适地,在该种药物组合物中使用的载体是具有合适防腐剂(例如苯酚和/或Tween 80™)的盐水溶液。

[0601] 本发明进一步提供了用于任意上述诊断方法的试剂盒。该种试剂盒通常包含进行诊断测定所必需的两种或更多种成分。该成分可以是化合物、试剂、容器和/或装置。

[0602] 例如,试剂盒内的一个容器可包含特异性结合蛋白的单克隆抗体或其片段。该种抗体或片段可连接至上述支持材料。一个或多个另外的容器可包含用于该测定的元素,如试剂或缓冲液。该种试剂盒还可以(或者替代性地)包含上述检测试剂,该检测试剂包含适用于直接或间接检测抗体结合的报告基团。

[0603] 替代性地,试剂盒可被设计为检测生物样本中编码蛋白的mRNA的水平。该种试剂盒通常包含至少一种与编码蛋白的多核苷酸杂交的上述寡核苷酸探针或引物。该种寡核苷酸可用于,例如,PCR或杂交测定中。可存在于该种试剂盒内的其它成分包括用于促进编码本发明蛋白的多核苷酸检测的第二寡核苷酸和/或诊断试剂或容器。

[0604] 其它的诊断试剂盒包括设计用于检测细胞介导的应答(其可以,例如,用于本发明的诊断方法)的试剂盒。该种试剂盒通常包含:

[0605] (i) 从对象获得合适细胞样本的设备;

[0606] (ii) 用Rv2386c多肽(或其变体、其免疫原性片段或编码该多肽的DNA)刺激所述细胞样本的装置;

[0607] (iii) 对刺激的细胞应答进行检测和定量的装置。

[0608] 用于定量细胞应答的合适装置包括B-细胞ELISPOT试剂盒或替代性地T-细胞ELISPOT试剂盒,其为本领域技术人员已知。

[0609] 一种可能的试剂盒包含:

[0610] (a) 本发明的多肽;和

[0611] (b) 适于直接或间接检测抗体结合的检测试剂。

[0612] 特别感兴趣的是对T细胞应答的定量所定制的诊断试剂盒:

[0613] 诊断试剂盒,其包含:

[0614] (a) 本发明的多肽;和

[0615] (b) 使所述多肽与个体的真皮细胞充分接触的设备。

[0616] 诊断试剂盒,其包含:

[0617] (a) 本发明的多肽;

[0618] (b) 使所述多肽与来自个体的样本(例如,全血或更合适的PBMC)充分接触的设备;和

[0619] (c) 对T细胞应答(例如,增殖或IFN-生成)定量的装置。

[0620] 实施例

[0621] 以下的实施例仅为阐述目的提供,并非对本发明构成限制。本领域技术人员将容易地理解,可改变或调整多种非关键参数以获得基本类似的结果。

[0622] 实施例1-作为潜伏TB疫苗靶标的Rv2386c的鉴别

[0623] 基因Rv2386c也称为MbtI,其编码参与含羟苯噁唑啉铁载体分枝杆菌素的生物起

源的蛋白,其中它将分支酸盐转化为水杨酸盐(分枝菌铁载体合成的起始单元)。

[0624] Rv2386c可基于与如Murphy和Brown BMC.Infect.Dis.2007 7:84-99中的休眠期维持和传染性相关的结核分枝杆菌基因的全基因组分析进行选择。结核分枝杆菌中的潜在休眠期基因靶标通过在模拟休眠条件下细菌基因表达的公开的全基因组DNA微阵列数据组的生物信息学meta分析中是优先的(prioritised)。然后在整个基因组进行由基因编码的结核分枝杆菌蛋白的亚细胞定位以鉴别疫苗靶标。

[0625] 简言之,休眠模型中的实验条件变化很大,因此开发了零至五的评分系统,以根据两个标准归一化这些数据:1) 实验条件与休眠状态的相关性和2) 表达的排位顺序。特定实验数据组中的最大评分可基于与休眠期结核分枝杆菌感染的临床发生率潜在关联性进行调整。表1显示了为步骤1采集的数据组以及对每个数据组调整的最大评分。对于生长必需的基因的附加数据组可采用基于转座子的敲除实验(TraSH)从公开研究获得。对于生长没有作用的基因的评分为零。

[0626] 表1-结核分枝杆菌DNA微阵列基因表达和全基因组敲除(生长期必需)的来源、实验模型和评分标准

[0627]

参考文献	实验模型	时间点: 最大评分 ^a
Betts JC 等人 Mol. Microbiol. 2002 43:717-731	在受控 O ₂ 下饥饿	96h: 3 24h: 2 4h: 1

[0628]

参考文献	实验模型	时间点: 最大评分 ^a
Hampshire T 等人 Tuberculosis (Edinb.) 2004 84:228-238	在受控 O ₂ 下营养耗竭	62 和 75d: 5 49d: 4 18d: 2
Muttucumaru DG 等人 Tuberculosis.(Edinb.) 2004 84:239-246	乏氧的 Wayne 模型 [#]	14d (NRP-2): 4 7d (NRP-1): 2
Voskuil MI 等人 Tuberculosis.(Edinb.) 2004 84:218-227	乏氧的 Wayne 模型 [#]	30 和 80d: 5 14 和 20d: 4 10 和 12d: 3 6 和 8d: 2
Schnappinger D 等人 J. Exp. Med. 2003 198:693-704	小鼠巨噬细胞的感 染, +/- γ -INF	24 和 48h: 5
Karakousis PC 等人 J. Exp. Med. 2004 200:647-657	小鼠皮下植入中空纤 维	10d: 3
Talaat AM 等人 Proc. Natl. Acad. Sci. U.S.A 2004, 101:4602-4607	小鼠感染。从肺部收 获 MTB ^b	28d: 3
Sasseti CM 等人 Mol. Microbiol. 2003 48:77-84	在固体介质上生长的 TraSH 突变库	14d:5
Rengarajan J 等人 Proc.Natl.Acad.Sci.U.S. A 2005, 102:8327-8332	小鼠巨噬细胞感染, 具有结核分枝杆菌 TraSH 突变库的 +/- γ -INF	7d:5
Sasseti CM 等人 Proc.Natl.Acad.Sci.U.S. A 2003 100:12989-12994	用结核分枝杆菌的 TraSH 突变库感染的 C57BL/6J 小鼠	7、14、28 和 56d:5

[0629] ^a基于休眠模型相关性的最大评分;h=小时;d=天。[0630] ^b在28d的充气培养物中来自Balb/c肺部的结核分枝杆菌与MTB的比例。[0631] [#]Wayne LG和Hayes LG Infect.Immun.1996 64:2062-2069

[0632] 步骤2-在应用第二个标准时,基因表达的等级次序、来自每个数据组的基因评分可在表达比例(实验条件表达相对于对数期液体培养物中的细胞的倍数)的基础上由最高到最低排序。最高评分的基因可得到特定基因组的最大评分(在表1第3列中列举(例如,5、4.....1点))。每个基因的评分依次降低0.005点直至零,或者达到数据组末。因此当最大评分是4点时,第100个等级的基因的评分为3.500。对于5点的最大评分,1000个基因或者结核分枝杆菌基因组的25%具有评分。对于采集多个时间点的数据的实验,在所有时间点中最大的评分被用作最终评分。

[0633] 在步骤3中,将每个实验条件下每个基因的评分采集进入Microsoft Access数据库。加入参考域(例如Refseq ID、Genbank功能、Genbank注解、Tuberculist分类以及KEGG和Sanger Center links)以促进优先排序。通过合并来自不同研究和来源的数据,对在休眠状态中存活最关键的特定基因和途径达成了共识。

[0634] 在步骤4,采用评分最高的400个基因(约基因组的10%)并辅以对生物化学途径、酶学、药物易处理性、与人基因的同源性以及其它已有知识的专家计算和人工分析得出治疗靶标的优先列表。高评分基因大部分来自于该组中的两个或三个的相交的子集。

[0635] 在步骤5,在整个基因组进行对基因编码的结核分枝杆菌蛋白的亚细胞定位的鉴定。用于膜蛋白预测的启发法描述于Chalker等人J.Bacteriol.2001 183:1259-1268。采用以梯形窗加权的GES亲水性值(Engelman DM等人Annu.Rev.Biophys.Biophys.Chem.1986 15:321-353)生成平均亲水性图(H)(von Heijne G J.Mol.Biol.1992 225:487-494)。采用与TopPred II算法(Claros MG等人Comput.Appl.Biosci.1994 10:685-686)的初始步骤类似的方法,通过选择集中在最高H值(MaxH)的19个氨基酸、使其不受进一步考虑并重复该方法直至不剩下H>0.5的峰,为每个肽序列预测螺旋跨膜区段(TMS)。基于峰MaxH值、H>1.0的区段的数量以及推定TMS的分布和峰H值给定亚细胞定位。选择1.15的MaxH截断以使分别包含跨膜和细胞质蛋白的两个SwissProtein release 34测试数据组之间的差别最大化(Boyd D等人Protein Sci.1998 7:201-205)。将MaxH<1.15的蛋白分类为细胞质蛋白,而MaxH>1.15以及至少三种可能的TMS分类为膜蛋白。锚蛋白被定义为具有正好两个TMS,一个在氨基酸(aa)35之前起始,且一个H>1.15,另一个的H不小于0.5。革兰氏阳性设置的SignalP被特别地用于M.Bacterium以在启发分析中被分类为细胞质蛋白或“未知”蛋白的那些蛋白中鉴定分泌的蛋白(Nielsen H等人Protein Eng.1997 10:1-6)。

[0636] 根据数个标准,Rv2386c作为疫苗抗原的评分很高:

[0637] (i) Rv2386c在所有休眠模型中均一致性地上调。在meta分析中评分的3999个基因的整个组合中,Rv2386c作为所有休眠模型中过量表达基因最高的10%中的一个,排位第120位。Rv2386c的上调评分为13.165,比最高基因评分22.28有利。Rv2386c最低限度地下调,评分仅略低于0(-2.316,与最大下调基因的-18.13相比)。

[0638] (ii) 从巨噬细胞清除铁对于结核分枝杆菌存活和感染性是关键。称为铁载体的小分子在细菌中能够结合铁以进行胞内摄取,MbtI催化第一步骤以形成唯一的结核分枝杆菌铁载体,分枝杆菌素T,其生物合成取决于从分支酸盐生成水杨酸盐(Harrison等人J.Bacteriol.2006 188:6081-6091)。在人THP-1巨噬细胞的结核分枝杆菌感染中诱导MbtI的表达(Gold等人Mol.Microbiol.2001 42:851-865)。该3D蛋白结构已知,且MbtI的酶作用机制已得到了充分研究(Zwahlen等人Biochemistry 2007 46:954-964)。Rv2386c在必要性

评分中排名很高,且为结核分枝杆菌在体外生长模型中的存活所需(在可能评分5中得分2.6)。

[0639] (iii) 亚细胞定位预测分泌Rv2386c蛋白,因此具有明显的胞外暴露,表明其作为疫苗靶标的合适性。

[0640] 实施例2-Rv2386c表位鉴别

[0641] 方法

[0642] 基于如下方法进行细胞T表位预测:

[0643]

预测	名称	URL/参考文献
----	----	----------

[0644]

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 Brusie,V. (2005) “MULTIPRED: 一种预测混杂 HLA 结合肽的计算机系统” (“MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides”) Nucleic Acids Res. 33, W172 - W179.
	SVMHC	网址: www-bs.informatik.uni-tuebingen.de/SVMHC 用 SVMHC 预测 I 类 MHC 结合肽 (“Prediction of MHC class I binding peptides, using SVMHC.”) Pierre Dönnes 和 Arne Elofsson: BMC Bioinformatics 2002 3: 25
CD4	ProPred	网址: www.imtech.res.in/raghava/propred/ Singh,H.和 Raghava,G.P.S.(2001) ProPred: 预测 HLA-DR 结合位点 (“ProPred: Prediction of HLA-DR binding sites.”) Bioinformatics,17(12), 1236-37.
	Tepitope2	本公司程序, 基于H. Bian, J. Hammer (2004) 用TEPITOPE发现混杂HLA – II – 限制的T细胞表位 (“Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE.”) Methods 34 : 468-75

[0645]

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) 预测混杂 I 类 MHC 限制的 T 细胞表位的杂交方法 (“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/ CTL表位预测的综合方法。综合MHC-I结合、TAP转运效率和蛋白体裂解的预测的联合算法 (“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: 多步骤 T 细胞表位预测的服务器 (“EpiJen: a server for multi-step T cell epitope prediction.”) BMC Bioinformatics, 2006, 7, 131.

[0646]

预测	名称	URL/参考文献
	Syfpethi	网址: www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm Hans-Georg Rammensee, Jutta Bachmann, Niels Nikolaus Emmerich, Oskar Alexander Bachor, Stefan Stevanovic: SYFPEITHI: MHC 配体和肽基序的数据库 ("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) PREDTAP: 预测肽结合与抗原加工相关的人转运子的系统 ("PREDTAP: a system for prediction of peptide binding to the human transporter associated with antigen processing.") Immunome Res. 2(1), 3.

[0647]

预测	名称	URL/参考文献
	PAPROC	网址: www.paproc2.de/paproc1/paproc1.html C. Kuttler, A.K. Nussbaum, T.P. Dick, H.-G. Rammensee, H. Schild, K.P. Haderler, 预测蛋白体裂解的算法 (“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: 在 WWW 上可获得的蛋白体裂解的预测算法 (“PAPROC: A Prediction Algorithm for Proteasomal Cleavages available on the WWW”), Immunogenetics 53 (2001), 87-94

[0648] 结果

[0649] 表2-推定Rv2386c人CD4+T细胞表位

[0650]

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	54	WVLAAGVQA	SEQ ID No: 30	DRB1_0401
2	55	VLAAGVQAM	SEQ ID No: 31	DRB1_0301, DRB1_1301
3	73	VIRDGVTTR	SEQ ID No: 32	DRB1_0301, DRB1_0401
4	125	LAPHTPLAR	SEQ ID No: 33	DRB1_1301
5	148	IRLFDAGIR	SEQ ID No: 34	DRB1_1101, DRB1_1301, DRB1_1501
6	155	IRHREAIDR	SEQ ID No: 35	DRB1_0801, DRB1_1301
7	164	LLATGVREV	SEQ ID No: 36	DRB1_1301
8	187	FRRRVAVAV	SEQ ID No: 37	DRB1_0101, DRB1_0801
9	216	FAIDFPLTY	SEQ ID No: 38	DRB1_0301, DRB1_0401
10	220	FPLTYRLGR	SEQ ID No: 39	DRB1_1101
11	224	YRLGRRHNT	SEQ ID No: 40	DRB1_0801, DRB1_1101, DRB1_1501
12	234	VRSFLLQLG	SEQ ID No: 41	DRB1_1301
13	238	LLQLGGIRA	SEQ ID No: 42	DRB1_1501
14	253	LVTAVRADG	SEQ ID No: 43	DRB1_1301
15	257	VRADGVVIT	SEQ ID No: 44	DRB1_0301, DRB1_0401

[0651]

推定的 CD4 表位号	氨基酸 位置	表位序列	SEQ ID No:	HLA 等位基因
16	262	VVITEPLAG	SEQ ID No: 45	DRB1_0301, DRB1_0401, DRB1_1301, DRB1_1501
17	295	IVEHAISVR	SEQ ID No: 46	DRB1_1301
18	321	IDFMTVRER	SEQ ID No: 47	DRB1_1301
19	375	FRLDECPRG	SEQ ID No: 48	DRB1_0301, DRB1_0401
20	384	LYSGAVVML	SEQ ID No: 49	DRB1_0301
21	385	YSGAVVMLS	SEQ ID No: 50	DRB1_0401, DRB1_1101
22	389	VVMLSADGG	SEQ ID No: 51	DRB1_0401, DRB1_1301
23	415	WLRAGAGII	SEQ ID No: 52	DRB1_0101

[0652] 表3-推定的Rv2386c人CD8+T细胞表位

[0653]

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	3	ELSVATGAV	SEQ ID No: 53	A_0201
2	5	SVATGAVST	SEQ ID No: 54	A_0201
3	18	IPMPAGVNP	SEQ ID No: 55	B_0702, B51, Cw_0401
4	19	PMPAGVNPA	SEQ ID No: 56	A_0201
5	21	PAGVNPADL	SEQ ID No: 57	A_2402, B_3501
6	23	GVNPADLAA	SEQ ID No: 58	A3
7	25	NPADLAAEL	SEQ ID No: 59	A24, B_0702, B44, Cw_0401, Cw_0602, B7, B_3501, B51
8	28	DLAAELAAV	SEQ ID No: 60	A_0201
9	33	LAADVTESV	SEQ ID No: 61	A2, A_0201, B51
10	37	VTESVDEDY	SEQ ID No: 62	A1, A_0101, A_0301
11	38	TESVDEDYL	SEQ ID No: 63	B44
12	40	SVDEDYLLY	SEQ ID No: 64	A1, A_0101, A3
13	48	YECDGQWVL	SEQ ID No: 65	A24, B_4403, B51, Cw_0401, B44
14	52	GQWVLAAGV	SEQ ID No: 66	A2, A_0201
15	55	VLAAGVQAM	SEQ ID No: 67	A2, B8
16	65	ELDSDELRV	SEQ ID No: 68	A_0201
17	67	DSDELRVIR	SEQ ID No: 69	A_0101, A_0301
18	82	QQWSGRPGA	SEQ ID No: 70	A_0201, A_0301
19	86	GRPGAALGE	SEQ ID No: 71	A3, B8
20	87	RPGAALGEA	SEQ ID No: 72	B7, B_0702, B_3501, B51
21	91	ALGEAVDRL	SEQ ID No: 73	A2, A_0201
22	93	GEAVDRLLL	SEQ ID No: 74	B44
23	106	AFGWVAFEF	SEQ ID No: 75	A24
24	124	RLAPHTPLA	SEQ ID No: 76	A2, A_0201, A3
25	125	LAPHTPLAR	SEQ ID No: 77	A_0101, A_0301
26	126	APHTPLARV	SEQ ID No: 78	B7, B_3501, B51
27	127	PHTPLARVF	SEQ ID No: 79	A24, B44
28	130	PLARVFSPR	SEQ ID No: 80	A_0101, A3, A_0301
29	133	RVFSPRTRI	SEQ ID No: 81	B7
30	134	VFSPRTRIM	SEQ ID No: 82	A24, B8

[0654]

推定的 CD8 表位号	氨基酸 位置	表位序列	SEQ ID No:	HLA 等位基因
31	143	VSEKEIRLF	SEQ ID No: 83	A1, A24
32	153	AGIRHREAI	SEQ ID No: 84	B8, B51
33	164	LLATGVREV	SEQ ID No: 85	A2, A_0201
34	170	REVPQSRVS	SEQ ID No: 86	B44
35	172	VPQSRSDV	SEQ ID No: 87	B7, B8
36	180	VSDDPSGFR	SEQ ID No: 88	A_0101, A_0301
37	183	DPSGFRRRV	SEQ ID No: 89	B_3501, B51
38	185	SGFRRRVAV	SEQ ID No: 90	B8, B51
39	186	GFRRRVAVA	SEQ ID No: 91	A_0301, B8
40	187	FRRRVAVAV	SEQ ID No: 92	A3, B8, B51
41	190	RVAVAVDEI	SEQ ID No: 93	A24, B7
42	198	IAAGRYHKV	SEQ ID No: 94	B8, B51
43	200	AGRYHKVIL	SEQ ID No: 95	B7, B8
44	202	RYHKVILSR	SEQ ID No: 96	A3, A24
45	206	VILSRCVEV	SEQ ID No: 97	A2, A_0201
46	214	VPFAIDFPL	SEQ ID No: 98	B7, B_3501, B51
47	216	FAIDFPLTY	SEQ ID No: 99	A1, A_0101, B_3501, B51
48	218	IDFPLTYRL	SEQ ID No: 100	B44
49	226	LGRRHNTVP	SEQ ID No: 101	B8, B51
50	228	RRHNTPVRS	SEQ ID No: 102	A3, B8
51	231	NTPVRSFLL	SEQ ID No: 103	A_0101, A24
52	233	PVRSFLLQL	SEQ ID No: 104	A_0201, B7
53	245	RALGYSPLE	SEQ ID No: 105	A2, A_0201, B_3501, B51
54	246	ALGYSPLELV	SEQ ID No: 106	A_0101, A2, A_0201
55	249	YSPLEVTAV	SEQ ID No: 107	A2
56	250	SPLEVTAVR	SEQ ID No: 108	B_0702, B_3501, B51
57	256	AVRADGVVI	SEQ ID No: 109	A3, A_0301, B7
58	264	ITEPLAGTR	SEQ ID No: 110	A_0101, A_0301
59	266	EPLAGTRAL	SEQ ID No: 111	B7, B8, B_3501, B51
60	270	GTRALGRGP	SEQ ID No: 112	A3, B8
61	272	RALGRGPAI	SEQ ID No: 113	A24, B7, B51
62	288	LESNSKEIV	SEQ ID No: 114	B44
63	294	EIVEHAISV	SEQ ID No: 115	A_0201
64	295	IVEHAISVR	SEQ ID No: 116	A_0101
65	298	HAISVRSSL	SEQ ID No: 117	B7, B8, B_3501
66	301	SVRSSLEEI	SEQ ID No: 118	B7
67	305	SLEEITDIA	SEQ ID No: 119	A2
68	311	DIAEPGSAA	SEQ ID No: 120	A_0101, A_0301
69	313	AEPGSAAVI	SEQ ID No: 121	B44
70	327	RERGSVQHL	SEQ ID No: 122	B7, B44
71	331	SVQHLGSTI	SEQ ID No: 123	A24, B7
72	342	RLDPSSDRM	SEQ ID No: 124	A2, A_0201
73	344	DPSSDRMAA	SEQ ID No: 125	B7, B_3501
74	346	SSDRMAALE	SEQ ID No: 126	A1, B8
75	348	DRMAALEAL	SEQ ID No: 127	B7, Cw_0401, Cw_0602
76	349	RMAALEALF	SEQ ID No: 128	A24
77	351	AALEALFPA	SEQ ID No: 129	A2, A_0201, B_3501
78	352	ALEALFPAV	SEQ ID No: 130	A_0101, A2, A_0201
79	353	LEALFPAVT	SEQ ID No: 131	B8, B_4403

[0655]

推定的 CD8 表位号	氨基酸 位置	表位序列	SEQ ID No:	HLA 等位基因
80	357	FPAVTASGI	SEQ ID No: 132	B7, B8, B_3501, B51
81	359	AVTASGIPK	SEQ ID No: 133	A3, A_0301, B7
82	360	VTASGIPKA	SEQ ID No: 134	A_0201
83	364	GIPKAAGVE	SEQ ID No: 135	A_0301, B8
84	365	IPKAAGVEA	SEQ ID No: 136	B7, B_3501
85	367	KAAGVEAIF	SEQ ID No: 137	A24, B_3501
87	369	AGVEAIFRL	SEQ ID No: 138	A_0201, B51
87	376	RLDECPRGL	SEQ ID No: 139	A2, A_0201
88	380	CPRGLYSGA	SEQ ID No: 140	B7, B_3501
89	383	GLYSGAVVM	SEQ ID No: 141	A2, A_0201, A_0301, A3
90	384	LYSGAVVML	SEQ ID No: 142	A24, B44
91	390	VMLSADGGL	SEQ ID No: 143	A2, A_0201, A24
92	397	GLDAALTLR	SEQ ID No: 144	A_0101, A3, A_0301
93	398	LDAALTLRA	SEQ ID No: 145	A_0201, B_4403
94	400	AALTLRAAY	SEQ ID No: 146	A_0101, A_0301, B_3501
95	402	LTLRAAYQV	SEQ ID No: 147	A_0201
96	407	AYQVGGRW	SEQ ID No: 148	A24
97	414	TWLRAGAGI	SEQ ID No: 149	A24
98	424	EESEPEREF	SEQ ID No: 150	B44
99	430	REFEETCEK	SEQ ID No: 151	B44
100	438	KLSTLTPYL	SEQ ID No: 152	A2, A_0201
101	440	STLTPYLVA	SEQ ID No: 153	A_0101, A_0201
102	441	TLTPYLVAR	SEQ ID No: 154	A3, A_0301

[0656] 由表2和3可见, Rv2386c包含一些预测的CD4⁺和CD8⁺的T细胞表位。此外, 该信息提示该蛋白携带了能被在全世界存在的HLAs所识别的表位(即, 来自白种人、非洲、亚洲或拉丁美洲个体的HLAs, 参见网址www.allele frequencies.net)。

[0657] 实施例3-H37Rv同源物

[0658] 利用GenBank的BLASTP检索鉴定来自一系列结核分枝杆菌菌株和BCG的Rv2386c序列(H37Rv参考序列的索引号YP_177877.1)。

[0659]

菌株	索引号	一致性%
CDC1551	NP_336935.1	100
F11	YP_001288342.1	100
Haarlem	ZP_02247771.1	100
C	ZP_00878160.1	99
BCG	CAL72388.1	100

[0660] 同源物序列的比对显示了很高水平的一致性。

[0661] 生物测定

[0662] 对Rv2386c的T细胞应答的定量

[0663] 可就激活T细胞(诱导增殖和/或生成细胞因子)的能力在感染(例如潜伏感染)个体的外周血单核细胞(PBMC)或全血制备物中筛选多肽。

[0664] 潜伏感染个体通常可通过直径大于10mm的皮试、无症状、无Mtb阳性培养物、痰菌

阴性且无病变(通过胸部X射线检测)进行鉴定。

[0665] 一系列的体外测定可基于PBMC样本或全血使用:在抗原(或其适当的变体/免疫原性片段)存在下再刺激后,可测定细胞的增殖(通过CFSE/流式细胞仪测定)或对细胞因子的生成进行定量(存在于培养细胞的上清液中并通过ELISA测定,或者在CD4和CD8T细胞胞内染色后通过流式细胞仪分析)。

[0666] 例如,PBMC样本可通过Ficoll-Hypaque密度梯度离心后通过标准步骤从肝素化全血中获得。可随后洗涤该细胞,并冷冻保存在液氮中直至测试(进一步的细节可参见Lalvani A等人J.Infect.Dis.1999 180:1656-1664)。

[0667] T细胞增殖

[0668] 特异性免疫应答可通过采用氘化胸腺嘧啶进行淋巴细胞增殖分析进行表征。该技术评估了对抗原体外刺激时的细胞扩增。事实上,细胞增殖可通过评估氘化胸腺嘧啶在DNA中的结合进行确定,该过程与细胞数量的根本变化紧密相关。

[0669] 更合适地,可采用二醋酸羧基荧光素的琥珀酰亚胺酯(CFSE)进行淋巴细胞增殖。CFSE通过与赖氨酸侧链和其它可获得胺基团的反应自发和不可逆地偶联至胞内和细胞表面蛋白。当淋巴细胞分裂时,CFSE标记在子细胞中平均分布,因此后者具有亲代细胞一半的荧光。结果,增殖细胞群体的每一个后代中的细胞荧光强度对分,并可容易地通过流式细胞仪跟踪(进一步的细节参见Hodgkins,PD等人J.Exp.Med.1996184:277-281)。

[0670] 实际上,在冻融后,可洗涤PBMC并用CFSE染色,然后以10 μ g/ml抗原在培养基(RPMI-1640,添加了谷氨酰胺、非必需氨基酸、丙酮酸盐和热灭活人AB血清)中培养(2 $\times$ 10⁵细胞)72小时。然后可收获细胞,通过表面染色表征其表型以鉴定记忆CD8和CD4+T-细胞。然后,流式细胞仪分析可用于显示响应每种抗原的淋巴细胞的增殖程度(在体外刺激时具有CFSE强度下降的细胞部分)。

[0671] 细胞因子生成

[0672] IFN- γ 生成(或者其它细胞因子的生成,例如IL2、TNF- α 、IL5、IL12等)可采用酶联免疫吸附测定(ELISA)进行测量。室温下可将含于PBS的针对人IFN- γ (PharMingen, San Diego,CA)的小鼠单克隆抗体涂覆ELISA板达四小时。然后用含5%(W/V)脱脂奶粉的PBS室温封闭孔1小时。然后洗涤板,例如,用PBS/0.2%TWEEN-20洗涤六次,并在ELISA板中以1:2稀释于培养基中的样本在室温下孵育过夜。再次洗涤该板,并向每个孔添加多克隆兔抗人IFN- γ 血清(例如,在PBS/10%普通山羊血清以1:3000稀释)。然后将该板室温孵育两小时,洗涤并可添加辣根过氧化物酶偶联的抗兔IgG(Sigma Chemical So.,St.Louis,MO),例如其在PBS/5%脱脂奶粉中以1:2000稀释。进一步室温孵育两小时后,洗涤该板并添加TMB底物。该反应可在20分钟后用1N硫酸终止。然后在450nm下以570nm作为参考波长测定光密度。典型地,使两个重复样本得到比单独在培养基中培养的细胞的平均OD大两倍的OD的部分可被认为阳性。

[0673] 实施例4-CB6F1小鼠中的免疫原性

[0674] 在CB6F1小鼠(BALB/c和C57BL/6小鼠杂交后的第一代)中评估抗原的免疫原性。

[0675] 用0.5 μ g或2 μ g蛋白抗原联合佐剂系统AS01E(包含3D-MPL和QS21的脂质体佐剂制剂)肌肉内三次(在第0天、第14天和第28天)免疫CB6F1小鼠。

[0676] 该实验设计如下:

[0677]

组	第0天	第14天	第28天
1	2 μ g Rv2386c/AS01E	2 μ g Rv2386c/AS01E	2 μ g Rv2386c/AS01E
2	0.5 μ g Rv2386c/AS01E	0.5 μ g Rv2386c/AS01E	0.5 μ g Rv2386c/AS01E

[0678] 每个方案组中使用了总共24只小鼠。

[0679] 在21天(即,第二次免疫后7天)和第35天(即,第三次免疫后7天)收集并汇集外周血淋巴细胞(PBL),并在用覆盖目标序列的15聚肽集合体外再刺激过夜后通过流式细胞仪测量抗原特异性CD4&CD8T细胞应答(通过CD4或CD8T细胞生成IL-2和/或IFN-gamma和/或TNF-alpha测定)。表达IL-2和/或IFN-gamma和/或TNF-alpha的小鼠T细胞的检测通过使用在细胞因子表达的短期抗原驱动体外扩增实现。

[0680] 简言之,将PharmLyse溶液(BD-Pharmingen)添加至肝素化小鼠外周血中以裂解红细胞。洗涤所得的PBLs(外周血淋巴细胞),然后在覆盖了目标抗原序列的15-聚肽(11个氨基酸重叠)集合和1 μ g/ml针对CD28和CD49d的抗体(BD-Pharmingen)存在下进行孵育。每种15-聚肽使用的终浓度为1 μ g/ml。培养基对照也以针对CD28和CD49d的抗体进行刺激。

[0681] 在37 $^{\circ}$ C,5%CO₂下,培养开始后2小时添加细胞因子分泌阻断化合物布雷菲德菌素-A(BD-Pharmingen),并继续在37 $^{\circ}$ C,5%CO₂下培养细胞4小时,然后在+4 $^{\circ}$ C下孵育过夜。

[0682] 然后收获细胞并用海洋蓝偶联的抗-CD4抗体(BD-克隆RM4-5,BD-Pharmingen)和多甲藻黄素叶绿素A蛋白(PerCp)花青苷(Cy5.5)-偶联的抗-CD8alpha抗体(克隆53-6.7,BD-Pharmingen)染色。

[0683] 然后洗涤细胞,固定,渗透(Cytofix-cytoperm试剂盒,BD-Pharmingen)并用别藻蓝蛋白偶联抗-IFN-g抗体(克隆XMG1.2,BDPharmingen)、异硫氰酸荧光素(FITC)-偶联抗-IL-2抗体(克隆JES 6-5H4,Beckman Coulter)和藻红蛋白(PE)-偶联抗-TNF alpha抗体(克隆MP6-XT22,BDPharmingen)染色。最后的洗涤后,在LSR II流式细胞仪(Beckton-Dickinson)上分析染色细胞。在CD8+子集中获得最少10,000个细胞。

[0684] 进一步的背景参见Walzer T等人Cell Immunol.2000206(1):16-25和Maecker HT等人J.Immunol.Methods 2001255(1-2):27-40。

[0685] 作为阴性对照,一些细胞也在培养基中体外培养过夜(未刺激)。抗原特异性应答可通过从肽刺激细胞生成的平均细胞因子应答减去未刺激细胞生成的平均细胞因子应答计算得到。

[0686] 在每个时间点,对于每个组,该数据可从4个集合(pool)中获得,每个集合具有6只小鼠。如下数据显示为生成IL-2和/或IFN-gamma和/或TNF-alpha的CD4或CD8T细胞的百分比。对每个单独小鼠集合(三角),以及该组的平均值(条形)绘图。

[0687] 图1显示了在第21天(即,第二次免疫后7天)在以任一Rv2386c/AS01E剂量免疫的小鼠中测得的Rv2386c-特异性CD4和CD8T细胞应答。与Rv2386的免疫剂量无关,Rv2386c-特异性CD4T细胞应答水平相近。不同地,用2 μ g Rv2386c/AS01E免疫的小鼠显示了比用0.5 μ g Rv2386c/AS01E免疫小鼠更高的Rv2386c-特异性CD8T细胞应答。

[0688] 图2显示了在第21天(即,第二次免疫后7天)从Rv2386c肽集合刺激PBL(未除去培养基)获得的CD4T细胞应答的细胞因子分布。

[0689] 图3显示了在第21天(即,第二次免疫后7天)从Rv2386c肽集合刺激PBL(未除去培

培养基)获得的CD8T细胞应答的细胞因子分布。

[0690] 图4显示了在第35天(即,第三次免疫后7天)在以任一Rv2386c/AS01E剂量免疫的小鼠中测得的Rv2386c-特异性CD4和CD8T细胞应答。与Rv2386的免疫剂量无关,Rv2386c-特异性CD4T细胞应答水平相近。不同地,用2 μ g Rv2386c/AS01E免疫的小鼠显示了比用0.5 μ g Rv2386c/AS01E免疫小鼠更高的Rv2386c-特异性CD8T细胞应答。第三次免疫剂量提高了CD4T细胞应答,但未提高CD8T细胞应答。

[0691] 图5显示了在第35天(即,第三次免疫后7天)从Rv2386c肽集合刺激PBL(未除去培养基)获得的CD4T细胞应答的细胞因子分布。

[0692] 图6显示了在第35天(即,第三次免疫后7天)从Rv2386c肽集合刺激PBL(未除去培养基)获得的CD8T细胞应答的细胞因子分布。

[0693] 实施例5-C57BL/6小鼠中的免疫原性

[0694] C57BL/6小鼠中的抗原免疫原性也进行了评估。

[0695] 用1 μ g或4 μ g蛋白抗原联合佐剂系统AS01E(包含3D-MPL和QS21的脂质体佐剂制剂)肌肉内三次(在第0天、第14天和第28天)免疫C57BL/6小鼠。

[0696] 该实验设计如下:

[0697]

组	第0天	第14天	第28天
1	4 μ g Rv2386c/AS01E	4 μ g Rv2386c/AS01E	4 μ g Rv2386c/AS01E
2	1 μ g Rv2386c/AS01E	1 μ g Rv2386c/AS01E	1 μ g Rv2386c/AS01E

[0698] 在21天(即,第二次免疫后7天)和第35天(即,第三次免疫后7天)收集并汇集外周血淋巴细胞(PBL),并在用覆盖目标序列的15聚肽集合体外再刺激过夜后通过流式细胞仪测量抗原特异性CD4&CD8T细胞应答(通过CD4或CD8T细胞生成IL-2和/或IFN- γ 和/或TNF- α 测定)。该步骤如上所述。

[0699] 作为阴性对照,一些细胞也在培养基中体外培养过夜(未刺激)。抗原特异性应答可通过从肽刺激细胞生成的平均细胞因子应答减去未刺激细胞生成的平均细胞因子应答计算得到。

[0700] 在每个时间点,对于每个组,该数据可从4个集合中获得,每个集合具有6只小鼠。如下数据显示为生成IL-2和/或IFN- γ 和/或TNF- α 的CD4或CD8T细胞的百分比。对每个单独小鼠集合(三角),以及该组的平均值(条形)绘图。

[0701] 图7显示了在第21天(即,第二次免疫后7天)在以任一Rv2386c/AS01E剂量免疫的小鼠中测得的Rv2386c-特异性CD4和CD8T细胞应答。与Rv2386c的免疫剂量无关,Rv2386c-特异性CD4T细胞应答水平相近。不同地,用4 μ g Rv2386c/AS01E免疫的小鼠显示了比用1 μ g Rv2386c/AS01E免疫的小鼠更高的Rv2386c-特异性CD8T细胞应答。由于技术上的障碍,在4 μ g剂量下的第一细胞集合的数据不可得。

[0702] 图8显示了在第21天(即,第二次免疫后7天)从Rv2386c肽集合刺激PBL(未除去培养基)获得的CD4T细胞应答的细胞因子分布。

[0703] 由于技术上的障碍,在4 μ g剂量下的第一细胞集合的数据不可得。

[0704] 图9显示了在第21天(即,第二次免疫后7天)从Rv2386c肽集合刺激PBL(未除去培养基)获得的CD8T细胞应答的细胞因子分布。

[0705] 由于技术上的障碍,在4 μ g剂量下的第一细胞集合的数据不可得。

[0706] 在准备本申请时,第35天的免疫学数据仍未得到。

[0707] 综上所述,应注意Rv2386c抗原能够同时诱发CB6F1和C57BL/6小鼠的免疫应答。此外,细胞因子生成的分布表明了大部分抗原特异性T细胞表达了多种Th1相关的细胞因子(即,诱发了多功能T-细胞应答)。重要的是,CD4和CD8抗原特异性T细胞均在免疫后存在,CD8细胞对于潜伏TB的情况特别重要。因此,可预期Rv2386c对预防、治疗和诊断潜伏结核病感染具有重要的价值。

[0708] 尽管前述发明已通过阐释目的的图示和实施例进行详细描述,本领域的普通技术人员根据本发明的教导应当能够容易地理解,可在不脱离附加权利要求的精神和范围的情况下对其进行一定的更改和调整。

[0709] 本申请中所涉及的所有参考文献,包括专利和专利申请,均最大程度引用作为参考,其引用程度如同每一个单独的公布和专利申请均为特定的单独的引用作为参考。

[0710] 在整个说明书和权利要求书中,除非另行指明,用语“包含”及其变体“包括”将理解为包含所述整数、节距、整数组或节距组,但不排除任意其它的整数、节距、整数组或节距组。

序列表

<110> 葛兰素史密丝克莱恩生物有限公司

葛兰素集团有限公司

P. 梅滕斯

J. 布朗

D. 墨菲

<120> 新型组合物和方法

<130> VB63088PCT

<150> US61/083699

<151> 2008-07-25

<160> 156

<170> PatentIn version 3.5

<210> 1

<211> 450

<212> PRT

<213> 结核分枝杆菌

<220>

<221> MISC_FEATURE

<223> 菌株H37Rv

<400> 1

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

450	
<210>	2
<211>	1353
<212>	DNA
<213>	结核分枝杆菌
<220>	
<221>	misc_feature
<223>	菌株H37Rv
<400>	2
atgtccgagc tcagcgtcgc gacaggcgcc gtcagcaccg cgtcgtcgtc catccccgatg	60
cccgccggtg tcaaccccgc cgacctggca gcggagctgg cggcggtggt taccgagtcc	120
gtcgacgagg attacctgct ctacgagtgc gacggccaat gggtcctggc cgccggtgtg	180
caggcgatgg tggagctaga cagcgacgaa ctgcgcgtca tccgtgatgg cgttacgcgg	240
cgacagcaat ggtcgggtcg cccgggagcg gccctgggcg aagccgtcga tcggctgttg	300
ctggaaaccg atcaagcttt tggtcgggtc gccttcgaat tcggcgtgca ccgctatggg	360
ttgcagcagc ggctggcgcc gcacacccca ctggcccggg tgttttcgcc ccgaaccg	420
atcatggtga gcgaaaagga gattcgctg ttcgatgctg ggattcgcca ccgcgaggcc	480
atcgaccgat tactcgccac cggggtgcga gagtgccgc agtcccgtc cgtcgacgtc	540
tccgacgate catccggctt ccgccgtcgg gtggcggtag ccgtcgatga aatcgctgcc	600
ggccgctacc acaaggtgat tctgtcccgt tgtgtcgaag tgcctttcgc gatcgacttt	660
ccgttgacct accggctggg gcgtcggcac aacaccccgg tgaggtcgtt tttgttcgag	720
ttgggcggaa tccgtgctct gggttacagc cccgaactcg tcacggcggt gcgcgccgac	780
ggagtggtag tcaccgagcc gttggccggt acccgcgctt tgggccgtgg tcccgccatt	840
gaccgactgg ctctgatga cctggaatca aactccaaag aaattgtcga gcacgccatt	900
tcagtgcgct cttcgcttga ggagattacc gacatcgccg aaccagggag tgctgcggtc	960
atcgatttca tgacggtgcg cgagcgcggc agtgtgcagc acctcggtc caccatcaga	1020
gcacggttgg atccatcgag cgaccggatg gccgccctgg aagccctttt tcctgctgtc	1080
actgcatccg gaatcccga agcagctggc gttgaggcca tctttcgct cgatgagtgc	1140
ccacgtgggc tgtattccgg tgcggtggtg atgctttcgg cggatggcgg gctagacgcc	1200
gcgctgacgc tgcgggcggc ataccaggtc ggccggcgga cttggctgcg ggccggcgcc	1260
ggcatcatcg aagaatcgga gccagagcgc gaattcgagg agacttgcga aaagctatcc	1320
acattgacgc cttatctggt tgcacgccag taa	1353
<210>	3
<211>	450
<212>	PRT
<213>	结核分枝杆菌
<220>	
<221>	MISC_FEATURE
<223>	菌株CDC1551

<400> 3
 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
 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
 Val Ala Val Asp Glu Ile Ala Ala Gly Arg Tyr His Lys Val Ile Leu
 195 200 205
 Ser Arg Cys Val Glu Val Pro Phe Ala Ile Asp Phe Pro Leu Thr Tyr
 210 215 220
 Arg Leu Gly Arg Arg His Asn Thr Pro Val Arg Ser Phe Leu Leu Gln
 225 230 235 240
 Leu Gly Gly Ile Arg Ala Leu Gly Tyr Ser Pro Glu Leu Val Thr Ala
 245 250 255
 Val Arg Ala Asp Gly Val Val Ile Thr Glu Pro Leu Ala Gly Thr Arg
 260 265 270
 Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu
 275 280 285
 Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser
 290 295 300

Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val
305 310 315 320
Ile Asp Phe Met Thr Val Arg Glu Arg Gly Ser Val Gln His Leu Gly
325 330 335
Ser Thr Ile Arg Ala Arg Leu Asp Pro Ser Ser Asp Arg Met Ala Ala
340 345 350
Leu Glu Ala Leu Phe Pro Ala Val Thr Ala Ser Gly Ile Pro Lys Ala
355 360 365
Ala Gly Val Glu Ala Ile Phe Arg Leu Asp Glu Cys Pro Arg Gly Leu
370 375 380
Tyr Ser Gly Ala Val Val Met Leu Ser Ala Asp Gly Gly Leu Asp Ala
385 390 395 400
Ala Leu Thr Leu Arg Ala Ala Tyr Gln Val Gly Gly Arg Thr Trp Leu
405 410 415
Arg Ala Gly Ala Gly Ile Ile Glu Glu Ser Glu Pro Glu Arg Glu Phe
420 425 430
Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala
435 440 445
Arg Gln
450
<210> 4
<211> 450
<212> PRT
<213> 结核分枝杆菌
<220>
<221> MISC_FEATURE
<223> 菌株F11
<400> 4
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
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	
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	
Val Ala Val Asp Glu Ile Ala Ala Gly Arg Tyr His Lys Val Ile Leu					
195		200		205	
Ser Arg Cys Val Glu Val Pro Phe Ala Ile Asp Phe Pro Leu Thr Tyr					
210		215		220	
Arg Leu Gly Arg Arg His Asn Thr Pro Val Arg Ser Phe Leu Leu Gln					
225		230		235	
Leu Gly Gly Ile Arg Ala Leu Gly Tyr Ser Pro Glu Leu Val Thr Ala					
245		250		255	
Val Arg Ala Asp Gly Val Val Ile Thr Glu Pro Leu Ala Gly Thr Arg					
260		265		270	
Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu					
275		280		285	
Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser					
290		295		300	
Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val					
305		310		315	
Ile Asp Phe Met Thr Val Arg Glu Arg Gly Ser Val Gln His Leu Gly					
325		330		335	
Ser Thr Ile Arg Ala Arg Leu Asp Pro Ser Ser Asp Arg Met Ala Ala					
340		345		350	
Leu Glu Ala Leu Phe Pro Ala Val Thr Ala Ser Gly Ile Pro Lys Ala					
355		360		365	
Ala Gly Val Glu Ala Ile Phe Arg Leu Asp Glu Cys Pro Arg Gly Leu					
370		375		380	
Tyr Ser Gly Ala Val Val Met Leu Ser Ala Asp Gly Gly Leu Asp Ala					
385		390		395	
				400	

Ala Leu Thr Leu Arg Ala Ala Tyr Gln Val Gly Gly Arg Thr Trp Leu
 405 410 415
 Arg Ala Gly Ala Gly Ile Ile Glu Glu Ser Glu Pro Glu Arg Glu Phe
 420 425 430
 Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala
 435 440 445
 Arg Gln
 450
 <210> 5
 <211> 450
 <212> PRT
 <213> 结核分枝杆菌
 <220>
 <221> MISC_FEATURE
 <223> 菌株Haarlem A
 <400> 5
 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
 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
Val Ala Val Asp Glu Ile Ala Ala Gly Arg Tyr His Lys Val Ile Leu		
195	200	205
Ser Arg Cys Val Glu Val Pro Phe Ala Ile Asp Phe Pro Leu Thr Tyr		
210	215	220
Arg Leu Gly Arg Arg His Asn Thr Pro Val Arg Ser Phe Leu Leu Gln		
225	230	235
Leu Gly Gly Ile Arg Ala Leu Gly Tyr Ser Pro Glu Leu Val Thr Ala		
245	250	255
Val Arg Ala Asp Gly Val Val Ile Thr Glu Pro Leu Ala Gly Thr Arg		
260	265	270
Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu		
275	280	285
Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser		
290	295	300
Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val		
305	310	315
Ile Asp Phe Met Thr Val Arg Glu Arg Gly Ser Val Gln His Leu Gly		
325	330	335
Ser Thr Ile Arg Ala Arg Leu Asp Pro Ser Ser Asp Arg Met Ala Ala		
340	345	350
Leu Glu Ala Leu Phe Pro Ala Val Thr Ala Ser Gly Ile Pro Lys Ala		
355	360	365
Ala Gly Val Glu Ala Ile Phe Arg Leu Asp Glu Cys Pro Arg Gly Leu		
370	375	380
Tyr Ser Gly Ala Val Val Met Leu Ser Ala Asp Gly Gly Leu Asp Ala		
385	390	395
Ala Leu Thr Leu Arg Ala Ala Tyr Gln Val Gly Gly Arg Thr Trp Leu		
405	410	415
Arg Ala Gly Ala Gly Ile Ile Glu Glu Ser Glu Pro Glu Arg Glu Phe		
420	425	430
Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala		
435	440	445
Arg Gln		
450		
<210> 6		
<211> 450		
<212> PRT		
<213> 结核分枝杆菌		

<220>

<221> MISC_FEATURE

<223> 菌株C

<400> 6

```

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
Asp Arg Leu Leu Leu Glu Thr Asp Gln Ala Phe Gly Trp Val Ala Leu
          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
Val Ala Val Asp Glu Ile Ala Ala Gly Arg Tyr His Lys Val Ile Leu
          195          200          205
Ser Arg Cys Val Glu Val Pro Phe Ala Ile Asp Phe Pro Leu Thr Tyr
          210          215          220
Arg Leu Gly Arg Arg His Asn Thr Pro Val Arg Ser Phe Leu Leu Gln
225          230          235          240
Leu Gly Gly Ile Arg Ala Leu Gly Tyr Ser Pro Glu Leu Val Thr Ala
          245          250          255
Val Arg Ala Asp Gly Val Val Ile Thr Glu Pro Leu Ala Gly Thr Arg
          260          265          270
Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu

```

275	280	285
Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser		
290	295	300
Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val		
305	310	315
Ile Asp Phe Met Thr Val Arg Glu Arg Gly Ser Val Gln His Leu Gly		
325	330	335
Ser Thr Ile Arg Ala Arg Leu Asp Pro Ser Ser Asp Arg Met Ala Ala		
340	345	350
Leu Glu Ala Leu Phe Pro Ala Val Thr Ala Ser Gly Ile Pro Lys Ala		
355	360	365
Ala Gly Val Glu Ala Ile Phe Arg Leu Asp Glu Cys Pro Arg Gly Leu		
370	375	380
Tyr Ser Gly Ala Val Val Met Leu Ser Ala Asp Gly Gly Leu Asp Ala		
385	390	395
Ala Leu Thr Leu Arg Ala Ala Tyr Gln Val Gly Gly Arg Thr Trp Leu		
405	410	415
Arg Ala Gly Ala Gly Ile Ile Glu Glu Ser Glu Pro Glu Arg Glu Phe		
420	425	430
Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala		
435	440	445
Arg Gln		
450		
<210> 7		
<211> 450		
<212> PRT		
<213> 牛分枝杆菌		
<220>		
<221> MISC_FEATURE		
<223> 菌株BCG		
<400> 7		
Met Ser Glu Leu Ser Val Ala Thr Gly Ala Val Ser Thr Ala Ser Ser		
1	5	10
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
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
Val Ala Val Asp Glu Ile Ala Ala Gly Arg Tyr His Lys Val Ile Leu			
	195	200	205
Ser Arg Cys Val Glu Val Pro Phe Ala Ile Asp Phe Pro Leu Thr Tyr			
210	215	220	
Arg Leu Gly Arg Arg His Asn Thr Pro Val Arg Ser Phe Leu Leu Gln			
225	230	235	240
Leu Gly Gly Ile Arg Ala Leu Gly Tyr Ser Pro Glu Leu Val Thr Ala			
	245	250	255
Val Arg Ala Asp Gly Val Val Ile Thr Glu Pro Leu Ala Gly Thr Arg			
	260	265	270
Ala Leu Gly Arg Gly Pro Ala Ile Asp Arg Leu Ala Arg Asp Asp Leu			
	275	280	285
Glu Ser Asn Ser Lys Glu Ile Val Glu His Ala Ile Ser Val Arg Ser			
290	295	300	
Ser Leu Glu Glu Ile Thr Asp Ile Ala Glu Pro Gly Ser Ala Ala Val			
305	310	315	320
Ile Asp Phe Met Thr Val Arg Glu Arg Gly Ser Val Gln His Leu Gly			
	325	330	335
Ser Thr Ile Arg Ala Arg Leu Asp Pro Ser Ser Asp Arg Met Ala Ala			
	340	345	350
Leu Glu Ala Leu Phe Pro Ala Val Thr Ala Ser Gly Ile Pro Lys Ala			
	355	360	365
Ala Gly Val Glu Ala Ile Phe Arg Leu Asp Glu Cys Pro Arg Gly Leu			

370	375	380
Tyr Ser Gly Ala Val Val Met Leu Ser Ala Asp Gly Gly Leu Asp Ala		
385	390	395
Ala Leu Thr Leu Arg Ala Ala Tyr Gln Val Gly Gly Arg Thr Trp Leu		400
	405	410
Arg Ala Gly Ala Gly Ile Ile Glu Glu Ser Glu Pro Glu Arg Glu Phe		415
	420	425
Glu Glu Thr Cys Glu Lys Leu Ser Thr Leu Thr Pro Tyr Leu Val Ala		430
	435	440
Arg Gln		445
450		
<210> 8		
<211> 110		
<212> PRT		
<213> 结核分枝杆菌		
<220>		
<221> mat_peptide		
<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
Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala Ala Leu		
5	10	15
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
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
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
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
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
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 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																
<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			
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
Pro His Ser Pro Ala Ala Gly			
385	390		
<210> 14			
<211> 392			
<212> PRT			
<213> 结核分枝杆菌			
<400> 14			
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		
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					
Lys	Val	Leu	Val	Arg	Asn	Val	Val										
385						390											
<210>	15																
<211>	423																

<212> PRT

<213> 结核分枝杆菌

<400> 15

```

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
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
 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> 16

<211> 95

<212> PRT

<213> 结核分枝杆菌

<220>

<221> INIT_MET

<222> (1) .. (1)

<220>

<221> mat_peptide

<222> (2) .. (95)

<400> 16

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
 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> 17
 <211> 338
 <212> PRT
 <213> 结核分枝杆菌
 <220>
 <221> mat_peptide
 <222> (43) .. (338)
 <400> 17
 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
 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
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
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> 18		
<211> 325		
<212> PRT		
<213> 结核分枝杆菌		
<220>		
<221> mat_peptide		
<222> (41) .. (325)		
<400> 18		
Met Thr Asp Val Ser Arg Lys Ile Arg Ala Trp Gly Arg Arg Leu Met		
-40	-35	-30
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
Gly Leu Arg Ala Gln Asp Asp Tyr Asn Gly Trp Asp Ile Asn Thr Pro		
45	50	55
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
 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> 19
 <211> 144
 <212> PRT
 <213> 结核分枝杆菌
 <220>
 <221> INIT_MET
 <222> (1) .. (1)
 <220>
 <221> mat_peptide
 <222> (2) .. (144)
 <400> 19
 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
 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> 20

<211> 228

<212> PRT

<213> 结核分枝杆菌

<220>

<221> mat_peptide

<222> (24) .. (228)

<400> 20

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
 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>	21		
<211>	355		
<212>	PRT		
<213>	结核分枝杆菌		
<220>			
<221>	mat_peptide		
<222>	(33) .. (355)		
<400>	21		
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	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 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	240
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>	323		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	成熟Mtb32A的Ser/Ala突变体		
<400>	22		
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
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
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
Pro Pro Ala		320
<210>	23	

<211> 96

<212> PRT

<213> 结核分枝杆菌

<400> 23

```

Met Ser Gln Ile Met Tyr Asn Tyr Pro Ala Met Leu Gly His Ala Gly
1           5           10           15
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> 24

<211> 723

<212> PRT

<213> 人工序列

<220>

<223> Mtb72f

<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
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
Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser		
	165	170
Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser		
	180	185
Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro Tyr		
	195	200
Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala		
	210	215
Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr		
225	230	235
Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu		
	245	250
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn		
	260	265
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe		
	275	280
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe		
	290	295
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala		
305	310	315
Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met		
	325	330
Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly		
	340	345
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro		
	355	360
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met		
	370	375
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
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly		
	420	425
Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala		
		430

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
Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly		480
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
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
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 Ser		
690	695	700
Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr		
705	710	715
Ala Ala Ser		720
<210> 25		
<211> 723		
<212> PRT		

<213> 人工序列

<220>

<223> M72

<400> 25

```

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
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
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
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
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			
<210>	26		
<211>	702		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	Mtb71f		
<400>	26		
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
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
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
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
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
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
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	Ala			
		450					455				460							
Thr	Gln	Ala	Ala	Gly	Ala	Gly	Ala	Val	Ala	Asp	Ala	Gln	Ala	Thr	Leu			
465					470					475					480			
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					
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>	27																	
<211>	920																	
<212>	PRT																	
<213>	人工序列																	
<220>																		

<223> M72-Mtb9.9-Mtb9.8

<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	
Ala	Asp	Ala	Leu	Asn	Gly	His	His	Pro	Gly	Asp	Val	Ile	Ser	Val	Thr
						100								105	
Trp	Gln	Thr	Lys	Ser	Gly	Gly	Thr	Arg	Thr	Gly	Asn	Val	Thr	Leu	Ala
						115								120	
Glu	Gly	Pro	Pro	Ala	Glu	Phe	Met	Val	Asp	Phe	Gly	Ala	Leu	Pro	Pro
						130								135	
Glu	Ile	Asn	Ser	Ala	Arg	Met	Tyr	Ala	Gly	Pro	Gly	Ser	Ala	Ser	Leu
145								150						155	
Val	Ala	Ala	Ala	Gln	Met	Trp	Asp	Ser	Val	Ala	Ser	Asp	Leu	Phe	Ser
						165								170	
Ala	Ala	Ser	Ala	Phe	Gln	Ser	Val	Val	Trp	Gly	Leu	Thr	Val	Gly	Ser
						180								185	
Trp	Ile	Gly	Ser	Ser	Ala	Gly	Leu	Met	Val	Ala	Ala	Ala	Ser	Pro	Tyr
						195								200	
Val	Ala	Trp	Met	Ser	Val	Thr	Ala	Gly	Gln	Ala	Glu	Leu	Thr	Ala	Ala
						210								215	
Gln	Val	Arg	Val	Ala	Ala	Ala	Ala	Tyr	Glu	Thr	Ala	Tyr	Gly	Leu	Thr
225								230						235	
Val	Pro	Pro	Pro	Val	Ile	Ala	Glu	Asn	Arg	Ala	Glu	Leu	Met	Ile	Leu
						245								250	
Ile	Ala	Thr	Asn	Leu	Leu	Gly	Gln	Asn	Thr	Pro	Ala	Ile	Ala	Val	Asn
						260								265	
Glu	Ala	Glu	Tyr	Gly	Glu	Met	Trp	Ala	Gln	Asp	Ala	Ala	Ala	Met	Phe
						275								280	
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
Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met		320
	325	330
Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly		335
	340	345
Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro		350
	355	360
His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met		365
	370	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		400
	405	410
Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly		415
	420	430
Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala		
	435	440
Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln		445
	450	460
Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser		
465	470	475
Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly		480
	485	490
Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val		495
	500	505
Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile		510
	515	520
Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu		525
	530	540
Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val		
545	550	555
Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr		560
	565	570
Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val		575
	580	585
Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln		590
	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 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
Ser Thr Tyr Thr Gly Phe Pro Trp			

				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			
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			
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
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
Ala Ala Ser Ser Gly Phe Ser Arg Pro Gly Leu Pro Val Glu Tyr Leu		
725	730	735
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
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
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> 29			
<211> 1148			
<212> PRT			
<213> 人工序列			
<220>			
<223> M114			
<400> 29			
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
Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu		
145	150	155
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
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
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	
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
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
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

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>	30		
<211>	9		
<212>	PRT		
<213>	结核分枝杆菌		
<400>	30		
Trp Val	Leu Ala Ala Gly Val	Gln Ala	
1		5	
<210>	31		
<211>	9		
<212>	PRT		
<213>	结核分枝杆菌		
<400>	31		
Val Leu	Ala Ala Gly Val	Gln Ala Met	
1		5	
<210>	32		
<211>	9		
<212>	PRT		
<213>	结核分枝杆菌		
<400>	32		
Val Ile	Arg Asp Gly Val	Thr Arg Arg	
1		5	
<210>	33		
<211>	9		

<212> PRT
<213> 结核分枝杆菌
<400> 33
Leu Ala Pro His Thr Pro Leu Ala Arg
1 5
<210> 34
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 34
Ile Arg Leu Phe Asp Ala Gly Ile Arg
1 5
<210> 35
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 35
Ile Arg His Arg Glu Ala Ile Asp Arg
1 5
<210> 36
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 36
Leu Leu Ala Thr Gly Val Arg Glu Val
1 5
<210> 37
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 37
Phe Arg Arg Arg Val Ala Val Ala Val
1 5
<210> 38
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 38
Phe Ala Ile Asp Phe Pro Leu Thr Tyr

1 5
<210> 39
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 39
Phe Pro Leu Thr Tyr Arg Leu Gly Arg
1 5
<210> 40
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 40
Tyr Arg Leu Gly Arg Arg His Asn Thr
1 5
<210> 41
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 41
Val Arg Ser Phe Leu Leu Gln Leu Gly
1 5
<210> 42
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 42
Leu Leu Gln Leu Gly Gly Ile Arg Ala
1 5
<210> 43
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 43
Leu Val Thr Ala Val Arg Ala Asp Gly
1 5
<210> 44
<211> 9
<212> PRT

<213> 结核分枝杆菌

<400> 44

Val Arg Ala Asp Gly Val Val Ile Thr

1 5

<210> 45

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 45

Val Val Ile Thr Glu Pro Leu Ala Gly

1 5

<210> 46

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 46

Ile Val Glu His Ala Ile Ser Val Arg

1 5

<210> 47

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 47

Ile Asp Phe Met Thr Val Arg Glu Arg

1 5

<210> 48

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 48

Phe Arg Leu Asp Glu Cys Pro Arg Gly

1 5

<210> 49

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 49

Leu Tyr Ser Gly Ala Val Val Met Leu

1 5

<210> 50
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 50
Tyr Ser Gly Ala Val Val Met Leu Ser
1 5
<210> 51
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 51
Val Val Met Leu Ser Ala Asp Gly Gly
1 5
<210> 52
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 52
Trp Leu Arg Ala Gly Ala Gly Ile Ile
1 5
<210> 53
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 53
Glu Leu Ser Val Ala Thr Gly Ala Val
1 5
<210> 54
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 54
Ser Val Ala Thr Gly Ala Val Ser Thr
1 5
<210> 55
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 55

Ile Pro Met Pro Ala Gly Val Asn Pro

1 5

<210> 56

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 56

Pro Met Pro Ala Gly Val Asn Pro Ala

1 5

<210> 57

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 57

Pro Ala Gly Val Asn Pro Ala Asp Leu

1 5

<210> 58

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 58

Gly Val Asn Pro Ala Asp Leu Ala Ala

1 5

<210> 59

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 59

Asn Pro Ala Asp Leu Ala Ala Glu Leu

1 5

<210> 60

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 60

Asp Leu Ala Ala Glu Leu Ala Ala Val

1 5

<210> 61

<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 61
Leu Ala Ala Val Val Thr Glu Ser Val
1 5
<210> 62
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 62
Val Thr Glu Ser Val Asp Glu Asp Tyr
1 5
<210> 63
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 63
Thr Glu Ser Val Asp Glu Asp Tyr Leu
1 5
<210> 64
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 64
Ser Val Asp Glu Asp Tyr Leu Leu Tyr
1 5
<210> 65
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 65
Tyr Glu Cys Asp Gly Gln Trp Val Leu
1 5
<210> 66
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 66

Gly Gln Trp Val Leu Ala Ala Gly Val

1 5

<210> 67

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 67

Val Leu Ala Ala Gly Val Gln Ala Met

1 5

<210> 68

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 68

Glu Leu Asp Ser Asp Glu Leu Arg Val

1 5

<210> 69

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 69

Asp Ser Asp Glu Leu Arg Val Ile Arg

1 5

<210> 70

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 70

Gln Gln Trp Ser Gly Arg Pro Gly Ala

1 5

<210> 71

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 71

Gly Arg Pro Gly Ala Ala Leu Gly Glu

1 5

<210> 72

<211> 9

<212> PRT
<213> 结核分枝杆菌
<400> 72
Arg Pro Gly Ala Ala Leu Gly Glu Ala
1 5
<210> 73
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 73
Ala Leu Gly Glu Ala Val Asp Arg Leu
1 5
<210> 74
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 74
Gly Glu Ala Val Asp Arg Leu Leu Leu
1 5
<210> 75
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 75
Ala Phe Gly Trp Val Ala Phe Glu Phe
1 5
<210> 76
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 76
Arg Leu Ala Pro His Thr Pro Leu Ala
1 5
<210> 77
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 77
Leu Ala Pro His Thr Pro Leu Ala Arg

1 5
<210> 78
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 78
Ala Pro His Thr Pro Leu Ala Arg Val
1 5
<210> 79
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 79
Pro His Thr Pro Leu Ala Arg Val Phe
1 5
<210> 80
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 80
Pro Leu Ala Arg Val Phe Ser Pro Arg
1 5
<210> 81
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 81
Arg Val Phe Ser Pro Arg Thr Arg Ile
1 5
<210> 82
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 82
Val Phe Ser Pro Arg Thr Arg Ile Met
1 5
<210> 83
<211> 9
<212> PRT

<213> 结核分枝杆菌

<400> 83

Val Ser Glu Lys Glu Ile Arg Leu Phe

1 5

<210> 84

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 84

Ala Gly Ile Arg His Arg Glu Ala Ile

1 5

<210> 85

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 85

Leu Leu Ala Thr Gly Val Arg Glu Val

1 5

<210> 86

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 86

Arg Glu Val Pro Gln Ser Arg Ser Val

1 5

<210> 87

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 87

Val Pro Gln Ser Arg Ser Val Asp Val

1 5

<210> 88

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 88

Val Ser Asp Asp Pro Ser Gly Phe Arg

1 5

<210> 89

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 89

Asp Pro Ser Gly Phe Arg Arg Arg Val

1 5

<210> 90

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 90

Ser Gly Phe Arg Arg Arg Val Ala Val

1 5

<210> 91

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 91

Gly Phe Arg Arg Arg Val Ala Val Ala

1 5

<210> 92

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 92

Phe Arg Arg Arg Val Ala Val Ala Val

1 5

<210> 93

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 93

Arg Val Ala Val Ala Val Asp Glu Ile

1 5

<210> 94

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 94
Ile Ala Ala Gly Arg Tyr His Lys Val
1 5
<210> 95
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 95
Ala Gly Arg Tyr His Lys Val Ile Leu
1 5
<210> 96
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 96
Arg Tyr His Lys Val Ile Leu Ser Arg
1 5
<210> 97
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 97
Val Ile Leu Ser Arg Cys Val Glu Val
1 5
<210> 98
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 98
Val Pro Phe Ala Ile Asp Phe Pro Leu
1 5
<210> 99
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 99
Phe Ala Ile Asp Phe Pro Leu Thr Tyr
1 5
<210> 100

<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 100
Ile Asp Phe Pro Leu Thr Tyr Arg Leu
1 5
<210> 101
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 101
Leu Gly Arg Arg His Asn Thr Pro Val
1 5
<210> 102
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 102
Arg Arg His Asn Thr Pro Val Arg Ser
1 5
<210> 103
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 103
Asn Thr Pro Val Arg Ser Phe Leu Leu
1 5
<210> 104
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 104
Pro Val Arg Ser Phe Leu Leu Gln Leu
1 5
<210> 105
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 105

Arg Ala Leu Gly Tyr Ser Pro Glu Leu

1 5

<210> 106

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 106

Ala Leu Gly Tyr Ser Pro Glu Leu Val

1 5

<210> 107

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 107

Tyr Ser Pro Glu Leu Val Thr Ala Val

1 5

<210> 108

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 108

Ser Pro Glu Leu Val Thr Ala Val Arg

1 5

<210> 109

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 109

Ala Val Arg Ala Asp Gly Val Val Ile

1 5

<210> 110

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 110

Ile Thr Glu Pro Leu Ala Gly Thr Arg

1 5

<210> 111

<211> 9

<212> PRT
<213> 结核分枝杆菌
<400> 111
Glu Pro Leu Ala Gly Thr Arg Ala Leu
1 5
<210> 112
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 112
Gly Thr Arg Ala Leu Gly Arg Gly Pro
1 5
<210> 113
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 113
Arg Ala Leu Gly Arg Gly Pro Ala Ile
1 5
<210> 114
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 114
Leu Glu Ser Asn Ser Lys Glu Ile Val
1 5
<210> 115
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 115
Glu Ile Val Glu His Ala Ile Ser Val
1 5
<210> 116
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 116
Ile Val Glu His Ala Ile Ser Val Arg

1 5
<210> 117
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 117
His Ala Ile Ser Val Arg Ser Ser Leu
1 5
<210> 118
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 118
Ser Val Arg Ser Ser Leu Glu Glu Ile
1 5
<210> 119
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 119
Ser Leu Glu Glu Ile Thr Asp Ile Ala
1 5
<210> 120
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 120
Asp Ile Ala Glu Pro Gly Ser Ala Ala
1 5
<210> 121
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 121
Ala Glu Pro Gly Ser Ala Ala Val Ile
1 5
<210> 122
<211> 9
<212> PRT

<213> 结核分枝杆菌

<400> 122

Arg Glu Arg Gly Ser Val Gln His Leu

1 5

<210> 123

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 123

Ser Val Gln His Leu Gly Ser Thr Ile

1 5

<210> 124

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 124

Arg Leu Asp Pro Ser Ser Asp Arg Met

1 5

<210> 125

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 125

Asp Pro Ser Ser Asp Arg Met Ala Ala

1 5

<210> 126

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 126

Ser Ser Asp Arg Met Ala Ala Leu Glu

1 5

<210> 127

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 127

Asp Arg Met Ala Ala Leu Glu Ala Leu

1 5

<210> 128
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 128
Arg Met Ala Ala Leu Glu Ala Leu Phe
1 5
<210> 129
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 129
Ala Ala Leu Glu Ala Leu Phe Pro Ala
1 5
<210> 130
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 130
Ala Leu Glu Ala Leu Phe Pro Ala Val
1 5
<210> 131
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 131
Leu Glu Ala Leu Phe Pro Ala Val Thr
1 5
<210> 132
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 132
Phe Pro Ala Val Thr Ala Ser Gly Ile
1 5
<210> 133
<211> 9
<212> PRT
<213> 结核分枝杆菌

<400> 133
Ala Val Thr Ala Ser Gly Ile Pro Lys
1 5
<210> 134
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 134
Val Thr Ala Ser Gly Ile Pro Lys Ala
1 5
<210> 135
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 135
Gly Ile Pro Lys Ala Ala Gly Val Glu
1 5
<210> 136
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 136
Ile Pro Lys Ala Ala Gly Val Glu Ala
1 5
<210> 137
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 137
Lys Ala Ala Gly Val Glu Ala Ile Phe
1 5
<210> 138
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 138
Ala Gly Val Glu Ala Ile Phe Arg Leu
1 5
<210> 139

<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 139
Arg Leu Asp Glu Cys Pro Arg Gly Leu
1 5
<210> 140
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 140
Cys Pro Arg Gly Leu Tyr Ser Gly Ala
1 5
<210> 141
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 141
Gly Leu Tyr Ser Gly Ala Val Val Met
1 5
<210> 142
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 142
Leu Tyr Ser Gly Ala Val Val Met Leu
1 5
<210> 143
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 143
Val Met Leu Ser Ala Asp Gly Gly Leu
1 5
<210> 144
<211> 9
<212> PRT
<213> 结核分枝杆菌
<400> 144

Gly Leu Asp Ala Ala Leu Thr Leu Arg

1 5

<210> 145

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 145

Leu Asp Ala Ala Leu Thr Leu Arg Ala

1 5

<210> 146

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 146

Ala Ala Leu Thr Leu Arg Ala Ala Tyr

1 5

<210> 147

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 147

Leu Thr Leu Arg Ala Ala Tyr Gln Val

1 5

<210> 148

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 148

Ala Tyr Gln Val Gly Gly Arg Thr Trp

1 5

<210> 149

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 149

Thr Trp Leu Arg Ala Gly Ala Gly Ile

1 5

<210> 150

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 150

Glu Glu Ser Glu Pro Glu Arg Glu Phe

1 5

<210> 151

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 151

Arg Glu Phe Glu Glu Thr Cys Glu Lys

1 5

<210> 152

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 152

Lys Leu Ser Thr Leu Thr Pro Tyr Leu

1 5

<210> 153

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 153

Ser Thr Leu Thr Pro Tyr Leu Val Ala

1 5

<210> 154

<211> 9

<212> PRT

<213> 结核分枝杆菌

<400> 154

Thr Leu Thr Pro Tyr Leu Val Ala Arg

1 5

<210> 155

<211> 1053

<212> PRT

<213> 结核分枝杆菌

<400> 155

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	
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
Asn Thr Gly Ser Tyr Asn Met Gly Asp Phe Asn Pro Gly Ser Ser Asn		
385	390	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	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	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
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	695	700	
Gly Leu Pro Ser Ile	Gly Trp Asp Val Phe	Leu Ser Thr Pro Arg	Ile
705	710	715	720
Thr Val Pro Ala Phe	Gly Ile Pro Phe Thr	Leu Gln Phe Gln Thr	Asn
725	730	735	
Val Pro Ala Leu Gln	Pro Pro Gly Gly Gly	Leu Ser Thr Phe Thr	Asn
740	745	750	
Gly Ala Leu Ile Phe	Gly Glu Phe Asp Leu	Pro Gln Leu Val Val	His
755	760	765	
Pro Tyr Thr Leu Thr	Gly Pro Ile Val Ile	Gly Ser Phe Phe Leu	Pro
770	775	780	
Ala Phe Asn Ile Pro	Gly Ile Asp Val Pro	Ala Ile Asn Val Asp	Gly
785	790	795	800
Phe Thr Leu Pro Gln	Ile Thr Thr Pro Ala	Ile Thr Thr Pro	Glu Phe
805	810	815	
Ala Ile Pro Pro Ile	Gly Val Gly Gly Phe	Thr Leu Pro Gln Ile	Thr
820	825	830	
Thr Gln Glu Ile Ile	Thr Pro Glu Leu Thr	Ile Asn Ser Ile Gly	Val
835	840	845	
Gly Gly Phe Thr Leu	Pro Gln Ile Thr Thr	Pro Pro Ile Thr Thr	Pro
850	855	860	
Pro Leu Thr Ile Asp	Pro Ile Asn Leu Thr	Gly Phe Thr Leu Pro	Gln
865	870	875	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
 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> 156
 <211> 324
 <212> PRT
 <213> 结核分枝杆菌
 <400> 156
 Met Ser Asp Gln Val Pro Lys Pro His Arg His His Ile Trp Arg Ile
 1 5 10 15
 Thr Arg Arg Thr Leu Ser Lys Ser Trp Asp Asp Ser Ile Phe Ser Glu
 20 25 30
 Ser Ala Gln Ala Ala Phe Trp Ser Ala Leu Ser Leu Pro Pro Leu Leu
 35 40 45
 Leu Gly Met Leu Gly Ser Leu Ala Tyr Val Ala Pro Leu Phe Gly Pro
 50 55 60
 Asp Thr Leu Pro Ala Ile Glu Lys Ser Ala Leu Ser Thr Ala His Ser
 65 70 75 80
 Phe Phe Ser Pro Ser Val Val Asn Glu Ile Ile Glu Pro Thr Ile Gly
 85 90 95
 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
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
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
Gly Thr Thr Thr Tyr Ser Ser Thr Ala Gln His Ser Ala Val Ala Ala		
305	310	315
Glu Pro Pro Ser		320

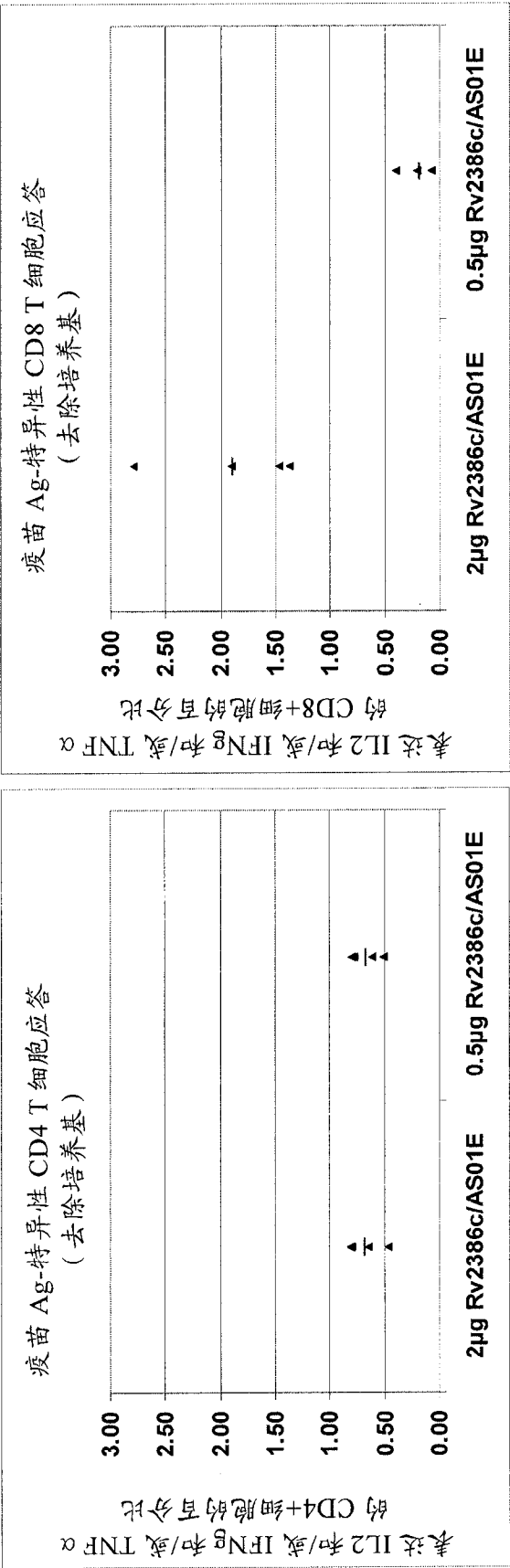


图 1

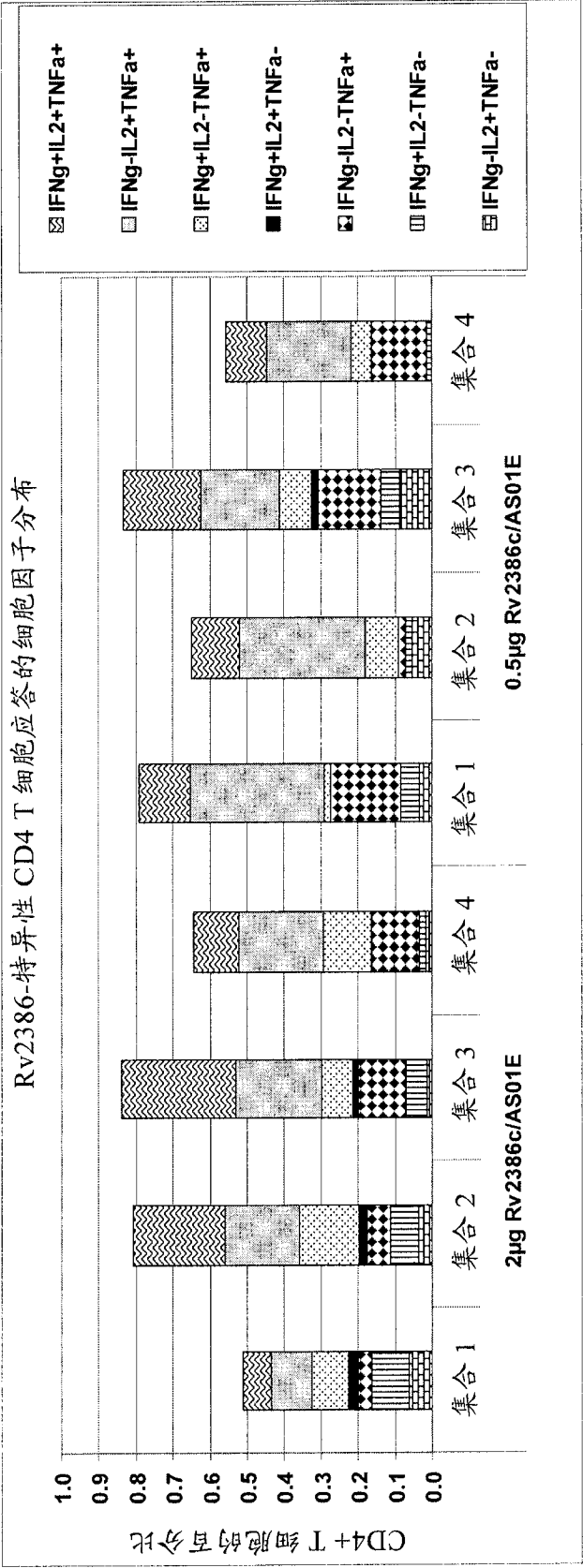


图 2

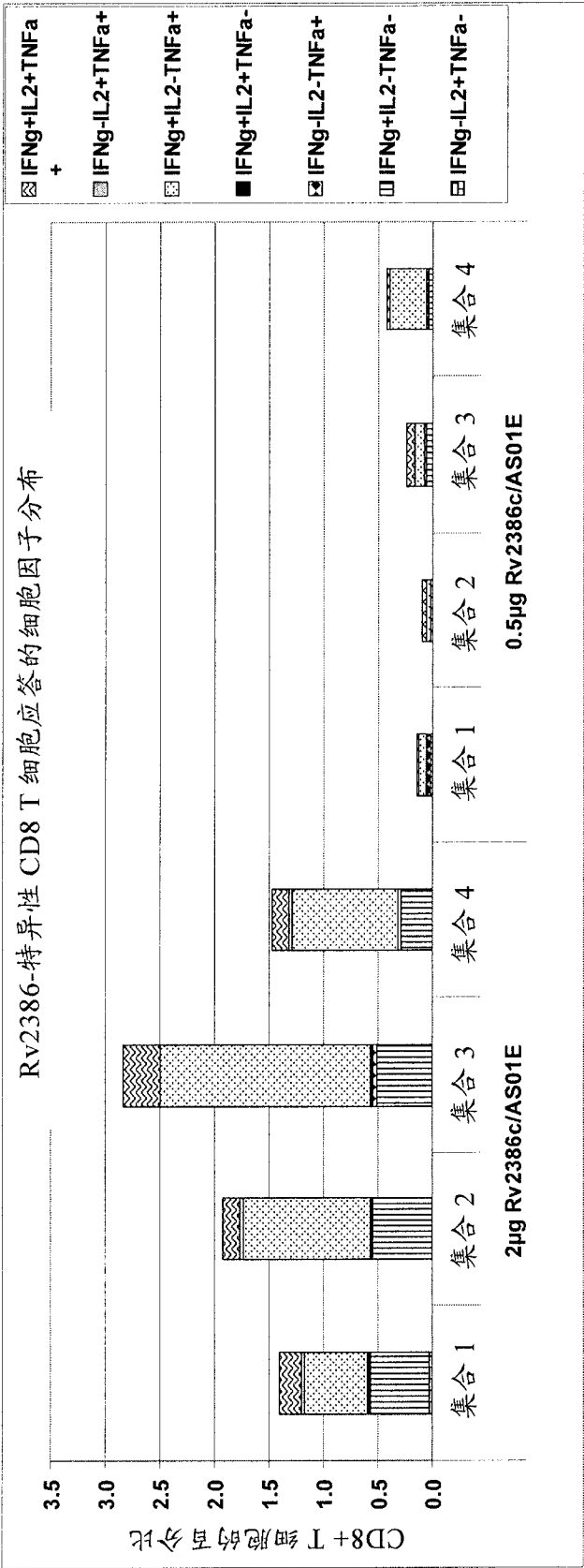


图 3

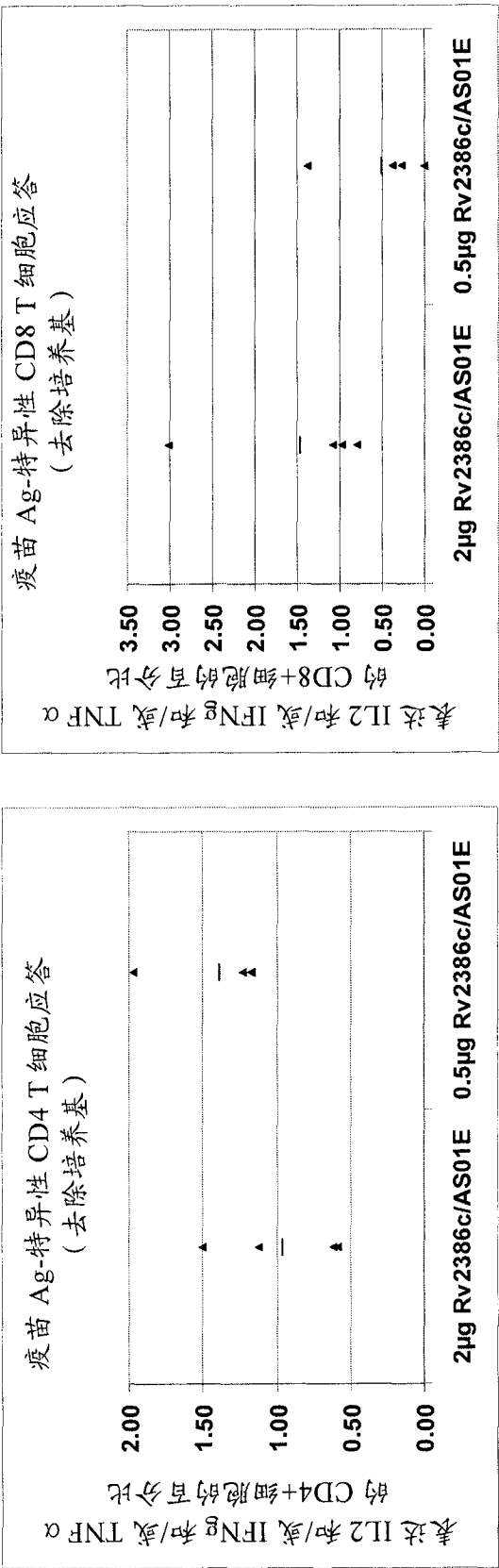


图 4

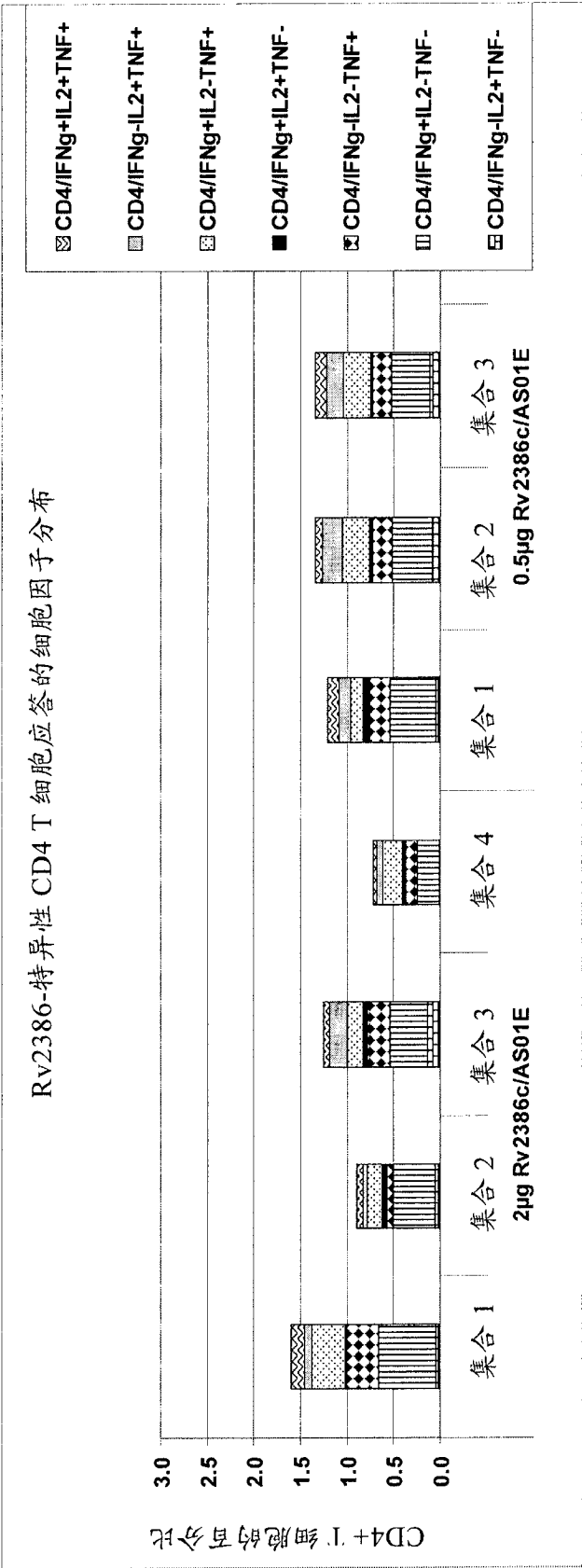


图 5

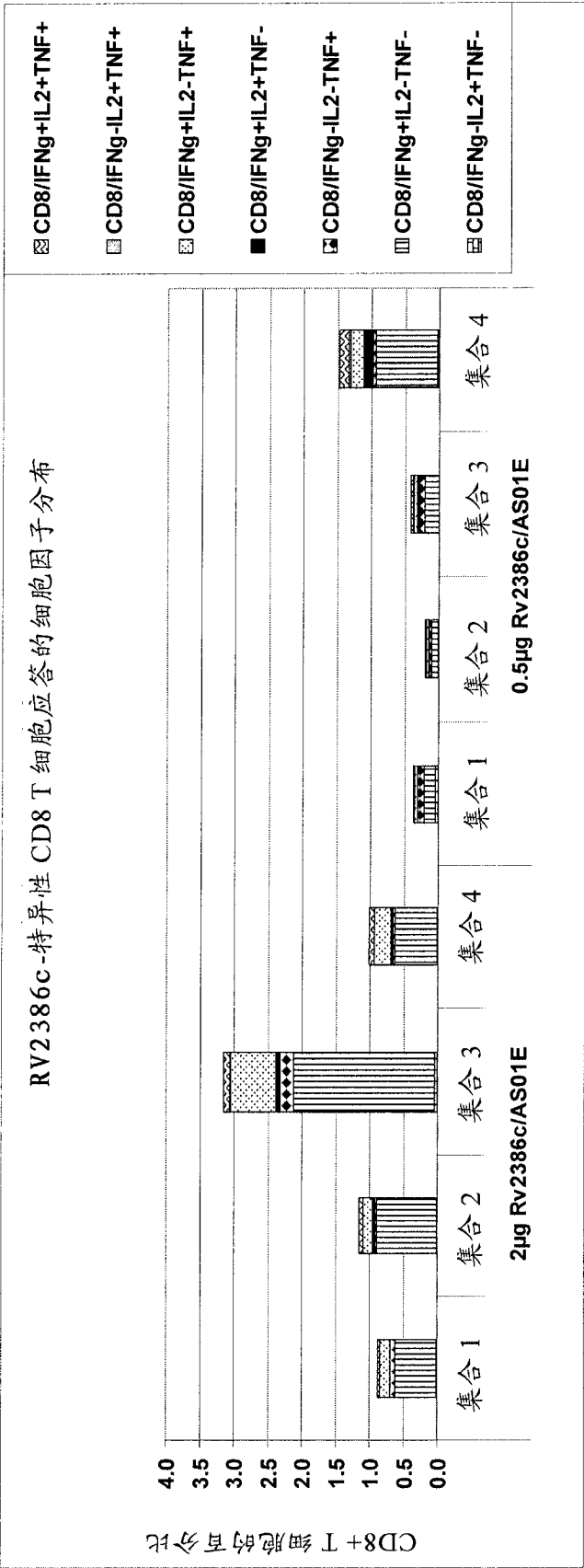


图 6

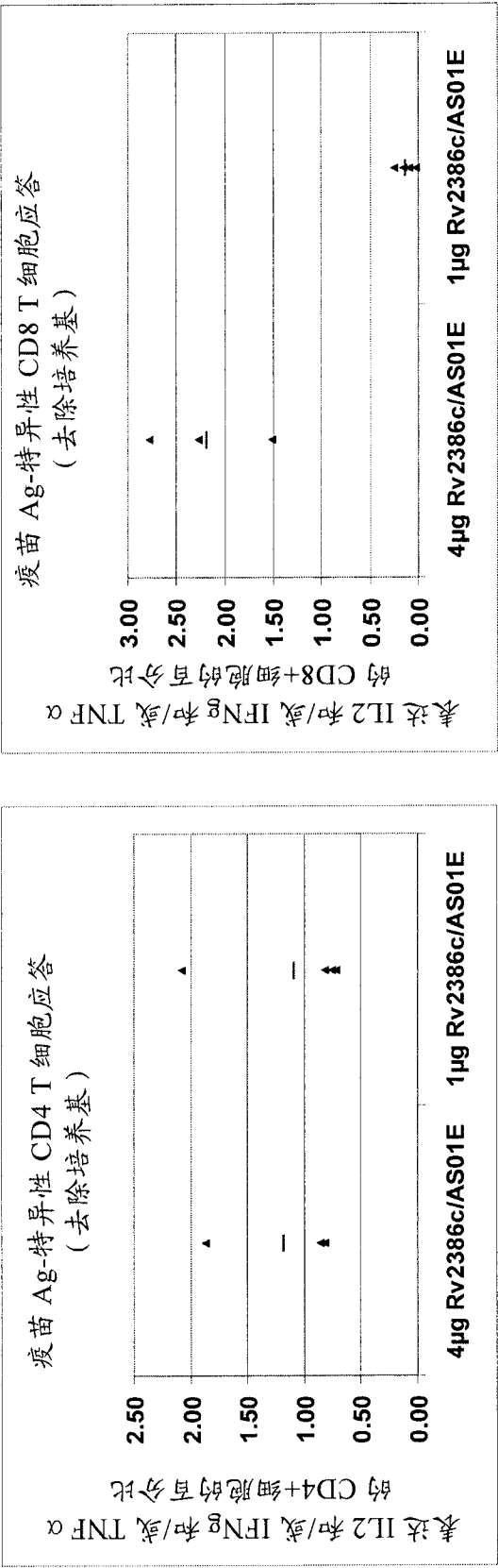


图 7

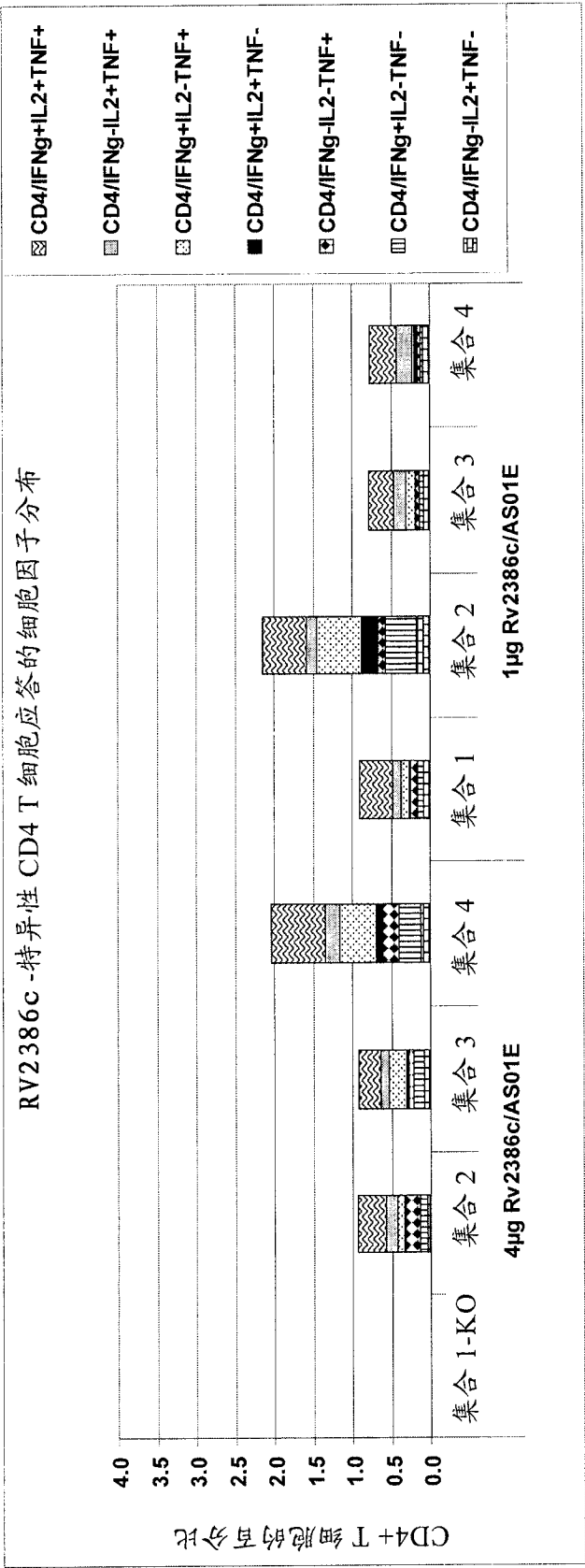


图 8

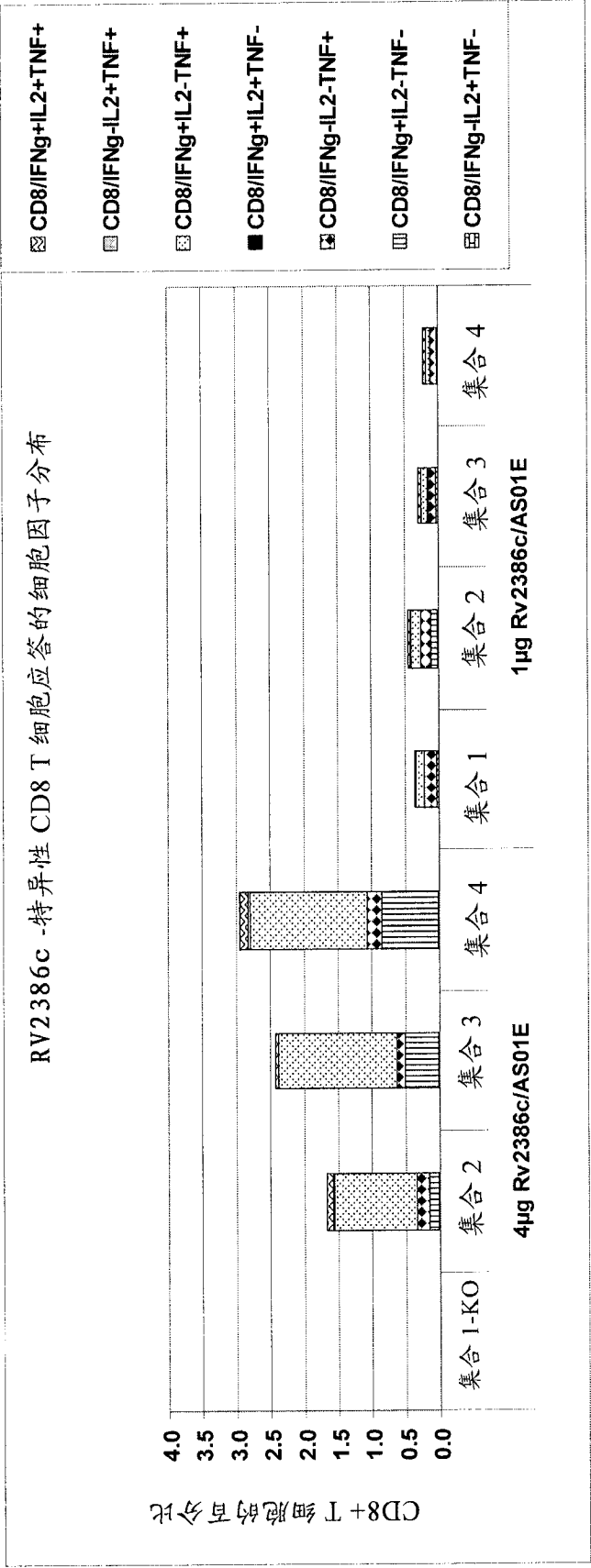


图 9