

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.



[12] 发明专利申请公布说明书

[21] 申请号 200680044019. X

A61K 39/015 (2006.01)

C07K 14/445 (2006.01)

C07K 16/20 (2006.01)

[43] 公开日 2009 年 4 月 22 日

[11] 公开号 CN 101415436A

[22] 申请日 2006.11.23

[21] 申请号 200680044019. X

[30] 优先权

[32] 2005.11.23 [33] US [31] 60/739,973

[86] 国际申请 PCT/IB2006/003967 2006.11.23

[87] 国际公布 WO2007/060550 英 2007.5.31

[85] 进入国家阶段日期 2008.5.23

[71] 申请人 巴斯德研究院

地址 法国巴黎

共同申请人 国立科学研究中心

[72] 发明人 S·朗埃克 H·波尔森 R·佩罗
F·纳托

[74] 专利代理机构 永新专利商标代理有限公司
代理人 林晓红

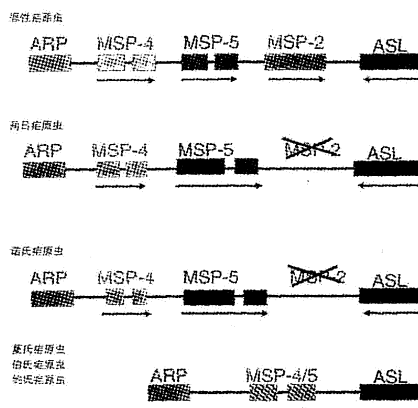
权利要求书 5 页 说明书 94 页 附图 35 页

[54] 发明名称

重组恶性疟原虫裂殖子表面蛋白 4 和 5 及其用途

[57] 摘要

因此，本发明提供了构建体，其中编码恶性疟原虫 MSP4 和 MSP5 的核酸和所形成的多肽已被修饰。更具体地，本发明提供了编码重组 MSP4 和 MSP5 多肽的构建体，这些多肽在杆状病毒-昆虫细胞表达体系中以可溶分泌多肽表达。意外地发现，所述重组多肽含有一 C 端 EGF 样结构域，其在所述多肽中适当地折叠。



几种不同疟原虫物种的msp4和msp5基因位点的基因组结构。

1. 一种纯化的核酸分子，其包含 SEQ ID NO: 1、2、3、4、5、6、7、28、29 或 8 的 DNA 序列。
2. 一种纯化的核酸分子，其编码包含 SEQ ID NO: 9、10、11、12、13、14、15、16、17、18、19、20、21、30 或 22 的序列的氨基酸序列。
3. 一种纯化的核酸分子，其与包含权利要求 1 或 2 任一项所述核酸序列的变性、双链 DNA 的任一链在 42°C、50%甲酰胺和 6XSSC 的中度严格性条件和 60°C、0.5XSSC、0.1 % SDS 的洗涤条件下杂交。
4. 如权利要求 3 所述的纯化的核酸分子，其中所述分离的核酸分子源自 SEQ ID NO:1-8 的体外诱变。
5. 一种纯化的核酸分子，其来自 SEQ ID NO: 1、2、3、4、5、6、7 或 8 的遗传密码简并。
6. 一种纯化的核酸分子，其编码重组 MSP4 或 MSP5 多肽，重组 MSP4 或 MSP5 多肽 DNA 的等位基因变体或重组 MSP4 或 MSP5 多肽 DNA 的同系物。
7. 一种重组载体，其指导选自由权利要求 1、2、5 和 6 所述的纯化的核酸分子组成的组中的核酸分子的表达。
8. 一种重组载体，其指导权利要求 3 的核酸分子的表达。
9. 一种重组载体，其指导权利要求 4 的核酸分子的表达。
10. 一种纯化的多肽，其由选自由权利要求 1、2、5 和 6 所述的纯化的核酸分子组成的组中的核酸分子编码。
11. 如权利要求 10 所述的纯化的多肽，其为非糖基化形式。
12. 一种纯化的多肽，其由权利要求 3 的核酸分子编码。
13. 如权利要求 12 所述的纯化的多肽，其为非糖基化形式。
14. 一种纯化的多肽，其由权利要求 4 的核酸分子编码。
15. 如权利要求 14 所述的纯化的多肽，其为非糖基化形式。
16. 纯化抗体，其结合权利要求 10 的多肽。
17. 如权利要求 16 所述的纯化抗体，其中所述抗体为单克隆抗体。
18. 纯化抗体，其结合权利要求 12 的多肽。

19. 如权利要求 18 所述的纯化抗体, 其中所述抗体为单克隆抗体。
20. 纯化抗体, 其结合权利要求 14 的多肽。
21. 如权利要求 20 所述的纯化抗体, 其中所述抗体为单克隆抗体。
22. 一种宿主细胞, 其以权利要求 7 的载体转染、转导或感染。
23. 用于生产重组 MSP4 或 MSP5 多肽的方法, 其包括在促进表达的条件下培养权利要求 22 的宿主细胞, 和从培养基回收所述多肽。
24. 如权利要求 23 所述的方法, 其中所述宿主细胞选自昆虫细胞、细菌细胞、酵母细胞、植物细胞和动物细胞。
25. 一种宿主细胞, 其以权利要求 8 所述载体转染、转导或感染。
26. 用于生产重组 MSP4 或 MSP5 多肽的方法, 其包括在促进表达的条件下培养权利要求 25 的宿主细胞, 和从培养基或细胞沉淀回收所述多肽。
27. 如权利要求 26 所述的方法, 其中所述宿主细胞选自昆虫细胞、细菌细胞、酵母细胞、植物细胞和动物细胞。
28. 一种宿主细胞, 其以权利要求 9 所述载体转染、转导或感染。
29. 用于生产重组 MSP4 或 MSP5 多肽的方法, 其包括在促进表达的条件下培养权利要求 28 的宿主细胞, 和从培养基或细胞聚集物回收所述多肽。
30. 如权利要求 29 所述的方法, 其中所述宿主细胞选自昆虫细胞、细菌细胞、酵母细胞、植物细胞和动物细胞。
31. 一种重组载体, 其选自 I-3512、I-3513、I-3514、I-3694、I-3695 和 I-3696。
32. 一种免疫复合物, 其包含重组 MSP4 和 MSP5 多肽和特异识别所述多肽的抗体。
33. 用于检测疟原虫感染的方法, 其中所述方法包括提供包含疑似疟原虫感染的生物材料的组合物, 和测定结合特异于重组 MSP4 和 MSP5 多肽的抗体的疟原虫天然 MSP4 和/或 MSP5 多肽的存在。
34. 如权利要求 33 所述的方法, 其中它被用于检测恶性疟原虫 (*Plasmodium falciparum*)。
35. 如权利要求 33 所述的方法, 其中它被用于检测间日疟原虫 (*Plasmodium vivax*)。

36. 如权利要求 33 所述的方法, 其中通过电泳法或使用与所述重组 MSP4 和 MSP5 多肽发生免疫反应的抗体的免疫测定法来检测所述天然 MSP4 或 MSP5 多肽。

37. 用于检测结合包含重组 MSP4 或 MSP5 多肽的抗原的抗体存在与否的体外诊断方法, 其中所述方法包括将所述抗原与生物学液体在足以使所述抗原和所述生物学液体中的抗体形成抗原-抗体复合物的条件下接触一段时间, 和检测所述复合物的形成。

38. 如权利要求 35 所述的方法, 其进一步包括测量所述抗原-抗体复合物的形成。

39. 如权利要求 35 所述的方法, 其中通过基于蛋白印迹技术、ELISA、间接免疫荧光测定或免疫沉淀测定的免疫测定来检测抗原-抗体复合物的形成。

40. 用于检测结合重组 MSP4 或 MSP5 多肽或它们的混合物的抗体存在与否的诊断试剂盒, 其中该试剂盒包括包含重组 MSP4 或 MSP5 多肽或重组 MSP4 和 MSP5 多肽混合物的抗原, 和用于检测所述抗原和抗体之间免疫复合物形成的试剂, 其中所述试剂的量足以进行所述检测。

41. 一种免疫原性组合物, 其含量足以诱发体内免疫原性或保护性反应的至少一种重组 MSP4 或 MSP5 多肽或它们的混合物, 和一种药学上可接受的载体。

42. 如权利要求 41 所述的免疫原性组合物, 其中所述组合物包含足量的至少一种重组 MSP4 或 MSP5 多肽以诱发体内中和抗体。

43. 如权利要求 41 所述的免疫原性组合物, 其至少包含重组多肽 MSP4p40。

44. 如权利要求 41 所述的免疫原性组合物, 其至少包含重组多肽 MSP4p30。

45. 如权利要求 41 所述的免疫原性组合物, 其至少包含重组多肽 MSP4p20。

46. 如权利要求 41 所述的免疫原性组合物, 其至少包含重组多肽 MSP5p45 和 MSP5p35。

47. 如权利要求 39 所述的免疫原性组合物, 其至少包含重组多肽 PvMSP4p20。

48. 一种多核苷酸, 其选自 SEQ ID NO: 1、2、3、4、5、6、7 或 8。

49. 一种多核苷酸, 其选自 SEQ ID NO: 9、10、11、12、13、14、15、16、17、18、19、20、21、22 或 30。

50. 如权利要求 41 所述的免疫原性组合物, 其至少包含选自 PfMSP4p21ss1 和 PfMSP4p21ss2 的重组多肽。

51. 如权利要求 39 所述的免疫原性组合物, 其至少包含重组多肽 MSP4p40 和 MSP4p20。

52. 如权利要求 39 所述的免疫原性组合物, 其还包含含有重组 MSP1p19 多肽的抗原。

53. 一种疫苗组合物, 其含量足以诱发体内保护性反应的至少一种重组 MSP4 或 MSP5 多肽或它们的混合物, 和一种药学上可接受的载体, 其中所述至少一种重组 MSP4 或 MSP5 多肽选自 MSP4p40、MSP4p30、MSP4p20、MSP4p21ss1、MSP4p21ss2、MSP5p45、MSP5p35 和 PvMSP4p20。

54. 用于体外诊断来自疑为疟原虫感染的人的生物材料样品的疟原虫感染的方法, 其中所述方法包括下列步骤:

- 将包含重组 MSP4 或 MSP5 多肽的抗原与所述样品在足以使所述抗原和所述样品中的抗体形成抗原-抗体复合物的条件下接触一段时间, 以及
- 检测所述复合物的形成。

55. 如权利要求 54 所述的方法, 其还包括测量所述抗原-抗体复合物的形成。

56. 如权利要求 54 所述的方法, 其中通过基于蛋白印迹技术、ELISA、间接免疫荧光测定或免疫沉淀测定的免疫测定来检测所述抗原-抗体复合物的形成。

57. 如权利要求 36 所述的方法, 其中以 2005 年 11 月 16 日以保藏号 I-3517、I-3518、I-3519 或 I-3520 保藏于 C.N.C.M.的杂交瘤分泌的抗体经免疫测定来测定所述天然 MSP4 或 MSP5 多肽。

58. 如权利要求 17 所述的纯化抗体,其由在 2005 年 11 月 16 日以保藏号 I-3517、I-3518、I-3519 或 I-3520 保藏于 C.N.C.M.的杂交瘤所产生。

59. 于 2005 年 11 月 16 日以保藏号 I-3517、I-3518、I-3519 或 I-3520 保藏于 C.N.C.M.的杂交瘤。

60. 用于检测疟原虫的方法,其中所述方法包括提供包含疑为疟原虫感染的生物材料的组合物,和测定疟原虫天然 MSP4 和/或 MSP5 多肽的存在,所述天然 MSP4 和/或 MSP5 多肽结合特异于重组 MSP4 或 MSP5 多肽的抗体或至少一种特异于重组 MSP4 多肽的抗体和至少一种特异于重组 MSP5 多肽的抗体的混合物。

61. 如权利要求 60 所述的方法,其中所述重组 MSP4 或 MSP5 多肽选自权利要求 10 所述的纯化的多肽。

62. 如权利要求 60 所述的方法,其中所述特异于重组 MSP4 或 MSP5 多肽的抗体选自权利要求 58 所述的纯化抗体。

63. 如权利要求 60 所述的方法,其用于检测恶性疟原虫或间日疟原虫。

重组恶性疟原虫裂殖子表面蛋白 4 和 5 及其用途

本发明涉及重组恶性疟原虫(*Plasmodium falciparum*)裂殖子表面蛋白 4 和 5 及其用途。

发明领域

本发明涉及重组分离多肽、编码所述多肽的核酸、生产所述多肽重组形式的方法、抗所述多肽生成的抗体以及所述多肽和抗体在诊断方法、试剂盒、疫苗或抗寄生虫治疗中的用途。

发明背景

疟疾每年在全球引起大约 200 万人的死亡,其中大多数为 5 岁以下的非洲儿童,因此给世界上许多最贫穷国家带来了巨大的公共卫生负担。随着寄生虫及其蚊媒的抗药性扩展,这种负担正在以令人担忧的速度增长,加剧了对有效疫苗的迫切需要。

迄今为止检测过的最有前景的红内期候选疫苗为裂殖子表面蛋白 1(MSP-1)和顶端膜抗原(AMA1)。据发现,针对这些表面抗原的体液免疫应答与发病率下降有关,而且此类抗体在体外可以抑制寄生虫重新侵入红细胞(RBC)[1-3]。然而,与编码不易接近免疫效应物的蛋白质的基因相比,这些抗原基因一般表现出高得不成比例的非同义单核苷酸多态性(nsSNP)[4-6],这些 nsSNP 中有一些编码根本性的氨基酸置换,这些氨基酸置换集中在该蛋白最易接近宿主免疫系统的区域。

此类氨基酸多态性可以通过改变 B 细胞和 T 细胞表位在免疫逃避中发挥功能[4,8]。现在一般认为任何有功能的疟疾疫苗需含有每个靶抗原的数个等位基因型以期诱导多等位基因应答和/或需含有数个靶抗原的保守区域。

更具体而言,本领域需要可用于疟疾特别是恶性疟和间日疟诊断和治疗的抗原。具体而言,需要与赋予流行区域疾病保护的特异免疫应答相关联的保守抗原,还需要评价它们作为多价疟疾疫苗的组分的适宜性。

发明概述

本发明辅助实现本领域的这些需求。本发明提供了包含 SEQ ID NO: 1, 2, 3, 4, 5, 6, 7 或 8 的 DNA 序列的纯化或重组或合成核酸分子和编码 SEQ ID NO: 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 或 22 的氨基酸序列的纯化或重组或合成核酸分子。本发明还包括了与这些序列互补的纯化核酸分子。

本发明还包括了由所述包含 SEQ ID NO: 1-8 的 DNA 序列的纯化或重组核酸分子编码的纯化多肽, 正如所述 DNA 序列所预测。

SEQ ID NO: 1-8 的 DNA 序列或 SEQ ID No. 9-22 的氨基酸序列是指这些序列中的任一序列和每个序列。

本发明包括了包含 SEQ ID NO: 1-8 的 DNA 序列的纯化双链核酸分子和编码 SEQ ID NO: 9-22 的氨基酸序列的纯化双链核酸分子。单链和双链 RNA 和 DNA 核酸分子均包括在本发明内。这些分子可用于检测编码本发明所包括的多肽的单链和双链 RNA 和 DNA 变体。双链 DNA 探针允许检测相当于所述核酸分子两条链中任意一条的核酸分子。

本发明包括与包含 SEQ ID NO: 1-8 的 DNA 序列以及编码 SEQ ID NO: 9-22 的氨基酸序列的变性双链 DNA 在 42°C, 50%甲醛和 6XSSC 的中度严格性条件和 60°C, 0.5XSSC, 0.1%SDS 的洗涤条件下杂交的纯化核酸分子。

本发明还包括了源自 SEQ ID NO: 1-8 体外诱变的纯化核酸分子。体外诱变包括许多为本领域所知的技术, 这些技术包括但不限于定点诱变、随机诱变和体外核酸合成。根据一个具体实施方式, 来自所述引用序列体外诱变的此类核酸分子具有与原核酸分子相同的长度或具有比原核酸分子更短的序列。

包括 DNA 和 RNA 的本发明的核酸分子在本文中被称作“重组 MSP4 和 MSP5 核酸”或“重组 MSP4 和 MSP5 DNA”, 由这些分子编码的氨基酸在本文中被称作“重组 MSP4 和 MSP5 多肽”或“本发明的多肽”。

本发明还包括由于遗传密码的简并性自 SEQ ID NO: 1-8 得到的纯化核酸分子, 重组 MSP4 和 MSP5 核酸的等位基因变体的纯化核酸分子或重组 MSP4 和 MSP5 核酸的物种同系物。本发明还包括指导这些核酸分子表达的重组载体和以这些载体转化、转染或感染的宿主细胞。

结合重组 MSP4 或 MSP5 多肽的纯化多克隆或单克隆抗体包括在本发明内, 在本文中被称为“本发明的抗体”。

本发明还包括生产重组 MSP4 和 MSP5 多肽的方法, 其包括在促进表达的条件下培养宿主细胞和从培养基或细胞沉淀回收所述多肽。具体而言, 重组 MSP4 和 MSP5 多肽在杆状病毒昆虫表达系统中的表达包括在本发明内。

本发明还提供了标记的重组 MSP4 和 MSP5 多肽。优选情况下, 所述标记多肽为纯化形式。还优选的是, 所述未标记或标记多肽可为含有抗疟疾抗体的人体液免疫识别。所述多肽可以用例如选自放射性标记、酶标记、荧光标记、化学发光标记和发色团的免疫测定标记进行标记。

本发明还提供了本发明的重组 MSP4 或 MSP5 多肽和识别这些多肽的抗体之间的免疫复合物。所述免疫复合物可以用例如选自放射性标记、酶标记、荧光标记、化学发光标记和发色团组的免疫测定标记进行标记。

此外, 本发明提供了检测疟原虫 MSP4 和/或 MSP5 多肽的体外方法。所述方法包括提供包含疑为含有疟原虫 MSP4 和/或 MSP5 多肽的生物材料的组合物, 以及检测疟原虫 MSP4 和/或 MSP5 多肽的存在。疟原虫的所述 MSP4 和 MSP5 多肽典型地由电泳测定或以本发明所述抗体进行免疫测定。该方法可用于生物样品中疟原虫的检测, 在一个优选实施方式中该方法可用于恶性疟原虫和间日疟原虫(*Plasmodium vivax*)的检测。

本发明还提供了检测抗体存在与否的体外诊断方法, 这些抗体结合包含本发明所述重组或纯化的 MSP4 或 MSP5 多肽或本发明所述重组或纯化的 MSP4 或 MSP5 多肽的混合物的抗原。所述方法包括将所述抗原与生物学液体在足够让所述抗原与所述生物学液体中抗体形成抗原-抗体复合物的条件下接触一段时间, 然后检测所述复合物的形成。所述检测步骤还可以包括测量所述抗原-抗体复合物的形成。所述抗原-抗体复合物的形成优选地通过基于蛋白印迹技术、ELISA(酶联免疫吸附测定)、间接免疫荧光

测定或免疫沉淀测定的免疫测定来测量。所述方法可用于检测对来自感染疟疾的动物或人患者的生物学液体中疟原虫的免疫应答。在一优选实施方式中,该方法可以用于检测对恶性疟原虫感染或间日疟原虫感染的免疫应答。

用于检测结合本发明所述 MSP4 或 MSP5 多肽或本发明所述 MSP4 或 MSP5 多肽混合物的抗体存在与否的诊断试剂盒含有包含所述重组 MSP4 和/或 MSP5 多肽或所述重组 MSP4 和/或 MSP5 多肽的混合物的抗原,以及检测所述抗原和抗体之间免疫复合物形成的试剂。所述抗原和所述试剂的量足以进行所述检测。

用于检测疟原虫 MSP4 和/或 MSP5 多肽存在与否的诊断试剂盒含有本发明的抗体和检测抗原和所述抗体之间免疫复合物形成的试剂(means)。所述抗体和所述试剂的量足以进行所述检测。

本发明还提供了一种免疫原性组合物,其包含量足以诱发体内免疫原性或保护性应答的本发明的重组 MSP4 或 MSP5 多肽或本发明的重组 MSP4 或 MSP5 多肽的混合物,以及包含一种药学上可接受的免疫刺激剂。本发明的疫苗组合物包含足量的重组 MSP4 和/或 MSP5 多肽和药学上可接受的免疫刺激剂以诱发中和抗体。

本发明所述 MSP4 和 MSP5 多肽因此可用作检测抗疟疾相关抗原蛋白的抗体存在的诊断组合物的一部分。

此外,所述重组 MSP4 和 MSP5 多肽可以用于制备检测疟疾相关抗原蛋白存在的抗体。

本发明的多肽还可以用于制备中和抗体,这些中和抗体或灭活寄生虫,减少寄生虫体内生存能力或抑制或防止寄生虫复制。当本发明所述多肽用于免疫或接种组合物时,诱发中和寄生虫抗体的能力尤为重要。

下面是一种疫苗,它包括(A)天然信号序列(B)C 端 His 标签以及(C)类似恶性疟原虫抗原酸性重复区域的酸性重复区域,该酸性重复区域与区域(MSP4p20)的保护性抗体应答关系最为密切。

MKVAYFLSVLDLLIFSLYFDGRRSAFAGIAACIRHGRILGEGGE/QNSTPGSGGQTG
DHSAEAENG DYNEQGDDHGDDHGDDHGDDHGDEQDGEDYDDAEDDDLYELSE

VDENANLCLDNNGGCGDDKICENLGKGIVKCLCKPGYKLVGTECVEHHHHHH
[SEQ ID NO: 30]

附图简述

本发明将参考附图进行描述，其中：

图 1 描述了几种不同疟原虫物种的 *m*sp4 和 *m*sp5 基因位点的基因组结构。这些基因的种间保守性指示了基本功能。

图 2 显示了 SALSA 序列与恶性疟原虫 MSP-4(PfMSP4)有 92%的相同性。已知 SALSA 序列具有 B 细胞和 T 细胞表位。PfMSP4 序列以黑色显示，SALSA 序列以灰色显示。

图 3 显示了 PfMSP4 的全长合成基因序列。该图显示的所有序列以寡核苷酸表示。PfMSP4 的基因序列由编号的重叠寡核苷酸序列(1-38)编码并交替以粗体和正常文本突出显示。限制性位点以小写文本显示。

图 4 描绘了 PCR 法制备所述合成基因。“基因装配反应”(5 或 10 μ l，分别为泳道 1 和泳道 2)和“基因扩增反应”(5 或 10 μ l，分别为泳道 3 和泳道 4)的样品在 1%琼脂糖凝胶上分离，DNA 大小标准品位于样品侧面并将样品分开。

图 5 显示了四个不同 MSP4 构建体的比对，这些构建体的生成是用于杆状病毒系统中的重组蛋白表达。推定信号序列(根据现有算法预测)以灰色突出显示，共有多态性位点以粗体字突出显示，最有可能处于平衡选择并因此参与免疫逃避的位点以“*”标出。

图 6 显示了恶性疟原虫 MSP5(PfMSP5)的全长合成基因序列。该图显示的所有序列以寡核苷酸表示。PfMSP5 的基因序列由编号的重叠寡核苷酸序列(1-38)编码并交替以粗体字和正常文本突出显示。限制性位点以小写文本显示。

图 7 描绘了所述合成基因的 PCR 制备。“基因扩增反应”(5 μ L)的样品在 1%琼脂糖凝胶上分离，其旁侧为 DNA 大小标准品。

图 8 显示了 PfMSP4 在两种不同昆虫细胞系中随时间变化的表达。通过将细胞悬浮液与 4%台盼蓝以 1: 1 混合并计数盖玻片下一方格内的总细胞和蓝色细胞，计算出细胞死亡百分比。感染后 24、30、36、42、48、

54、60 和 72 h 时, 从受感染的培养 SN 中纯化蛋白(分别为 1-8 泳道), 20 uL 的洗脱蛋白以 4-12% Bis-Tris 凝胶(Invitrogen)分离并以 SimplyBlue safe stain (Invitrogen)染色。这些蛋白种类在未感染细胞培养物中未观察到(第 9 泳道)。

图 9 为突出显示每个 MSP4 产物 p40、p30 和 p20 N 端序列的序列比对。该研究中鉴定出的信号序列以灰色突出显示。每个产物的 N 端序列以蓝色粗体字显示。

图 10 显示了三种已知疟疾抗原已报道的信号切割位点附近约 80 个氨基酸残基的疏水曲线。成熟蛋白的首个残基以黑色菱形显示。S-抗原信号序列切割的曲线类似于杆状病毒系统中 MSP4 信号序列切割的曲线。

图 11 显示了 PfMSP5 在两种不同昆虫细胞系中随时间变化的表达。通过将细胞悬浮液与 4%台盼蓝以 1: 1 混合并计数盖玻片下一方格内的总细胞和蓝色细胞, 计算出细胞死亡百分比。感染后 24, 30, 36, 42, 48, 54, 60 和 72 h 时, 从受感染的培养 SN 中纯化蛋白(分别为 1-8 泳道), 20 uL 的洗脱蛋白以 4-12% Bis-Tris 凝胶(Invitrogen)分离并以 SimplyBlue safe stain (Invitrogen)染色。这些蛋白种类在未感染细胞培养物中未观察到(第 9 泳道)。

图 12 涉及在兔中制备的多克隆抗血清。针对(A)全长 MSP4p40(R6-7)和(B)MSP4 降解产物 p20(R8-9)而制备的血清的 ELISA 终点效价, 采用天然(NR)或不可逆还原的(R) MSP4-p40 作为包被抗原。当 MSP4 抗原被还原后, 血清 R8 失去了 50%以上的结合活性。免疫前血清标记为 NEG。图(C)显示了抗 MSP5 血清(56-57)的效价, 采用 MSP5-p45 和 p35 作为包被抗原进行分析。

图 13 描述了采用针对杆状病毒表达的抗原而制备的多克隆血清和来自 Dielmo 人血清库或来自抗 MSP4 兔血清库的 MSP4p40 亲和力纯化的血清对寄生虫材料进行探测而得到的结果。使用 Wang 等人(2003) [9]的方案, 从 iRBC 提取并粗略分部不同步生长的寄生虫。各免疫印迹的每个泳道 1 加入非还原的 Triton X100 可溶寄生虫部分, 泳道 2 加入非还原的膜部分, 泳道 3 加入还原的 Triton X100 可溶寄生虫部分, 泳道 4 加入还原

的膜部分(图 A, B 和 D)。图 C 显示了以风干后感染有恶性疟原虫的 RBC 而获得的 IFA 数据。

图 14 显示了采集自 1990 年雨季肆虐期间 Dielmo 的 9 个免疫成年人 (18-49 岁) 血清的 ELISA 分析。培养板以 MSP4p40 (A) 或 MSP5p45 和 p35 (B) 包被。

图 15 是使用来自 3 号和 9 号个体的人免疫血清的重组 MSP4 p40, p30 和 p20 (分别为 1, 2, 3 泳道) 的免疫印迹(A) 和使用来自 8 号个体的人免疫血清的重组 MSP5 的免疫印迹分析(B)。

每份血清生成的 ELISA 曲线显示在图 16A 和 16B 的每块印迹上, 图 16A 和 16B 描绘了两个特异识别 PfMSP4, p40 和 p20 表位的单克隆抗体 mAb L11-16 和 mAb F12-7 的蛋白印迹。图例说明: nR = 非还原, R = 被 invitrogen 商用缓冲液还原, iR = 被 DTT 和丙烯酰胺不可逆还原。

图 17A 和 17B 描述了以稀释因子(x-轴) 分别作为图 16A 和 16B 中的 mAb 的 OD 或吸光度(y-轴) 的函数。

图 18A 和 18B 描述了两个特异识别 PfMSP5, p45 和 p35 的表位的单克隆抗体 mAb G21-2 和 mAb J18-4 的蛋白印迹。图例说明: nR = 非还原, R = 被 invitrogen 商用缓冲液还原, iR = 被 DTT 和丙烯酰胺不可逆还原。

图 19A 和 19B 描述了以稀释因子(x-轴) 分别作为图 18A 和 18B 中 mAb 的 OD 或吸光度(y-轴) 的函数。

图 20A-B 描述了用以优化所述蛋白包被浓度的 ELISA 分析。培养板以 PBS 中 0.5 或 1 $\mu\text{g}/\text{ml}$ 蛋白质过夜包被, 加入 100 $\mu\text{L}/\text{孔}$ 。图(A)显示了以来自 MSP4 的三个不同抗原 MSP4p40, MSP4p30(本文称 MSP4p40/2) 和 MSP4p20 获得的数据。图(B)显示了以 MSP5 获得的数据, 根据来自图(A) 的 MSP4p20 数据绘制以展示应答的不同强度。由于使用任一种浓度所观测到的光密度近乎相同, 于是在后续所有分析中均使用了较低的包被浓度 (0.5 $\mu\text{g}/\text{ml}$)。

图 21A-D 描绘了来自首次接受试验的个体(naïf)、超免疫成年人(shi) 和来自 2005 年 Dielmo 队列的个体的对照血清的 ELISA 分析。培养板以 0.5 $\mu\text{g}/\text{mL}$ 的 MSP5 (A), MSP4p40 (B), MSP4p20, MSP4p20 (C) 和 MSP4p30

(D)包被,并以不同稀释度的抗血清(3 倍稀释系列)进行反应。所有抗原为免疫血清识别,但不为未感染疟疾的对照(malaria naive control)识别。

图 22A-D 描绘了来自 2000 年 Ndiop 队列个体的抗体应答,根据 15 岁以上和 15 岁以下的年龄分层。图(A)显示了抗原 MSP4p20, MSP4p40, MSP4p30 和 MSP5(从左至右)光密度形式的 ELISA 数据。图(C)显示了 OD 比形式的相同数据。图(B)中个体对 MSP1p19 的抗体应答绘制为以 OD 比作为 2000 年 Ndiop 队列年龄的函数,而图(D)显示了以相同血清抗 MSP5 而获得的 ELISA OD 比。

图 23 描绘了来自 Ndiop 2000 队列的血清对 MSP4 及其衍生物的识别差异。左边的图显示了将平均应答作为 MSP4p20, MSP4p40 和 MSP4p30 (等同于 MSP4md2)OD 比的函数进行作图。右边的图显示了上四分位和下四分位以及将抗体对每个 MSP4 抗原的应答作为年龄函数的数据范围。

图 24 描绘了抗体依赖性的单核细胞介导的细胞毒性,根据特异血清存在下诱发的氧化爆发所测量。根据 OD 比高于或低于队列中位反应将血清分组。图(A)显示了来自划分为 MSP5 反应性函数(高于或低于 2.4 的 OD 比)的血清的数据,图(B)显示了来自划分为 MSP4p40 反应性函数(高于或低于 20 的 OD 比)的血清的数据。采用化学发光法分析它们在 PBMC(外周血单核细胞)和恶性疟原虫裂殖子存在下诱发氧化爆发的能力。在两组中,观察到表现出高 OD 比的个体发生氧化爆发的增加。

图 25 描绘了采用 IgG 特异二抗试剂分析每个抗原的抗体同种型曲线。显示了 MSP4p40(图 A)和 MSP5(图 B)的同种型曲线。数据作图为年龄的函数(高于和低于 15 岁)。

图 26 为 PfMSP4 和 PvMSP4 氨基酸序列的比对。突出显示了 Pf 序列的几个特点(i)分泌和 GPI 锚定信号以绿色显示,(ii)已知为 SALSA 的序列加以下划线,(iii)所有带负电的残基(D 和 E)以红色显示,以及(iv) PfMSP4p20 N 端序列和重要半胱氨酸残基以蓝色突出显示。

图 27 描绘了杆状病毒 PfMSP5。(A)纯化的重组 PfMSP5 在 SDS-PAGE 凝胶上迁移,转移到纤维素膜并以人免疫血清探测。(B)在存在胍化豆蔻酸的情况下表达 PfMSP5。纯化蛋白在 SDS-PAGE 凝胶上迁移,蓝染和进行 7 周的放射自显影。

图 28 描绘了 PfMSP4 在 RBC 各期随时间变化的表达。(A)采用单克隆抗体 G17.12 探测丙酮固定的环状体期寄生虫和风干的后期和游离裂殖子,该单克隆抗体特异于 PfMSP1p19 或以重组 PfMSP4p20 (Rp20)亲和纯化的多克隆兔血清。抗体染色以 Alexa Fluor® 488 羊抗鼠或抗兔的缀合抗体显示,寄生虫 DNA 以 Hoechst 33342 染色。(B)采用亲和纯化的兔(Rp20)和人血清(Ndiop 和 Dielmo 库: Hp20 为 1:1)对来自成熟裂殖体(S)、游离裂殖子(M)和环状体(R)的寄生虫提取物进行蛋白印迹分析。

图 29 描绘了由用于重组 PfMSP4 表达的构建体所编码的蛋白序列。显示了设计用于促进 PfMSP4p20 直接表达的所述构建体。构建体的名称列在序列文本的左侧,每个分泌蛋白的 N 端序列在文本内以粗体突出显示,所述重组蛋白名称列在右侧。

图 30 描绘了 PfMSP4p20 随时间变化的直接表达。感染 PfMSP4p40, PfMSP4p21, PfMSP4p21ss1 或 PfMSP4p21ss2 杆状病毒的昆虫细胞培养上清的样品在感染后 24h 到 66h 之间每隔 6h 进行收集,透析并以 TALON 树脂分批纯化。纯化蛋白样品在 NuPAGE 4-12%梯度凝胶上分离,蓝染,并在右边标注蛋白的大小。采用感染后 66h 时剩下的最后 130 ml 培养上清计算大致蛋白产量。以 IMAC 和 HPLC 纯化蛋白,并采用 BCA 蛋白定量试剂盒(PIERCE)计算该蛋白产量。

图 31 描绘了 PfMSP4p40 和 PvMSP4/His 的蛋白序列比对。杆状病毒间日疟原虫 MSP4 和恶性疟原虫 MSP4 表达构建体的氨基酸(单字母符号)序列以 clustalx 进行比对。氨基酸相同性表示为(*),保守置换表示为(:),半保守置换表示为(.)和根本性改变表示为空格。鉴定的 N 端序列以下划线和粗体显示。不同蛋白质产物的 N 端序列以下划线和粗体显示。

图 32 描绘了 PvMSP4/His 随时间变化的表达。简而言之,感染后 24 到 66 h 之间每隔 6h 收集 8 ml 旋动培养 SN 样品,透析,然后在 TALON 树脂上分批纯化。纯化蛋白样品在 NuPAGE 4-12%梯度凝胶上分离,以 SimplyBlue SafeStain 染色,并在左边和右边标注蛋白的大小。

图 33 描绘了人免疫血清对 PvMSP4/His 的反应性。(A)ELISA 培养板以杆状病毒表达系统中表达的 PvMSP4/His 包被。检测了 3 倍稀释系列或全部 24 份血清。(B)与此同时,以不可逆还原的 PvMSP4 检测每个稀释系

列。该图在每个血清的下方图中显示了稀释为 1/2700 时的 OD, 得到高于阴性对照抗天然抗原的 OD。上方图显示了血清以 1/2700 血清稀释时抗原还原所观察到的 OD 下降。

图 34 和 35 描绘了血清(去补体、除去 IgG 或纯化 IgG)处理在基于 PMN 对裂殖子(APDm)抗体依赖性吞噬的功能检测中的结果。实施例 19 描述了所述测定。

发明详述

两个近来鉴定的裂殖子表面抗原恶性疟原虫裂殖子表面蛋白 4 和 5 (PfMSP4 和 PfMSP5)有望成为潜在多组分抗疟疾疫苗的蛋白组分。msp4 和 msp5 基因均编码 272 个残基的蛋白, 每个蛋白具有单个 C 端 EGF 样结构域和 GPI 锚定基序[10,11]并相邻位于 2 号染色体上, 恰好处于 msp2 的上游。已表明两种蛋白都具有裂殖子表面处的膜连接, 而且已表明人血清可与大肠杆菌表达的重组 MSP4 反应[12,13]。

3 种鼠类疟原虫即约氏疟原虫(*P. yoelii*)、夏氏疟原虫(*P. chabaudi*)和伯氏疟原虫(*P. berghei*)在 MSP4 和 MSP5 位点处(MSP4/5)仅有一个基因, 相互之间表现出一定程度的同源性[14-16]。该基因被称为 MSP4/5, 并已被用于研究约氏疟原虫致死剂量攻击模型(lethal challenge model)的保护性免疫[17]。

已经表明, MSP4/5 使用多种免疫策略来提供保护, 当与 MSP1p19 一起输送时, 效力达到最大[17-20]。此外, 鼠 MSP4/5 蛋白诱发的免疫应答似乎没有菌株特异性[21]。在恶性疟原虫中, msp4 和 msp5 均在同源位点处具有一个内含子[22]。

恶性疟原虫裂殖子表面蛋白 4 (PfMSP4)的蛋白序列包括一个分泌信号序列, 一个 C 端 EGF 样结构域和 GPI 锚定信号[10]。

Pfmsp4 基因序列长 960bp, 它包括一个 144bp 的内含子并编码 272 个氨基酸残基。Pfmsp5 基因序列长 955bp, 它包括一个 136bp 的内含子并编码 272 个氨基酸残基。至于 PfMSP4, PfMSP5 的蛋白序列由一个分泌信号序列, 一个 C 端 EGF 样结构域和一个 GPI 锚定信号组成[11,22]。

这簇 MSP 基因的下游是高度保守的腺苷酸琥珀酸裂解酶(ASL)基因, 该基因已被证明为有利于鉴定其他物种中的该位点(图 1)。

当前许多数据支持这一概念, 即 PfMSP4 和 PfMSP5 是良好的候选疫苗。尽管如此, 几项发表的研究结果表明两种不同表达系统(大肠杆菌和酵母)产生的 PfMSP4 的重组类似物具有不同的抗原性, 并诱发针对该蛋白质的 EGF 样结构域的不依赖构象的应答, 使用人免疫血清时未观察到这一现象[13]。如果识别所述蛋白质的该区域的构象表位对于保护是重要的, 正如 MSP1-19 一样, 则有必要产生一种忠实复制所有表位的产物。

因此, 本发明提供了构建体, 其中编码恶性疟原虫 MSP4 和 MSP5 的所述核酸和由此形成的多肽已被修饰以在昆虫细胞中实现最佳表达。更具体地, 本发明提供了编码重组 MSP4 多肽的构建体, 这些多肽在杆状病毒昆虫细胞表达体系中以可溶分泌多肽表达。所述重组多肽在 C 端含有一个似已适宜折叠的 EGF 样结构域。这一点表现在当该蛋白在 ELISA 培养板上被不可逆还原时多克隆兔血清对 MSP4 的识别发生了明显下降 (50-60%) (参见图 12)。

一种本发明的重组 MSP4 多肽是 MSP4 菌表抗原(不含 C 端 GPI 锚定位点的疏水残基, 因此可实现蛋白分泌), 其 N 端附近多态性区域有 30 个氨基酸的缺失。本发明的多肽被称为 MSP4p30。

另一种本发明的重组 MSP4 多肽是 MSP4 菌表抗原(不含 C 端 GPI 锚定位点的疏水残基, 因此可实现蛋白分泌), 其 N 端附近多态性区域没有 30 个氨基酸的缺失。本发明的该多肽被称为 MSP4p40, 它是全长的基因产物。

再一种本发明的重组 MSP4 多肽是大致对应于 MSP4 C 端一半的 20kDa 多肽, 开始于 KSPKE 基序附近并包括 EGF 结构域。该多肽可以包括上游补充(supplementary)氨基酸残基, 尤其是位于 KSPKE 基序上游的 MSP4p40 序列的残基。本发明的该重组多肽被称为 MSP4p20。

此外, 本发明提供了包括同时产生的 p35 和 p45 形式的重组 MSP5 多肽。二者均含有翻译后修饰, 可能参与了脂肪酸残基的共价锚定(豆蔻酰化), 该共价锚定可以促进免疫原性。本发明的这些重组多肽被称为 MSP5p45 和 MSP5p35。

尽管如此,本发明的重组 MSP4 和 MSP5 多肽还可以表达为 C 端 GPI 锚定实体,或使用其天然 GPI 信号序列或使用标志 GPI 修饰的来自其他 GPI 锚定的蛋白质序列。此类 GPI 修饰实体在没有任何免疫佐剂下预计可显著增加重组 MSP4 和 MSP5 多肽的免疫原性。

本发明的重组多肽参照其对应如下 SEQ ID NO 进行更详细地描述:

核酸

MSP4p20 SEQ ID NO: 1
MSP4p30 SEQ ID NO: 2
MSP4p40 SEQ ID NO: 3
MSP5 SEQ ID NO: 4
MSP5p10 SEQ ID NO: 5
MSP4p21 SEQ ID NO: 6
MSP4p21ss1 SEQ ID NO: 7 和 SEQ ID NO: 28
MSP4p21ss2 SEQ ID NO: 8 和 SEQ ID NO: 29

多肽

MSP4p20 (降解产物) SEQ ID NO: 9
MSP4p30 (构建体中有 ORF=该构建体编码的多肽) SEQ ID NO: 10
MSP4p30 (表达终产物=昆虫细胞产生的多肽) SEQ ID NO: 11
MSP4p40 (构建体中有 ORF) SEQ ID NO: 12
MSP4p40 (表达终产物) SEQ ID NO: 13
MSP5 (构建体中有 ORF) SEQ ID NO: 14
MSP5p10 SEQ ID NO: 15
PvMSP4/His SEQ ID NO: 16
MSP4p21 (构建体中有 ORF) SEQ ID NO: 17
MSP4p21 (表达终产物) SEQ ID NO: 18
MSP4p21ss1 (构建体中有 ORF) SEQ ID NO: 19
MSP4p21ss1 (表达终产物) SEQ ID NO: 20
MSP4p21ss2 (构建体中有 ORF) SEQ ID NO: 21
MSP4p21ss2 (表达终产物) SEQ ID NO: 22

PvMSP4p20(基于“PfMSP4 和 PvMSP4 氨基酸序列的比对”中所给出的间日疟原虫 MSP4 序列的推荐疫苗构建体序列) SEQ ID NO: 30

这些多肽在本文分别称为和合称为本发明的“重组 MSP4 和 MSP5 多肽”。与之类似,编码这些多肽的核酸被称为本发明的“重组 MSP4 和 MSP5 核酸”。

本发明应用广泛。所述重组 MSP4 和 MSP5 多肽的发现能够构建包含编码本发明的重组 MSP4 和 MSP5 多肽的核酸的表达载体;用所述表达载体转染或转化的宿主细胞;有生物活性的重组 MSP4 和 MSP5 多肽和作为分离或纯化蛋白的重组 MSP4 和 MSP5 多肽;与重组 MSP4 和 MSP5 多肽发生免疫反应的抗体;所述重组 MSP4 和 MSP5 多肽和抗所述重组 MSP4 和 MSP5 多肽的抗体在检测疟原虫和疟疾感染中的诊断用途以及所述重组 MSP4 和 MSP5 多肽保护免受疟原虫感染的疫苗用途。

如本文所使用,术语“重组 MSP4 和 MSP5 多肽”也系指一类还包括具有 SEQ ID NO: 9-22 氨基酸序列的多肽,以及具有与此类氨基酸序列高度相似性(至少 90%的同源性)的蛋白质和多肽以及具有免疫反应性的蛋白质和多肽。此外,重组 MSP4 和 MSP5 多肽系指 SEQ ID NO: 1-8 的核苷酸的基因产物。

本文使用的术语“纯化”意为所述重组 MSP4 和 MSP5 多肽基本不与例如作为重组宿主细胞培养物的纯化产物或作为来自非重组来源的纯化产物的其他蛋白质或多肽关联。本文使用的术语“基本纯化”系指含有重组 MSP4 和 MSP5 多肽,且除了存在使用特异抗体可以除去的已知蛋白,基本不与其他蛋白质或多肽关联的混合物,而且基本纯化的重组 MSP4 和 MSP5 多肽可以用作抗原。

本文所称的重组 MSP4 和 MSP5 多肽“变体”意为与重组 MSP4 和 MSP5 多肽基本同源但因一个或多个缺失、插入或置换而具有与重组 MSP4 和 MSP5 多肽不同的氨基酸序列的多肽。所述变体氨基酸序列与重组 MSP4 和 MSP5 多肽的氨基酸序列优选具有至少 80%的相同性,最优选具有至少 90%的相同性。百分比相同性可以例如根据 Devereux 等人提出的 6.0 版 GAP 计算程序比较序列信息进行确定(Nucl. Acids Res. 12:387,

1984), 该程序可从威斯康辛大学遗传计算组(University of Wisconsin Genetics Computer Group, UWGCG)获得。所述 GAP 程序使用 Needleman 和 Wunsch 的比对方法(J. Mol. Biol. 48:443, 1970), 该方法由 Smith 和 Waterman 修订(Adv. Appl. Math 2:482, 1981)。GAP 程序的优选默认参数包括: (1)核苷酸的一元比较矩阵(其中, 值为 1 表示相同和值为 0 表示相同)和 Gribskov 和 Burgess 加权比较矩阵(Nucl. Acids Res. 14:6745, 1986), 由著者 Schwartz 和 Dayhoff 描述(Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, pp. 353-358, 1979); (2) 每一缺口罚分为 3.0, 每一缺口中的每一符号加罚 0.10 以及(3)末缺口没有罚分。

变体可以包含保守置换的序列, 这意味着某个氨基酸残基为具有类似生理化学特点的残基置换。保守置换的实例包括脂肪酸残基之间的置换, 例如 Ile, Val, Leu 或 Ala 之间的置换, 或极性残基之间的置换, 例如 Lys 和 Arg 之间、Glu 和 Asp 之间或 Gln 和 Asn 之间的置换。其他此类保守置换, 例如具有类似疏水特性的整个区域的置换, 是众所周知的。本发明的 MSP4 和 MSP5 多肽变体的实例为产生自交替 mRNA 剪切事件的蛋白质或产生自所述重组 MSP4 和 MSP5 多肽的蛋白酶切割的蛋白质。归因于蛋白酶解的变异包括例如在不同类型宿主细胞中表达后末端的差异, 这是由于一个或多个末端氨基酸残基从重组 MSP4 和 MSP5 多肽上蛋白水解除去。归因于移码的变异包括例如在不同类型宿主细胞中表达后末端的差异。

正如上文指出, 本发明提供了分离和纯化或同种的重组 MSP4 和 MSP5 多肽。可以用作抗原的重组 MSP4 和 MSP5 多肽的变体和衍生物可以通过编码重组 MSP4 和 MSP5 多肽的核苷酸序列的突变而获得。所述氨基酸序列的变化可通过许多常规方法中的任意一种而完成。通过合成含有突变序列的寡核苷酸可以在特定位点引入突变, 该突变序列两旁为限制性位点, 可使该突变序列连接到天然序列的片段上。连接后, 所形成的重建序列编码具有所需氨基酸插入、置换或缺失的类似物。

或者, 可以使用寡核苷酸引导的定点特异性诱变方法来提供改变的基因, 其中预先确定的密码子可通过置换、缺失或插入而改变。产生上述变化的示例性方法由 Walder 等人(Gene 42:133, 1986); Bauer 等人(Gene

37:73, 1985); Craik (BioTechniques, January 1985, 12-19); Smith 等人 (Genetic Engineering: Principles and Methods, Plenum Press, 1981); Kunkel (Proc. Natl. Acad. Sci. USA 82:488, 1985); Kunkel 等人 (Methods in Enzymol. 154:367, 1987) 和美国专利 4,518,584 和 4,737,462 号披露, 它们全部通过引用的方法并入。

在本发明的一个方面, 重组 MSP4 和 MSP5 多肽可以被用于制备特异结合重组 MSP4 和 MSP5 多肽的抗体。术语“抗体”用于包括多克隆抗体、单克隆抗体和它们的片段, 例如 F(ab')₂ 和 Fab 片段以及任何重组产生的结合配偶体。如果它们以大于或等于约 10^7 M⁻¹ 的 K_a 结合重组 MSP4 和 MSP5 多肽, 抗体则被定义为特异结合的。结合配偶体或抗体的亲和力可使用常规技术容易地测定, 例如 Scatchard 等人 Ann N.Y Acad. Sci., 51 :660 (1949) 描述的技术。使用本领域公知的方法可从许多来源例如马、牛、山羊、绵羊、狗、鸡、兔、小鼠或大鼠容易地生成多克隆抗体。

本发明还包括源自 SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7 和 8 的核苷酸序列的分离片段和寡核苷酸。本发明还包括了由这些片段和寡核苷酸编码的多肽。

本发明范围内的核酸序列包括分离的 DNA 和 RNA 序列, 它们在中度或高度严格性条件下可以杂交到本文公开的所述天然重组 MSP4 和 MSP5 核酸上, 而且这些 DNA 和 RNA 序列编码重组 MSP4 和 MSP5 多肽。如本文所使用, 为本领域普通技术人员知晓和为 Sambrook 等人 Molecular Cloning: A Laboratory Manual, 2 ed. Vol. 1, pp. 1.101-104, Cold Spring Harbor Laboratory Press, (1989) 确定的中度严格性条件包括使用 5X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0) 的预洗涤溶液用于硝酸纤维素滤膜, 42°C, 50% 甲酰胺, 6X SSC 的杂交条件(或其他类似杂交溶液, 例如以 50% 甲酰胺配制的 42°C Stark's 溶液)和约 60°C, 0.5X SSC, 0.1 % SDS 的洗涤条件。高度严格性条件定义为上面的杂交条件, 以及 68°C, 0.2X SSC, 0.1% SDS 的洗涤条件。本领域技术人员会认识到根据探针长度等因素需要时可以调节温度和洗涤溶液的盐浓度。

由于已知的遗传密码简并性, 其中一个以上密码子可以编码同一个氨基酸, 因此一个 DNA 序列可以不同于 SEQ ID NO: 1-8 显示的序列但仍可

编码具有 SEQ ID NO: 9-22 的氨基酸序列的重组 MSP4 和 MSP5 多肽。此类变体 DNA 序列可以来自沉默突变(例如发生在 PCR 扩增中), 或可以为天然序列故意诱变(deliberate mutagenesis)的产物。

本发明因此提供了编码重组 MSP4 和 MSP5 多肽的等效分离 DNA 序列, 这些 DNA 序列选自于(a); (a) 包含 SEQ ID NO: 1-8 的核苷酸序列的 DNA; (b)可与(a)中的 DNA 在中度严格性条件下杂交并编码重组 MSP4 和 MSP5 多肽的 DNA; 以及(c)因遗传密码与(a)或(b)定义的 DNA 简并且编码重组 MSP4 和 MSP5 多肽的 DNA。由此类 DNA 等效序列编码的所述多肽包括在本发明内。

等效于 SEQ ID NO: 1-8 的 DNA 序列的 DNA 将在中度严格性条件下与编码包含 SEQ ID NO: 9-22 的氨基酸序列的多肽的 DNA 序列杂交。由此类 DNA 编码的重组 MSP4 和 MSP5 多肽的实例包括但不限于重组 MSP4 和 MSP5 多肽片段以及包含失活 N-糖基化位点、失活蛋白酶处理位点或上文所述的保守氨基酸置换的重组 MSP4 和 MSP5 多肽。源自其他疟原虫物种的 DNA 编码的所述多肽也包含在内, 其中所述 DNA 将与 SEQ ID NO: 1-8 的 DNA 的互补序列杂交。

含有编码重组 MSP4 和 MSP5 多肽的核酸序列的重组表达载体可以使用公知的方法制备。所述表达载体包括可操作地连接到适宜转录或翻译调节核苷酸序列的重组 MSP4 和 MSP5 DNA 序列, 这些调节核苷酸序列例如来自哺乳动物、微生物、病毒或昆虫基因的核苷酸序列。调节序列的实例包括转录启动子、操纵基因或增强子、mRNA 核糖体结合位点和调控转录和翻译起始和终止的适宜序列。当所述调节序列的功能与所述重组 MSP4 和 MSP5 DNA 序列有关时, 核苷酸序列被“可操作地连接”。因此, 如果所述启动子核苷酸序列调控重组 MSP4 和 MSP5 DNA 序列的转录, 启动子核苷酸序列被可操作地连接到重组 MSP4 和 MSP5 DNA 序列。此外, 在所需宿主细胞内通常由复制起点赋予的复制能力和可鉴定转化体的选择基因可整合到所述表达载体。

此外, 编码适宜信号肽的序列可以整合进表达载体, 这些信号肽在自然情况下不与重组 MSP4 和 MSP5 多肽关联。例如, 用于信号肽(分泌前导)的 DNA 序列可以通过符合读框的方式融合到所述重组 MSP4 和 MSP5

核苷酸序列,使得所述重组 MSP4 和 MSP5 多肽开始翻译为包含该信号肽的融合蛋白。在所需宿主细胞中有功能的信号肽增强重组 MSP4 和 MSP5 多肽的细胞外分泌。所述信号肽在重组 MSP4 和 MSP5 多肽从所述细胞分泌出后可以从所述重组 MSP4 和 MSP5 多肽上切割下来。

用于原核宿主细胞的表达载体一般包含一个或更多表型可选择标记基因。表型可选择标记基因是例如编码赋予抗生素抗性 or 提供自养需求的蛋白质的基因。用于原核宿主细胞的有用表达载体的实例包括源自市售质粒的表达载体。市售载体包括为蛋白质的表达专门设计的载体。这些载体包括 pMAL-p2 和 pMAL-c2 载体,它们用于融合到麦芽糖结合蛋白的蛋白质的表达(New England Biolabs, Beverly, MA, USA)。

通常用于重组原核宿主细胞表达载体的启动子序列包括 β -内酰胺酶(青霉素酶),乳糖启动子系统(Chang 等人, Nature 275:615, 1978;和 Goeddel 等人, Nature 281 :544, 1979), 色氨酸(trp)启动子系统(Goeddel 等人, Nucl. Acids Res. 8:4057, 1980; 和 EP-A-36776) 和 tac 启动子(Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, p. 412, 1982)。

用于重组 MSP4 和 MSP5 多肽表达的适宜宿主细胞包括原核细胞、酵母或更高等的真核细胞。优选为昆虫细胞。用于细菌、真菌、酵母和哺乳动物细胞宿主的适宜克隆和表达载体被例如 Pouwels 等人 Cloning Vectors: A Laboratory Manual, Elsevier, New York, (1985)描述。也可以使用无细胞翻译系统利用源自本文公开的 DNA 构建体的 RNA 来产生重组 MSP4 和 MSP5 多肽。

应理解的是,本发明旨在包括之前描述的分离或纯化形式的蛋白质或多肽,不论是使用本文描述的技术还是其他方法而获得。在本发明的一个优选实施方式中,所述重组 MSP4 和 MSP5 多肽基本不含人组织和人组织组分,核酸,外来蛋白和脂类,和外来微生物例如细菌和病毒。还应理解的是,本发明包括了具有基本相同的生物学和免疫原性的等效蛋白。

取决于本发明所述重组 MSP4 和 MSP5 多肽的用途,理想情况下可对它们进行标记。适宜标记的实例为放射性标记、酶标记、荧光标记、化学发光标记和发色团。标记本发明的蛋白质的方法与广泛用于标记免疫球蛋白

白的方法没有本质差异。通过使用抗本发明的抗原的标记抗体或使用抗所述抗原的抗体的免疫球蛋白作为间接标记可避免标记的需要。

一旦得到本发明所述重组 MSP4 和 MSP5 多肽,它们用于产生与之反应的多克隆和单克隆抗体。因此,可以使用本发明的蛋白质或多肽按照本领域已知技术来免疫动物宿主。此类技术通常涉及接种,但它们也可涉及其他给药方式。施用足量的所述多肽以在动物宿主中引起免疫原性应答。可以使用产生抗本发明的抗原的抗体的任何宿主。一旦动物被免疫并经过开始产生抗该抗原的抗体的足够时间,可以回收多克隆抗体。一般方法包括从所述动物取出血液并从取出的血液分离血清。可以使用含有抗该抗原的抗体的血清作为抗该抗原的抗血清。或者,可从所述血清中回收所述抗体。亲和纯化是从血清中回收抗该抗原的纯化多克隆抗体的优选技术。

也可以制备抗本发明的抗原的单克隆抗体。生产与所述抗原反应的单克隆抗体的方法包括下列步骤:以所述抗原免疫宿主;从宿主脾脏回收产生抗体的细胞;将所述产生抗体的细胞与缺乏次黄嘌呤-鸟嘌呤磷酸核糖基转移酶的骨髓瘤细胞融合以形成杂交瘤;通过在包含次黄嘌呤、氨基蝶呤和胸腺嘧啶的培养基上培养来选择至少一个杂交瘤;鉴定产生抗所述抗原的抗体的至少一个杂交瘤;培养所述鉴定的杂交瘤以产生可回收量的抗体;以及回收由所述培养的杂交瘤产生的所述抗体。

这些多克隆或单克隆抗体可以用于许多场合。其中有相应蛋白的中和。它们还可用于检测生物制品中的疟原虫抗原,或用于纯化相应蛋白、糖蛋白或它们的混合物,例如用于亲和层析柱时。

所述重组 MSP4 和 MSP5 多肽可以用作抗原来检测生物样品中疟原虫特异性抗体的存在和评价此类抗体在这些生物样品中的水平,这些抗体水平组成当前或之前感染的信号。此类生物样品自然包括人组织和人细胞,以及生物学液体例如包括人血清的人体液。当在免疫测定中用作试剂测定抗疟疾抗体的存在或浓度时,本发明的所述抗原提供了一种方便、快速、敏感和特异的检测。

更具体而言,可以使用本发明的抗原通过熟知用于检测或定量液体中体液组分的免疫测定法来检测疟疾。因此,通过二级反应例如沉淀或凝集可以直接观察或测定抗原-抗体相互作用。此外,还可以使用免疫电泳技

术。例如，可以使用琼脂电泳和随后抗血清反应的经典组合，以及双向电泳、火箭电泳、和聚丙烯酰胺凝胶谱的免疫标记(蛋白印迹或免疫印迹)。可使用本发明所述抗原的其他免疫测定法包括但不限于放射免疫测定法、竞争性免疫沉淀测定法、酶免疫测定法和免疫荧光测定法。应理解的是还可以使用比浊法(turbidimetric)、比色法和比浊法(nephelometric)技术。基于蛋白印迹技术的免疫测定法是优选的。

通过将所述免疫试剂中的一种，或本发明的抗原或抗该抗原的本发明所述抗体，固定在载体表面但保留所述试剂的免疫反应性可进行免疫测定。所述相互免疫试剂可以不标记或标记，使得免疫反应性也得到保留。这些技术特别适合用于酶免疫测定，例如酶联免疫吸附测定(ELISA)和竞争性抑制酶免疫测定(CIEIA)。

当本发明所述抗原或抗所述抗原的抗体吸附到固相支持物上时，所述支持物通常为玻璃或塑性材料。塑造成板、管、珠或盘形的塑性材料是优选的。适宜塑性材料的实例为聚苯乙烯和聚氯乙烯。如果所述免疫试剂不容易结合到所述固相支持物上，可在所述试剂和所述支持物之间插入载体材料。适宜载体材料的实例为蛋白质(例如牛血清白蛋白)或化学试剂(例如戊二醛或尿素)。所述固相的包被可使用常规技术进行。

本发明提供了免疫原性的重组 MSP4 和 MSP5 多肽，更具体而言，提供了用于制备抗疟疾的疫苗组合物的保护性多肽。通过给易受疟疾感染的哺乳动物施用所述多肽，这些多肽因此可以用作疫苗。可以使用常规的给药方式。例如，可通过口服、舌下、呼吸道或肠胃外途径给药。当所述疫苗经肠胃外给药时，皮内、皮下、肌内和静脉内给药途径是优选的。

感染疟疾哺乳动物的免疫应答的主要目的是灭活疟原虫和便于杀灭疟原虫以及清除感染寄生虫的红细胞。免疫应答中 B 细胞应答主要负责灭活红内期的疟原虫。实现该目标的主要方式是通过中和感染性(抑制红细胞侵入)和抗体依赖性细胞毒性(ADCC)。靶抗原必须是保守的，以有效防御其他寄生虫菌株的再感染，还必须可诱导 T 助细胞活性(CD4+)以产生长期记忆应答。破坏感染寄生虫细胞的 T 细胞介导机制是通过细胞毒性(CD8+)T 淋巴细胞(CTL)提供的。该淋巴细胞可以在感染的红前期识别肝细胞表面与 I 类组织相容性抗原结合表达的重组 MSP4 和 MSP5 抗原。

下面是一种疫苗，它包括(A)天然信号序列(B)C 端 His 标签以及(C)类似恶性疟原虫抗原酸性重复区域的酸性重复区域，其与区域(MSP4p20)的保护性抗体应答关系最为密切。

MKVAYFLSVLDLLIIFSLYFDGRRSAFAGIAACIRHGRILGEGGE/QNSTPGSGGQTG
DHSAEAENG DYNEQGDDHGDDHGDDHGDDHGDEQDGEDYDDAEDDDL YELSE
VDENANLCLDNNGGCGDDKICEN LGKGIVKCLCKPGYKLVGTECVEHHHHHHH
[SEQ ID NO: 30]

本发明所述重组 MSP4 和 MSP5 多肽和疫苗在宿主体内诱发中和抗体的保护性水平(即体液应答期间针对所述抗原所引发的抗体，体液应答直接阻断了病原体感染红细胞的能力)的能力可以通过佐剂(免疫刺激剂)乳化、整合入脂质体、偶联适宜载体或这些技术的组合得到增强。例如，本发明所述重组 MSP4 和 MSP5 多肽可以与常规佐剂施用，例如磷酸铝和氢氧化铝凝胶，其量足以加强宿主体内的体液或细胞介导的免疫应答。

所述免疫程序将取决于几个因素，例如宿主的感染易感性和宿主的年龄。可以给宿主施用本发明的单一剂量疫苗或可以进行间隔施用几种剂量的初次免疫过程。需要时，在初次过程后可以施用用作加强剂量的随后剂量。

本发明的 MSP4 和 MSP5 蛋白、多肽和疫苗可以足够的量给宿主施用以诱发阻止或抑制寄生虫在体内感染和复制的免疫应答，以便降低宿主体内的寄生虫负荷并减轻临床症状。通过将本发明的多肽以每剂量为 10 到 500 毫克、优选每剂量为约 50 到 100 毫克的量给宿主施用可以获得免疫原性应答。本发明的蛋白质和疫苗可与生理学上可接受的载体一同施用。例如可以使用稀释剂如水或盐溶液。

本发明的另一方面提供了 DNA 接种的方法。所述方法还包括在带或不带载体分子情况下给个体施用编码重组 MSP4 和 MSP5 多肽的核酸、所述蛋白质和多肽本身的任意组合。在实施方式中，所述个体为动物，优选地为哺乳动物。更优选地，所述哺乳动物选自人、狗、猫、牛、猪和马。在一特别优选的实施方式中，所述哺乳动物为人。

本领域技术人员知晓核酸疫苗(例如 DNA 疫苗)和核酸疫苗技术以及基于蛋白质和多肽技术的概念、应用和效果。基于核酸的技术允许将编码

重组 MSP4 和 MSP5 多肽的核酸(裸露或封装)直接施用于组织和细胞,而无需在施用前生产被编码的蛋白质。该技术基于这些核酸被受体生物细胞摄取和表达以产生受体免疫系统对之应答的免疫原决定簇的能力。一般情况下,所述表达抗原呈现在业已摄取和表达该核酸的细胞表面,但所编码抗原的表达和输出到受体个体的循环系统也在本发明的范围内。此类核酸疫苗技术包括但不限于裸露 DNA 和 RNA 的输送和编码 MSP4 和 MSP5 多肽的表达载体的输送。尽管该技术名为“疫苗”,但它同样适用于不会引起保护性应答的免疫原性组合物。此类不诱导保护作用的组合物和方法包括在本发明内。

尽管将编码重组 MSP4 和 MSP5 多肽的核酸和载体分子以裸露核酸输送在本发明以内,但本发明还包括作为更大或更复杂组合物一部分的核酸的输送。包括在这些输送体系内的有病毒、病毒样颗粒或含有编码重组 MSP4 和 MSP5 多肽的核酸的细菌。此外,本发明的核酸和载体分子与可透过细胞的化合物(例如脂质体)的复合物包括在本发明的范围内。其他化合物,例如分子载体(EP 696,191, Samain 等人)和核酸疫苗输送系统为技术人员知晓,并例示在例如 WO 93 06223 和 WO 90 11092, U.S. 5,580,859 和 U.S. 5,589,466 (Vical 公司专利),它们通过引用的方式并入本文,并且这些化合物可以制备和使用而无需不适当或过多的实验。

本发明将在下列具体实施方式中得到更为详细的描述。

长期以来人们已知道使用杆状病毒表达系统产生的抗原忠实地含有参与半胱氨酸键形成的复合物结构,且适合于结晶研究。这些抗原通常被认为是检测其他系统产生的抗原的“金标准” [24,25]。由于 PfMSP4 和 PfMSP5 均含有由二硫键形成的 EGF 样结构域,它们可能利于在杆状病毒系统内生成。但是,正如通常使用的所有表达系统一样,蛋白质产量受到了恶性疟原虫异常密码子使用的有害影响[26,27]。此外,这些高度富含 A+T 的序列经常在大肠杆菌内发生突变[28],且很少含有独特的限制性酶切位点。出于这些原因,采用 Withers-Martinez (1999) [29]描述的方法设计并构建了合成基因。

设计合成基因

基于当时得自 GenBank 的序列资料和预期将来用途包括人攻击模型(challenge model), 选择常见的氨基酸序列 PfMSP4 NF54 序列(保藏号 AF295318)和 PfMSP5 3D7 序列(保藏号 AF106476)用于修饰。与此同时, 为了丰富序列数据集和进行种群间和种群内分析以及物种间分析, 使用取自如下具有不同地方性的位点的样本进行了多态性研究。

msp4 和 msp5 基因多态性概况

研发抗恶性疟原虫的疫苗时, 抗原多态性是一项重要的考虑因素。迄今为止检测过的最有前景的红内期候选疫苗是裂殖子表面蛋白 1(MSP1)和顶端膜蛋白(AMA1 [1-3])。但是, 与编码不易被免疫效应物机制接近的抗原的基因相比, 这些和其他表面抗原基因表现出高得不成比例的非同步单核苷酸多态性(nsSNP)的频率[4-6]。作为一个混杂因素(compounding factor), 这些 nsSNP 经常导致根本性的氨基酸置换, 这些氨基酸置换主要聚集在多肽最易被宿主免疫系统接近的区域。据信, 所形成的氨基酸置换可通过改变重要 B 细胞和 T 细胞表位在免疫侵入中发挥作用[34,38]。因此, 以此类抗原完整重组形式进行接种可能会引起菌株特异性的保护[4,8]。普遍认为通用有效的疫苗会包括源自多阶段靶点(multi-stage target)的几种表面蛋白的保守区域[30]。

到目前为止, 仅有有限量的关于 PfMSP4 和 PfMSP5 多态性的资料[11,31,32]。为了充分探究来自疾病高度流行地区的 PfMSP4 和 PfMSP5 二者的多态性, 开展了序列分析研究, 其结果在 Poison 等人 2005[33]推出。总之, 发现 *Pfmsp5* 基因序列高度保守, 并潜在地处于纯化选择中。据发现, *Pfmsp4* 基因序列对一种恶性疟原虫表面抗原而言相对保守[5]并含有多态性位点的 N 端聚簇, 该 N 端聚簇包括两个潜在地处于平衡选择中的位点(N52 和 G74)。该发现可解释所观察到的旁侧序列内部显而易见的中性(或搭载)多态性的聚集, 它可以与 B 细胞或 T 细胞表位的变异联系在一起。典型情况下平衡选择起于携带影响分子免疫识别的“差异”的两个等位基因序列的存在。当等位基因 A 在群体中最为常见时, 成为等位基因 B 是有利的, 这是因为该等位基因不会被免疫系统以及显性等位基因所识别。因此, 等位基因 B 变得更加普遍, 直至其成为最常见类型, 免

疫系统也将识别等位基因 B，此时成为等位基因 A 是有利的。首先，平衡选择的存在表明所述 MSP4 蛋白是有效免疫机制的靶点。其次，这表明为了保护疫区的人们，你需要用这两个等位基因来接种，或排除该蛋白质的此区域，迫使免疫系统攻击保守区域。至于近来鉴定的可特异结合肝细胞的两个源自 PfMSP4 的肽[34]，第一个肽(代表 MSP4p40 76-92 位的残基)携带一个半保守的多态性位点 (A81)，而第二个肽(代表 MSP4p40 113-135 位的残基)含有一个缺失(115-119 位)和/或一个半保守的多态性位点(G119)。在现有的 PfMSP4 数据集内，这些序列内部差异出现的频率相对低，从而支持了这一观点，即它们是重要的功能位点(注意：这些序列也由 SALSA 所代表，图 2)。最后，还有一个潜在地处于平衡选择的位点 V190，它位于 EGF 样结构域首个半胱氨酸上游 17 个残基处，该位点可能也参与了 B 细胞或 T 细胞表位的修饰。

合成基因的构建：PfMSP4

首先，为了便于蛋白质分泌，除去从 EGF 结构域下游的 3 个残基起的全部序列。这包括了 GPI 锚定信号，因为已知杆状病毒系统可以使用此序列来整合昆虫细胞 GPI 部分并导致该蛋白在细胞表面定位(Bonnet 等人 2006)。此外，保留了天然信号 MSP4 序列，这是因为已知杆状病毒表达系统会正确切割重组类似物中的天然 MSP1 信号序列，尽管 MSP4 的天然切割 N 端尚未被确定。为了防止 N-糖基化(尚不知疟原虫会发生 N-糖基化，但该反应确实会发生在杆状病毒 PfMSP4 序列的两个潜在位点上)，用丙氨酸取代丝氨酸残基 S₆₅ 和 S₇₃。两个残基均具有小的侧链(Ser = HO-CH₂-, Ala = CH₃-)，尽管没有报道表明这些位点在自然状态下是多态性的，但据信这些变化会对该蛋白的局部或宏观结构产生微小影响。一旦加入余下包括 C 端六个组氨酸标签和终止密码子的特点，则使用 CODOP 程序设为粉纹夜蛾(*Trichoplusia ni*)(High Five)细胞的密码子用法对该序列进行回译(表 1)。

NF54 和合成基因的 MSP4 密码子用法

密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5		
GCA	Ala	2	4	GAA	Glu	36	25	AAA	Lys	24	5	ACA	Thr	1	3		
GCC		0	1	GAG		4	15	AAG		7	26	ACC		1	3		
GCG		1	0	总计		40	40	总计		31	31	ACG		0	0		
GCT		3	3	GGA		Gly	11	13		TTC	Phe	0		4	ACT	5	1
总计		6	8	GGC			1	2		TTT		4		0	总计	7	7
AGA	Arg	4	0	GGG	3		0	总计	4	4		TGG	Trp	1	1		
AGG		0	1	GGT	4	4	CCA	Pro	4	2	总计	1		1			
CGA		0	0	总计	19	19	CCC		1	1	TAC	Tyr	0	6			
CGC		0	1	CAC	His	1	12		CCG	0	0		TAT	6	0		
CGG		0	0	CAT		6	1		CCT	2	4		总计	6	6		
CGT		0	2	总计		7	13	总计	7	7	GTA	Val	5	3			
总计		4	4	ATA	Ile	5	1	AGC	Ser	1	3		GTC	0	5		
AAC		Asn	2	17		ATC	2	5		AGT	6		4	GTG	1	3	
AAT	15		0	ATT		3	4	TCA		4	1		GTT	10	5		
总计	17		17	总计		10	10	TCC		3	1		总计	16	16		
GAC	Asp	7	22	CTA	Leu	2	1	TCG		0	1	密码子		p.f.	H5		
GAT		18	3	CTC		0	2	TCT		5	7						
总计		25	25	CTG		0	0	总计		19	17						
TGC	Cys	0	4	CTT		0	3	TAA		终止	0	1	%GC	245	252		
TGT		7	3	TTA		6	0	TAG			0	0					
总计		7	7	TTG	2	4	TGA	0	0								
CAA	Gln	5	3	总计	10	10	总计	0	1		30.6	47.5					
CAG		0	2	ATG	Met	4	4										
总计	5	5	总计	4		4											

CODOP 是执行密码子优化的 Unix perl 脚本, 该程序由 Hale 和 Thompson 提出[35]并在其他文献中详细描述[29]。随后采用手动干预来优化重新编码密码子的核苷酸序列以生成具有 60-65°C T_m 的 38 x 40 寡聚核苷酸(mer)并含有数个独特的限制性位点以便于亚克隆(图 3)。

该基因的总 GC 含量增加了 15%, 显著增加了亚克隆的易操作性和效率并降低大肠杆菌在构建体操作中犯下的复制错误的频率。所述 38 个重叠寡核苷酸得自 Eurogentec(仅有标准纯化), 它们每个长 40 个碱基。根据前述方法实现基因装配和扩增[29]。

简而言之, 混合等摩尔量(各 25 μM)的基因寡核苷酸, 并被稀释 10 倍到含 1 μl 的 pfu DNA 聚合酶(Stratagene), 5 μl 10x 反应缓冲液, 200 μM dNTP 和 2 μl 25 mM MgSO₄ 的 50 μl PCR 反应液中。PCR 程序由 94°C, 1min 的变性步骤和随后的 25 个循环(94°C, 30 s, 52°C, 30 s 和 72°C, 2 min)组成。为了扩增全长基因产物, 将 5 μl 基因装配反应液稀释 10 倍达到 50 μL PCR 反应液, 其含有 1 μl pfu DNA 聚合酶(Stratagene), 5 μl 10x 反应

缓冲液, 200 μ M dNTP, 2 μ l 25 mM $MgSO_4$ 和各为 500nM 的第 1 和 20 号外基因装配寡核苷酸。PCR 程序由 94°C, 1min 的 1 个循环, 25 个循环 (94°C, 45 s, 68°C, 45 s 和 72°C, 2 min) 和 72°C, 10 min 的最终延伸步骤组成。凝胶提取图 4 所示的主要和大小正确的片段, 并将其克隆进 pMOSBlue 进行测序, 最后将其转移到载体 pVL1393 以整合进杆状病毒基因组(Baculogold, PharMingen)。该表达构建体被称为 MSP4p40/His。

修饰的 PfMSP4 构建体

还组装了另外两个 MSP4 构建体, 分别名为 MSP4-EGF/His 和 MSP4p30/His (图 5)。所述构建体 MSP4-EGF/His 设计用于仅表达 MSP4 EGF 样结构域, 主要用于结晶研究。通过高保真 PCR 生成含有 PlasmidDB 所述的预测信号序列(1-20 残基)和 2 个下游残基的 PCR 片段。反应含有基因装配寡核苷酸 1 和反向引物 MSP4modI (5'-TAT-AGC-AGA-TCT-TTG-TCG-AAG-TTG-ATG-GTG-CA-3') [SEQ ID NO: 23], 该反向引物含有一个 *Bgl* II 限制性位点。所述合成基因克隆 pMosMSP4 被用作模板。所形成的 86 bp PCR 产物以限制性内切酶 *Bam* HI 和 *Bgl* II 切割并凝胶提取。然后将该产物连接(T4 DNA 连接酶; NEB)到之前制备好的 pMosMSP4 载体, 同样以 *Bam* HI 和 *Bgl* II 消化以去除 EGF 样结构域上游的全部序列, 只余下 D205 到 H252 之间残基。

所述 MSP4p30/His 构建体设计用于去除 MSP4 C 端的 30 个残基(45 到 74 位残基), 其中报道的多态性大部分位于此处(在图 5 中突出显示)。使用基因装配寡核苷酸 1, 反向引物 MSP4modII (5'-ATA TGG CTG CAG CCA AGA TCC TCA TGT TAA GCA T-3') [SEQ ID NO: 24]和 pMosMSP4 载体作为模板来产生编码 1 到 44 位残基的 PCR 片段。该片段大小为 154bp, 分别在 3'和 5'最末端具有 *Bam* HI 和 *Pst* I 的限制性位点。使用限制性内切酶 *Pst* I 和 *Spe* I 从 pMosMSP4 载体切下编码 A75 残基下游全部序列的 DNA 片段(它位于 pMOSSBlue 载体序列内的 MSP4 ORF 下游)。这两个 DNA 片段在 T4 DNA 连接酶(NEB)反应中以等摩尔的量结合并于 4°C 下温育 3 天以上。将 2 μ l 连接酶反应液在溴化乙啶染色的 1%琼脂糖凝胶上分离, 结果揭示了几种产物的存在, 包括长 724 bp 的所需产物。

使用基因装配寡核苷酸 1 和 20 和 1 μ l 连接反应物作为模板通过高保真 PCR 扩增该产物, 凝胶提取该产物并将其克隆回 pMOSSBlue。与合成基因一样, 在被克隆到表达载体 pVL1393 前对 MSP4-EGF/His 和 MSP4p30/His 基因序列进行证实。

从前述合成 MSP4 基因还设计和构建了另一个表达构建体。该表达盒设计允许该蛋白质的非多态性、抗蛋白酶和结构完整的 C 端区域(称为 p20, 并在下面部分描述)的表达。该构建体包括 MSP4p40/His 的 *xba* I 克隆位点下游的全部核苷酸序列, 因此它编码 Leu₁₃₀(不包括在内)下游的所有残基。为了便于蛋白分泌, 使用基因装配寡核苷酸 1 和反向引物 MSP4p21 rev (ATT AAT CTA GAG GCT TTT CTT CAC CCA AGA TCC TCA TG) SEQ ID NO: 25 生成编码 MSP4 信号序列以及下游 9 个残基 (Met₁- Pro₄₉) 的 PCR 片段并将其连接到 Bam HI/Xba I 双消化的 pMosMSP4p40/His。该构建体被称为 MSP4p21/His, 其序列在连接到 pVL1319 以整合入杆状病毒基因组前进行证实。从上述合成 MSP4 基因还设计和构建了另外两个补充类似表达构建体。这些构建体命名为 MSP4p21 ss1 和 MSP4p21 ss2, 二者均编码 Leu₁₃₀ 下游的所有残基。为了便于蛋白分泌, 加入与 MSP4p21/His 相同的编码 MSP4 信号序列的序列 (Met₁- Pro₄₉)。此外还加入了 MSP4p40/His 构建体的一些密码子, 它们编码 MSP4p21 ss1 构建体中的序列 (Asn₅₀-Ser₅₇ 和 Leu₁₃₀) 和编码 MSP4p21 ss2 中的氨基酸残基 Leu₁₃₀。

PfMSP5 合成基因的构建

使用与 *Pfmsp4* 相同的方法来设计 *Pfmsp5* 合成基因。从 3D7 菌株的 GenBank 蛋白序列开始, 除去 GPI 锚定信号上游 3 个残基起的全部序列并替换为六组氨酸标签, 3 个糖基化位点被丝氨酸到丙氨酸的突变所破坏 (S83, S102 和 S126), N 端信号序列保持完整。CODOP 所执行的密码子修饰记录在表 2 中, 图 6 显示了整条合成基因序列。

3D7 和合成基因的 MSP5 密码子用法

密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5
GCA	Ala	2	1	GAA	Glu	23	23	AAA	Lys	15	5	ACA	Thr	2	0
GCC		2	2	GAG		6	6	AAG		3	13	ACC		1	2
GCG		0	1	总计		29	29	总计		18	18	ACG		2	0
GCT		0	3	GGA	Gly	6	5	TTC	Phe	1	6	ACT		1	4
总计		4	7	GGC		0	3	TTT		5	0	总计		6	6
AGA	Arg	6	2	GGG		0	0	总计		6	6	TGG	Trp	0	0
AGG		0	2	GGT		6	4	CCA	Pro	4	3	总计		0	0
CGA		0	0	总计		12	12	CCC		0	3	TAC	Tyr	0	7
CGC		0	2	CAC	His	0	9	CCG		0	0	TAT		7	0
CGG		0	0	CAT		4	1	CCT		5	3	总计		7	7
CGT		1	1	总计		4	10	总计		9	9	GTA	Val	3	2
总计		7	7	ATA	Ile	10	0	AGC	Ser	3	5	GTC		1	4
AAC	Asn	2	44	ATC		1	10	AGT		7	8	GTG		1	2
AAT		42	0	ATT		5	6	TCA		8	3	GTT		4	1
总计		44	44	总计		16	16	TCC		1	4	总计		9	9
GAC	Asp	2	13	CTA	Leu	3	2	TCG		1	1	密码子 247 254 % GC 45.8			
GAT		13	2	CTC		0	7	TCT		5	1				
总计		15	15	CTG		0	2	总计		25	25				
TGC	Cys	0	4	CTT		2	2	TAA	终止	0	1				
TGT		7	3	TTA		10	1	TAG		0	0				
总计		7	7	TTG		2	3	TGA		0	0				
CAA	Gln	5	4	总计		17	17	总计		0	1				
CAG		1	2	ATG	Met	6	6								
总计		6	6	总计		6	6								

按照前述方法进行 PfMSP5 合成基因的 PCR 装配, 由于该序列内缺乏重复基序, 因而生成了高纯的终产物(图 7)。凝胶提取该 PCR 产物并将其克隆进 pMOSBlue 并在转移到 pVL1393 载体以整合入杆状病毒基因(Baculo-gold, Stratagene)组前进行测序。

评价根据此方法制备的构建体, 其结果记录在下列实施例中。

实施例 1

PfMSP4 在昆虫细胞中的表达

使用标准方法[36]生成病毒后, 在 150 cm² 培养瓶(Corning)中进行试感染。病毒以感染复数为 10 进行 1 小时的感染。3 天后收集上清并分批在 talon 树脂上纯化。使用 MSP4p40/His (40 kDa)和 MSP4p30/His (30 kDa)构建体观察到预测大小的蛋白产物, 但使用 MSP4-EGF/His 构建体没有观察到蛋白质(将在下文进一步讨论)。MSP4p21/His 的表达仍处于初步分析阶段。

MSP4p40 和 MSP4p30 的生产被放大并从箱转移到旋转瓶培养。为了优化蛋白产量和限制产物降解(常见于更为剧烈的旋转瓶条件下),进行了时间过程实验。在草地夜蛾(*Spodoptera frugiperda*)(SF9) 昆虫细胞和 High Five 昆虫细胞中建立起感染 MSP4p40/His 或 MSP4p30/His 的旋动培养(210 mL)。在 24 h, 30h, 36 h, 42 h, 48 h, 54 h, 60 h 和 72 h 时收集样品(8 ml), 显微镜下分析样品进行细胞死亡统计, 然后以 20 mM Tris HCl (pH8); 500 mM NaCl 对样品进行透析。使用试管方案在 talon 树脂上纯化蛋白, 并将蛋白洗脱到 100 μ L 100mM 咪唑; 20 mM Tris HCl (pH8); 500 mM NaCl。绘制每次感染随时间变化的细胞死亡百分比并平行分析每次洗脱 20 μ L 样品的构建体和细胞类型(图 8)。

正如预计一样, High five 细胞的蛋白产量远高于 SF9 细胞。但是, 观察到在长表达期(超过 48 h)后, 此时蛋白开始降解, MSP4 蛋白没有完全降解, 而是降解成约为 20 kDa 的更小稳定产物(MSP4p20)。在 MSP4p30 感染的后期也观察到该产物。由于所有蛋白都是通过 C 端 his 标签纯化的, 该产物必须包括 EGF 样结构域和一些上游序列, 以形成抵抗蛋白酶的紧密结构。这与 Wang 等人(2001) [13]的发现吻合, 即 EGF 样结构域的上游区域携带有还原敏感性的 B 细胞表位。

实施例 2

N 端测序和质谱

使用得自时程实验的信息, 通过在不同时间点收集培养 SN 生成三种不同的纯 MSP4 蛋白制品。首先, 全长 MSP4 即 MSP4p40, 其次, 全长修饰产物即 MSP4p30, 第三, 降解产物 MSP4p20。标准 High Five 旋动培养中每个产物的平均蛋白产量分别为 8 mg/L, 6 mg/L 和 15 mg/L, 这些平均蛋白产量反映了收集时间, 即产物收集越早蛋白产量越低。根据近来市售杆状病毒 (Henogen)性能的进展, 据信这些产量是最保守估算。

每个产物的 N 端测序生成了一个强而清晰的信号(p40:从残基 41 开始-MRILG, p30: 从残基 41 开始-MRILA, p20: 从残基 132 开始-KSPKE), 并揭示体内恶性疟原虫的信号序列可能实际上是 PlasmoDB 预测长度的两倍(40 个残基比 20 个残基) (<http://plasmodb.org/plasmodb/servlet/sv?>

page=gene&source_id=PFB0310c)。MSP4p40 和 MSP4p20 的特性也已为质谱证实，这些蛋白分别具有 23.554 kDa 和 13.714 kDa 的实际分子量。

杆状病毒表达的 MSP4 的 N 端序列可能识别寄生虫使用的信号序列。首先，MSP4-EGF/His 构建体含有 PlasmODB 预测的信号序列和 3 个下游残基，并且从未在培养 SN 中检测到。其次，在 *Pfmsp4* 基因内部鉴定的多态性位点的 N 端聚簇从第 45 个密码子开始。正如 Polson 等人 2005 [33] 报道的分析表明，密码子 45 下游的基因区域是有效免疫应答的靶点，该基因区域似已通过群体内平衡等位基因型(涉及残基 52 和 74，参见图 5)进化出对免疫应答的逃避机制。因此有可能加工后的 PfMSP4N 端由 4 个构成信号序列切割位点(MRIL)的保守残基和随后紧接的多态性位点的聚簇组成。这些多态性位点能够降低优先针对该蛋白更易接近区域的免疫应答的有效性。

实施例 3

信号序列疏水性

为了进一步证实杆状病毒系统对 MSP4 构建体使用的信号切割位点复制了天然切割，进行了一项针对之前报道的天然恶性疟原虫信号序列的生物信息学研究。2001 年发表的一项研究对已知恶性疟原虫信号序列切割位点进行了比较[37]。由于获得此类数据存在技术上难题，仅有 10 个此类序列存在。出于本研究目的，选择和提交了(通过 Pasteur 服务器)蛋白序列的 80 个残基以计算局部疏水性的数值，这些残基主要集中在每个蛋白报道的信号序列切割位点周围。图 10 显示了 Nacer 于 2001 年报道的三个蛋白切割位点的疏水性曲线和杆状病毒系统的 MSP4p40 的疏水性曲线。在上方图中，KAHRP 和 MSP1 信号序列表现出经典的疏水性曲线，其切割刚好出现在疏水性数值跨过代表中性的直线后(正对负)。已知 PfMSP1 信号的切割是由杆状病毒系统忠实地进行的(Bonnet 等人 2005)。局部疏水性的相同显著变化没有出现在 MSP4p40 蛋白序列 N 端预测的切割位点，切割实际上发生在疏水性最大的局部变化后的下游(负对正)。该曲线符合使用 S-抗原所观察到的结果，因此并非独特的，这为 MSP4p40 信号序列长 40 个残基这一发现增加了分量。

实施例 4

PfMSP5 在昆虫细胞中的表达

使用前述方法生成 MSP5/His 病毒原液并根据用于 MSP4p40/His 的方法进行试感染和时间过程研究。如图 11 所示,检测到 SDS-PAGE 凝胶电泳上分别以 45、35 和 10kDa 迁移的源自 PfMSP5 的三个不同产物。有趣的是,观察到两个较大的产物同时出现在时间过程实验的早期,而不是生成较大产物并随后降解形成较小产物。使用 PfMSP4 观察到的另一明显差异是研究的两个细胞系相反的蛋白质生产能力,即 SF9 细胞产生蛋白质的能力远超过生长更快和更脆弱的 High Five 细胞。这一发现本身是表明 PfMSP5 具有显著不同于 PfMSP4 性质的首个证据。第二个证据来自 N 端测序数据。

实施例 5

N 端测序和质谱

目前尚不能获得 MSP5 表达的两个最大产物的 N 端信号,这表明它们均携带有阻断多肽骨架降解的 N 端修饰,而这正是 N 端测序中必不可少的事件。较小的 10kDa 产物的测序取得成功,该测序确定该产物的 N 端起自残基 189: YNKVE [SEQ ID NO: 26]。该产物的身份已为质谱所证实,它具有 7.746 kDa 的实际分子量。此外, MSP5 病毒基因组的序列已被证实, YNKVE [SEQ ID NO:27]残基只位于正确的 ORF 内,因此这些尚未被确定的产物一定是 PfMSP5。

MALDI-MS 获得的两个最大 MSP5 产物的质量数值分别为 24.679 和 20.551 kDa,不匹配于任何只源自 MSP5 蛋白序列的多肽,因而支持这一观点即这些蛋白携带转录后修饰。为了研究可能的 N 端封闭修饰,将 MSP5 序列通过 ExPASy 主页提交到 prosite。这表明了两个 N-豆蔻酰化位点的存在,它们可以指导 C14 脂肪酸的共翻译加入(残基 42:GGFTSK 处和残基 66:GSLPTK 处) [38]。然而,尽管这些修饰只位于 N 端并可阻止 N 端测序,它们不会产生适宜的质量数值(分别为 24,015 和 21,301)。但是,基序引导的 N 端甘氨酸残基修饰不是真核细胞进行的豆蔻酰化的唯一形式。它们还可转录后修饰任何具有豆蔻酰基的 N 端甘氨酸和任何

赖氨酸(下文评述)。带着这些可能性, 假设每个 MSP5 产物 (p45 和 p35) 携带两个脂肪酸修饰, 得到了 24.673 kDa 和 20.553 kDa 的近似匹配, 但目前尚不确定假设位点。

实施例 6

脂肪酸修饰和疟原虫

真核细胞使用了几种不同类型的额外翻译修饰, 包括糖基化、棕榈酰化、N-豆蔻酰化和豆蔻酰化。其中此处表达的 MSP5 蛋白不具有第一种翻译修饰。在合成基因的构建过程中从该蛋白序列除去所有的糖基化位点, 因为已知恶性疟原虫不会进行该类修饰。真核生物的第二类修饰尚未得到很好阐明, 据信棕榈酰化可通过依赖酶和不依赖酶的途径发生。迄今为止尚无任何棕榈酰化的恶性疟原虫蛋白的报道。但是, 已经表明蛋白质的 N-豆蔻酰化在恶性疟原虫中的重要性。已鉴定出一个 PfN-豆蔻酰转移酶基因的同系物(PlasmoDB PF14_10127), 并发现其与人类和酵母基因直向同源物高度同源[40,41]。如果该酶在疟原虫的作用与其在其他寄生原生动物的例如布氏锥虫(*Trypanosoma brucei*)和硕大利什曼原虫(*Leishmania major*)一样重要, 则多个蛋白质的 N-豆蔻酰化预计对寄生虫生长起着必不可少的作用[41]。事实上, 已表明 PfADF(核糖体因子)的 N-端豆蔻酰化对介导 GTP 结合和发挥酶活性起着重要作用[42]。已表明该脂肪酸修饰(N-端甘氨酸和内部赖氨酸上)在较高等的真核生物中起着许多作用, 其中涉及蛋白膜定位和蛋白-蛋白相互作用的稳定[39, 43-45]。

实施例 7

杆状病毒系统中

由于昆虫细胞的杆状病毒感染是更高等的真核生物系统, 它可以执行许多更为进化类型的蛋白修饰。事实上, 该系统已被专门用来生产几种 N-豆蔻酰化的蛋白, 其中了解最多的蛋白是 NAP-22[46]。

实施例 8

制备抗体

通过 EUROGENTEC 使用标准免疫方法生成重组 MSP4p40、MSP4p20 和 MSP5 (p45 和 p35)的兔多克隆抗血清。免疫前所有兔对 MSP4 和 MSP5 呈阴性(每种抗原两只兔子), 3 至 4 次免疫后经 ELISA 测定产生了高效价的特异血清抗体。图 12 显示了抗还原和非还原重组蛋白的终点效价。如图 B 所示, 当所述抗原被还原时, 来自 8 号兔的抗体对 MSP4p20 的识别减少了 50%以上, 这表明 MSP4p20 蛋白的许多表位是构象型的。由于还原主要影响具备二硫键的结构, 该数据表明 MSP4p20 的 EGF 样结构域保持完整并参与了形成至少一个 B 细胞表位。

实施例 9

源自寄生虫的 MSP4

使用前述的多克隆免疫血清生成免疫印迹的初步数据, 图 13 显示了该数据。MSP4 的数据与 R. Coppel 小组[9]发表的结果一致, 即只在后期寄生虫制品的溶于 Triton X100 的部分中检测到一个 40 kDa 的蛋白质(图 13A)。为了研究 MSP4 C 端区域 MSP4p20 的性质, 亲和纯化了抗固定的 MSP4p20(Amersham:NHS-活化 Sepharose 4 快速流动)的兔血清库(R6-9)和 2002 年采集自塞内加尔 Dielmo 的人血清库。当用这些血清探测非同步寄生虫材料时(以 SDS-PAGE 分离并转移到硝酸纤维素膜上), 除之前在 40 kDa 处观察到的条带, 还观察到一条约为 18 kDa 的条带(图 13B), 这表明自然感染中 MSP4 抗原可能经历了类似于 MSP1p42 的蛋白酶解加工或可能被蛋白酶以自限方式降解, 如杆状病毒感染后期所获得的。为了进一步证实恶性疟原虫裂殖子上 MSP4p20 表位的自然存在, 对风干的非同步生长的寄生虫培养物进行了间接免疫荧光测定(IFA)。使用人抗 MSP4p20 可见裂殖子特异性染色, 而单用二抗或使用未感染疟疾的血清未观察到有染色(图 13C)。

实施例 10

源自寄生虫的 MSP5

使用抗杆状病毒表达的 MSP5 制备的多克隆血清得到了免疫印迹数据, 该数据表明其曲线与其他研究组发表的结果迥异, 这些研究组显然面

临检测 MSP5 的难题并暗示它是低水平表达的。本发明的免疫印迹数据显示了抗 MSP5 特异性血清的强烈信号(图 13D)。与 MSP4 不同的是,源自寄生虫的 MSP5 也存在于粗制备的寄生虫提取物的膜部分,这表明它没有采用与仅携带一个 GPI 锚点的蛋白质相同的方法来分部(fractionate)。事实上,当考虑到还原的寄生虫材料时,有可能得出结论,即针对假设 N-豆蔻酰化的 MSP5 而制备的多克隆血清可以与其他 N-豆蔻酰化的蛋白质发生交叉反应,进一步支持这一观点,即这是恶性疟原虫中一种重要且经常使用的翻译后修饰。

实施例 11

人免疫血清的分析

为了证实本发明的蛋白产物为受疟疾感染个体的抗体所识别,采用 ELISA 分析 9 名随机选择的半免疫成人的血清,他们生活在塞内加尔一个疾病流行村庄(Dielmo)(图 14)。所有个体对所有蛋白产物均呈阳性(具有大于对照血清的 OD 值以及 3 倍的标准差)。个体间和产物间的相对效价各异,但 MSP4 和 MSP5 的总终点效价分别为 1×10^5 和 1×10^4 。使用来自 2 例 MSP4 呈高度阳性的个体(个体 3 和 9)和 1 例 MSP5 呈高度阳性的个体(个体 8)的血清进行蛋白印迹分析(图 15)。正如前面使用抗 MSP1 p19 观察到的结果一样,抗含有 EGF 的裂殖子表面蛋白的人抗体对聚集材料具有特别的亲和力(图 15)。

实施例 12

塞内加尔的抗 MSP4 抗体;保护作用

为了进一步研究自然感染中抗 MSP4 抗体的假设作用,进行了一项流行病学研究。一项大规模前瞻性研究评价了三种不同纯化形式的 MSP4 即 p40, p30 和 p20, 该研究已发表(参见参考文献[47])并在下面的实施例 16 中得到详细描述。

总之,每种形式的重组 MSP4 蛋白均被强烈识别,其血清阳性率 >90%,但表现出显著不同的平均 IgG OD 比($p40 = 20 \pm 10$, $p30 = 12.3 \pm 7.3$ 和 $p20 17.3 \pm 10$)。与观察的所有疟疾抗原一样,抗体效价随年龄增长,必

须使用年龄校正渐进模型对抗体效价与保护之间的关联性进行研究。使用这种形式的分析,除了抗 MSP4p40 的抗体,抗 MSP4-p30 和抗 MSP4-p20 的抗体与临床疟疾发作的发生率减少呈正相关(P 分别为 0.036, 0.018+0.067)。这些发现支持 MSP4 蛋白 N 端区域内潜在平衡多态性的鉴定[33],具体而言支持在基于 MSP4 的疫苗构建体中删除所述蛋白质的该区域这一思想。

实施例 13

塞内加尔的抗 MSP5 抗体;保护作用

为了进一步研究自然感染中抗 MSP5 抗体的作用,进行了一项流行病学研究。根据别处描述方法(参见参考文献[47])对含有所述两个高分子量形式的 MSP5 的制品进行分析。

尽管 MSP5 的血清阳性率低于 MSP4(59%),使用年龄校正渐进模型观察到其与减少的临床疟疾发作具有极为显著的统计学相关性($P=0.0028$)。正如上文讨论,该抗原可能携带脂肪酸修饰,该修饰可能参与脂类诱导的免疫原性和抗感染免疫的免疫应答,正如 GPI 修饰所观察到[48];Bonnet 等人 2005,待发表)。

* * *

总之, PfMSP4 具有一个 C 端 EGF 样结构域并通过 GPI 锚定子连接到裂殖子表面。尽管尚不知晓其功能,但它似乎对于寄生虫生存是必不可少的,这是因为尚未观察到存活的 PfMSP4“敲除”突变体。三个重组 PfMSP4 构建体在杆状病毒昆虫细胞表达系统中以可溶、分泌蛋白表达,以优化天然抗原的复制,包括 EGF 样结构域的适宜折叠。两个构建体分别对应于具有或不具有来自 N 端(MSP4p30)和 MSP4p40 附近多态性区域的 30 个氨基酸缺失的 MSP4 菌表抗原(不含 C 端疏水残基以允许分泌)。第三个 20kDa 的蛋白大致对应于含 EGF 结构域的 MSP4 的 C 端部分(MSP4p20)。使用前述构建体 MSP4p21/His 正在对该抗原更为直接的表达进行评价。一项纳入 205 名个体并针对暴露于疟疾的人的自然获得性免疫的横断面研究对抗 MSP4 抗体对 3 种不同抗原的应答进行比较。这些个体居住在疟疾中度流行的村庄,其疟疾传播为季节性的。

传播季节前, 所述 3 个构建体被强烈识别, 其血清阳性率超过 90%, 但却具有显著不同($P < 0.01$)的平均 IgG OD 比, MSP4p40, MSP4p30 和 MSP4p20 分别为 20 ± 10 , 12.3 ± 7.3 和 17.3 ± 10 (效价约为 5×10^{-3})。抗体对 MSP4 构建体的应答与 15 岁以下个体的年龄相关, 其 IgG 水平显著低于成年人($P < 0.001$, Rho 0.25-0.33)。在随后包括传播季节的 6 个月内, 年龄校正模型的临床发作分析表明 IgG 的存在对缺乏多态性区域的 MSP4p30 和 MSP4p20 作出应答而不对具有多态性区域的 MSP4p40 作出应答, 这一特点与疟疾临床发作发生率下降呈显著相关($P < 0.05$, rate # 0.75)。总之, 这些结果表明: (i)使用杆状病毒重组抗原监测到 IgG 对 PfMSP4 存在强烈的自然获得性应答; (ii)抗 MSP4 的 IgG 与抗临床疟疾的保护有关, 以及; (iii)IgG 对可变区域的应答可能会干扰本来与保护相关的抗 MSP4 IgG。

实施例 14

PfMSP4 单克隆抗体

四个 PfMSP4 特异性单克隆抗体(mAb)识别 3 个不同的表位。

单克隆抗体 L11-16 代表了其他 2 个由 p40(完整 PfMSP4 多态性胞外结构域)诱导的单克隆抗体。其还原非敏感表位(17B)位于 PfMSP4 的 N 端部分(non-p20; 16A)。它与 PfMSP4 的异质聚集物反应尤为强烈, 该聚集物的形成是还原敏感的(图 16A)并可表明非还原(nR)和部分还原(R)SDS-PAGE 中的高分子量物质的确是 PfMSP4(正如之前的杆状病毒 PfMSP1 p19)。

单克隆抗体 F12-7 由 PfMSP4 p30(缺少多态性区域)诱导并识别 PfMSP4 的 p20 C 端部分的一个还原非敏感的表位(图 16B 和 17B)。与 L11-16 相比, 它与聚集物反应性明显降低可能是因为它小得多的亲和力:

克隆	同种型	KD/MSP5	表位
L11-16	IgG1	3.2×10^{-9} M	A
F12-7	IgG1	1.0×10^{-7} M	C

实施例 15

PfMSP5 单克隆抗体

7 个 PfMSP5 特异性单克隆抗体(mAb)识别 5 个不同的表位。

单克隆抗体 G21-2 识别 p45 和 p35 PfMSP5 产物(图 18A), 其表位不是还原敏感的(图 18A 和 19A)。它与不相关联的较高分子量的聚集物反应(图 18A, nR), 这些聚集物的形成是还原敏感的(图 18A, R)。

单克隆抗体 J18-14 仅识别 p45, 其表位是还原敏感的(图 19B)。它也与聚集物反应。

由于这些抗原是通过 C 端的六组氨酸标签而纯化的, p45 和 p35 被认为具有不同的 N 端。这些特点表明 J18-14 表位可能对应于涉及 C 端 EGF 结构域和 N 端的构象结构。两个 mAb 具有类似的亲和力:

克隆	同种型	KD/PfMSP5	表位
G21-2	IgG1	$2.8 \cdot 10^{-8} \text{ M}$	B
J18-14	IgG1	$4.2 \cdot 10^{-8} \text{ M}$	D

这两个 mAb 的反应性可以用于确定 PfMSP5 产物。

实施例 16

暴露个体的免疫血清和队列

下列结果是使用(A)取自生活在疟疾流行区域且未表现出疾病症状个体的超免疫血清的对照库, (B)2005 年 6 月收集自 Dielmo 村(全地方病传播)的个体血清和(C)2000 年 7 月到 8 月于传播季节开始前收集自 Ndiop 村(局部地方病传播)的 205 份血清的队列而获得的。该队列包括 108 名男性和 97 名女性, 年龄介于 3 岁到 75 岁之间。积极记录随后 5 个月内的临床发作, 共治疗 278 例临床发作(Perraut 等人, JID 2005 191 264-271)。

分析方案

(1) 抗原包被的优化 (2)待测试标准血清稀释度的选择(3)队列血清的系统分析(4)中位 OD 比值的计算(5)按年龄对结果分层(6)入选血清用于侵入和生长抑制的功能分析(7)选定组血清的同种型分析(8)中位 OD 比作为临床发作数据的函数, 以此函数对结果分层并进行统计分析。

(1) 校正：使用 $1\mu\text{g}/\text{m}$ 和 $0.5\mu\text{g}/\text{ml}$ 的 2 个抗原包被浓度测试来自 Dielmo 村(2005)的“血清超免疫”(SHI)和入选个体的稀释系列。

结果：发现以 $0.5\mu\text{g}/\text{ml}$ 包被是最佳的，因为其结果与对来自村庄和 SHI 个体使用 $1\mu\text{g}/\text{ml}$ 抗原的结果一样。Dielmo 入选个体的应答强于 Ndiop，且 MSP5 对 Ab 的应答比 MSP4 更低。参见图 20A 和 20B。

(2) 校正：来自 Dielmo 的 SHI 和阳性血清的滴定

结果：参看图 21 A, 21 B, 21 C 和 21 D，当使用来自 Dielmo(60605)的血清时，抗 MSP4 的反应为高度阳性(>SHI)。但是，与之后 2000 年 Ndiop 对列得到的 OD 相比，该 OD 相对低。尽管如此，SHI 的效价是可比较的。初一看检测的 Ab 具有相对较低的亲和力，正如曲线的锐角所示(与抗之前研究的 MSP1 的曲线相比)

SHI 抗 MSP5 的效价=1/1600

SHI 抗 MSP4-40 - MSP4-30 - MSP4-20 的效价=1/3200

(3) 2000 年 Ndiop 队列分析：主要结果

SHI 在每一检测中进行系统滴定。根据每块培养板上 SHI 对照系列计算 OD 比。对该队列而言，以 1/200 稀释的 OD 比都是高的，尽管其实际效价为平均的(4-6 之间的 OD 比值给出了约 1/2000 的终点效价)(结果如下)

效价/OD 比的结果

MSP5				MSP4 K40		MSP4 K20		MSP4 MOD II	
SHI 60605				SHI 60605		SHI 60605		SHI 60605	
calib2	OD_r	4.0	13.6	31.0	36.1	14.9	30.0	19.1	29.5
	效价	1600	6400	6400	25600	6400	25600	6400	25600
calib 1	OD_r	5.1		36.8		22.5		26.1	
	效价	3200		12800		3200		12800	
Ndiop_00	OD_r	6.6		28.4		24.4		18.6	
	效价	1600		12800		6400		12800	

(4) 应答者的发病率

下列表格显示了来自 Ndiop 2000 的个体队列的 Ab 应答特点。

阳性应答者被认为具有大于 2 的 OD 比。

如下所示, MSP4 几乎被所有的个体(>90%)所识别, MSP5 的识别则显著较低(约为 60%)。

统计学描述	n=205 MPS4-20		MSP4-40		MSP4-mod2		MSP5	
	DO	rtDO	DO	rtDO	DO	rtDO	DO	rtDO
统计学描述	n=205 MPS4-20		MSP4-40		MSP4-mod2		MSP5	
	DO	rtDO	DO	rtDO	DO	rtDO	DO	rtDO
中值	1.30	16.8	1.50	19.9	1.28	12.3	0.38	3.9
中位数	1.33	17.3	1.83	24.3	1.35	12.9	0.21	2.4
标准差	0.82	10.0	0.78	10.0	0.80	7.3	0.45	3.8
最小值	0.02	1	0.02	1	0.01	1	0.01	1
最大值	2.63	32.4	2.47	32.2	2.72	26.9	2.63	23.2
报告发病率	94%		97%		90%		59%	
发病率<15岁	89%		93%		80%		49%	
发病率>15岁	98%		100%		97%		66%	

(5) 与年龄和其他抗 MSP 应答的相关性

参看图 22 A, 22 B, 22 C 和 22D, 与其他抗 MSP 应答有显著相关性。

-MSP1 显著 Rho # 0.27-0.36

-各种 MSP4 抗原之间 Rho >0.9

-MSP4 和 MSP5 之间 Rho # 0.44-0.47

与 Ndiop 队列个体年龄显著相关($P < 0.001$, Rho 为 0.33 至 0.22), 但是略不如 MSP1 显著。抗 MSP5 的 Ab 应答与年龄相关性最低, 但小于 15 岁的个体具有显著更低的 Ab 水平。

聚焦 MSP4: MSP4 抗原的差异

参看图 23, MSP4 不同片段的识别水平有极为显著的差异($P < 0.001$)。矛盾的是, 最小和最大的抗原(p20 和 p40)比抗原 MSP4-30 (或 MSP4md2) 被更好的识别, 后者本身仅比全长 MSP4(MSP4-40)少 30 个残基。这一结果很有趣, 因为它似乎对应于免疫逃避策略的存在, 其中抗体应答集中针对可变结构域。这可以解释 IgG 应答和侵入抑制之间相关性的缺乏(见下文)。

(6) 与功能测试的关系

参看图 24 A 和 24 B, 根据 MSP4p40, MSP4p20, MSP4_mod (p30) (<20>)或 MSP5 (<2.4>)的中位 OD 比值的二分法比较, 生长或侵入抑制没有显著差异。(这一性质似乎只对本队列的抗 MSP1p19 Ab 表现出显著差异)。但是裂殖子的吞噬作用存在显著差异。这种与裂殖子吞噬作用的功能相关性是预期的, 因为该现象与特异 Ab 效价(因此 OD 比)的总体上升有关, 已知特异 Ab 效价会以疾病流行地区的年龄函数而增长。参见下列附图:

(7)同种型

分析限制于用于 MSP4-40 和 MSP5 的有限数目的血清。结果显示在图 25A 和 25B。对于 MSP4-p40, 其同种型曲线为 IgG1 +, IgG3+和不可忽视量的 IgG4:

- 效价和近来流行的寄生虫血症没有相关性。

- 15 岁以上个体的 IgG3 水平更高($P=0.006$), 并与近来疟疾发作的个体相同($P=0.03$), 这是预计的结果, 因为发作在疾病流行地区是与年龄相关的事件。

- IgG1 和 IgG3 是 IgG 应答的重要组分, 分别代表了 40%和 30%的 IgG。

- 对于更强的应答者, IgG 应答和侵入抑制之间存在着显著关系(独立 IgG 亚类)。

对于 MSP-5, 其同种型曲线主要为 IgG1

- 与近来流行寄生虫血症(接近显著性差异 $P=0.06$)或年龄的关系没有差异。

- IgG1 是 IgG 应答的重要主要组分, 总共占 75%的应答。

- IgG 应答和侵入或生长抑制之间没有相关性(与特定的同种型没有关联)。

(8)与疟疾发作的相关性

使用年龄校正泊松回归模型分析随访期间抗 Ag 构建体的 Ab 应答和疟疾发作的发生率之间的关系。在纳入 205 例来自 Ndiop 村的个体队列的前瞻性研究中(2000), 这些抗原表现出与抗临床发作的保护之间的显著相关性。

关于 MSP4

与之前指出的观察结果一致，其中 IgG 对 MSP4-p40 和 MSP4-p30 的效价存在显著差异，抗全长抗原(MSP4-p40) 的 Ab 应答与保护没有关联。但是，抗缺乏所述抗原 30 个残基的可变结构域的构建体(MSP4-p30) 和最小抗原(MSP4p20)的 Ab 应答与保护呈显著相关 ($P < 0.05$)。这些发现支持这一观点，即所述蛋白质 N 端的残基多态性参与免疫逃避，而且针对该蛋白质这一区域的抗体具有(A)有限亲和力(B)表现出菌株特异性，因此对大多数菌株具有有限的效应，或/和(C)由于 APL 的存在，从未实现亲和力成熟或缺少所需的 T 细胞辅助。无论涉及何种机制，该数据支持这一观点，即保守蛋白亚基诱导广泛特异性免疫的可能性比全长实体更高，全长实体经常含有认为在免疫逃避中发挥作用的多态性区域。

MSP4-40	系数	标准误差	p值	速率比	下限 (C. I.)	上限 (C. I.)
%GM	-6.049	0.1935	< 0.001	0.00236	0.001615	0.003449
AgeCI3	1.215	0.1021	< 0.001	3.37	2.759	4.116
m4_40cl2	-0.2279	0.1245	0.067	0.7962	0.6238	1.016

MSP4-30	系数	标准误差	p值	速率比	下限 (C. I.)	上限 (C. I.)
%GM	-6.031	0.1931	< 0.001	0.002403	0.001646	0.003508
AgeCI3	1.218	0.1017	< 0.001	3.38	2.769	4.125
M4md cl2	-0.2546	0.1216	0.036	0.7752	0.6108	0.9839

MSP4-20	系数	标准误差	p值	速率比	下限 (C. I.)	上限 (C. I.)
%GM	-6.017	0.1914	< 0.001	0.002438	0.001676	0.003548
AgeCI3	1.216	0.1013	< 0.001	3.373	2.766	4.114
m4_20cl2	-0.2888	0.1223	0.018	0.7491	0.5895	0.9521

发现有发作。鉴于预测可由所述抗原进行的脂肪酸修饰，这些结果对疾病严重程度和抗脂质反应可能具有有趣的提示。

MSP5	系数	标准误差	p值	速率比	下限 (C. I.)	上限 (C. I.)
%GM	-5.964	0.1917	< 0.001	0.00257	0.001765	0.003741
AgeCI3	1.201	0.1016	< 0.001	3.324	2.724	4.057
m5p5cl2	-0.3706	0.1243	0.0028	0.6904	0.541	0.8809

从上述介绍数据可得出的主要结论是(1)杆状病毒表达的 MSP4 和 MSP5 可用作疫苗，其活性与 MSP1p19 相当，(2)这些抗原被免疫个体强

烈识别(3)高于 MSP4(p30, 而 p20 更是如此)保守结构域的 IgG 识别中位值和高度保守的 MSP5 抗原的 IgG 识别中位值与抗临床疟疾的自然保护呈相关性。

实施例 17

MSP4p20 表达构建体

根据恶性疟原虫 MSP4 合成基因序列设计两个构建体用于 MSP4p20 的表达。一个构建体鉴定为 MSP4p21ss1, 别名 PfMSP4p21ss1 [SEQ ID NO:28]。下面描述了由构建体 MPS4p21ss1 编码的 *orf* [SEQ ID NO:19]和预测的氨基酸序列[SEQ ID NO:20]:

MSP4p21 ss1 构建体的核苷酸序列

>MSP4p21 ss1

```
ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
CGTATGCTTAACATGAGGATCTTGGGTGAAGAAAAGCCTAACGTTGACGGTGT
GTCAACATCTctGAAAAGAGTCCCAAGGAGAGTCAAATGGTCGACGACAAGAA
GAAGACCGAGGCCATTCCAAAGAAAGTCGTGCAGCCAAGCTCGAGCAACTCTG
GAGGTCACGTCCGTGAAGAAGAAGACCACAACGAAGGAGAGGGAGAGCACGA
AGAGGAGGAAGAACACGAAGAAGACGATGACGACGAGGACGACGACACATA
CAACAAAGACGACTTGGAGGACGAAGATCTTTGCAAGCACAACAACGGAGGA
TGTGGAGATGACAAGCTCTGCGAGTACGTTGGAACCGTCGCGTAAAATGTAA
ATGTAAGGAAGGATACAAGTTGGAAGGAATTGAGTGCGTTGAACACCACCACC
ACCATCACTAA [SEQ ID NO: 28]
```

构建体 MSP4p21ss1 编码的 ORF

>MSP4p21ss1

```
MWIVKFLIWHFFIICTINFDKLYISYSYNIVPENGRMLNMRILGEEKPNVDGVSTSJS
EKSPKESQMVDKTKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEHEHEHEHEHE
```

EDDDDEDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLE
E GIECVEHHHHHH [SEQ ID NO: 19]

浅灰色：加入的氨基酸直接跟在 MSP4p40 序列的残基 MRIL 之后

深灰色：加入的氨基酸直接位于 MSP4p40 序列中 MSP4p20 序列的
起点之前

预计蛋白产物 MSP4p21 ss1

MRILGEEKPNVDGVSTSLKSPKESQMVDKTKTEAIPKKVVQPSSSNSGGHVGEE
EDHNEGEGEHEEEEEHEEDDDDEDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEY
VGNRRVKCKCKEGYKLEGIECVEHHHHHH [SEQ ID NO: 20]

另外一个构建体鉴定为 MSP4p21ss2，别名 PfMSP4p21 ss2 [SEQ ID
NO: 29]。本文描述了由构建体 MPS4p21ss2 编码的 *orf* [SEQ ID NO:21]和
预测的氨基酸序列[SEQ ID NO: 22]。

构建体 MSP4p21ss2 的核苷酸序列

>MSP4p21ss2

ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
CGTATGCTTAACATGAGGATCTTGGGTGAAGAAAAGCCTCTAGAAAAGAGTCC
CAAGGAGAGTCAAATGGTCGACGACAAGAAGAAGACCGAGGCCATTCCAAAG
AAAGTCGTGCAGCCAAGCTCGAGCAACTCTGGAGGTCACGTCGGTGAAGAAGA
AGACCACAACGAAGGAGAGGGAGAGCACGAAGAGGAGGAAGAACACGAAGA
AGACGATGACGACGAGGACGACGACACATACAACAAAGACGACTTGGAGGAC
GAAGATCTTTGCAAGCACAACAACGGAGGATGTGGAGATGACAAGCTCTGCGA
GTACGTTGGAAACCGTCGCGTAAAATGTAAATGTAAGGAAGGATACAAGTTGG
AAGGAATTGAGTGCGTTGAACACCACCACCACCATCACTAA [SEQ ID NO: 29]

构建体 MSP4p21ss2 编码的 ORF

>MSP4p21ss2

MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLNMRILGEEKPiEKSPKESQ
 MVDDKKKTEAIPKKVVQPSSSNSGGHVGEEEDHNEGEHEHEEEEEHEEDDDDED
 DTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEGIECVEH
 H HHHH [SEQ ID NO: 21]

浅灰色：加入的氨基酸直接跟在 MSP4p40 序列的切割位点 MRIL 之后

深灰色：加入的氨基酸 L 直接位于 MSP4p40 序列中 MSP4p20 序列的起点之前

预计蛋白产物 MSP4p21ss2

MRILGEEKPLEKSPKESQMVDKDKKTEAIPKKVVQPSSSNSGGHVGEEEDHNEGE
 GEHEEEEEHEEDDDDEDDEDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRV
 KCKCKEGYKLEGIECVEHHHHHH [SEQ ID NO: 22]

实施例 18

MSP4 特征的进化保守

使用一个常用比对程序 (ClustalX)对来自恶性疟原虫和间日疟原虫的 MSP4 氨基酸序列进行比对时,可观察到几个明显结果。参见图 26。首先,恶性疟原虫信号序列切割位点下游的残基(依本文确定)在物种间几乎 100%保守(Pf MRILGE 和 Pv GRILGE)。这一发现再次强烈支持这一观点,即成熟 PfMSP4 蛋白起自 41 位残基,直接位于其前的残基是保守的切割位点的一部分。其次,间日疟原虫蛋白中没有被称为 SALSA(下划线)的序列元件。这表明肝细胞结合(如果由 PfMSP4 体内介导)不是保守的,因此也不是该蛋白质的重要功能。相反,间日疟原虫蛋白的该区域与恶性疟原虫蛋白的该区域完全不同,且含有极低复杂度的序列。虽然与恶性疟原虫的序列没有明显的可比性,我们可以想象其对免疫系统的效应是类似的。恶性疟原虫中,该蛋白质 N 端最外侧具有 N 端高度可变区域。多态性频率的统计学分析表明该区域内的两个位点可能参与了免疫逃避。这可以通过 B 细胞或 T 细胞表位的修饰或去除实现,因此当暴露给变异序列时,可导致菌株特异性的应答和/或免疫耐受。这种类型的免疫调节

也是重复序列的一种性质。(A) 它们含有其自身改变的肽配体以及(B)它们可以诱导不依赖 T 细胞的 B 细胞活化, 一般引起短期和低亲和力应答的生成。从比对序列可以得出的第三个观察结果是 MSP4p20 的主要序列元件存在于两个物种中。不令人震惊的是, 核心 EFG 结构域和 GPI 锚定信号均得到了保留, 更有趣的是所述酸性重复序列的排列方式似乎差不多。这增加了这一观点的分量, 即自然感染中 MSP4 或是被降解或是被切割, 使得结合膜的小结构实体位于裂殖子的表面。该实体是否会像 MSP1p19 进入刚被侵入的 RBC 尚未被证实, 但该模型似乎符合这一可能。基于这一比较, 可推荐下列蛋白质序列作为抗间日疟原虫的疟疾候选疫苗。

实施例 19

抗 MSP4p20 抗体在多形核嗜中性白细胞的裂殖子抗体依赖性吞噬中的作用

嗜中性白细胞内化病原体并使用活性氧和颗粒水解蛋白将其破坏。活化的嗜中性白细胞通过一个由 NADPH 氧化酶催化的称为呼吸爆发的过程高效地生成 ROS 例如 O_2^- , H_2O_2 等。可以通过化学发光染料来检测 ROS, ROS 可用于测量抗体诱导的吞噬活性。所述化学发光染料是鲁米诺, 它在存在 ROS 的情况下发光, 借助连接到电脑上的光度计(Microlumat+)可对光进行定量。

使用基于多形核嗜中性白细胞(PMN)的裂殖子抗体依赖性吞噬(ADPm)的新型功能性检测, 可以测量由天然获得和接种诱发的抗体所诱发的恶性疟原虫裂殖子的吞噬作用。使用来自疟疾流行区域个体的免疫血清和冷冻裂殖子制品, 显示该测试为高度特异性的, 而且具有可忽略的非免疫本底信号和良好的批内测定重复性。使用标准阳性对照进行批间测定的比较。

根据 ELISA 测量, 来自生活在季节性传播区域的个体的数据表明 ADPm 活性与年龄和 IgG 抗体对裂殖子和 MSP 重组抗原的应答显著相关。特异针对来自塞内加尔 Dielmo 和 Ndiop 村居民的流行免疫血清的杆状病毒重组 MSP1p19 和 MSP4p20 的抗体消耗表明抗这些抗原的抗体是 ADPm 活性的重要组分。

具体而言,通过离心培养物上清液收集恶性疟原虫裂殖子并将其冷冻分装储存。取自新鲜捐血的多形核嗜中性白细胞以简单 Ficoll-Hystopaque 1077 梯度分离并旋即用于 ADPm 测定。从 Ndiop (局部人口地方病)和 Dielmo (全人口地方病)的居民获得地方病免疫血清。在杆状病毒表达系统内生成具有 C 端六组氨酸标签的重组 MSP1 p19 和 MSP4p20 抗原并通过固定化金属亲和层析纯化这些抗原。使用载有相应重组抗原的 TALON 金属亲和树脂去除血清中抗 MSP1p19 或/和抗 MSP4p20 的抗体。加入鲁米诺和 PMN(每孔 5×10^6 细胞)启动该反应前将裂殖子与血清在 96 孔板上温育。使用 Berthold MicroLumat Plus 96 孔板测量发光量(luminescence output)1 小时。使用阳性标准人免疫血清对照(HIS)来定量该反应:

在本测定中阳性读数需要裂殖子的抗原特异性 IgG 调理(结合)(具或不具补体),这是因为去除补体的血清(56°C 处理 30 分钟)或以蛋白 G 纯化的总 IgG 诱导了对初始血清类似的应答,而且去除总 IgG 的血清不再产生应答(图 34)。但是,由特异 IgG 产生的化学发光效应的大小是可变的。图 35 显示了特异于杆状病毒 MSP4p20 和 MSP1p19 的抗体在本检测中均有功能,但其程度不一样。耗竭 MSP4p20 抗体后所述化学发光信号减少了 67%,相比之下 MSP1p19 抗体耗竭使化学发光信号减少 34%。特异于识别 MSP4p20 和 MSP1p19 抗体的耗竭减少了 78%的化学发光信号。

这些结果表明识别杆状病毒 MSP4p20 和 MSP1p19 重组蛋白的天然抗体是介导裂殖子吞噬作用和 PMN 免疫效应物的破坏的重要组分,这为使用这些抗原作为疫苗提供了强有力支持。

实施例 20

PfMSP5 修饰

为了考量位于杆状病毒表达的 PfMSP5 N 端封闭修饰的性质,进行了放射性标记实验。由于质谱表明该实体共价结合于豆蔻酰基,因此首先检查了豆蔻酰基的掺入。如之前实施例 1 所描述,使用 SF9 细胞在 T25 细胞培养瓶中进行昆虫细胞感染。感染后 24h 时,向培养上清加入 $200\ \mu\text{Ci}$ $[9,10_{(n)}\text{-}^3\text{H}]$ 豆蔻酸(Amersham)和不含脂肪酸的 BSA 的复合物。感染后 70h 时,收集培养上清,以 20 mM Tris pH8; 500 mM NaCl 对其进行透析,并

以 Talon®树脂对其进行分批纯化。在蓝染的 SDS-PAGE 凝胶上, 在存在放射性标记的情况下观察到未标记对照和培养物具有相同的蛋白质表达。当以放射自显影分析时, 杆状病毒表达的两个 PfMSP5 蛋白(35 和 45 kDa)均已掺入携带氘化的豆蔻酰基(图 27)。未标记蛋白未观察到有本底活性, 且空病毒感染未观察到这两个 PfMSP5 蛋白条带。

这些结果表明杆状病毒表达的重组 PfMSP5 蛋白是豆蔻酰化的。这些数据证实先前实施例 4 到 7 有关 PfMSP5 在昆虫细胞中表达的数据。

实施例 21

PfMSP4 的去向

为了跟踪 PfMSP4 的自然去向, 产生了几种不同的纯化血清。首先, 以 NHS 活化 sepharose 固定的杆状病毒表达的 PfMSP4p20 对收集自 4 只接种 MSP4 的兔的抗血清库进行亲和纯化。使用已知对来自 Ndiop 或 Dielmo 的 PfMSP4 呈阳性的 5 个人免疫血清的 2 个不同库来重复该步骤。使用这些抗体试剂, 通过 IFA 和蛋白印迹跟踪源自菌株 3D7 培养物中 PfMSP4 的去向, 并将其结果与使用特异于 PfMSP1p19 (G17.12) 的单克隆抗体获得的结果进行比较。如图 28A 所示, 在后期寄生虫和游离裂殖子表面检测出 PfMSP4, 但在可鉴定出 PfMSP1p19 的丙酮固定的环状体期寄生虫没有检测出 PfMSP4。此外, 当通过蛋白印迹平行分析等量的寄生虫提取物时(图 28B), 收集自寄生虫培养上清的游离裂殖子的 PfMSP4 水平似远少于成熟裂殖体, 而且在环状体期寄生虫提取物中没有检测出该蛋白。该数据符合这一观点, 即 PfMSP4 以类似于顶端膜抗原-1 的方式从裂殖子表面排出, 而不是像 MSP1p19 一样被蛋白酶解加工并携带进刚被侵入的 RBC。

有趣的是, 两种亲和纯化血清(兔和人)均识别仅来自成熟裂殖体提取物的多个高分子量蛋白条带。与之前 MSP3 和 MSP6 一样, 这可能是含有交叉反应表位的成熟裂殖体中存在的抗原, 或者这有可能是天然 PfMSP4 蛋白三聚物, 因为该寄生虫提取物没有被还原。

为了研究 PfMSP4 在肝前期寄生虫中的表达, 使用抗 PfMSP4p20 亲和纯化的兔血清对恶性疟原虫子孢子进行 IFA。尽管阳性对照抗 CSP 单

克隆抗体给出了强而清晰的表面染色，使用抗 PfMSP4p20 抗体未见有染色。

这些结果表明天然 MSP4 蛋白定位于裂殖子的表面，且寄生虫重新侵入红细胞后该蛋白不再存在(参见环状体期感染的红细胞的 IFA, 图 28A)。图 28B 的蛋白印迹显示了使用抗天然 MSP4(人血清)抗体和使用抗重组 MSP4(兔血清)抗体同样的模式。因此可从这一结果推断得出，抗杆状病毒重组抗原的抗体能够识别天然形式的抗原。

实施例 22

p20 亚结构域独立于 p40 的表达

为了便于 p20 亚结构域的直接表达，构建(图 29 和实施例 17)和测试了一系列新的表达构建体。如图 30 所示，编码最小信号序列的构建体(PfMSP4p21)没有引起蛋白分泌，但检测到高水平的胞内蛋白(数据未显示)。使用 PfMSP4p21ss1 和 PfMSP4p21ss2 构建体实现了更高的表达水平，正如图 30 所示和 N 端测序所揭示(参看图 29)，两种蛋白质均很快失去了余下的 N 端残基以产生与杆状病毒 PfMSP4p40 相同的稳定 p20 实体。但是，使用杆状病毒 PfMSP4p40 和 PfMSP4p21ss1 对平行感染后 66h 时的蛋白表达进行比较时，原有构建体产生高得多的蛋白产量(2 倍)。这些结果补充了前述的实施例 17。

实施例 23

PvMSP4 在杆状病毒表达系统中的表达

为了获得 PvMSP4 开放阅读框(ORF)，对来自寄生虫菌株 Belem 的基因组 DNA 进行巢式 PCR。由于间日疟原虫的基因组 DNA 大约富含 50% 的 GC 含量，因此没有必要制备合成基因。将整个 ORF 克隆进 pMosBLUE 并进行测序。然后，以 PCR 扩增两个外显子，其中外显子 2 的 3 端的 GPI 锚定信号被六组氨酸标签取代。此外，将一个适宜限制性位点分别引入到外显子 1 和 2 的 3' 或 5' 端，继而将这些序列克隆到 pMosBLUE 载体以形成以组氨酸标签结尾的连续 ORF(图 31)。

使用标准方案生成编码 PvMSP4 的杆状病毒，并根据前述方法使用 High Five 昆虫细胞对随时间变化的蛋白质表达进行评测。如图 32 所示，感染后约 30h 时可检测到蛋白质表达。

对该蛋白质条带进行的 N 端序列分析揭示了 GIAAC 的 N 端。感染后约 54h 时蛋白质表达达到最高，此时可见一些蛋白质降解。但是，PvMSP4 降解的方式与 PfMSP4 不同。终产物的 N 端测序表明该蛋白质已降解为 13 或 30-32 个氨基酸，生成 EGGEQ 的 N 端序列和产生 GDSSG、DSSGG 和 SSGGL 的混合信号。检测到的最小蛋白质产物大小约为 8kDa，它携带 N 端序列 LDNNG。

实施例 24

PvMSP4 的人免疫血清识别

时程分析结束时(如图 32 所示的 66h)，以 HPLC 在 Talon 树脂上纯化 130 mL 的培养上清。所述纯化产物由图 32 所示的 30, 25 kDa 和 8 kDa 条带所组成。

为了考量人血清识别，以 0.5 μg /mL 杆状病毒表达的 PvMSP4 包被 ELISA 培养板。

同时检测了共 24 例血清样本。这些样本收集自斯里兰卡至少经历一次间日疟原虫疟疾发作的个体。这些样本中有 17 例为 PvMSP4 血清阳性，其终点效价位于 1/25,000 和 1/200,000 之间(图 33A)。这些 17 例阳性血清以 1/2700 稀释在 ELISA 培养板上检测，其中 ELISA 培养板以用 DTT 和碘乙酰胺处理而不可逆还原的杆状病毒 PvMSP4 包被。所述 17 例阳性血清中，有 11 例表现出介于 12-80%之间的还原敏感性程度，不论是以固定的 1/2700 稀释度计算还是使用落在每条曲线中点的稀释度计算(图 33B)。

涉及间日疟原虫 MSP4(PvMSP4)的实施例 23 和 24 补充了实施例 18。如图 33B 所示，尽管 2、3、4、6、7、8、9、10、11、14 和 17 号血清对 DTT 还原的抗原比对非还原抗原具有更小的亲和力，但 12 和 15 号血清对还原抗原比对非还原抗原具有更高的亲和力。可以得出结论，抗原构象影响抗体-抗原结合，但由于尚未确定每份受检血清抗疟原虫感染的保护

能力,抗体对还原或非还原抗原的亲合力不能与两种类型抗体的保护价值联系起来。

下列 PfMSP5/His、PfMSP4p30/His 和 PfMSP4p40/His 杆状病毒于 2005 年 11 月 10 日保藏在巴斯德研究所的法国微生物保藏中心(C.N.C.M.)(28, rue du Docteur Roux, 75724 Paris, Cedex 15, FRANCE)并分别分配给下列保藏号。

材料	保藏号.
PfMSP5/His	I-3512
PfMSP4p30/His	I-3513
PfMSP4p40/His	I-3514

下列 F12-7、G21-2、J18-14 和 L11-6 杂交瘤于 2005 年 11 月 16 日保藏在巴斯德研究所的法国微生物保藏中心(C.N.C.M.)(28, rue du Docteur Roux, 75724 Paris, Cedex 15, FRANCE)并分别分配给下列保藏号。

材料	保藏号
F12-7	I-3517
G21-2	I-3518
J18-14	I-3519
L11-16	I-3520

下列 PfMSP4p21ss1/His、PfMSP4p21ss2/His 和 PvMSP4/His 于 2006 年 11 月 21 日存放在巴斯德研究所的法国微生物保藏中心(C.N.C.M.)(28, rue du Docteur Roux, 75724 Paris, Cedex 15, FRANCE)并分别分配给下列保藏号。

材料	保藏号
PfMSP4p21ss1/His	I-3695
PfMSP4p21ss2/His	I-3696
PvMSP4/His	I-3694

下列是上述肽和多核苷酸的氨基酸和核酸序列:

MSP5p10 (MSP5 降解产物, 186-253 位残基)

YNKVEKNVTDEMLLYNMMSDQNRKSCAINNGGCSDDQICININNIGVKCICKDGY
LLGTKCIHHHHHH [SEQ ID NO: 16]

编码 MSP5p10 的核酸序列

>readseq-48009_tmp_1 207 bp

TACAACAAGGTCGAGAAGAACGTAACCGACGAGATGCTCTTGTACAACATGAT
GTCCGACCAAAACCGCAAAAGCTGTGCTATCAACAACGGTGGCTGCAGTGACG
ACCAGATCTGCATCAACATCAACAACATCGGTGTGAAGTGCATTTGTAAGGAT
GGATACCTACTTGGTACCAAGTGCATTACACCACCACCACCACCTGA [SEQ
ID NO: 5]

SEQ ID NO: 1=编码 SEQ ID NO: 9 的降解产物 MSP4p20(132-251 位残基)
的核酸

>readseq-41112_tmp_1 363 bp

AAGAGTCCCAAGGAGAGTCAAATGGTCGACGACAAGAAGAAGACCGAGGCCA
TTCCAAAGAAAGTCGTGCAGCCAAGCTCGAGCAACTCTGGAGGTCACGTCGGT
GAAGAAGAAGACCACAACGAAGGAGAGGGAGAGCACGAAGAGGAGGAAGAA
CACGAAGAAGACGATGACGACGAGGACGACGACACATACAACAAAGACGACT
TGGAGGACGAAGATCTTTGCAAGCACAACAACGGAGGATGTGGAGATGACAA
GCTCTGCGAGTACGTTGGAAACCGTCGCGTAAAATGTAAATGTAAGGAAGGAT
ACAAGTTGGAAGGAATTGAGTGCGTTGAACACCACCACCACCATCACTAA
[SEQ ID NO:1]

MSP5

>readseq.input(1), 762 个碱基, 6761 DE71 校验和.

ATGAACATTCTCTGTATTCTCAGCTACATTTACTTCTTCGTCATCTTCTACAGTT
TAAACCTCAACAACAAAAACGAGAACTTCTTGGTGGTCCGCAGACTCATGAAC
GACGAAAAGGGAGAAGGTGGCTTCACTAGTAAGAACAAGGAAAACGGAAACA
ACAACAGGAACAACGAGAACGAACTCAAAGAAGAAGGATCTTTGCCCACTAA
GATGAACGAGAAAAACAGTAACTCCGCGGATAAGCAACCAAACGACATCTCC

CACGACGAAAGCAAGAGCAACAGTAACAACGCCCAAACATCCAAAAGGAAC
CTGAAGAGAAGGAAAACCTCAAACCCCAACCTCGACTCGAGTGAAAACCTCCGCT
GAAAGTGCTACTAGAAGCGTCGACATCAGTGAACACAACCTCAAACAACCCCGA
AACTAAAGAAGAAAACGGAGAAGAACCTCTAGACCTGGAAATTAACGAAAAC
GCAGAAATCGGCCAGGAACCTCCAAACCGTCTTCACTTCGACAACGTTGACGA
CGAAGTACCACATTACTCAGCCCTGAGGTACAACAAGGTTCGAGAAGAACGTAA
CCGACGAGATGCTCTTGTACAACATGATGTCCGACCAAACCGCAAAGCTGT
GCTATCAACAACGGTGGCTGCAGTGACGACCAGATCTGCATCAACATCAACAA
CATCGGTGTGAAGTGCATTTGTAAGGATGGATACCTACTTGGTACCAAGTGCAT
TCACCACCACCACCACCCTGA [SEQ ID NO: 4]

ORF 编码的蛋白质

>readseq.input(1), 253 个碱基, 7870B947 校验和.

MNILCILSYIYFFVIFYSLNLNKNNENFLVVRRLMNDEKGEGGFTSKNKENGNNNR
NNENELKEEGLPTKMNEKNSNSADKQPN DISHDESKSNSNNAQNIQKEPEEKENS
NPNLDSSSENSAESATRSVDISEHNSNNPETKEENGEEPLDLEINENAEIGQEPPNRLH
FDNVDDEVPHYSALRYNKVEKNVTDEMLLYNMMSDQNRKSCAINNGGCSDDQIC
ININNIGVKCICKDGYLLGTKCIHHHHHH [SEQ ID NO: 14]

MSP4p40

>readseq.input(1), 756 碱基, 733F5E48 校验和.

ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
CGTATGCTTAACATGAGGATCTTGGGTGAAGAAAAGCCTAACGTTGACGGTGT
GTCAACATCTAACACACCTGGCGGAAACGAGGCATCTAGTGCTTCTCCTAACCT
TGCTGACGCTGCAGAAAAGAAGGACGAAAAGGAAGCAAGCGAGCAAGGCGAA
GAATCCCACAAGAAGGAAAACCTCTCAGGAATCTGCAAACGGAAAAGACGACG
TTAAGGAGGAGAAGAAGACCAACGAGAAGAAGGACGACGGAAAGACTGACA
AGGTACAAGAAAAGGTTCTAGAAAAGAGTCCCAAGGAGAGTCAAATGGTCGA
CGACAAGAAGAAGACCGAGGCCATTCCAAAGAAAGTCGTGCAGCCAAGCTCG

AGCAACTCTGGAGGTCACGTCGGTGAAGAAGAAGACCACAACGAAGGAGAGG
GAGAGCACGAAGAGGAGGAAGAACACGAAGAAGACGATGACGACGAGGACG
ACGACACATACAACAAAGACGACTTGGAGGACGAAGATCTTTGCAAGCACAA
CAACGGAGGATGTGGAGATGACAAGCTCTGCGAGTACGTTGGAACCGTCGCG
TAAAATGTAAATGTAAGGAAGGATACAAGTTGGAAGGAATTGAGTGCGTTGAA
CACCACCACCACCATCACTAA [SEQ ID NO: 3]

ORF 编码的蛋白质

>readseq.input(1), 251 个碱基, 1EFAEF35 校验和.

MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLNMRILGEEKPNVDGVSTSN
TPGGNEASSASPNLADAAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKKT
NEKKDDGKTDKVQEKVLEKSPKESQMVDDKKKTEAIPKKVVQPSSSNSGGHVGE
EEDHNEGEGEHEEEEEHEEDDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCE
YVGNRRVKCKCKEGYKLEGIECVEHHHHHH [SEQ ID NO: 12]

最终产生的蛋白质

>readseq.input(1), 211 个碱基, 67EAFE8C 校验和.

MRILGEEKPNVDGVSTSNTPGGNEASSASPNLADAAEKKDEKEASEQGEESHKKE
NSQESANGKDDVKEEKKTNEKKDDGKTDKVQEKVLEKSPKESQMVDDKKKTEAI
PKKWQPSSSNSGGHVGEEDHNEGEGEHEEEEEHEEDDDDDDDDTYNKDDLEDE
DLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEGIECVEHHHHHH [SEQ ID
NO: 13]

与最终产生的蛋白质一起分泌的 P20 蛋白

>readseq.input(1), 120 个碱基, 9358BD28 校验和.

KSPKESQMVDDKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEGEHEEEEEHEE
DDDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLE
GIECVEHHHHHH [SEQ ID NO: 9]

MSP4p30

>readseq.input(1), 666 个碱基, A9B2DCBC 校验和.

```
ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
CGTATGCTTAACATGAGGATCTTGGCTGCAGAAAAGAAGGACGAAAAGGAAG
CAAGCGAGCAAGGCGAAGAATCCCACAAGAAGGAAAACCTCTCAGGAATCTGC
AAACGGAAAAGACGACGTTAAGGAGGAGAAGAAGACCAACGAGAAGAAGGA
CGACGGAAAGACTGACAAGGTACAAGAAAAGGTTCTAGAAAAGAGTCCCAAG
GAGAGTCAAATGGTCGACGACAAGAAGAAGACCGAGGCCATTCCAAAGAAAG
TCGTGCAGCCAAGCTCGAGCAACTCTGGAGGTCACGTCGGTGAAGAAGAAGAC
CACAAACGAAGGAGAGGGAGAGCACGAAGAGGAGGAAGAACACGAAGAAGAC
GATGACGACGAGGACGACGACACATACAACAAAGACGACTTGGAGGACGAAG
ATCTTTGCAAGCACACAACGGAGGATGTGGAGATGACAAGCTCTGCGAGTAC
GTTGGAAACCGTCGCGTAAAATGTAAATGTAAGGAAGGATACAAGTTGGAAG
GAATTGAGTGCGTTGAACACCACCACCACCATCACTAA [SEQ ID NO: 2]
```

ORF 编码的蛋白质

>readseq.input(1), 221 个碱基, F982D7F0 校验和.

```
MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLNMRILAAEKKDEKEASEQ
GEESHKKENSQESANGKDDVKEEKKTNEKKDDGKTDKVQEKVLEKSPKESQMVD
DKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEGEHEEEEEHEEDDDDEDDDTY
NKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEGIECVEHHHH
HH [SEQ ID NO: 10]
```

最终产生的蛋白质

>readseq.input(1), 181 个碱基, FB2C49DD 校验和.

```
MRILAAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKKTNEKKDDGKTDK
VQEKVLEKSPKESQMVDKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEGEHE
EEEEHEEDDDDEDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKC
KEGYKLEGIECVEHHHHHHH [SEQ ID NO: 11]
```

MSP4p21

>readseq.input(1), 498 个碱基, C2453A47 校验和.

```
ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
CGTATGCTTAACATGAGGATTCTAGAAAAGAGTCCCAAGGAGAGTCAAATGGT
CGACGACAAGAAGAAGACCGAGGCCATTCCAAGAAAGTCGTGCAGCCAAGC
TCGAGCAACTCTGGAGGTCACGTTCGGTGAAGAAGAAGACCACAACGAAGGAG
AGGGAGAGCACGAAGAGGAGGAAGAACACGAAGAAGACGATGACGACGAGG
ACGACGACACATACAACAAAGACGACTTGGAGGACGAAGATCTTTGCAAGCA
CAACAACGGAGGATGTGGAGATGACAAGCTCTGCGAGTACGTTGGAAACCGTC
GCGTAAAATGTAAATGTAAGGAAGGATACAAGTTGGAAGGAATTGAGTGC GTT
GAACACCACCACCACCATCACTAA [SEQ IDNO: 6]
```

ORF 编码的蛋白质

>readseq.input(1), 165 个碱基, 27BB4F2E 校验和.

```
MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLNMRILEKSPKESQMVD DK
KKTEAIPKKVVQPSSSNSGGHVGEEEDHNEGEGEHEEEEEHEEDDDDEDDDTYNK
DDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEGIECVEHHHHHH
[SEQ ID NO: 17]
```

最终产生的蛋白质

readseq.input(i), 125 个碱基, 50D64DD2 校验和.

```
MRILEKSPKESQMVD DKKKTEAIPKKVVQPSSSNSGGHVGEEEDHNEGEGEHEEE
EEHEEDDDDEDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKE
GYKLEGIECVEHHHHHHH [SEQ ID NO: 18]
```

MSP4p21ss1

>readseq.input(1), 540 个碱基, 96DADE5 校验和.

```
ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
```

CGTATGCTTAACATGAGGATCTTGGGTGAAGAAAAGCCTAACGTTGACGGTGT
GTCAACATCTctaGAAAAGAGTCCCAAGGAGAGTCAAATGGTCGACGACAAGAA
GAAGACCGAGGCCATTCCAAAGAAAGTCGTGCAGCCAAGCTCGAGCAACTCTG
GAGGTCACGTCGGTGAAGAAGAAGACCACAACGAAGGAGAGGGAGAGCACGA
AGAGGAGGAAGAACACGAAGAAGACGATGACGACGAGGACGACGACACATA
CAACAAAGACGACTTGGAGGACGAAGATCTTTGCAAGCACAACAACGGAGGA
TGTGGAGATGACAAGCTCTGCGAGTACGTTGGAAACCGTCGCGTAAAATGTAA
ATGTAAGGAAGGATACAAGTTGGAAGGAATTGAGTGC GTTGAACACCACCACC
ACCATCACTAA [SEQ ID NO: 7]

ORF 编码的蛋白质

>readseq.input(1), 179 个碱基, F0D6C58 校验和.

MWIVKFLIVVHFFHICTINFDKLYISYSYNIVPENGRMLNMRILGEEKPNVDGVSTSL
EKSPKESQMVDKDKKTEAIPKKVVQPSSSNSGGHVGEEEDHNEGEHEHEEEEEHE
EDDDDEDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKL
EGIECVEHHHHHH [SEQ ID NO: 19]

最终产生的蛋白质

>readseq.input(1), 139 个碱基, 4D5C3E0F 校验和.

MRILGEEKPNVDGVSTSLKSPKESQMVDKDKKTEAIPKKVVQPSSSNSGGHVGEE
EDHNEGEHEHEEEEEHEEDDDDEDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEY
VGNRRVKCKCKEGYKLEGIECVEHHHHHH [SEQ ID NO: 20]

MSP4P21ss2

>readseq.input(1), 516 个碱基, 1 B3CABBA 校验和.

ATGTGGATCGTAAAGTTCTTGATTGTGGTCCACTTCTTCATCATATGCACCATC
AACTTCGACAAGCTCTACATTAGTTACTCTTACAACATCGTCCCTGAAAACGGA
CGTATGCTTAACATGAGGATCTTGGGTGAAGAAAAGCCTCTAGAAAAGAGTCC
CAAGGAGAGTCAAATGGTCGACGACAAGAAGAAGACCGAGGCCATTCCAAAG
AAAGTCGTGCAGCCAAGCTCGAGCAACTCTGGAGGTCACGTCGGTGAAGAAGA

AGACCACAACGAAGGAGAGGGAGAGCACGAAGAGGAGGAAGAACACGAAGA
AGACGATGACGACGAGGACGACGACACATACAACAAAGACGACTTGGAGGAC
GAAGATCTTTGCAAGCACAACAACGGAGGATGTGGAGATGACAAGCTCTGCGA
GTACGTTGGAAACCGTCGCGTAAAATGTAAATGTAAGGAAGGATACAAGTTGG
AAGGAATTGAGTGCGTTGAACACCACCACCACCATCACTAA [SEQ ID NO: 8]

ORF 编码的预测蛋白质

>readseq.input(1), 171 个碱基, EB09D214 校验和.

MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLNMRILGEEKPLEKSPKESQ
MVDDKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEHEHEEEEEHEEDDDDED
DTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEGIECVEH
HHHHH [SEQ ID NO: 21]

最终产生的蛋白质

>readseq.input(1), 131 个碱基, 6B6B091D 校验和.

MRILGEEKPLEKSPKESQMVDDKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGE
GEHEEEEEHEEDDDDEDDEDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRV
KCKCKEGYKLEGIECVEHHHHHH [SEQ ID NO: 22]

PvMSP4/His 蛋白

MKVAYFLSVLDLLIIFSLYFDGRRSAFAGIAACIRHGRILGEGGEQSGGASGGSSGG
SSGDSSGGLSGGSSGGPSPAGSSSGSGSDPANSATGPQNSTPGSGGQTGDHSAEA
ENG DYNEQGDDHGDDHGDDHGDDHGDEQDGEDYDDAEDDDL YELSEVDENAN
LCLDNNGGCGDDKICENLGKGIVKCLCKPGYKLVGTECVESHHHHHH [SEQ ID
NO: 16]

本申请描述了、尤其是在实施例中描述了公开为 PfMSP5/His、
PfMSP4p30/His、PfMSP4p40/His、PfMSP4p21ss1/His、PfMSP4p21ss2/His
和 PvMSP4/His 的生物材料。它满足以下要求：

- 增殖病毒的要求：

细胞类型：草地夜蛾(Sf9)

培养基：

添加 L-谷氨酰胺(GIBCO, Invitrogen Corporation) Cat N° 10902-088 的 SF-900 II 培养基

200 mM L-谷氨酰胺(GIBCO, Invitrogen Corporation): 2mM 终浓度 Cat N° 25030-024

庆大霉素(50 mg/ml; GIBCO, Invitrogen Corporation):

50 µg/ml 终浓度 Cat N° 15750-037

血清

胎牛血清 5%

pH 6,2

温度 27-28°C

气相：环境(无 CO₂ 富集)

- 用于获得病毒悬浮液的方法：

增殖的细节：在一容器(T-25, T-75 或 T-150) 中以新鲜完全培养基(添加 5%FCS)培养单层 Sf9 细胞, 该培养物接种少量病毒悬浮液(或噬斑或通过有限稀释取自克隆方法的一块微量滴定板的孔)。所述培养物于 27-28°C 温育, 直至 5-6 天左右时所有细胞裂解。

感染证据的标准：细胞具延展细胞核, 随后细胞核呈粒状外观, 数天后细胞裂解。

制备病毒悬浮液的细节

细胞上清以 4000 g 离心 10 分钟以除去细胞碎片。

- 用于滴定的方法

有限稀释(从稀释起感染一个孔, 导致 96 孔微量滴定板约 1/10 的阳性)。6 天后读取结果, 预计效价为 10⁷-10⁸ pfu/ml。

- 储存条件

-80°C 冷冻长期保存。

悬浮液为标准的培养基。

小鼠骨髓瘤细胞(浆细胞瘤 Ag8 X 63 6.5.3)和取自以指定抗原免疫动物脾脏的小鼠免疫细胞之间进行细胞融合获得上述杂交瘤 F12-7、G21-2、J18-14 和 L11-16。

- 细胞的特点和产物

骨髓瘤命名: X63-Ag8-6.5.3

抗原特异性: MSP4 II 用于 F 12-7, MSP5 用于 G21-2 和 518-14, MSP4 用于 L11-16。

抗体亚类: IgG1k

分泌的稳定性: 良好, 不受限制

- 培养条件

培养基

Eagle MEM- Eurobio

丙酮酸钠-Gibco BRL

谷氨酰胺-Gibco BRL

抗生素: 青链霉素(péni-streptomycine)-Gibco BRL

马血清 10%

血清碳酸氢根 2,2 g/l

血清: 马 10%

pH7,4

最佳温度: 37°C

气相: 7% CO₂

冻融注意事项: 37°C 快速解冻, 立即以培养基+血清洗涤。培养于 24 孔板上, 每孔 1ml。

- 培养条件: 细胞悬浮

群体倍增时间: 15h

最佳分裂比: 3 天

预计细胞密度: $5 \cdot 10^6$

无限生命期

常规传代技术: 分裂

- 储存条件

悬浮液：95%马血清-5%DM50

细胞收集技术：取细胞悬浮液，离心和吸取冰冻培养基中的沉淀。

冷冻技术：含 5×10^6 个细胞的 1 ml 冰冻培养基于 -20°C 放置 2h, 于 -80°C 和液氮中放置 3 至 4 天。悬浮液为标准的培养基。

参考文献

基于下列每篇出版物的全部公开内容并通过引用的方式将其并入本文中。

- 1 Genton, B., Al-Yaman, F., Betuela, I. et al. Safety and immunogenicity of a three-component blood-stage malaria vaccine (MSP1, MSP2, RESA) against *Plasmodium falciparum* in Papua New Guinean children. *Vaccine* 2003, 22(1), 30-41.
- 2 Stowers, A. W., Kennedy, M.C., Keegan, B. P., Saul, A., Long, CA. & Miller, L.H. Vaccination of monkeys with recombinant *Plasmodium falciparum* apical membrane antigen 1 confers protection against blood-stage malaria. *Infect Immun* 2002, 70(12), 6961-6967.
- 3 O'Donnell, R.A., de Koning-Ward, T.F., Burt, R.A. et al. Antibodies against merozoite surface protein (MSP)-I (19) are a major component of the invasion-inhibitory response in individuals immune to malaria. *J Exp Med* 2001, 193(12), 1403-1412.
- 4 Hughes, M. K. & Hughes, A.L. Natural selection on *Plasmodium* surface proteins. *Mol Biochem Parasitol* 1995, 71(1), 99-113.
- 5 Escalante, A.A., LaI, A.A. & Ayala, FJ. Genetic polymorphism and natural selection in the malaria parasite *Plasmodium falciparum*. *Genetics* 1998, 149(1), 189-202.
- 6 Volkman, S.K., Hartl, D. L., Wirth, D. F. et al. Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*. *Science* 2002, 298(5591), 216-218.
- 7 Rayner, J. C, Corredor, V., Feldman, D. et al. Extensive polymorphism in the *Plasmodium vivax* merozoite surface coat protein MSP-3alpha is limited to specific domains. *Parasitology* 2002, 125(Pt 5), 393-405.

- 8 Taylor, R.R., Smith, D.B., Robinson, V.J., McBride, J.S. & Riley, E.M. Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass. *Infect Immun* 1995, 63(11), 4382-4388.
- 9 Wang, L., Mohandas, N., Thomas, A. & Coppel, R.L. Detection of detergent-resistant membranes in asexual blood-stage parasites of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2003, 130(2), 149-153.
- 10 Marshall, V.M., Silva, A., Foley, M. et al. A second merozoite surface protein (MSP-4) of *Plasmodium falciparum* that contains an epidermal growth factor- like domain. *Infect Immun* 1997, 65(11), 4460-4467.
- 11 Wu, T., Black, CG. , Wang, L., Hibbs, A.R. & Coppel, R.L. Lack of sequence diversity in the gene encoding merozoite surface protein 5 of *Plasmodium falciparum*. *Mol Biochem Parasitol* 1999, 103(2), 243-250.
- 12 Wang, L., Black, CG. , Marshall, V.M. & Coppel, R.L. Structural and antigenic properties of merozoite surface protein 4 of *Plasmodium falciparum*. *Infect Immun* 1999, 67(5), 2193-2200.
- 13 Wang, L., Richie, T.L., Stowers, A., Nhan, D.H. & Coppel, R.L. Naturally acquired antibody responses to *Plasmodium falciparum* merozoite surface protein 4 in a population living in an area of endemicity in Vietnam. *Infect Immun* 2001 , 69(7), 4390-4397.
- 14 Black, CG., Wang, L., Hibbs, A.R., Werner, E. & Coppel, R.L. Identification of the *Plasmodium chabaudi* homologue of merozoite surface proteins 4 and 5 of *Plasmodium falciparum*. *Infect Immun* 1999, 67(5), 2075-2081.
- 15 Black, CG. & Coppel, R.L. Synonymous and non-synonymous mutations in a region of the *Plasmodium chabaudi* genome and evidence for selection

acting on a malaria vaccine candidate. *Mol Biochem Parasitol* 2000, 111(2), 447- 451.

16 Kedzierski, L., Black, CG. & Coppel, R.L. Characterization of the merozoite surface protein 4/5 gene of *Plasmodium berghei* and *Plasmodium yoelii*. *Mol Biochem Parasitol* 2000, 105(1), 137-147.

17 Kedzierski, L., Black, CG., Goschnick, M.W., Stowers, A.W. & Coppel, R.L. Immunization with a Combination of Merozoite Surface Proteins 4/5 and 1 Enhances Protection against Lethal Challenge with *Plasmodium yoelii*. *Infect Immun* 2002, 70(12), 6606-6613.

18 Kedzierski, L., Black, CG. & Coppel, R.L. Immunization with recombinant *Plasmodium yoelii* merozoite surface protein 4/5 protects mice against lethal challenge. *Infect Immun* 2000, 68(10), 6034-6037.

19 Kedzierski, L., Black, CG. , Stowers, A. W., Goschnick, M.W., Kaslow, D.C & Coppel, R.L. Comparison of the protective efficacy of yeast-derived and *Escherichia coli*-derived recombinant merozoite surface protein 4/5 against lethal challenge by *Plasmodium yoelii*. *Vaccine* 2001, 19(32), 4661-4668.

20 Wang, L., Goschnick, M.W. & Coppel, R.L. Oral Immunization with a Combination of *Plasmodium yoelii* Merozoite Surface Proteins 1 and 4/5 Enhances Protection against Lethal Malaria Challenge. *Infect Immun* 2004, 72(10), 6172-6175.

21 Goschnick, M.W., Black, CG. , Kedzierski, L., Holder, A.A. & Coppel, R.L. Merozoite Surface Protein 4/5 Provides Protection against Lethal Challenge with a Heterologous Malaria Parasite Strain. *Infect Immun* 2004, 72(10), 5840-5849.

- 22 Marshall, V.M., Tieqiao, W. & Coppel, R.L. Close linkage of three merozoite surface protein genes on chromosome 2 of *Plasmodium falciparum*. *Mol Biochem Parasitol* 1998, 94(1), 13-25.
- 23 Nishimura, Y., Chen, Y.Z., Uemura, Y. et al. Degenerate recognition and response of human CD4⁺ Th cell clones: implications for basic and applied immunology. *Mol Immunol* 2004, 40(14-15), 1089-1094.
- 24 Brady, C.P., Shimp, R.L., Miles, A.P., Whitmore, M. & Stowers, A.W. High- level production and purification of P30P2MSP1(19), an important vaccine antigen for malaria, expressed in the methylotropic yeast *Pichia pastoris*. *Protein Expr Purif* 2001, 23(3), 468-475.
- 25 Epp, C, Kauth, C.W., Bujard, H. & Lutz, R. Expression and purification of *Plasmodium falciparum* MSP-1(42): A malaria vaccine candidate. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003, 786(1-2), 61-72.
- 26 Withers-Martinez, C, Saldanha, J.W., Ely, B. et al. Expression of recombinant *Plasmodium falciparum* subtilisin-like protease-1 in insect cells. Characterization, comparison with the parasite protease, and homology modeling PCR-based gene synthesis as an efficient approach for expression of the A+T-rich malaria genome. *J Biol Chem* 2002, 277(33), 29698-29709. Epub 22002 Jun 29696.
- 27 Zhou, Z., Schnake, P., Xiao, L. & LaI, A.A. Enhanced expression of a recombinant malaria candidate vaccine in *Escherichia coli* by codon optimization. *Protein Expr Purif* 2004, 34(1), 87-94.
- 28 Weber, J. L. Molecular Biology of Malaria Parasites. *Exp. Parasitol.* 1988, 66, 143-170.

- 29 Withers-Martinez, C, Carpenter, E.P., Hackett, F. et al. PCR-based gene synthesis as an efficient approach for expression of the A+T-rich malaria genome. *Protein Eng* 1999, 12(12), 1113-1120.
- 30 Ballou, W.R., Arevalo-Herrera, M., Carucci, D. et al. Update on the clinical development of candidate malaria vaccines. *Am J Trop Med Hyg* 2004, 71(2 Suppl), 239-247.
- 31 Wang, L., Marshall, V.M. & Coppel, R.L. Limited polymorphism of the vaccine candidate merozoite surface protein 4 of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2002, 120(2), 301-303.
- 32 Jongwutiwes, S., Putaporntip, C, Friedman, R. & Hughes, A.L. The Extent of Nucleotide Polymorphism is Highly Variable Across a 3-kb Region on *Plasmodium falciparum* Chromosome 2. *Mol Biol Evol* 2002, 19(9), 1585-1590.
- 33 Polson, H. E., Conway, DJ. , Fandeur, T., Mercereau-Puijalon, O. & Longacre, S. Gene polymorphism of *Plasmodium falciparum* merozoite surface proteins 4 and 5. *Mol Biochem Parasitol* 2005, 142(1), 110-115.
- 34 Puentes, A.G., J. Vera, R. Oopez, R. Suarez, J. Rodriguez, L. Curtidor, H. Sporozoite and Liver Stage Antigen *Plasmodium falciparum* peptides bind specifically to human hepatocytes. *Vaccine* 2004, 22, 1150-1156.
- 35 Hale, R.S. & Thompson, G. Codon optimization of the gene encoding a domain from human type 1 neurofibromin protein results in a threefold improvement in expression level in *Escherichia coli*. *Protein Expr Purif* 1998, 12(2), 185-188.
- 36 Longacre, S., Mendis, K.N. & David, P.H. *Plasmodium vivax* merozoite surface protein 1 C-terminal recombinant proteins in baculovirus. *Mol Biochem Parasitol* 1994, 64(2), 191-205.

- 37 Nacer, A., Berry, L., Slomianny, C. & Mattei, D. Plasmodium falciparum signal sequences: simply sequences or special signals? *Int J Parasitol* 2001, 31(12), 1371-1379.
- 38 Utsumi, T., Nakano, K., Funakoshi, T. et al. Vertical-scanning mutagenesis of amino acids in a model N-myristoylation motif reveals the major amino-terminal sequence requirements for protein N-myristoylation. *Eur J Biochem* 2004, 271(4), 863-874.
- 39 Resh, M. D. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim Biophys Acta* 1999, 1451(1), 1-16.
- 40 Gunaratne, R.S., Sajid, M., Ling, IT., Tripathi, R., Pachebat, J.A. & Holder, A.A. Characterization of N-myristoyltransferase from Plasmodium falciparum. *Biochem J* 2000, 348 Pt 2, 459-463.
- 41 Price, H.P., Menon, M.R., Panethymitaki, C, Goulding, D., McKean, P.G. & Smith, D. F. Myristoyl-CoA:protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites. *J Biol Chem* 2003, 278(9), 7206-7214.
- 42 Stafford, W.H., Stockley, R.W., Ludbrook, S. B. & Holder, A.A. Isolation, expression and characterization of the gene for an ADP-ribosylation factor from the human malaria parasite, Plasmodium falciparum. *Eur J Biochem* 1996, 242(1), 104-113.
- 43 Stevenson, F.T., Bursten, S. L., Locksley, R.M. & Lovett, D.H. Myristyl acylation of the tumor necrosis factor alpha precursor on specific lysine residues. *J Exp Med* 1992, 176(4), 1053-1062.
- 44 Farazi, T.A., Waksman, G. & Gordon, J.I. The biology and enzymology of protein N-myristoylation. *J Biol Chem* 2001, 276(43), 39501-39504.

45 Hayashi, N., Nakagawa, C, Ito, Y. et al. Myristoylation-regulated direct interaction between calcium-bound calmodulin and N-terminal region of pp60v-src. *J Mol Biol* 2004, 338(1), 169-180.

46 Maekawa, S., Matsuura, Y. & Nakamura, S. Expression and myristoylation of NAP-22 using a baculovirus transfer vector system. *Biochim Biophys Acta* 1994, 1218(1), 119-122.

47 Perraut, R., Marrama, L., Diouf, B. et al. Distinct surrogate markers for protection against *Plasmodium falciparum* infection and clinical malaria identified in a Senegalese community after radical drug cure. *J Infect Dis* 2003, 188(12), 1940-1950.

48 Schofield, L., Hewitt, M. C, Evans, K., Siomos, M.A. & Seeberger, P. H. Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria. *Nature* 2002, 418(6899), 785-789.

49 Bonnet S. et al, Vaccine 2006, vol. 24, p. 5997-6008, soluble and glyco-lipid modified Baculovirus *Plasmodium falciparum* C-terminal Merozoite Surface Protein 1 , two forms of leading malaria vaccine candidate.

序列表

<110> 巴斯德研究院
国立科学研究中心

<120> 重组恶性疟原虫裂殖子表面蛋白 4 和 5 及其用途

<130> B6989A_AD/CAL

<140> PCT/IB2006/003967

<141> 2006-11-23

<150> US 60/739, 973

<151> 2005-11-23

<160> 57

<170> PatentIn Ver. 3.3

<210> 1

<211> 363

<212> DNA

<213> Plasmodium falciparum

<400> 1

```
aagagtccca aggagagtca aatgggtcgac gacaagaaga agaccgaggc cattccaaag 60
aaagtcgtgc agccaagctc gagcaactct ggaggtcacg tcggtgaaga agaagaccac 120
aacgaaggag agggagagca cgaagaggag gaagaacacg aagaagacga tgacgacgag 180
gacgacgaca catacaacaa agacgacttg gaggacgaag atctttgcaa gcacaacaac 240
ggaggatgtg gagatgacaa gctctgcgag tacgttgaa accgtcgcgt aaaatgtaaa 300
tgtaaggaag gatacaagtt ggaaggaatt gagtgcgttg aacaccacca ccaccatcac 360
taa 363
```

<210> 2

<211> 666

<212> DNA

<213> Plasmodium falciparum

<400> 2

```
atgtggatcg taaagttctt gattgtggtc cacttcttca tcatatgcac catcaacttc 60
gacaagctct acattagtta ctcttacaac atcgtccctg aaaacggacg tatgcttaac 120
atgaggatct tggtgcaga aaagaaggac gaaaaggaag caagcgagca aggcgaagaa 180
tcccacaaga aggaaaactc tcaggaatct gcaaacggaa aagacgacgt taaggaggag 240
aagaagacca acgagaagaa ggacgacgga aagactgaca aggtacaaga aaaggttcta 300
gaaaagagtc ccaaggagag tcaaatggtc gacgacaaga agaagaccga ggccattcca 360
aagaaagtcg tgcagccaag ctcgagcaac tctggaggtc acgtcgggtga agaagaagac 420
cacaacgaag gagagggaga gcacgaagag gaggaagaac acgaagaaga cgatgacgac 480
gaggacgacg acacatacaa caaagacgac ttggaggacg aagatctttg caagcacaac 540
aacggaggat gtggagatga caagctctgc gagtacgttg gaaaccgtcg cgtaaaatgt 600
aaatgtaagg aaggatacaa gttggaagga attgagtgcg ttgaacacca ccaccaccat 660
cactaa 666
```

<210> 3

<211> 756

<212> DNA

<213> Plasmodium falciparum

<400> 3

```

atgtggatcg taaagttctt gatttggtgc cacttcttca tcatatgcac catcaacttc 60
gacaagctct acattagtta ctcttacaac atcgtccttg aaaacggacg tatgcttaac 120
atgaggatct tgggtgaaga aaagcctaac gttgacggtg tgtcaacatc taacacacct 180
ggcggaaacg aggcattctag tgcttctcct aaccttgctg acgctgcaga aaagaaggac 240
gaaaaggaag caagcgagca aggcgaagaa tcccacaaga aggaaaactc tcaggaatct 300
gcaaacggaa aagacgacgt taaggaggag aagaagacca acgagaagaa ggacgacgga 360
aagactgaca aggtacaaga aaaggttcta gaaaagagtc ccaaggagag tcaaatggtc 420
gacgacaaga agaagaccga ggccattcca aagaaagtcg tgcagccaag ctcgagcaac 480
tctggaggtc acgtcgggtga agaagaagac cacaacgaag gagagggaga gcacgaagag 540
gaggaagaac acgaagaaga cgatgacgac gaggacgacg acacatacaa caaagacgac 600
ttggaggacg aagatctttg caagcacaac aacggaggat gtggagatga caagctctgc 660
gagtacgttg gaaaccgtcg cgtaaaatgt aaatgtaagg aaggatacaa gttggaagga 720
attgagtgcg ttgaacacca ccaccacat cactaa 756

```

<210> 4

<211> 762

<212> DNA

<213> Plasmodium falciparum

<400> 4

```

atgaacattc tctgtattct cagctacatt tacttcttcg tcatcttcta cagtttaaac 60
ctcaacaaca aaaacgagaa ctcttgggtg gtccgcagac tcatgaacga cgaaaaggga 120
gaagggtgct tcactagtaa gaacaaggaa aacggaaaca acaacaggaa caacgagaac 180
gaactcaaag aagaaggatc tttgccact aagatgaacg agaaaaacag taactccgcg 240
gataagcaac caaacgacat ctcccacgac gaaagcaaga gcaacagtaa caacgccaac 300
aacatccaaa aggaacctga agagaaggaa aactcaaac ccaacctcga ctcgagtga 360
aactccgctg aaagtgtctac tagaagcgtc gacatcagtg aacacaactc aaacaacccc 420
gaaactaaag aagaaaacgg agaagaacct ctagacctgg aaattaacga aaacgcagaa 480
atcgccagg aacctccaaa ccgtcttcac ttcgacaacg ttgacgacga agtaccacat 540
tactcagccc tgaggtaaca caaggtcgag aagaacgtaa ccgacgagat gctcttgtac 600
aacatgatgt ccgaccaaaa ccgcaaaagc tgtgctatca acaacgggtg ctgcagtgac 660
gaccagatct gcatcaacat caacaacatc ggtgtgaagt gcatttgtaa ggatggatac 720
ctacttggtg ccaagtgcac tcaccaccac caccaccact ga 762

```

<210> 5

<211> 207

<212> DNA

<213> Plasmodium falciparum

<400> 5

```

tacaacaagg tcgagaagaa cgtaaccgac gagatgctct tgtacaacat gatgtccgac 60
caaaaccgca aaagctgtgc tatcaacaac ggtggctgca gtgacgacca gatctgcatc 120
aacatcaaca acatcgggtg gaagtgcatt tgtaaggatg gatacctact tggtaccaag 180
tgcattcacc accaccacca ccactga 207

```

<210> 6

<211> 498

<212> DNA

<213> Plasmodium falciparum

<400> 6

```

atgtggatcg taaagttctt gatttggtgc cacttcttca tcatatgcac catcaacttc 60
gacaagctct acattagtta ctcttacaac atcgtccttg aaaacggacg tatgcttaac 120
atgaggattc tagaaaagag tccaaggag agtcaaatgg tcgacgacaa gaagaagacc 180
gaggccattc caaagaaagt cgtgcagcca agctcgagca actctggagg tcacgtcggt 240
gaagaagaag accacaacga aggagaggga gagcacgaag aggaggaaga acacgaagaa 300

```

gacgatgacg acgaggacga cgacacatac aacaaagacg acttggagga cgaagatctt 360
 tgcaagcaca acaacggagg atgtggagat gacaagctct gcgagtacgt tggaaccgt 420
 cgcgtaaaat gtaaatgtaa ggaaggatac aagttggaag gaattgagtg cgttgaacac 480
 caccaccacc atcactaa 498

<210> 7

<211> 540

<212> DNA

<213> Plasmodium falciparum

<400> 7

atgtggatcg taaagttctt gatttgtggtc cacttcttca tcatatgcac catcaacttc 60
 gacaagctct acattagtta ctcttacaac atcgtccctg aaaacggacg tatgcttaac 120
 atgaggatct tgggtgaaga aaagcctaac gttgacggtg tgtcaacatc tctagaaaag 180
 agtcccaagg agagtcaaat ggtcgacgac aagaagaaga ccgaggccat tccaaagaaa 240
 gtcgtgcagc caagctcgag caactctgga ggtcacgtcg gtgaagaaga agaccacaac 300
 gaaggagagg gagagcacga agaggaggaa gaacacgaag aagacgatga cgacgaggac 360
 gacgacacat acaacaaaga cgacttgag gacgaagatc ttgcaagca caacaacgga 420
 ggatgtggag atgacaagct ctgcgagtac gttggaaacc gtcgcgtaaa atgtaaatgt 480
 aaggaaggat acaagttgga aggaattgag tgcgttgaac accaccacca ccatcactaa 540

<210> 8

<211> 516

<212> DNA

<213> Plasmodium falciparum

<400> 8

atgtggatcg taaagttctt gatttgtggtc cacttcttca tcatatgcac catcaacttc 60
 gacaagctct acattagtta ctcttacaac atcgtccctg aaaacggacg tatgcttaac 120
 atgaggatct tgggtgaaga aaagcctcta gaaaagagtc ccaaggagag tcaaatggtc 180
 gacgacaaga agaagaccga ggccattcca aagaaagtcg tgcagccaag ctcgagcaac 240
 tctggaggtc acgtcgggtga agaagaagac cacaacgaag gagagggaga gcacgaagag 300
 gaggaagaac acgaagaaga cgatgacgac gaggacgacg acacatacaa caaagacgac 360
 ttggaggacg aagatctttg caagcacaaac aacggaggat gtggagatga caagctctgc 420
 gagtacgttg gaaaccgtcg cgtaaaatgt aaatgtaagg aaggatacaa gttggaagga 480
 attgagtgcg ttgaacacca ccaccacat cactaa 516

<210> 9

<211> 120

<212> PRT

<213> Plasmodium falciparum

<400> 9

Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp Lys Lys Lys Thr Glu
 1 5 10 15

Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser Ser Asn Ser Gly Gly
 20 25 30

His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly Glu His Glu
 35 40 45

Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp Asp Asp Thr
 50 55 60

Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys His Asn Asn
 65 70 75 80

Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly Asn Arg Arg
85 90 95

Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly Ile Glu Cys
100 105 110

Val Glu His His His His His His
115 120

<210> 10

<211> 221

<212> PRT

<213> Plasmodium falciparum

<400> 10

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
1 5 10 15

Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
20 25 30

Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Ala Ala Glu Lys
35 40 45

Lys Asp Glu Lys Glu Ala Ser Glu Gln Gly Glu Glu Ser His Lys Lys
50 55 60

Glu Asn Ser Gln Glu Ser Ala Asn Gly Lys Asp Asp Val Lys Glu Glu
65 70 75 80

Lys Lys Thr Asn Glu Lys Lys Asp Asp Gly Lys Thr Asp Lys Val Gln
85 90 95

Glu Lys Val Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp
100 105 110

Lys Lys Lys Thr Glu Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser
115 120 125

Ser Asn Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly
130 135 140

Glu Gly Glu His Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp
145 150 155 160

Glu Asp Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu
165 170 175

Cys Lys His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr
180 185 190

Val Gly Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu
195 200 205

Glu Gly Ile Glu Cys Val Glu His His His His His His
210 215 220

<210> 11
 <211> 181
 <212> PRT
 <213> Plasmodium falciparum

<400> 11
 Met Arg Ile Leu Ala Ala Glu Lys Lys Asp Glu Lys Glu Ala Ser Glu
 1 5 10 15
 Gln Gly Glu Glu Ser His Lys Lys Glu Asn Ser Gln Glu Ser Ala Asn
 20 25 30
 Gly Lys Asp Asp Val Lys Glu Glu Lys Lys Thr Asn Glu Lys Lys Asp
 35 40 45
 Asp Gly Lys Thr Asp Lys Val Gln Glu Lys Val Leu Glu Lys Ser Pro
 50 55 60
 Lys Glu Ser Gln Met Val Asp Asp Lys Lys Lys Thr Glu Ala Ile Pro
 65 70 75 80
 Lys Lys Val Val Gln Pro Ser Ser Ser Asn Ser Gly Gly His Val Gly
 85 90 95
 Glu Glu Glu Asp His Asn Glu Gly Glu Gly Glu His Glu Glu Glu Glu
 100 105 110
 Glu His Glu Glu Asp Asp Asp Asp Glu Asp Asp Asp Thr Tyr Asn Lys
 115 120 125
 Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys His Asn Asn Gly Gly Cys
 130 135 140
 Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly Asn Arg Arg Val Lys Cys
 145 150 155 160
 Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly Ile Glu Cys Val Glu His
 165 170 175
 His His His His His
 180

<210> 12
 <211> 251
 <212> PRT
 <213> Plasmodium falciparum

<400> 12
 Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
 1 5 10 15
 Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
 20 25 30
 Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys
 35 40 45
 Pro Asn Val Asp Gly Val Ser Thr Ser Asn Thr Pro Gly Gly Asn Glu
 50 55 60

Ala Ser Ser Ala Ser Pro Asn Leu Ala Asp Ala Ala Glu Lys Lys Asp
 65 70 75 80
 Glu Lys Glu Ala Ser Glu Gln Gly Glu Glu Ser His Lys Lys Glu Asn
 85 90 95
 Ser Gln Glu Ser Ala Asn Gly Lys Asp Asp Val Lys Glu Glu Lys Lys
 100 105 110
 Thr Asn Glu Lys Lys Asp Asp Gly Lys Thr Asp Lys Val Gln Glu Lys
 115 120 125
 Val Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp Lys Lys
 130 135 140
 Lys Thr Glu Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser Ser Asn
 145 150 155 160
 Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly
 165 170 175
 Glu His Glu Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp
 180 185 190
 Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys
 195 200 205
 His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly
 210 215 220
 Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly
 225 230 235 240
 Ile Glu Cys Val Glu His His His His His His
 245 250

<210> 13

<211> 211

<212> PRT

<213> Plasmodium falciparum

<400> 13

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

85										90					95				
Ser	Gln	Met	Val	Asp	Asp	Lys	Lys	Lys	Thr	Glu	Ala	Ile	Pro	Lys	Lys				
			100						105					110					
Val	Val	Gln	Pro	Ser	Ser	Ser	Asn	Ser	Gly	Gly	His	Val	Gly	Glu	Glu				
		115					120					125							
Glu	Asp	His	Asn	Glu	Gly	Glu	Gly	Glu	His	Glu	Glu	Glu	Glu	Glu	Glu	His			
	130					135					140								
Glu	Glu	Asp	Asp	Asp	Asp	Glu	Asp	Asp	Asp	Thr	Tyr	Asn	Lys	Asp	Asp				
145						150				155					160				
Leu	Glu	Asp	Glu	Asp	Leu	Cys	Lys	His	Asn	Asn	Gly	Gly	Cys	Gly	Asp				
				165					170						175				
Asp	Lys	Leu	Cys	Glu	Tyr	Val	Gly	Asn	Arg	Arg	Val	Lys	Cys	Lys	Cys				
			180						185				190						
Lys	Glu	Gly	Tyr	Lys	Leu	Glu	Gly	Ile	Glu	Cys	Val	Glu	His	His	His				
		195					200					205							
His	His	His																	
		210																	
<210> 14																			
<211> 253																			
<212> PRT																			
<213> Plasmodium falciparum																			
<400> 14																			
Met	Asn	Ile	Leu	Cys	Ile	Leu	Ser	Tyr	Ile	Tyr	Phe	Phe	Val	Ile	Phe				
1				5					10					15					
Tyr	Ser	Leu	Asn	Leu	Asn	Asn	Lys	Asn	Glu	Asn	Phe	Leu	Val	Val	Arg				
			20					25					30						
Arg	Leu	Met	Asn	Asp	Glu	Lys	Gly	Glu	Gly	Gly	Phe	Thr	Ser	Lys	Asn				
		35					40					45							
Lys	Glu	Asn	Gly	Asn	Asn	Asn	Arg	Asn	Asn	Glu	Asn	Glu	Leu	Lys	Glu				
	50					55					60								
Glu	Gly	Ser	Leu	Pro	Thr	Lys	Met	Asn	Glu	Lys	Asn	Ser	Asn	Ser	Ala				
65					70					75					80				
Asp	Lys	Gln	Pro	Asn	Asp	Ile	Ser	His	Asp	Glu	Ser	Lys	Ser	Asn	Ser				
				85					90					95					
Asn	Asn	Ala	Gln	Asn	Ile	Gln	Lys	Glu	Pro	Glu	Glu	Lys	Glu	Asn	Ser				
			100					105					110						
Asn	Pro	Asn	Leu	Asp	Ser	Ser	Glu	Asn	Ser	Ala	Glu	Ser	Ala	Thr	Arg				
		115					120					125							
Ser	Val	Asp	Ile	Ser	Glu	His	Asn	Ser	Asn	Asn	Pro	Glu	Thr	Lys	Glu				
						135						140							

Glu Asn Gly Glu Glu Pro Leu Asp Leu Glu Ile Asn Glu Asn Ala Glu
145 150 155 160

Ile Gly Gln Glu Pro Pro Asn Arg Leu His Phe Asp Asn Val Asp Asp
165 170 175

Glu Val Pro His Tyr Ser Ala Leu Arg Tyr Asn Lys Val Glu Lys Asn
180 185 190

Val Thr Asp Glu Met Leu Leu Tyr Asn Met Met Ser Asp Gln Asn Arg
195 200 205

Lys Ser Cys Ala Ile Asn Asn Gly Gly Cys Ser Asp Asp Gln Ile Cys
210 215 220

Ile Asn Ile Asn Asn Ile Gly Val Lys Cys Ile Cys Lys Asp Gly Tyr
225 230 235 240

Leu Leu Gly Thr Lys Cys Ile His His His His His His
245 250

<210> 15

<211> 68

<212> PRT

<213> Plasmodium falciparum

<400> 15

Tyr Asn Lys Val Glu Lys Asn Val Thr Asp Glu Met Leu Leu Tyr Asn
1 5 10 15

Met Met Ser Asp Gln Asn Arg Lys Ser Cys Ala Ile Asn Asn Gly Gly
20 25 30

Cys Ser Asp Asp Gln Ile Cys Ile Asn Ile Asn Asn Ile Gly Val Lys
35 40 45

Cys Ile Cys Lys Asp Gly Tyr Leu Leu Gly Thr Lys Cys Ile His His
50 55 60

His His His His
65

<210> 16

<211> 212

<212> PRT

<213> Plasmodium vivax

<400> 16

Met Lys Val Ala Tyr Phe Leu Ser Val Leu Asp Leu Leu Ile Ile Phe
1 5 10 15

Ser Leu Tyr Phe Asp Gly Arg Arg Ser Ala Phe Ala Gly Ile Ala Ala
20 25 30

Cys Ile Arg His Gly Arg Ile Leu Gly Glu Gly Gly Glu Gln Ser Gly
35 40 45

Gly Ala Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Asp Ser Ser Gly

50	55	60
Gly Leu Ser Gly Gly Ser Ser Gly Gly Pro Ser Pro Pro Ala Gly Ser		
65	70	75 80
Ser Gly Ser Gly Gly Ser Asp Pro Ala Asn Ser Ala Thr Gly Pro Gln		
	85	90 95
Asn Ser Thr Pro Gly Ser Gly Gly Gln Thr Gly Asp His Ser Ala Glu		
	100	105 110
Ala Glu Asn Gly Asp Tyr Asn Glu Gln Gly Asp Asp His Gly Asp Asp		
	115	120 125
His Gly Asp Asp His Gly Asp Asp His Gly Asp Glu Gln Asp Gly Glu		
	130	135 140
Asp Tyr Asp Asp Ala Glu Asp Asp Asp Leu Tyr Glu Leu Ser Glu Val		
	145	150 155 160
Asp Glu Asn Ala Asn Leu Cys Leu Asp Asn Asn Gly Gly Cys Gly Asp		
	165	170 175
Asp Lys Ile Cys Glu Asn Leu Gly Lys Gly Ile Val Lys Cys Leu Cys		
	180	185 190
Lys Pro Gly Tyr Lys Leu Val Gly Thr Glu Cys Val Glu Ser His His		
	195	200 205
His His His His		
	210	

<210> 17

<211> 165

<212> PRT

<213> Plasmodium falciparum

<400> 17

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

Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys His Asn Asn Gly Gly Cys
115 120 125

Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly Asn Arg Arg Val Lys Cys
130 135 140

Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly Ile Glu Cys Val Glu His
145 150 155 160

His His His His His
165

<210> 18

<211> 125

<212> PRT

<213> Plasmodium falciparum

<400> 18

Met Arg Ile Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp
1 5 10 15

Lys Lys Lys Thr Glu Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser
20 25 30

Ser Asn Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly
35 40 45

Glu Gly Glu His Glu Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp
50 55 60

Glu Asp Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu
65 70 75 80

Cys Lys His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr
85 90 95

Val Gly Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu
100 105 110

Glu Gly Ile Glu Cys Val Glu His His His His His His
115 120 125

<210> 19

<211> 179

<212> PRT

<213> Plasmodium falciparum

<400> 19

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
1 5 10 15

Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
20 25 30

Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys
35 40 45

Pro Asn Val Asp Gly Val Ser Thr Ser Leu Glu Lys Ser Pro Lys Glu

50	55	60
Ser Gln Met Val Asp Asp Lys Lys Lys Thr Glu Ala Ile Pro Lys Lys		
65	70	75 80
Val Val Gln Pro Ser Ser Ser Asn Ser Gly Gly His Val Gly Glu Glu		
	85	90 95
Glu Asp His Asn Glu Gly Glu Gly Glu His Glu Glu Glu Glu Glu His		
	100	105 110
Glu Glu Asp Asp Asp Asp Glu Asp Asp Asp Thr Tyr Asn Lys Asp Asp		
	115	120 125
Leu Glu Asp Glu Asp Leu Cys Lys His Asn Asn Gly Gly Cys Gly Asp		
	130	135 140
Asp Lys Leu Cys Glu Tyr Val Gly Asn Arg Arg Val Lys Cys Lys Cys		
	145	150 155 160
Lys Glu Gly Tyr Lys Leu Glu Gly Ile Glu Cys Val Glu His His His		
	165	170 175
His His His		

<210> 20
 <211> 139
 <212> PRT
 <213> Plasmodium falciparum

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

<210> 21

<211> 171

<212> PRT

<213> Plasmodium falciparum

<400> 21

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
 1 5 10 15

Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
 20 25 30

Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys
 35 40 45

Pro Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp Lys Lys
 50 55 60

Lys Thr Glu Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser Ser Asn
 65 70 75 80

Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly
 85 90 95

Glu His Glu Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp
 100 105 110

Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys
 115 120 125

His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly
 130 135 140

Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly
 145 150 155 160

Ile Glu Cys Val Glu His His His His His His
 165 170

<210> 22

<211> 131

<212> PRT

<213> Plasmodium falciparum

<400> 22

Met Arg Ile Leu Gly Glu Glu Lys Pro Leu Glu Lys Ser Pro Lys Glu
 1 5 10 15

Ser Gln Met Val Asp Asp Lys Lys Lys Thr Glu Ala Ile Pro Lys Lys
 20 25 30

Val Val Gln Pro Ser Ser Ser Asn Ser Gly Gly His Val Gly Glu Glu
 35 40 45

Glu Asp His Asn Glu Gly Glu Gly Glu His Glu Glu Glu Glu His
 50 55 60

Glu Glu Asp Asp Asp Asp Glu Asp Asp Asp Thr Tyr Asn Lys Asp Asp

65	70	75	80
Leu Glu Asp Glu Asp	Leu Cys Lys His Asn Asn Gly Gly Cys Gly Asp		
	85	90	95
Asp Lys Leu Cys Glu Tyr Val Gly Asn Arg Arg Val Lys Cys Lys Cys			
	100	105	110
Lys Glu Gly Tyr Lys Leu Glu Gly Ile Glu Cys Val Glu His His His			
	115	120	125
His His His			
	130		

<210> 23
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 23
 tatagcagat cttgtcgaa gttgatggtg ca 32

<210> 24
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 24
 atatggctgc agccaagatc ctcatgttaa gcat 34

<210> 25
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 25
 attaatctag aggcttttct tcaccaaga tcctcatg 38

<210> 26
 <211> 5
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic peptide

<400> 26

Tyr Asn Lys Val Glu
1 5

<210> 27

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic peptide

<400> 27

Tyr Asn Lys Val Glu
1 5

<210> 28

<211> 540

<212> DNA

<213> Plasmodium falciparum

<400> 28

```
atgtggatcg taaagttctt gattgtggtc cacttcttca tcatatgcac catcaacttc 60
gacaagctct acattagtta ctcttacaac atcgtccttg aaaacggacg tatgcttaac 120
atgaggatct tgggtgaaga aaagcctaac gttgacgggtg tgtcaacatc tctagaaaag 180
agtccaagg agagtcaa at ggtcgacgac aagaagaaga ccgaggccat tccaaagaaa 240
gtcgtgcagc caagctcgag caactctgga ggtcacgtcg gtgaagaaga agaccacaac 300
gaaggagagg gagagcacga agaggaggaa gaacacgaag aagacgatga cgacgaggac 360
gacgacacat acaacaaaga cgacttggag gacgaagatc tttgcaagca caacaacgga 420
ggatgtggag atgacaagct ctgcgagtac gttggaaacc gtcgcgtaaa atgtaa atgt 480
aaggaaggat acaagttgga aggaattgag tgcgttgaac accaccacca ccatcactaa 540
```

<210> 29

<211> 516

<212> DNA

<213> Plasmodium falciparum

<400> 29

```
atgtggatcg taaagttctt gattgtggtc cacttcttca tcatatgcac catcaacttc 60
gacaagctct acattagtta ctcttacaac atcgtccttg aaaacggacg tatgcttaac 120
atgaggatct tgggtgaaga aaagcctcta gaaaagagtc ccaaggagag tcaaatggtc 180
gacgacaaga agaagaccga ggccattcca aagaaagtcg tgcagccaag ctcgagcaac 240
tctggaggtc acgtcgggtga agaagaagac cacaacgaag gagagggaga gcacgaagag 300
gaggaagaac acgaagaaga cgatgacgac gaggacgacg acacatacaa caaagacgac 360
ttggaggacg aagatctttg caagcacaac aacggaggat gtggagatga caagctctgc 420
gagtacgttg gaaaccgtcg cgtaaaatgt aaatgtaagg aaggatacaa gttggaagga 480
attgagtgcg ttgaacacca ccaccacct cactaa 516
```

<210> 30

<211> 161

<212> PRT

<213> Plasmodium falciparum

<400> 30

Met Lys Val Ala Tyr Phe Leu Ser Val Leu Asp Leu Leu Ile Ile Phe
 1 5 10 15

Ser Leu Tyr Phe Asp Gly Arg Arg Ser Ala Phe Ala Gly Ile Ala Ala
 20 25 30

Cys Ile Arg His Gly Arg Ile Leu Gly Glu Gly Gly Glu Gln Asn Ser
 35 40 45

Thr Pro Gly Ser Gly Gly Gln Thr Gly Asp His Ser Ala Glu Ala Glu
 50 55 60

Asn Gly Asp Tyr Asn Glu Gln Gly Asp Asp His Gly Asp Asp His Gly
 65 70 75 80

Asp Asp His Gly Asp Asp His Gly Asp Glu Gln Asp Gly Glu Asp Tyr
 85 90 95

Asp Asp Ala Glu Asp Asp Asp Leu Tyr Glu Leu Ser Glu Val Asp Glu
 100 105 110

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

Ile Cys Glu Asn Leu Gly Lys Gly Ile Val Lys Cys Leu Cys Lys Pro
 130 135 140

Gly Tyr Lys Leu Val Gly Thr Glu Cys Val Glu His His His His His
 145 150 155 160

His

<210> 31

<211> 83

<212> PRT

<213> Plasmodium falciparum

<400> 31

Arg Val Ser Thr Ser Asp Thr Pro Gly Gly Asn Glu Ser Ser Ser Ala
 1 5 10 15

Phe Pro Gln Phe Ile Trp Ser Ala Glu Lys Lys Asp Glu Lys Glu Ala
 20 25 30

Ser Glu Gln Gly Glu Glu Ser His Lys Lys Glu Asn Ser Gln Glu Ser
 35 40 45

Ala Asn Gly Lys Asp Asp Val Lys Glu Glu Lys Lys Thr Asn Glu Lys
 50 55 60

Lys Asp Asp Gly Lys Thr Asp Lys Val Gln Glu Lys Val Leu Glu Lys
 65 70 75 80

Ser Pro Lys

<210> 32

<211> 272

<212> PRT

<213> Plasmodium falciparum

<400> 32

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
 1 5 10 15

Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
 20 25 30

Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys
 35 40 45

Pro Asn Val Asp Gly Val Ser Thr Ser Asp Thr Pro Gly Gly Asn Glu
 50 55 60

Ser Ser Ser Ala Ser Pro Asn Leu Ser Asp Ala Ala Glu Lys Lys Asp
 65 70 75 80

Glu Lys Glu Ala Ser Glu Gln Gly Glu Glu Ser His Lys Lys Glu Asn
 85 90 95

Ser Gln Glu Ser Ala Asn Gly Lys Asp Asp Val Lys Glu Glu Lys Lys
 100 105 110

Thr Asn Glu Lys Lys Asp Asp Gly Lys Thr Asp Lys Val Gln Glu Lys
 115 120 125

Val Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp Lys Lys
 130 135 140

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

Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly
 165 170 175

Glu His Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp
 180 185 190

Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys
 195 200 205

His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly
 210 215 220

Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly
 225 230 235 240

Ile Glu Cys Val Glu Leu Leu Ser Leu Ala Ser Ser Ser Leu Asn Leu
 245 250 255

Ile Phe Asn Ser Phe Ile Thr Ile Phe Val Val Ile Leu Leu Ile Asn
 260 265 270

<210> 33

<211> 760
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 construct

<220>
 <221> CDS
 <222> (10).. (759)

<400> 33

```
ccaggatcc atg tgg atc gta aag ttc ttg att gtg gtc cac ttc ttc atc 51
      Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile
        1             5             10

ata tgc acc atc aac ttc gac aag ctc tac att agt tac tct tac aac 99
Ile Cys Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn
  15             20             25             30

atc gtc cct gaa aac gga cgt atg ctt aac atg agg atc ttg ggt gaa 147
Ile Val Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu
      35             40             45

gaa aag cct aac gtt gac ggt gtg tca aca tct aac aca cct ggc gga 195
Glu Lys Pro Asn Val Asp Gly Val Ser Thr Ser Asn Thr Pro Gly Gly
      50             55             60

aac gag gca tct agt gct tct cct aac ctt gct gac gct gca gaa aag 243
Asn Glu Ala Ser Ser Ala Ser Pro Asn Leu Ala Asp Ala Ala Glu Lys
      65             70             75

aag gac gaa aag gaa gca agc gag caa ggc gaa gaa tcc cac aag aag 291
Lys Asp Glu Lys Glu Ala Ser Glu Gln Gly Glu Glu Ser His Lys Lys
      80             85             90

gaa aac tct cag gaa tct gca aac gga aaa gac gac gtt aag gag gag 339
Glu Asn Ser Gln Glu Ser Ala Asn Gly Lys Asp Asp Val Lys Glu Glu
      95             100            105            110

aag aag acc aac gag aag aag gac gac gga aag act gac aag gta caa 387
Lys Lys Thr Asn Glu Lys Lys Asp Asp Gly Lys Thr Asp Lys Val Gln
      115            120            125

gaa aag gtt cta gaa aag agt ccc aag gag agt caa atg gtc gac gac 435
Glu Lys Val Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp
      130            135            140

aag aag aag acc gag gcc att cca aag aaa gtc gtg cag cca agc tcg 483
Lys Lys Lys Thr Glu Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser
      145            150            155

agc aac tct gga ggt cac gtc ggt gaa gaa gaa gac cac aac gaa gga 531
Ser Asn Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly
      160            165            170

gag gga gag cac gaa gag gag gaa gaa cac gaa gaa gac gat gac gac 579
Glu Gly Glu His Glu Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp
      175            180            185            190
```

gag gac gac gac aca tac aac aaa gac gac ttg gag gac gaa gat ctt 627
 Glu Asp Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu
 195 200 205

tgc aag cac aac aac gga gga tgt gga gat gac aag ctc tgc gag tac 675
 Cys Lys His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr
 210 215 220

gtt gga aac cgt cgc gta aaa tgt aaa tgt aag gaa gga tac aag ttg 723
 Val Gly Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu
 225 230 235

gaa gga att gag tgc gtt gaa cac cac cac cac cat c 760
 Glu Gly Ile Glu Cys Val Glu His His His His His
 240 245 250

<210> 34

<211> 250

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
 construct

<400> 34

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
 1 5 10 15

Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
 20 25 30

Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys
 35 40 45

Pro Asn Val Asp Gly Val Ser Thr Ser Asn Thr Pro Gly Gly Asn Glu
 50 55 60

Ala Ser Ser Ala Ser Pro Asn Leu Ala Asp Ala Ala Glu Lys Lys Asp
 65 70 75 80

Glu Lys Glu Ala Ser Glu Gln Gly Glu Glu Ser His Lys Lys Glu Asn
 85 90 95

Ser Gln Glu Ser Ala Asn Gly Lys Asp Asp Val Lys Glu Glu Lys Lys
 100 105 110

Thr Asn Glu Lys Lys Asp Asp Gly Lys Thr Asp Lys Val Gln Glu Lys
 115 120 125

Val Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp Lys Lys
 130 135 140

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

Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly
 165 170 175

Glu His Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp
180 185 190

Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys
195 200 205

His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly
210 215 220

Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly
225 230 235 240

Ile Glu Cys Val Glu His His His His His
245 250

<210> 35

<211> 760

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
construct

<400> 35

```
ggggcgctctg aattcttagt gatggtggtg gtggtgttca acgcactcaa ttccttccaa 60
cttgatctct tccttacatt tacattttac gcgacggttt ccaacgtact cgcagagctt 120
gtcatctcca catectccgt tgttgtgctt gcaaagatct tcgtcctcca agtcgtcttt 180
gttgatgtg tcgtcgctct cgtcgtcctc gtcttcttcg tgttcttctt cctcttcgtg 240
ctctccctct ccttcgttgt ggtcttcttc ttcaccgacg tgacctccag agttgctcga 300
gcttggtgc acgactttct ttggaatggt ctcggtcttc ttcttgctgt cgaccatttg 360
actctccttg ggactctttt ctagaacctt ttcttgtagc ttgtcagctt ttccgtcgtc 420
cttcttctcg ttggtcttct tctctctctt aacgtcgtct ttccggttg cagattcctg 480
agagttttcc ttcttggtgg attcttcgcc ttgtcgtctt gcttctttt cgtccttctt 540
ttctgcagcg tcagcaaggt taggagaagc actagatgcc tcgtttccgc caggtgtgtt 600
agatgttgac acaccgtcaa cgtaggctt ttcttcaccc aagatcctca tgttaagcat 660
acgtccgttt tcaggacga tgttgtaaga gtaactaat tagagcttgt cgaagttagt 720
ggtgcatatg atgaagaagt ggaccacaat caagaacttt 760
```

<210> 36

<211> 69

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
construct

<400> 36

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
1 5 10 15

Thr Ile Asn Phe Asp Lys Asp Leu Cys Lys His Asn Asn Gly Gly Cys
20 25 30

Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly Asn Arg Arg Val Lys Cys

35	40	45
Lys Cys Lys Glu Gly Tyr	Lys Leu Glu Gly Ile	Glu Cys Val Glu His
50	55	60
His His His His His		
65		
<210> 37 <211> 171 <212> PRT <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Synthetic construct <400> 37 Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys 1 5 10 15 Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val 20 25 30 Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys 35 40 45 Pro Leu Glu Lys Ser Pro Lys Lys Val Val Gln Met Val Asp Asp Lys Lys 50 55 60 Lys Thr Glu Ala Ile Pro Lys Lys Val Val Gln Pro Ser Ser Ser Asn 65 70 75 80 Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly 85 90 95 Glu His Glu Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp 100 105 110 Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys 115 120 125 His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly 130 135 140 Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly 145 150 155 160 Ile Glu Cys Val Glu His His His His His His 165 170		

<210> 38
 <211> 760
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS

<222> (10).. (759)

<220>

<223> Description of Artificial Sequence: Synthetic construct

<400> 38

```

cccggatcc atg aac att ctc tgt att ctc agc tac att tac ttc ttc gtc 51
      Met Asn Ile Leu Cys Ile Leu Ser Tyr Ile Tyr Phe Phe Val
          1              5              10

atc ttc tac agt tta aac ctc aac aac aaa aac gag aac ttc ttg gtg 99
Ile Phe Tyr Ser Leu Asn Leu Asn Asn Lys Asn Glu Asn Phe Leu Val
    15              20              25              30

gtc cgc aga ctc atg aac gac gaa aag gga gaa ggt ggc ttc act agt 147
Val Arg Arg Leu Met Asn Asp Glu Lys Gly Glu Gly Gly Phe Thr Ser
          35              40              45

aag aac aag gaa aac gga aac aac aac agg aac aac gag aac gaa ctc 195
Lys Asn Lys Glu Asn Gly Asn Asn Asn Arg Asn Asn Glu Asn Glu Leu
          50              55              60

aaa gaa gaa gga tct ttg ccc act aag atg aac gag aaa aac agt aac 243
Lys Glu Glu Gly Ser Leu Pro Thr Lys Met Asn Glu Lys Asn Ser Asn
          65              70              75

tcc gcg gat aag caa cca aac gac atc tcc cac gac gaa agc aag agc 291
Ser Ala Asp Lys Gln Pro Asn Asp Ile Ser His Asp Glu Ser Lys Ser
          80              85              90

aac agt aac aac gcc caa aac atc caa aag gaa cct gaa gag aag gaa 339
Asn Ser Asn Asn Ala Gln Asn Ile Gln Lys Glu Pro Glu Glu Lys Glu
          95              100              105              110

aac tca aac ccc aac ctc gac tcg agt gaa aac tcc gct gaa agt gct 387
Asn Ser Asn Pro Asn Leu Asp Ser Ser Glu Asn Ser Ala Glu Ser Ala
          115              120              125

act aga agc gtc gac atc agt gaa cac aac tca aac aac ccc gaa act 435
Thr Arg Ser Val Asp Ile Ser Glu His Asn Ser Asn Asn Pro Glu Thr
          130              135              140

aaa gaa gaa aac gga gaa gaa cct cta gac ctg gaa att aac gaa aac 483
Lys Glu Glu Asn Gly Glu Glu Pro Leu Asp Leu Glu Ile Asn Glu Asn
          145              150              155

gca gaa atc ggc cag gaa cct cca aac cgt ctt cac ttc gac aac gtt 531
Ala Glu Ile Gly Gln Glu Pro Pro Asn Arg Leu His Phe Asp Asn Val
          160              165              170

gac gac gaa gta cca cat tac tca gcc ctg agg tac aac aag gtc gag 579
Asp Asp Glu Val Pro His Tyr Ser Ala Leu Arg Tyr Asn Lys Val Glu
          175              180              185              190

aag aac gta acc gac gag atg ctc ttg tac aac atg atg tcc gac caa 627
Lys Asn Val Thr Asp Glu Met Leu Leu Tyr Asn Met Met Ser Asp Gln
          195              200              205

aac cgc aaa agc tgt gct atc aac aac ggt ggc tgc agt gac gac cag 675

```

Asn Arg Lys Ser Cys Ala Ile Asn Asn Gly Gly Cys Ser Asp Asp Gln
 210 215 220

atc tgc atc aac atc aac aac atc ggt gtg aag tgc att tgt aag gat 723
 Ile Cys Ile Asn Ile Asn Asn Ile Gly Val Lys Cys Ile Cys Lys Asp
 225 230 235

gga tac cta ctt ggt acc aag tgc att cac cac cac c 760
 Gly Tyr Leu Leu Gly Thr Lys Cys Ile His His His
 240 245 250

<210> 39
 <211> 250
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 construct

<400> 39
 Met Asn Ile Leu Cys Ile Leu Ser Tyr Ile Tyr Phe Phe Val Ile Phe
 1 5 10 15

Tyr Ser Leu Asn Leu Asn Asn Lys Asn Glu Asn Phe Leu Val Val Arg
 20 25 30

Arg Leu Met Asn Asp Glu Lys Gly Glu Gly Gly Phe Thr Ser Lys Asn
 35 40 45

Lys Glu Asn Gly Asn Asn Asn Arg Asn Asn Glu Asn Glu Leu Lys Glu
 50 55 60

Glu Gly Ser Leu Pro Thr Lys Met Asn Glu Lys Asn Ser Asn Ser Ala
 65 70 75 80

Asp Lys Gln Pro Asn Asp Ile Ser His Asp Glu Ser Lys Ser Asn Ser
 85 90 95

Asn Asn Ala Gln Asn Ile Gln Lys Glu Pro Glu Glu Lys Glu Asn Ser
 100 105 110

Asn Pro Asn Leu Asp Ser Ser Glu Asn Ser Ala Glu Ser Ala Thr Arg
 115 120 125

Ser Val Asp Ile Ser Glu His Asn Ser Asn Asn Pro Glu Thr Lys Glu
 130 135 140

Glu Asn Gly Glu Glu Pro Leu Asp Leu Glu Ile Asn Glu Asn Ala Glu
 145 150 155 160

Ile Gly Gln Glu Pro Pro Asn Arg Leu His Phe Asp Asn Val Asp Asp
 165 170 175

Glu Val Pro His Tyr Ser Ala Leu Arg Tyr Asn Lys Val Glu Lys Asn
 180 185 190

Val Thr Asp Glu Met Leu Leu Tyr Asn Met Met Ser Asp Gln Asn Arg
 195 200 205

Lys Ser Cys Ala Ile Asn Asn Gly Gly Cys Ser Asp Asp Gln Ile Cys
210 215 220

Ile Asn Ile Asn Asn Ile Gly Val Lys Cys Ile Cys Lys Asp Gly Tyr
225 230 235 240

Leu Leu Gly Thr Lys Cys Ile His His His
245 250

<210> 40

<211> 760

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
construct

<400> 40

```
tctgaattct cagtgggtgt ggtgggtgtg aatgcacttg gtaccaagta ggtatccatc 60
cttacaaatg cacttcacac cgatgttggt gatgttgatg cagatctggt cgtcactgca 120
gccaccgttg ttgatagcac agcttttgcg gttttggtcg gacatcatgt tgtacaagag 180
catctcgteg gttacgttct tctcgacctt gttgtacctc agggctgagt aatgtggtag 240
ttcgtcgtea acgttgtcga agtgaagacg gtttgagggt tccctggccga tttctgcgtt 300
ttcgtaaatt tccaggtcta gaggttcttc tccgttttct tctttagttt cggggttggt 360
tgagtgtgtg tcaatgatgt cgacgttctt agtagcactt tcagcggagt tttcactcga 420
gtcgagggtg ggggttgagt tttccttctc ttcaggttcc ttttggtatg tttgggcgtt 480
gttactgttg cttctgcttt cgtcgtggga gatgtcggtt ggttgcctat ccgcggagtt 540
actgttttct tcttctcatc tagtgggcaa agatccttct tctttgagtt cgttctcgtt 600
gttcctgttg ttgtttccgt tttccttggt cttactagtg aagccacctt ctccttttct 660
gtcgttcatg agtctgcgga ccaccaagaa gtctcgtttt ttgttggtga ggtttaaact 720
gtagaagatg acgaagaagt aatgttagct gagaatacag 760
```

<210> 41

<211> 272

<212> PRT

<213> Plasmodium falciparum

<400> 41

Met Trp Ile Val Lys Phe Leu Ile Val Val His Phe Phe Ile Ile Cys
1 5 10 15

Thr Ile Asn Phe Asp Lys Leu Tyr Ile Ser Tyr Ser Tyr Asn Ile Val
20 25 30

Pro Glu Asn Gly Arg Met Leu Asn Met Arg Ile Leu Gly Glu Glu Lys
35 40 45

Pro Asn Val Asp Gly Val Ser Thr Ser Asn Thr Pro Gly Gly Asn Glu
50 55 60

Ser Ser Ser Ala Ser Pro Asn Leu Ser Asp Ala Ala Glu Lys Lys Asp
65 70 75 80

Glu Lys Glu Ala Ser Glu Gln Gly Glu Glu Ser His Lys Lys Glu Asn
85 90 95

Ser Gln Glu Ser Ala Asn Gly Lys Asp Asp Val Lys Glu Glu Lys Lys
100 105 110

Thr Asn Glu Lys Lys Asp Asp Gly Lys Thr Asp Lys Val Gln Glu Lys
115 120 125

Val Leu Glu Lys Ser Pro Lys Glu Ser Gln Met Val Asp Asp Lys Lys
130 135 140

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

Ser Gly Gly His Val Gly Glu Glu Glu Asp His Asn Glu Gly Glu Gly
165 170 175

Glu His Glu Glu Glu Glu His Glu Glu Asp Asp Asp Asp Glu Asp
180 185 190

Asp Asp Thr Tyr Asn Lys Asp Asp Leu Glu Asp Glu Asp Leu Cys Lys
195 200 205

His Asn Asn Gly Gly Cys Gly Asp Asp Lys Leu Cys Glu Tyr Val Gly
210 215 220

Asn Arg Arg Val Lys Cys Lys Cys Lys Glu Gly Tyr Lys Leu Glu Gly
225 230 235 240

Ile Glu Cys Val Glu Leu Leu Ser Leu Ala Ser Ser Ser Leu Asn Leu
245 250 255

Ile Phe Asn Ser Phe Ile Thr Ile Phe Val Val Ile Leu Leu Ile Asn
260 265 270

<210> 42

<211> 230

<212> PRT

<213> Plasmodium falciparum

<400> 42

Met Lys Val Ala Tyr Phe Leu Ser Val Leu Asp Leu Leu Ile Ile Phe
1 5 10 15

Ser Leu Tyr Phe Asp Gly Arg Arg Ser Ala Phe Ala Gly Ile Ala Ala
20 25 30

Cys Ile Arg His Gly Arg Ile Leu Gly Glu Gly Gly Glu Gln Ser Gly
35 40 45

Gly Ala Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Asp Ser Ser Gly
50 55 60

Gly Leu Ser Gly Gly Ser Ser Gly Gly Pro Ser Pro Pro Ala Gly Ser
65 70 75 80

Ser Gly Ser Gly Gly Ser Asp Pro Ala Asn Ser Ala Thr Gly Pro Gln
85 90 95

Asn Ser Thr Pro Gly Ser Gly Gly Gln Thr Gly Asp His Ser Ala Glu

100	105	110
Ala Glu Asn Gly Asp Tyr Asn Glu Gln Gly Asp Asp His Gly Asp Asp 115 120 125		
His Gly Asp Asp His Gly Asp Asp His Gly Asp Glu Gln Asp Gly Glu 130 135 140		
Asp Tyr Asp Asp Ala Glu Asp Asp Asp Leu Tyr Glu Leu Ser Glu Val 145 150 155 160		
Asp Glu Asn Ala Asn Leu Cys Leu Asp Asn Asn Gly Gly Cys Gly Asp 165 170 175		
Asp Lys Ile Cys Glu Asn Leu Gly Lys Gly Ile Val Lys Cys Leu Cys 180 185 190		
Lys Pro Gly Tyr Lys Leu Val Gly Thr Glu Cys Val Glu Ser Ser Lys 195 200 205		
Ser Ser Ser Leu Asn Ser Phe Phe Cys Trp Phe Leu Leu Val Ile Ile 210 215 220		
Val Leu Ala Ser Ile Asn 225 230		

<210> 43

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic peptide

<400> 43

Lys Ser Pro Lys Glu
1 5

<210> 44

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic peptide

<400> 44

Met Arg Ile Leu Gly
1 5

<210> 45

<211> 5

<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 45
Met Arg Ile Leu Ala
1 5

<210> 46
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 46
Met Arg Ile Leu
1

<210> 47
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 47
Gly Gly Phe Thr Ser Lys
1 5

<210> 48
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 48
Gly Ser Leu Pro Thr Lys
1 5

<210> 49
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 49

Gly Ile Ala Ala Cys
1 5

<210> 50

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 50

Glu Gly Gly Glu Gln
1 5

<210> 51

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 51

Gly Asp Ser Ser Gly
1 5

<210> 52

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 52

Asp Ser Ser Gly Gly
1 5

<210> 53

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 53

Ser Ser Gly Gly Leu
1 5

<210> 54
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 54
Leu Asp Asn Asn Gly
1 5

<210> 55
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 55
Met Arg Ile Leu Gly Glu
1 5

<210> 56
<211> 6
<212> PRT
<213> Artificial Sequence

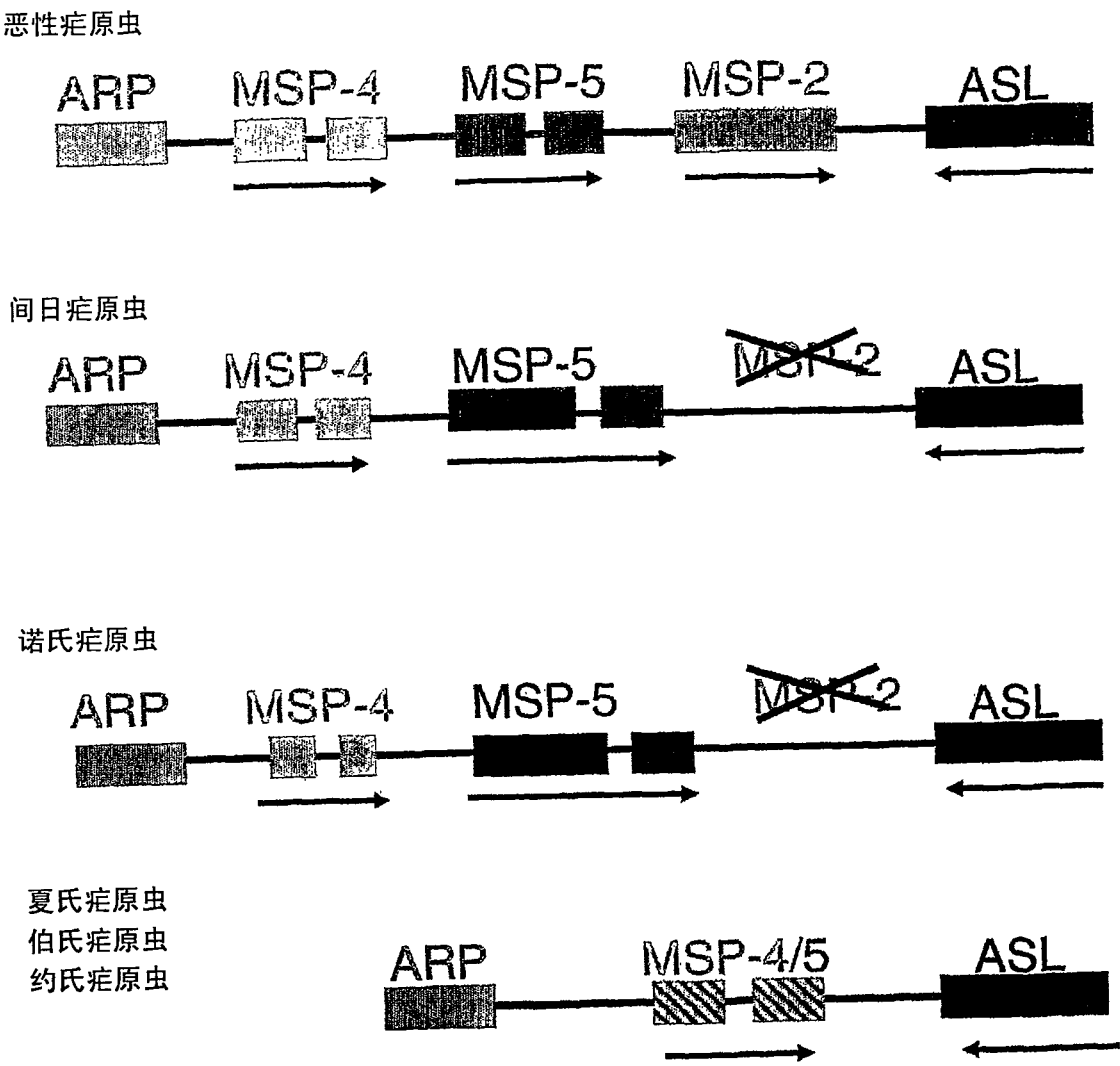
<220>
<223> Description of Artificial Sequence: Synthetic
peptide

<400> 56
Gly Arg Ile Leu Gly Glu
1 5

<210> 57
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic 6x
His tag

<400> 57
His His His His His His
1 5



几种不同疟原虫物种的msp4和msp5基因位点的基因组结构。

图1

SALSA	RVSTSDTPGGN
MSP4	MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLNMRILGEEKPNVDGVSTSDTPGGN
SALSA	ESSSAFPQFIWSAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKTNERKDDGKTDKVQ
MSP4	ESSSASPNSDAAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKTNERKDDGKTDKVQ
SALSA	EKVLEKSPK
MSP4	EKVLEKSPKESQMVDKDKKTEAIPKKVVQPSSSNSG-GHVGEEDHNEGEGEHEEEEEHEEDD
MSP4	DDEDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEGIECVELLSLAS
MSP4	SSDLNLFNSFITIFVVILLIN

SALSA序列与MSP-4有92%的相同性。
MSP4序列以黑色显示，SALSA序列以灰色显示。

图2

NF54和合成基因的MSP4密码子用法

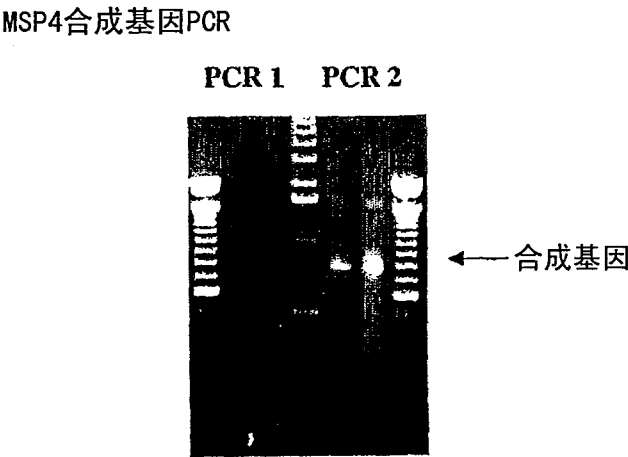
密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5
GCA	Ala	2	4	GAA	Glu	36	25	AAA	Lys	24	5	ACA	Thr	1	3
GCC		0	1	GAG		4	15	AAG		7	26	ACC		1	3
GCG		1	0	总计	40	40	总计	31		31	ACG	0		0	
GCT		3	3	GGA	Gly	11	13	TTC	Phe	0	4	ACT	5	1	
总计		6	8	GGC		1	2	TTT		4	0	总计	7	7	
AGA	Arg	4	0	GGG		3	0	总计		4	4	TGG	Trp	1	1
AGG		0	1	GGT	4	4	CCA	Pro	4	2	总计	1	1		
CGA		0	0	总计	19	19	CCC		1	1	TAC	Tyr	0	6	
CGC		0	1	CAC	His	1	12		CCG	0	0		TAT	6	0
CGG		0	0	CAT		6	1	CCT	2	4	总计	6	6		
CGT	0	2	总计	7	13	总计	7	7	GTA	Val	5	3			
总计	4	4	ATA	Ile	5	1	AGC	Ser	1		3	GTC	0	5	
AAC	Asn	2	17		ATC	2	5		AGT		6	4	GTG	1	3
AAT		15	0		ATT	3	4		TCA		4	1	GTT	10	5
总计		17	17		总计	10	10		TCC		3	1	总计	16	16
GAC		Asp	7	22	CTA	Leu	2	1	TCG	0	1	密码子		p.f.	H5
GAT	18		3	CTC	0		2	TCT	5	7					
总计	25		25	CTG	0		0	总计	19	17					
TGC	Cys		0	4	CTT		0	3	TAA	终止	0				
TGT		7	3	TTA	6	0	TAG	0	0						
总计		7	7	TTG	2	4	TGA	0	0						
CAA	Gln	5	3	总计		10	10	总计	0		1	%GC	30.6	47.5	
CAG		0	2	ATG	Met	4	4								
总计	5	5	总计			4	4								

表1 CODON实行的密码子变化

(下页)

MSP4的全长合成基因序列。该图显示的所有序列以寡核苷酸表示。
重叠寡核苷酸序列编号为1-38，并交替以粗体和正常文本突出显示。
限制性位点以小写文本显示。

图3



合成基因的PCR制备。“基因装配反应”(5或10 μ L, 分别为泳道1和泳道2)和“基因扩增反应”(5或10 μ L, 分别为泳道3和4)的样品在1%琼脂糖凝胶上分离, DNA大小标准品位于样品侧面并将样品分开。

图4

MSP4-EGF/His	MWIVKFLIVVHFFIICTINFDK-----
MSP4p21/His	MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRLNMRILGEEKP--*-----
MSP4p40/His	MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRLNMRILGEEKPNVDGVSTSNTP
MSP4p30/His	MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRLNMRIL-----
MSP4-EGF/His	-----
MSP4p21/His	-----*
MSP4p40/His	GGNEASSASPNLADAAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKKTNEKKDDG
MSP4p30/His	-----AAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKKTNEKKDDG
MSP4-EGF/His	-----
MSP4p21/His	-----LEKSPKESQMVDKTKTEAI PKKVVPSSSSNSGGHVGEEDHNEGEGEHEE
MSP4p40/His	KTDKVQEKVLEKSPKESQMVDKTKTEAI PKKVVPSSSSNSGGHVGEEDHNEGEGEHEE
MSP4p30/His	KTDKVQEKVLEKSPKESQMVDKTKTEAI PKKVVPSSSSNSGGHVGEEDHNEGEGEHEE
MSP4-EGF/His	-----*-----DLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEG
MSP4p21/His	EEEHEEDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEG
MSP4p40/His	EEEHEEDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEG
MSP4p30/His	EEEHEEDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLEG
MSP4-EGF/His	IECVEHHHHHH
MSP4p21/His	IECVEHHHHHH
MSP4p40/His	IECVEHHHHHH
MSP4p30/His	IECVEHHHHHH

图5

3D7和合成基因的MSP5密码子用法

密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5	密码子	aa	p.f	H5
GCA	Ala	2	1	GAA	Glu	23	23	AAA	Lys	15	5	ACA	Thr	2	0
GCC		2	2	GAG		6	6	AAG		3	13	ACC		1	2
GCG		0	1	总计		29	29	总计		18	18	ACG		2	0
GCT		0	3	GGA	Gly	6	5	TTC	Phe	1	6	ACT		1	4
总计		4	7	GGC		0	3	TTT		5	0	总计		6	6
AGA	Arg	6	2	GGG		0	0	总计		6	6	TGG	Trp	0	0
AGG		0	2	GGT		6	4	CCA	Pro	4	3	总计		0	0
CGA		0	0	总计		12	12	CCC		0	3	TAC	Tyr	0	7
CGC		0	2	CAC	His	0	9	CCG		0	0	TAT		7	0
CGG		0	0	CAT		4	1	CCT		5	3	总计		7	7
CGT		1	1	总计		4	10	总计		9	9	GTA	Val	3	2
总计		7	7	ATA	Ile	10	0	AGC	Ser	3	5	GTC		1	4
AAC	Asn	2	44	ATC		1	10	AGT		7	8	GTG		1	2
AAT		42	0	ATT		5	6	TCA		8	3	GTT		4	1
总计		44	44	总计		16	16	TCC		1	4	总计		9	9
GAC	Asp	2	13	CTA	Leu	3	2	TCG		1	1	密码子 247 254			
GAT		13	2	CTC		0	7	TCT		5	1				
总计		15	15	CTG		0	2	总计		25	25	% GC 45.8			
TGC	Cyc	0	4	CTT		2	2	TAA	终止	0	1				
TGT		7	3	TTA		10	1	TAG		0	0				
总计		7	7	TTG		2	3	TGA		0	0				
CAA	Gln	5	4	总计		17	17	总计		0	1				
CAG		1	2	ATG	Met	6	6								
总计		6	6	总计		6	6								

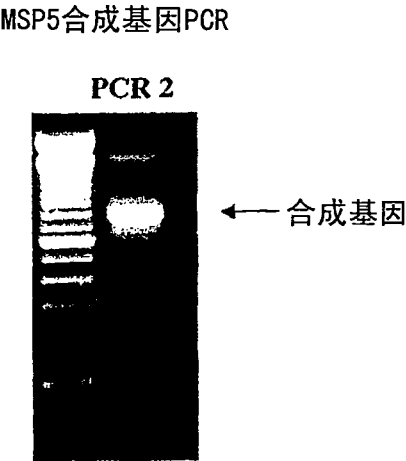
表2 CODOP实行的密码子变化

(下页)

MSP5的全长合成基因序列。该图显示的所有序列以寡核苷酸表示。
重叠寡核苷酸序列编号为1-38，并交替以粗体和正常文本突出显示。
限制性位点以小写文本显示。

图6

[illegible]



合成基因的PCR制备。“基因扩增反应”的样品(5 μ L)在1%琼脂糖凝胶上分离，其旁侧为DNA大小标准品。

图7

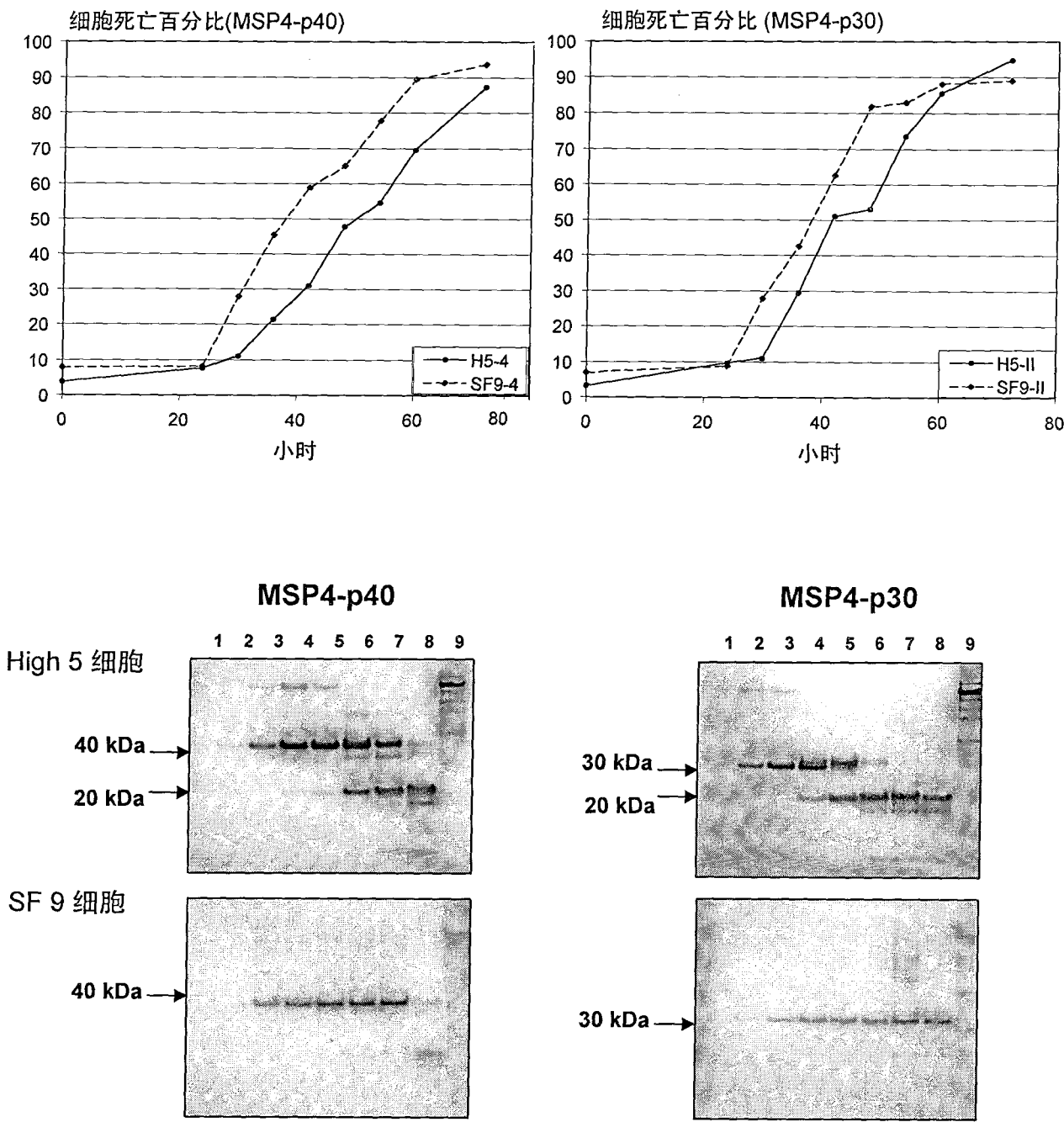


图8

```

MSP4-p40  MWIVKFLIVVHFFIICTINFDRLYISYSYNIVPENGRMLNMRILGEEKPNVDGVSTSTNP
MSP4-p30  MWIVKFLIVVHFFIICTINFDRLYISYSYNIVPENGRMLNMRIL-----

GGNEASSASPNLADAAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKKTNEKKDDG
-----AAEKKDEKEASEQGEESHKKENSQESANGKDDVKEEKKTNEKKDDG

MSP4-p20
KTDKVQEKVLEKSPKESQMVDDKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEGEHEE
KTDKVQEKVLEKSPKESQMVDDKKKTEAIPKKVVQPSSSNSGGHVGEEDHNEGEGEHEE

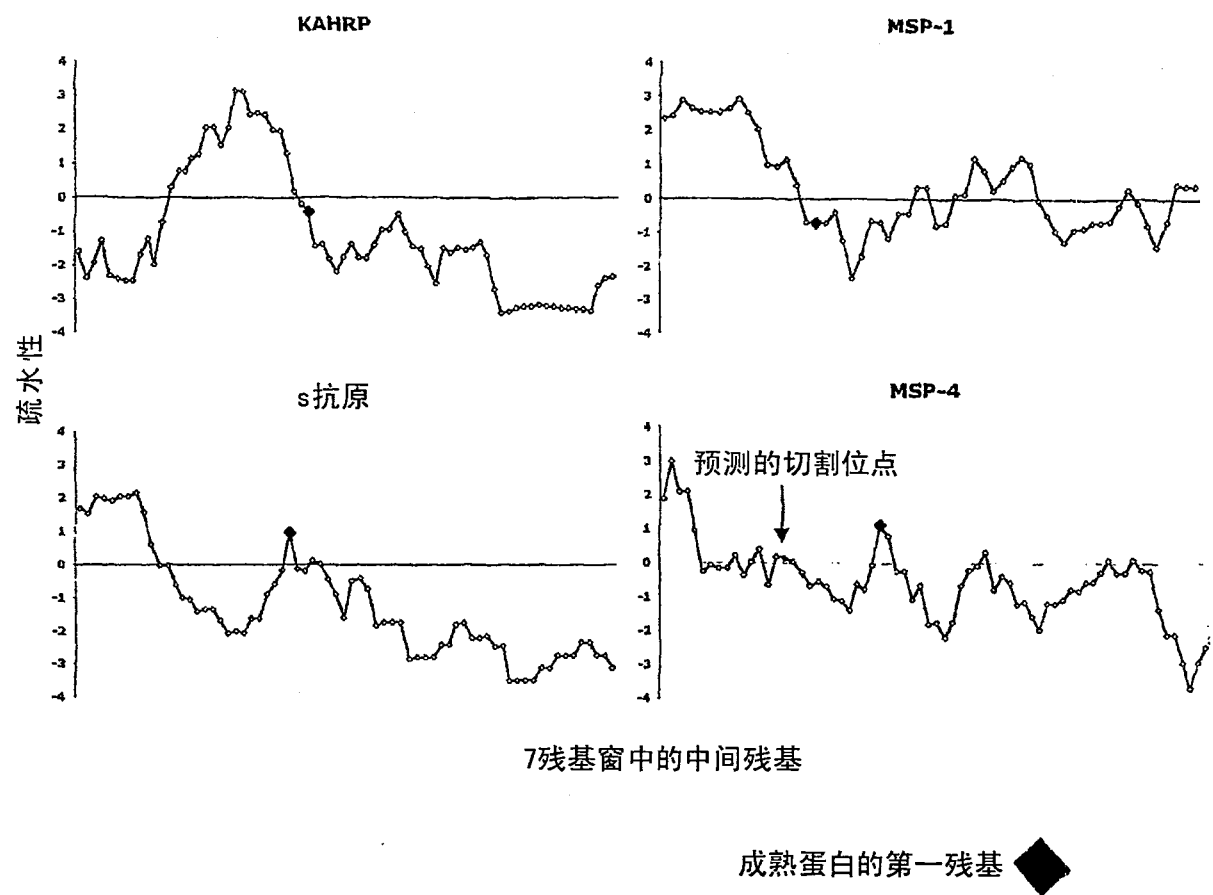
EEEHEEDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKL EYVGNRRVKCKCKEGYKLEG
EEEHEEDDDDDDDTYNKDDLEDEDLCKHNNGGCGDDKL EYVGNRRVKCKCKEGYKLEG

IE VEHHHHHH
IECVHHHHHH

```

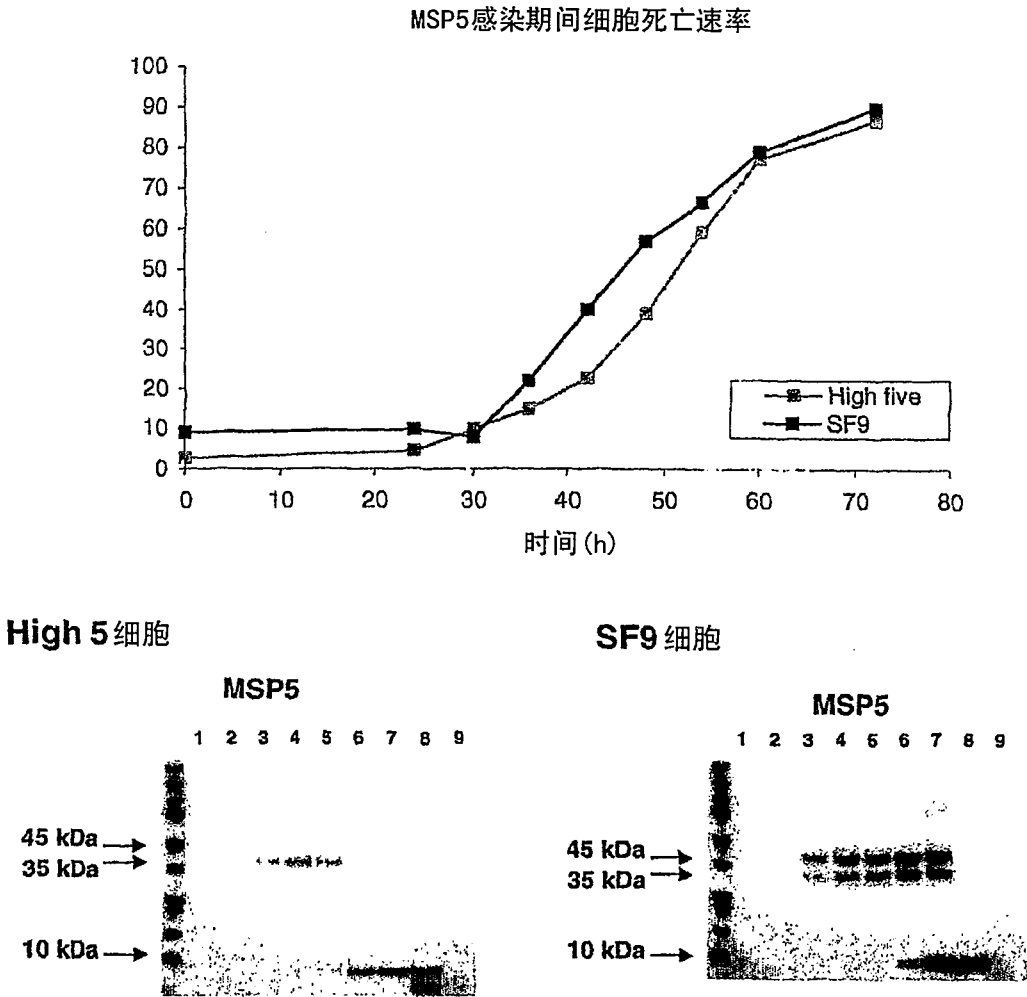
MSP4产物p40, p30和p20的序列比对和N端测序。
 在该研究中鉴定出的信号序列以灰色突出显示。
 各产物的N端序列以粗体显示。

图9



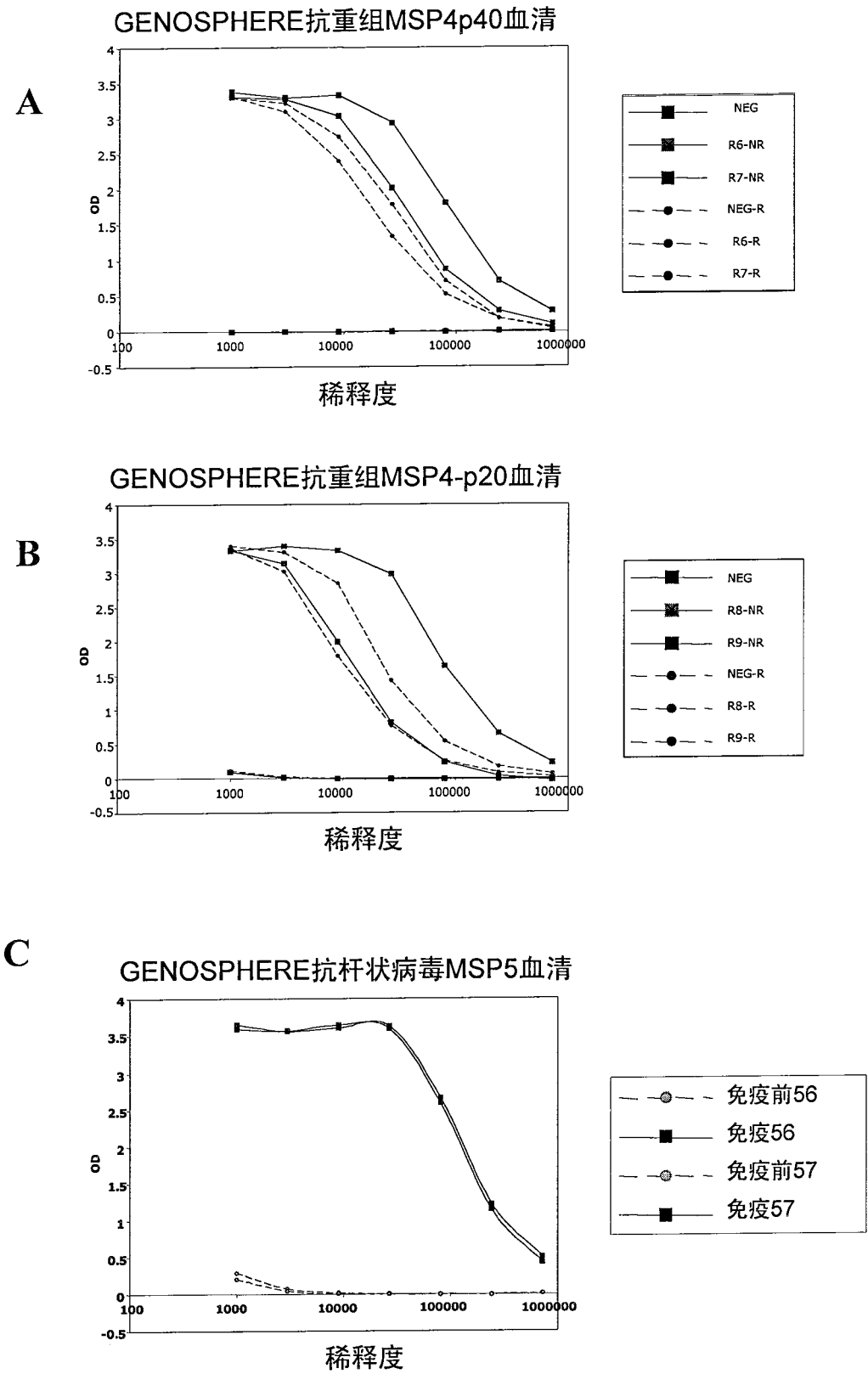
将约80个残基的序列(位于报道的切割位点中心附近)递交到疏水性区域分析,并计算7残基窗的疏水性。
恶性疟原虫的信号序列
Nacer等(2001) Inter. J. Parasitol. 31

图10



PfMSP5在两种不同昆虫细胞系中随时间变化的表达。通过将细胞悬浮液与4%台盼蓝以1: 1混合并计数盖玻片下一方格内的总细胞和蓝色细胞，细胞死亡。感染后24, 30, 36, 42, 48, 54, 60和72 h时(分别为1-8泳道)，从培养SN中纯化蛋白质，20 μ L的洗脱蛋白以4-12% Bis-Tris凝胶(Invitrogen)分离并以SimplyBlue safe stain (Invitrogen)染色。

图 11



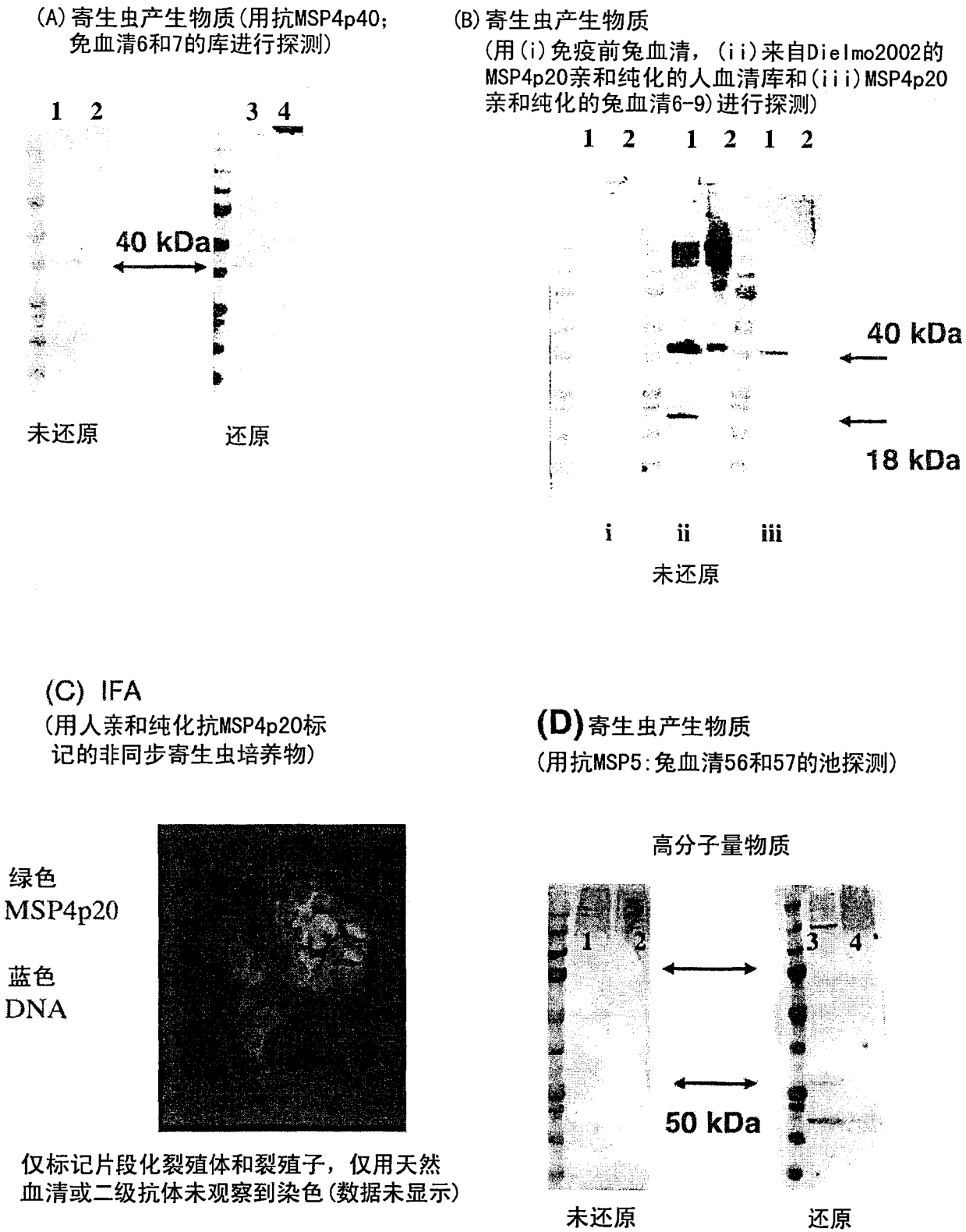
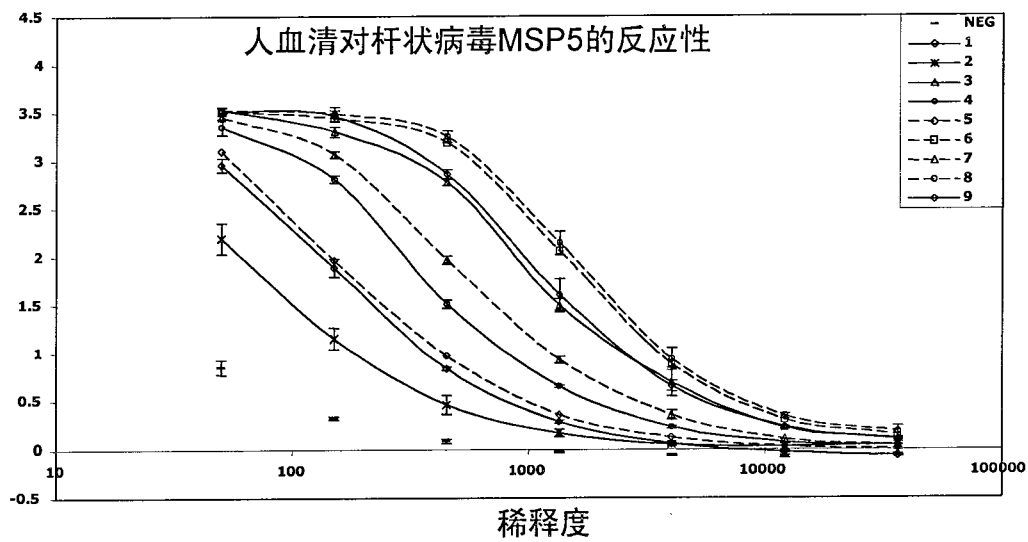
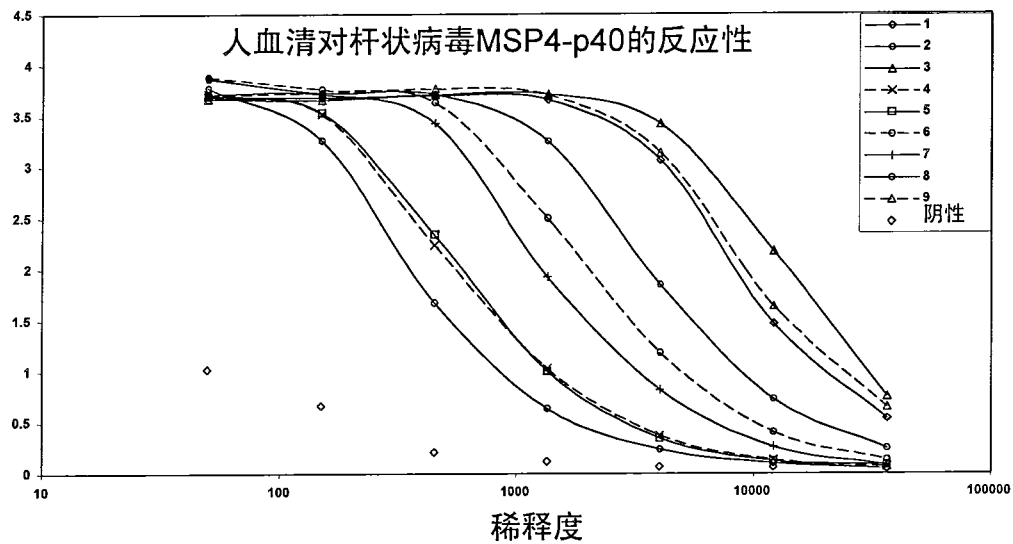


图13



通过ELISA观察采集自1990年雨季肆虐期间
Dielmo的9个免疫成年人(18-49岁)的血清。

图14

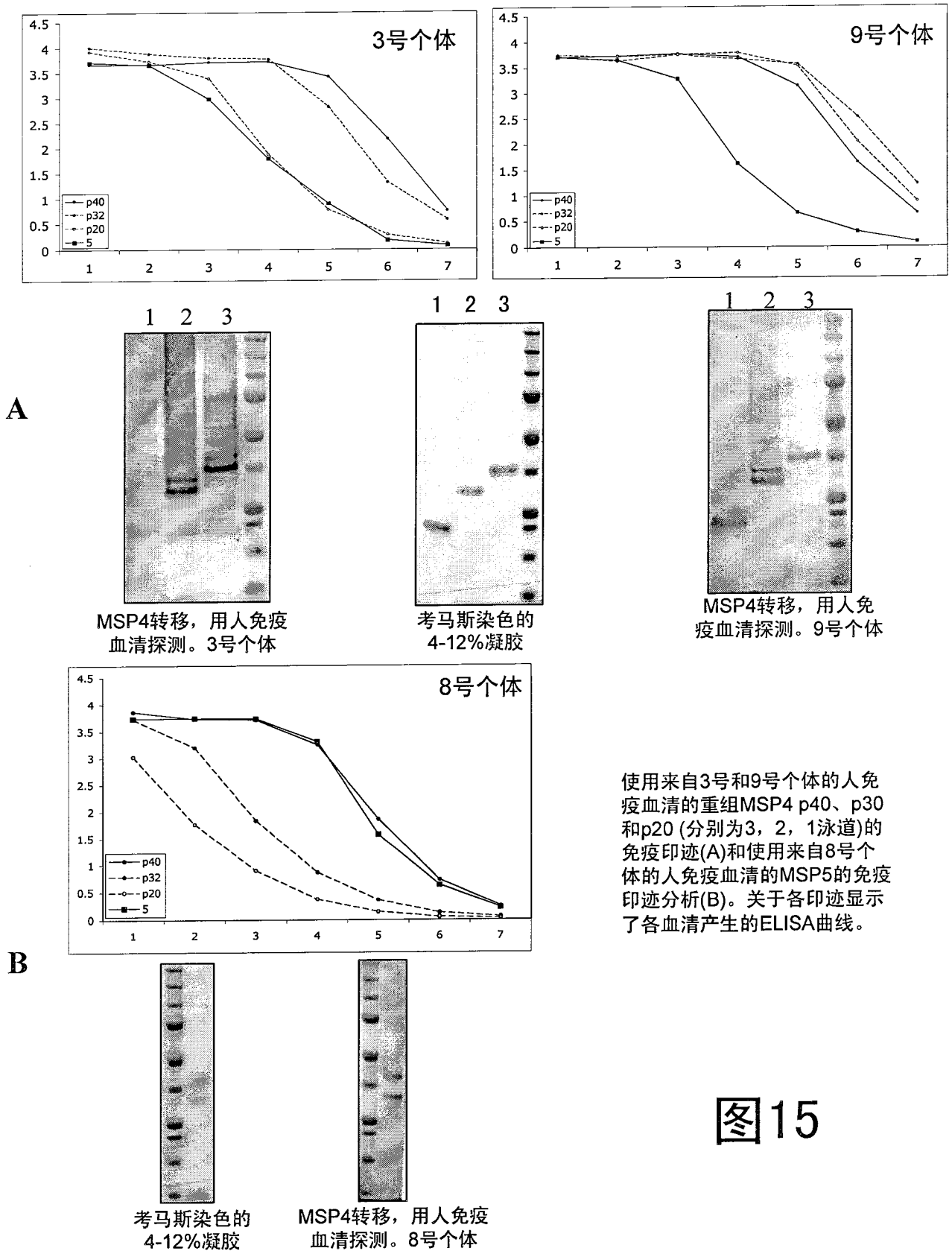
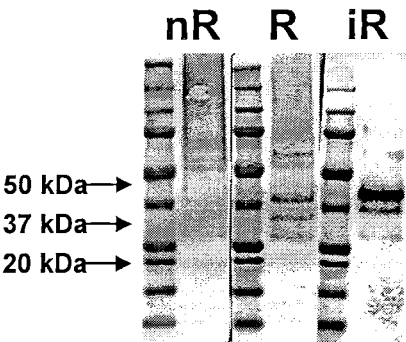
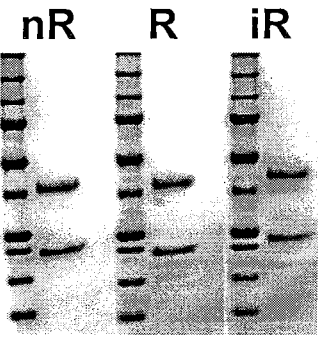


图15



蛋白印迹
PfMSP4 mAb L11-16

图16A



蛋白印迹
PfMSP4 mAb F12-7

图16B

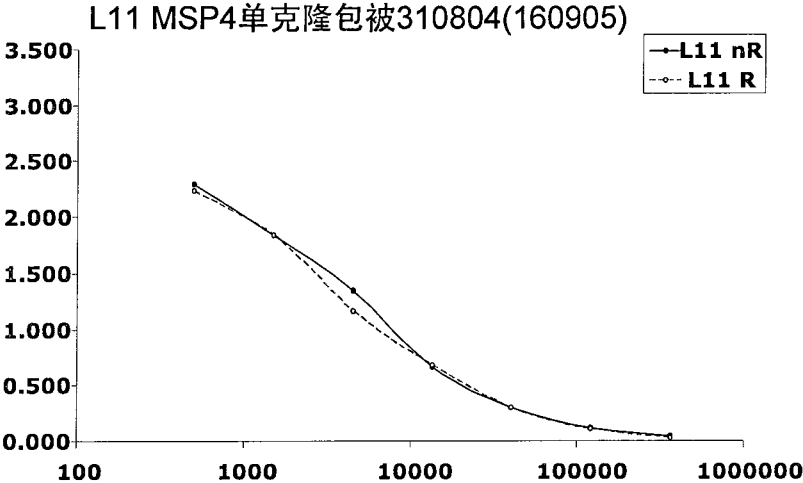


图17A

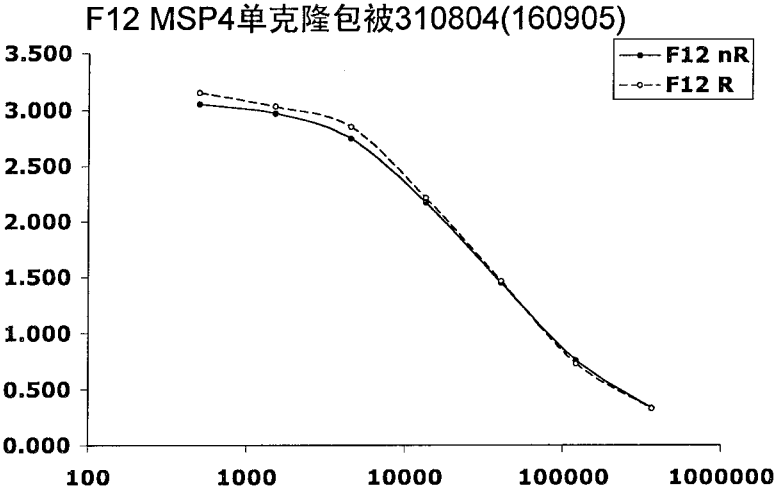


图17B

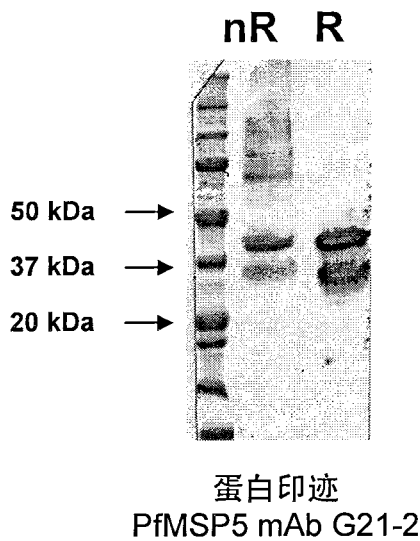


图18A

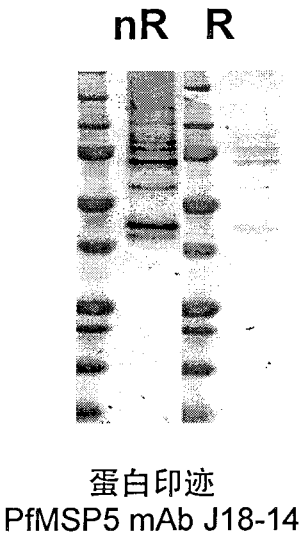


图18B

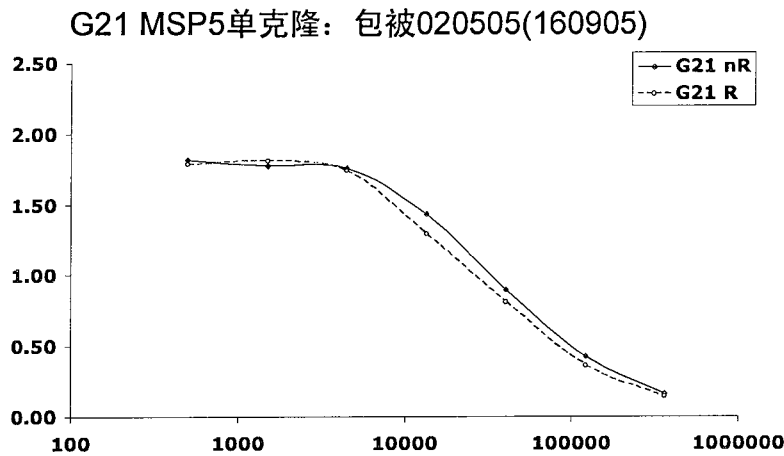


图19A

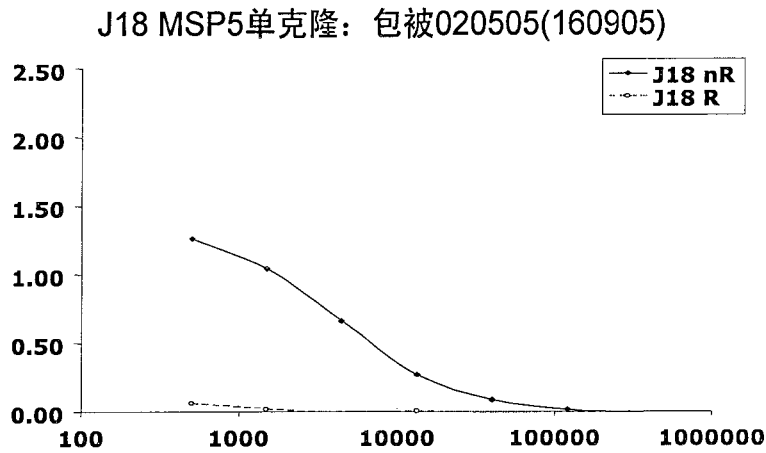


图19B

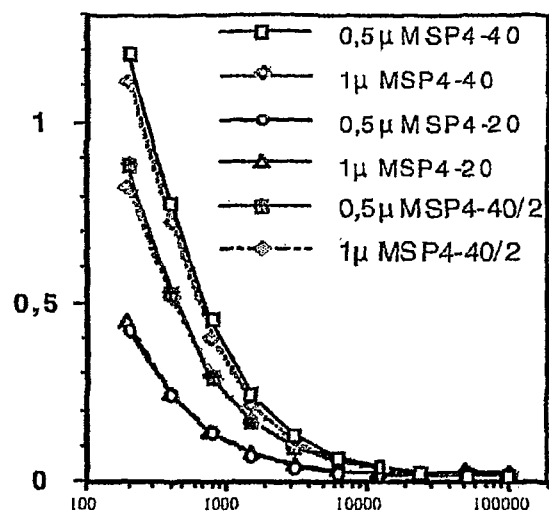


图20A

SHI的稀释

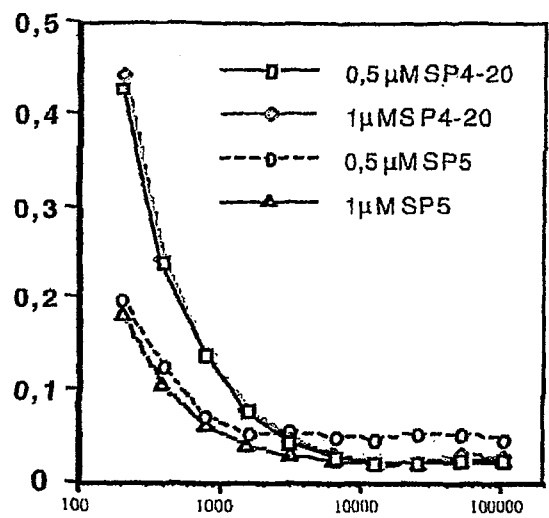


图20B

SHI 滴定 (以0.5 μ g/ml 包被)

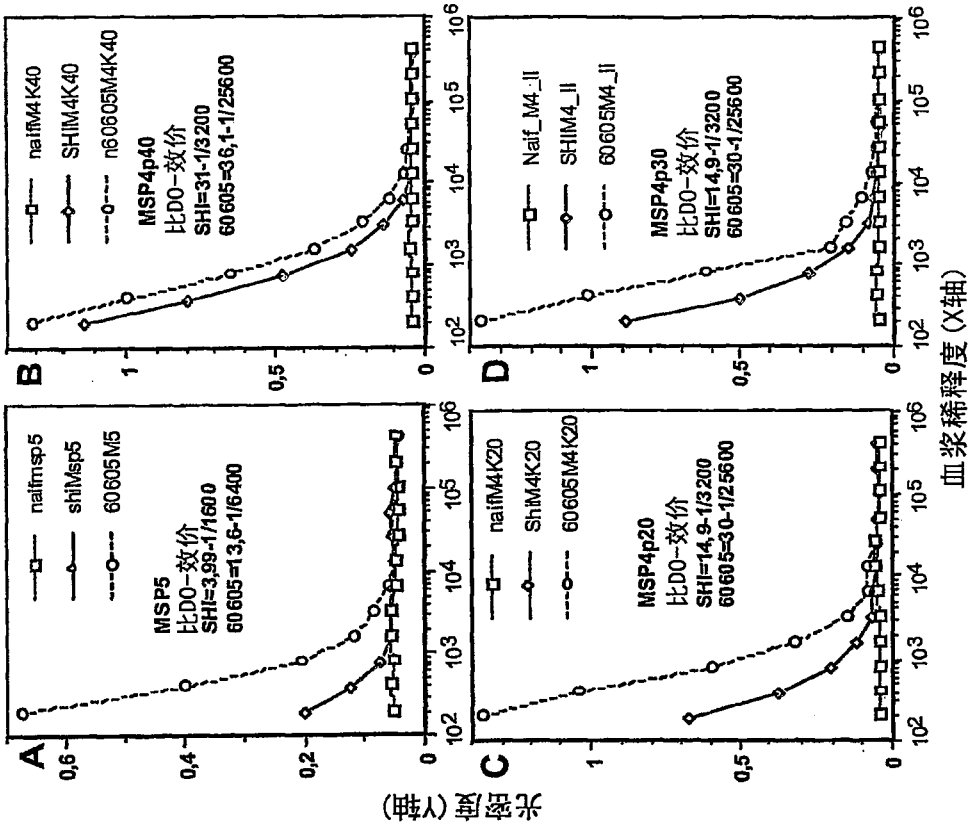


图21

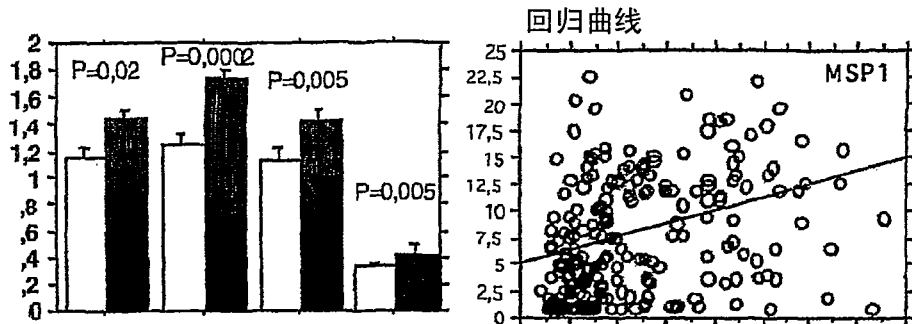


图22A

图22B

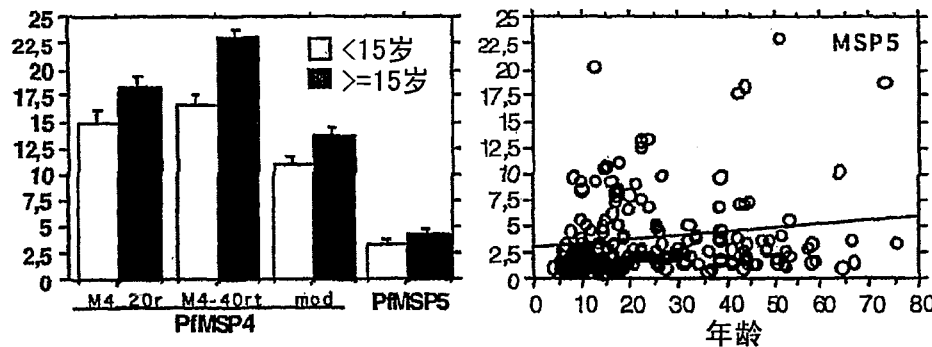


图22C

图22D

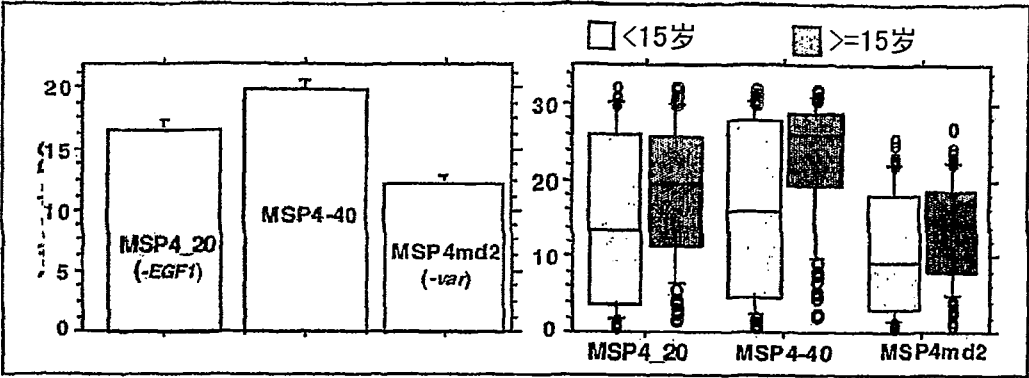


图23

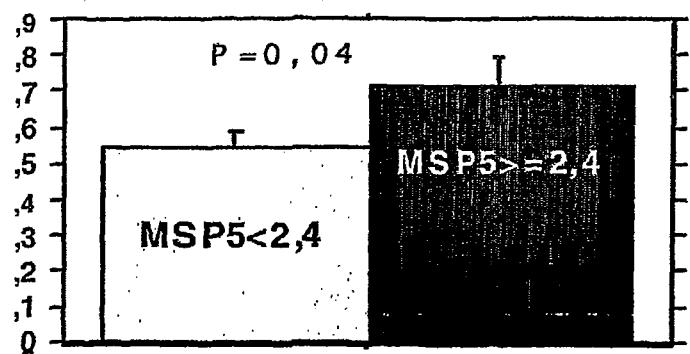


图 24A

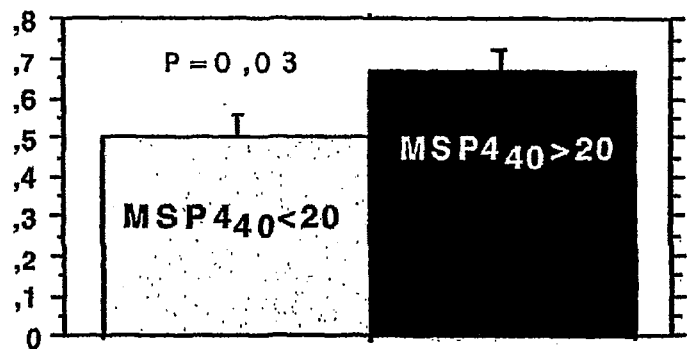


图 24B

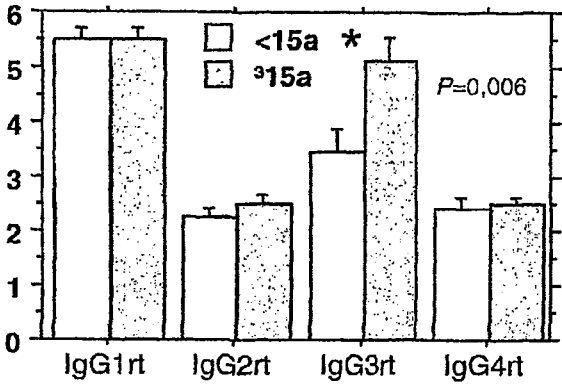


图25A

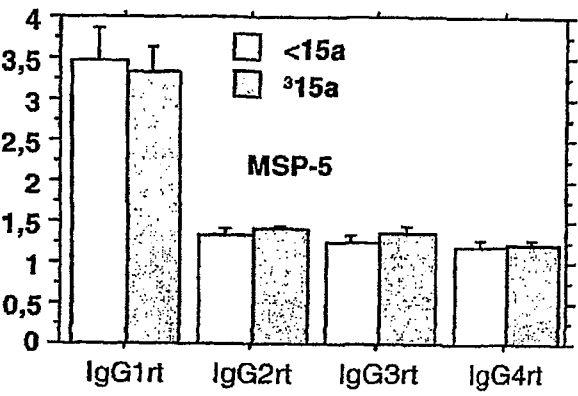


图25B

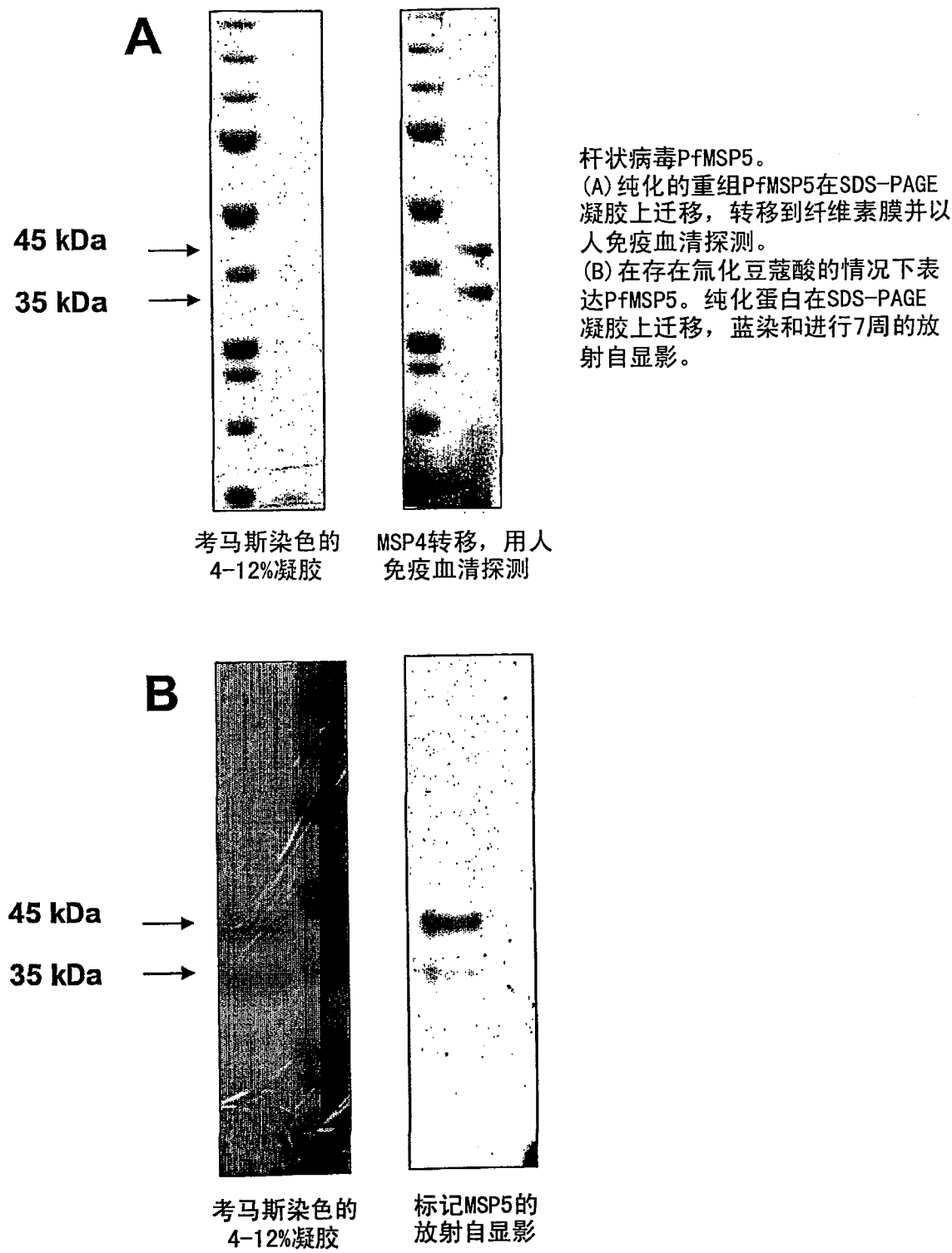


图27

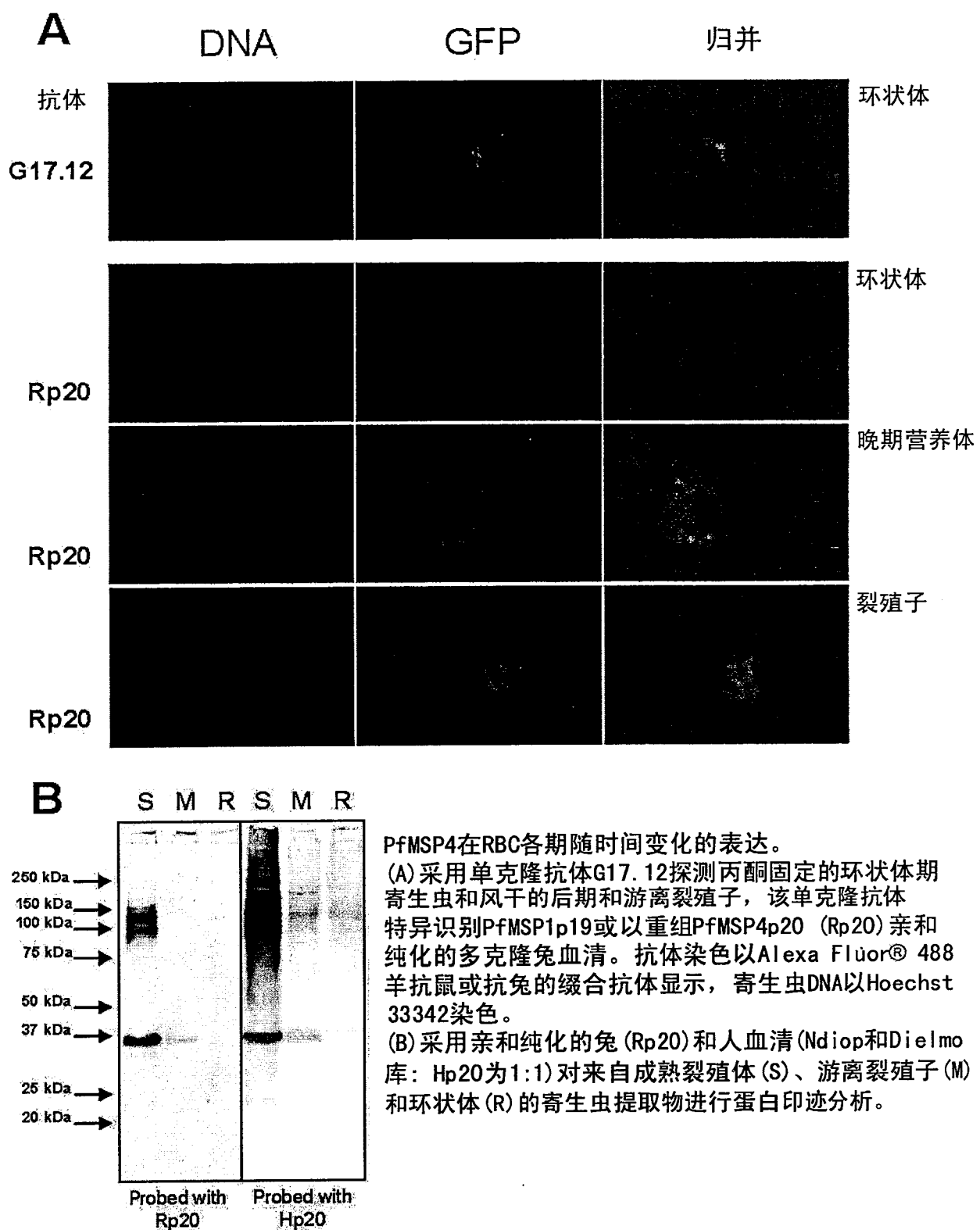
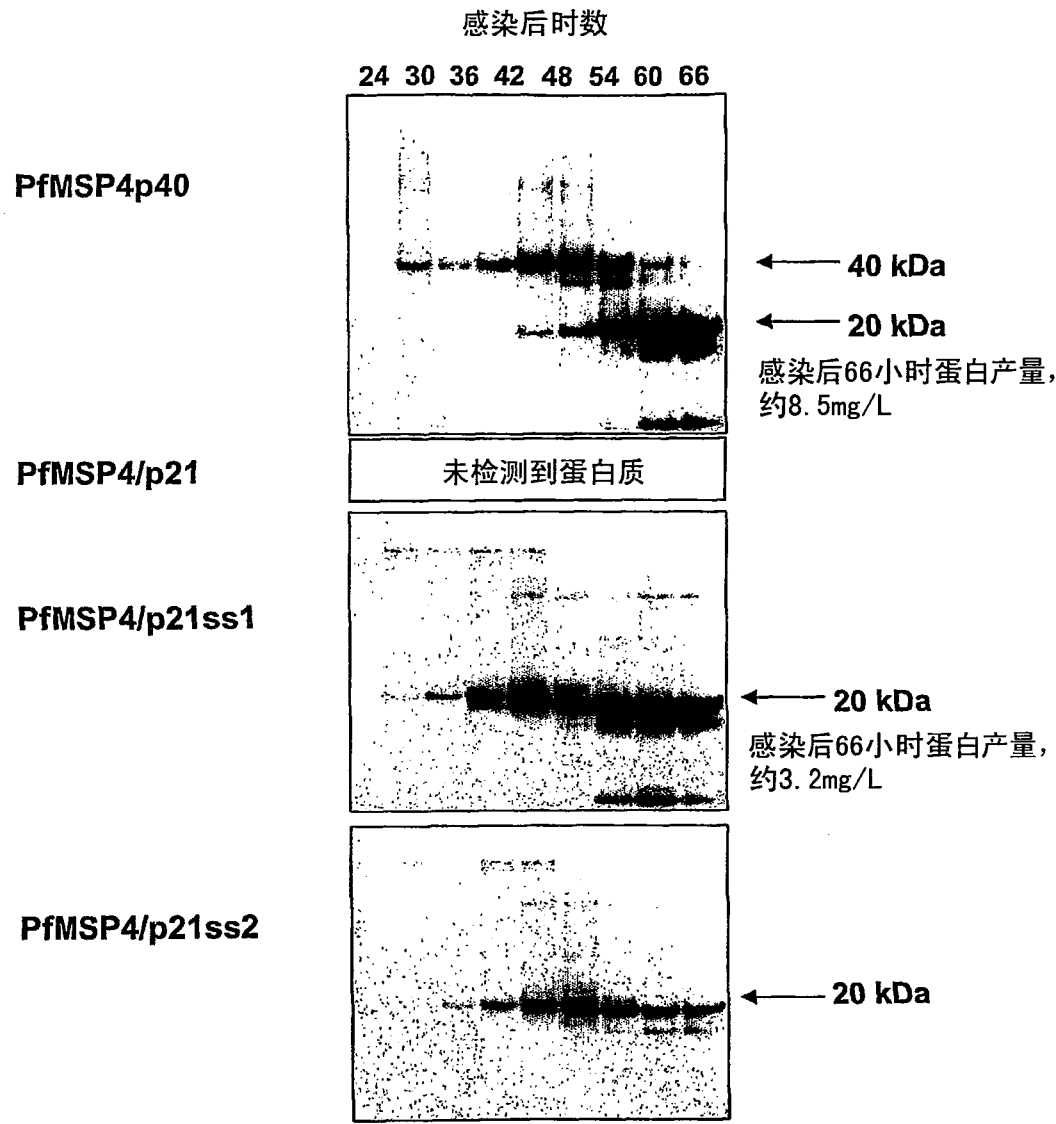


图28



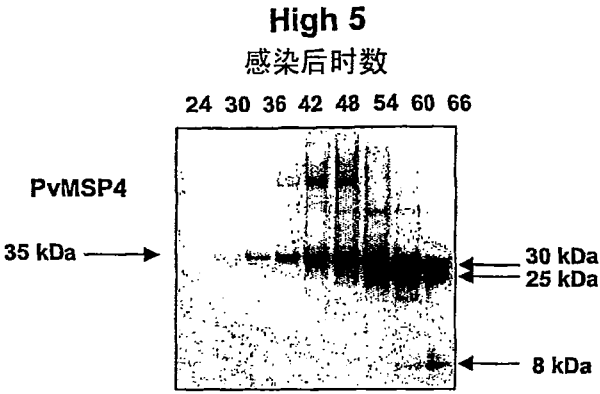
PfMSP4p20随时间变化的直接表达。培养上清的样品在感染后24h到66h之间每隔6h进行收集，透析并以TALON树脂分批纯化。纯化蛋白样品在NuPAGE4-12%梯度凝胶上分离，蓝染，并在右边标注蛋白的大小。采用感染后66h时剩下的最后130 mL培养上清计算大致蛋白产量。以IMAC和HPLC纯化蛋白质，并采用BCA蛋白定量试剂盒(PIERCE)计算该蛋白质产量。

图30

PfMSP4p40	MWIVKFLIVVHFFIICTINFDKLYISYSYNIVPENGRMLN MRILGEEKPNVDGVSTSNTP
PvMSP4/His	MKVAYFLSVLDLLIIFSLYFDGRRSAFAG--- IAACIRHGRILGEGGEQSGGASGGSSG
	* :. ** *:.:** :. ** :.: :. : ***** : .*.* ..:
PfMSP4p40	GGNEASSASPNLADAAEK KDEKEASEQGEESHKKENSQESANGKDDVKEEKTNEKKDDG
PvMSP4/His	GSS GDSGG ----- LSGGSSGGPSP PAGSSGGSGSDPANSATGPQNSTPG
	*.. **.. . *... . : .*..* .* :. . :.: *
PfMSP4p40	KTDKVQEKVLE KSPKES QMVDK KKTEA IPKKVVQPSSNSGGHVGEEDHNEGEGEHEE
PvMSP4/His	SGGQTGDHSAEAENG-----YNEQGDDHGDDHGDDHG-DDHGD
	. :.: :. * . : .:.*.. *:.....* .:* :
PfMSP4p40	EEEHEEDDDDDDDTYNKDDLED-EDLCKHNNGGCGDDKLCEYVGNRRVKCKCKEGYKLE
PvMSP4/His	EQDGEDYDDAEDDDLYELSEVDENANL CLDNNGGCGDDKI CENLGKGI VKCLCKPGYKLV
	*.: *: ** ***** *: :.: :. :* .*****:.* :*: *** ** *****
PfMSP4p40	GIECVE HHHHHH
PvMSP4/His	GTECVESHHHHHH
	* **** *****

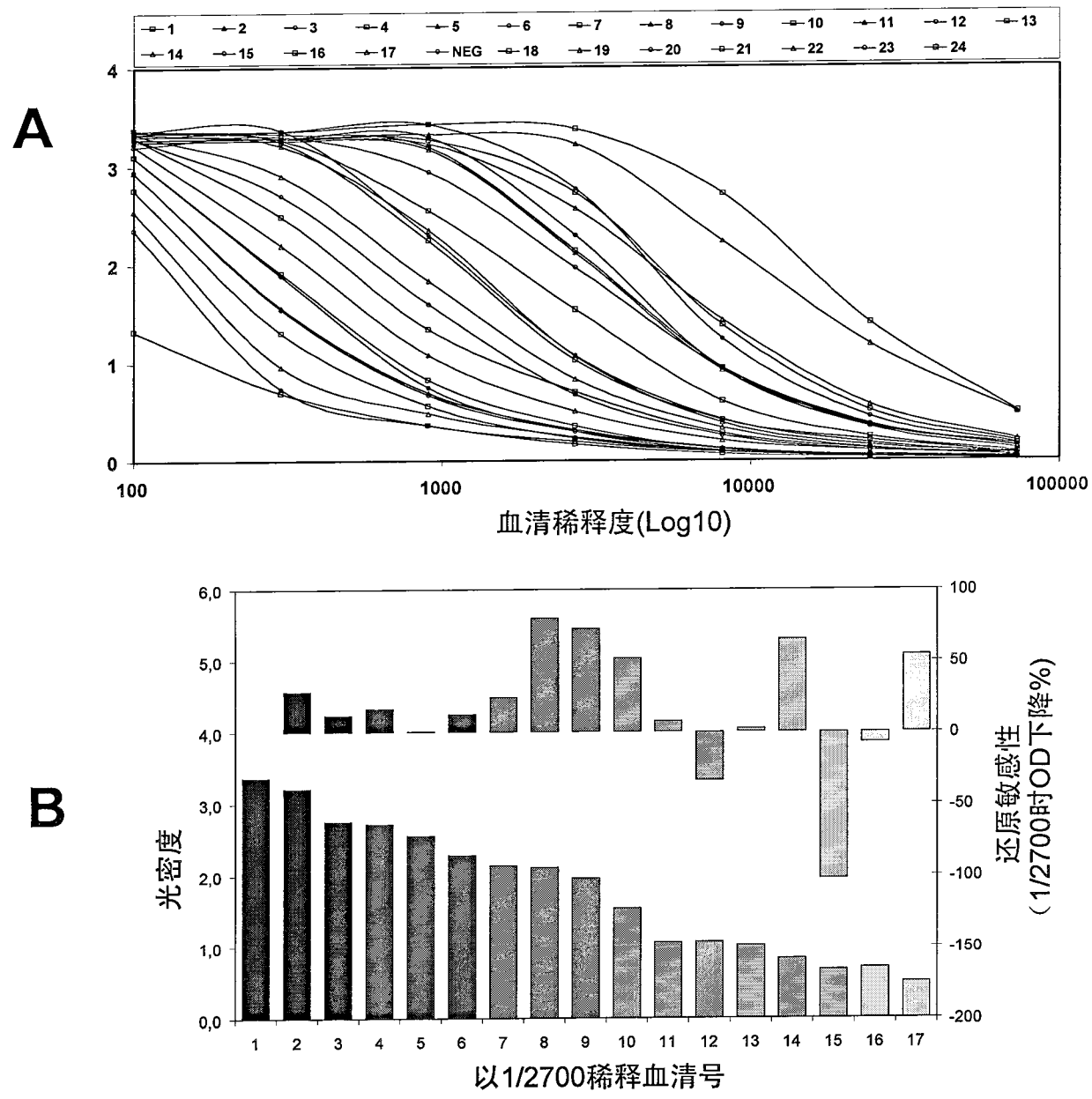
PfMSP4p40和PvMSP4/His的蛋白序列比对。杆状病毒间日疟原虫MSP4和恶性疟原虫MSP4表达构建体的氨基酸(单字母符号)序列以clustalx进行比对。氨基酸相同性表示为(*),保守置换表示为(:),半保守置换表示为(.)和根本性改变表示为空格。鉴定的N端序列以下划线和粗体显示。不同蛋白质产物的N端序列以下划线和粗体显示。

图31



PvMSP4/His随时间变化的表达。简而言之，感染后24到66 h之间每隔6h收集8ml 旋动培养SN样品，透析，然后在TALON树脂上分批纯化。纯化蛋白样品在NuPAGE 4-12%梯度凝胶上分离，以SimplyBlue SafeStain染色，并在左边和右边标注蛋白质的大小。

图32



人免疫血清对PvMSP4/His的反应性。(A)ELISA培养板以杆状病毒表达系统中表达的PvMSP4/His包被。检测了3倍稀释系列或全部24份血清。(B)与此同时，以不可逆还原的PvMSP4检测每个稀释系列。该图在每个血清的下方图中显示了稀释为1/2700时的OD，得到高于阴性对照天然抗原的OD。上方图显示了血清以1/2700血清稀释时抗原还原所观察到的OD下降。

图33

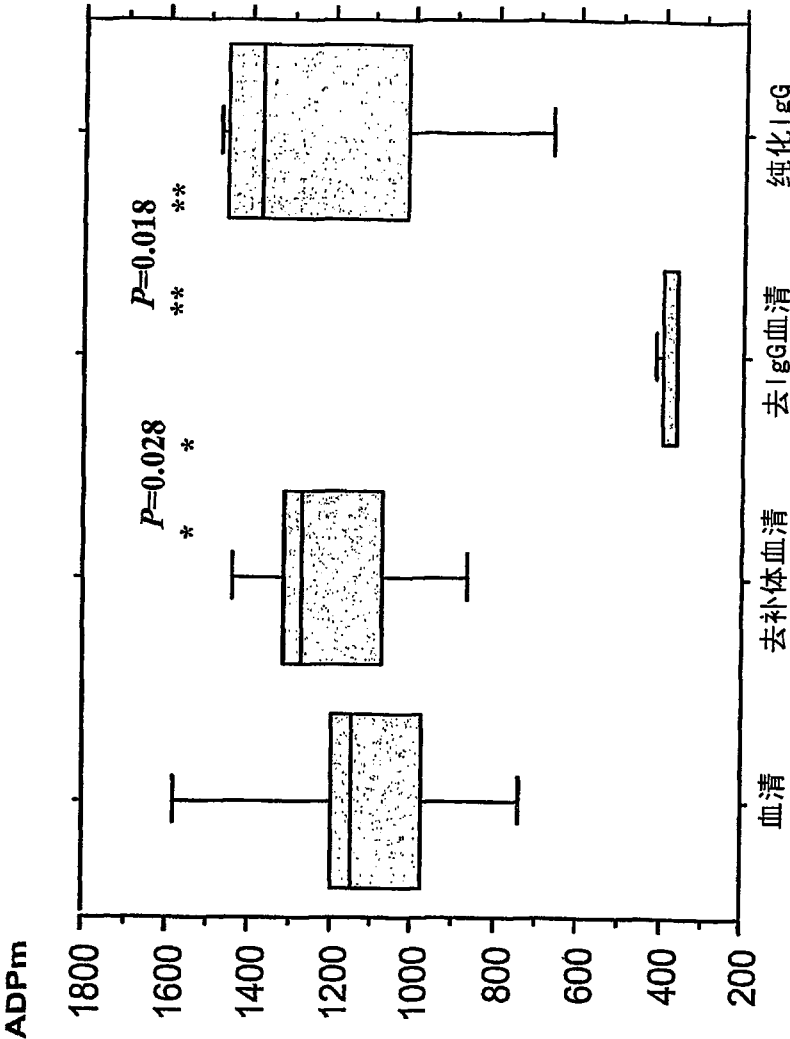


图34

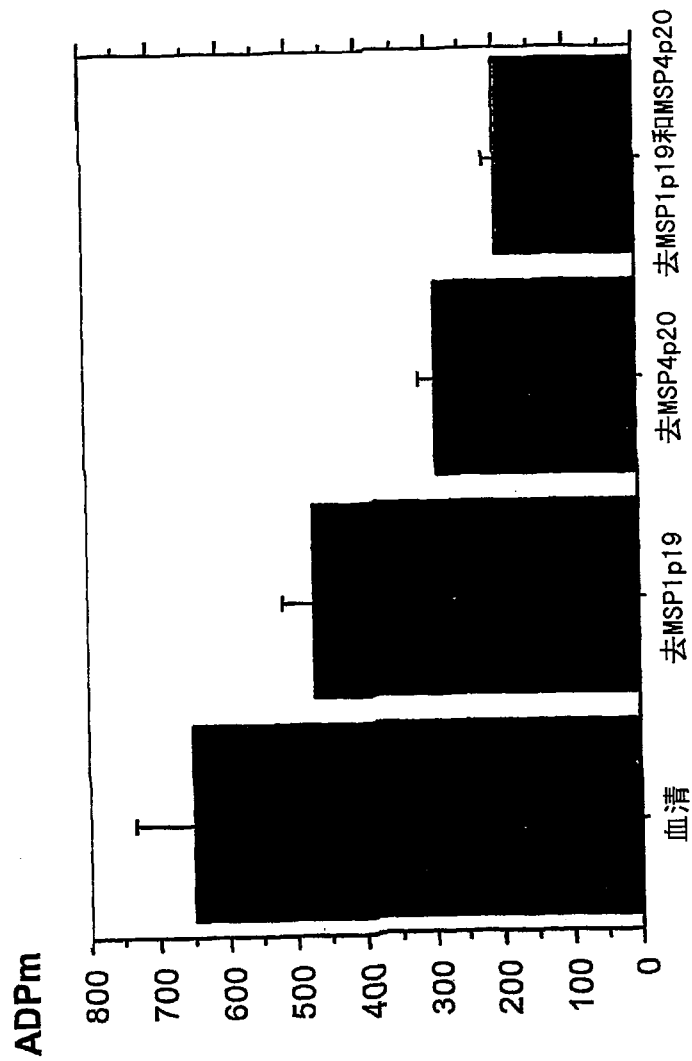


图35