

(12) 发明专利

(10) 授权公告号 CN 1694971 B

(45) 授权公告日 2010.05.26

(21) 申请号 03824965.0

(22) 申请日 2003.09.08

(30) 优先权数据

60/409,515 2002.09.09 US

(85) PCT申请进入国家阶段日

2005.05.08

(86) PCT申请的申请数据

PCT/US2003/028071 2003.09.08

(87) PCT申请的公布数据

W02004/021871 EN 2004.03.18

(73) 专利权人 诺华疫苗和诊断公司

地址 美国特拉华州

(72) 发明人 P·阿坎吉尔 D·齐恩

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 徐迅

(51) Int. Cl.

C12Q 1/70 (2006.01)

G01N 33/576 (2006.01)

(56) 对比文件

W0 0196870 A2, 2001.12.20, 说明书第7、
23、29、33-36 行.

US 6306579 B1, 2001.10.23, 摘要、第4 栏
16-25 行.

US 6428792 B1, 2002.06.08, 附图1、第2-4
栏.

Chien D. Y. et al., Use of a novel
hepatitis C virus (HCV) major-epitope chim
eric polypeptide..... Journal of Clinical
Microbiology vol. 37 no. 5. 1999, 摘要、图1、
1395 页左栏.

Chien D. Y. et al. Diagnosis of
hepatitis C virus (HCV) infection
using an immunodominant chimeric
polyprotein to..... Proc. Natl. Acad. Sci.
USA vol. 89. 1992, 89 摘要、图1、10012 页.

审查员 田园

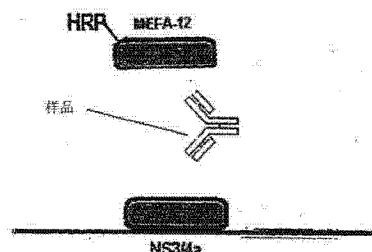
权利要求书 3 页 说明书 51 页 附图 22 页

(54) 发明名称

HCV 检测

(57) 摘要

提供一种 HCV 抗原 / 抗体 / 抗原检测。该检测使用得自 HCV 多蛋白区域的分离的第一抗原，和含有得自与该第一抗原相同的多蛋白区域的表位的 HCV 多表位融合抗原。该第一抗原和该多表位融合抗原都与受 HCV 感染的样品中存在的抗体结合。



1. 一种用于检测生物样品中丙型肝炎病毒 (HCV) 感染的试剂盒, 其特征在于, 所述试剂盒包括:

(a) 免疫检测固相支持物, 该支持物含有结合于其上的 HCV 抗原, 其中, 所述 HCV 抗原由一种或多种得自 HCV 多蛋白的第一区域的分离抗原组成;

(b) 可检测的标记 HCV 多表位融合抗原 (MEFA), 其中, 该标记的 MEFA 含有至少一个得自与所述一个或多个分离的抗原相同的 HCV 多蛋白区域的表位, 其中, 所述 MEFA 能与存在于生物样品中的与该 HCV 抗原结合的 HCV 抗体结合。

2. 如权利要求 1 所述的试剂盒, 其特征在于, 所述得自 HCV 多蛋白的第一区域的一种或多种分离的抗原是一种或多种分离的 NS3/4a 构象表位, 所述 MEFA 含有至少一种得自 NS3/4a 区域的表位。

3. 如权利要求 2 所述的试剂盒, 其特征在于, 所述 NS3/4a 构象表位含有得自 HCV 多蛋白的 NS3/4a 蛋白酶区域的表位。

4. 如权利要求 2 所述的试剂盒, 其特征在于, 所述 NS3/4a 构象表位含有得自 HCV 多蛋白的 NS3/4a 解旋酶区域的表位。

5. 如权利要求 2 所述的试剂盒, 其特征在于, 所述 NS3/4a 构象表位含有 SEQ ID NO:2 所示的氨基酸序列。

6. 如权利要求 1-5 中任一项所述的试剂盒, 其特征在于, 所述 MEFA 含有得自 HCV 多蛋白的 NS3/4a 蛋白酶区域的表位。

7. 如权利要求 1-5 中任一项所述的试剂盒, 其特征在于, 所述 MEFA 含有得自 HCV 多蛋白的 NS3/4a 解旋酶区域的表位。

8. 如权利要求 7 所述的试剂盒, 其特征在于, 所述 MEFA 含有氨基酸 1193-1657, 相对于 HCV-1 序列编号。

9. 如权利要求 1-5 中任一项所述的试剂盒, 其特征在于, 所述 MEFA 含有得自 HCV 多蛋白的 c33c 区域的表位。

10. 如权利要求 9 所述的试剂盒, 其特征在于, 所述 MEFA 含有氨基酸 1211-1457, 相对于 HCV-1 编号。

11. 如权利要求 9 所述的试剂盒, 其特征在于, 所述 MEFA 含有氨基酸 1192-1457, 相对于 HCV-1 编号。

12. 如权利要求 1-5 中任一项所述的试剂盒, 其特征在于, 所述 MEFA 含有得自 HCV 多蛋白的 5-1-1 区域的表位。

13. 如权利要求 2 所述的试剂盒, 其特征在于, 所述 MEFA 含有氨基酸 1689-1735, 相对于 HCV-1 编号。

14. 如权利要求 1-5 中任一项所述的试剂盒, 其特征在于, 所述 MEFA 含有 SEQ ID NO:4 所示的氨基酸序列。

15. 如权利要求 1-5 中任一项所述的试剂盒, 其特征在于, 所述 MEFA 含有 SEQ ID NO:6 所示的氨基酸序列。

16. 一种用于检测生物样品中丙型肝炎病毒 (HCV) 感染的试剂盒, 其特征在于, 所述试剂盒包括:

(a) 免疫检测固相支持物, 该支持物包含结合于其上的 HCV 抗原, 其中, 该 HCV 抗原由一

种或多种多表位融合抗原 (MEFA) 组成；

(b) 可检测的标记的分离 HCV 抗原,它得自存在于所述一种或多种 MEFA 中的 HCV 多蛋白区域,其中所述分离 HCV 抗原能与存在于所述生物样品中的与所述 MEFA 结合的 HCV 抗体结合。

17. 如权利要求 16 所述的试剂盒,其特征在于,所述分离的 HCV 抗原是分离的 NS3/4a 构象表位,所述 MEFA 含有至少一种得自 NS3/4a 区域的表位。

18. 如权利要求 17 所述的试剂盒,其特征在于,所述 NS3/4a 构象表位含有得自 HCV 多蛋白的 NS3/4a 蛋白酶区域的表位。

19. 如权利要求 17 所述的试剂盒,其特征在于,所述 NS3/4a 构象表位含有得自 HCV 多蛋白的 NS3/4a 解旋酶区域的表位。

20. 如权利要求 17 所述的试剂盒,其特征在于,所述 NS3/4a 构象表位含有 SEQ ID NO : 2 所示的氨基酸序列。

21. 如权利要求 16-20 中任一项所述的试剂盒,其特征在于,所述 MEFA 含有得自 HCV 多蛋白的 NS3/4a 蛋白酶区域的表位。

22. 如权利要求 16-20 中任一项所述的试剂盒,其特征在于,所述 MEFA 含有得自 HCV 多蛋白的 NS3/4a 解旋酶区域的表位。

23. 如权利要求 22 所述的试剂盒,其特征在于,所述 MEFA 含有氨基酸 1193-1657,相对于 HCV-1 序列编号。

24. 如权利要求 16-20 中任一项所述的试剂盒,其特征在于,所述 MEFA 含有得自 HCV 多蛋白的 c33c 区域的表位。

25. 如权利要求 24 所述的试剂盒,其特征在于,所述 MEFA 含有氨基酸 1211-1457,相对于 HCV-1 编号。

26. 如权利要求 24 所述的试剂盒,其特征在于,所述 MEFA 含有氨基酸 1192-1457,相对于 HCV-1 编号。

27. 如权利要求 16-20 中任一项所述的试剂盒,其特征在于,所述 MEFA 含有得自 HCV 多蛋白的 5-1-1 区域的表位。

28. 如权利要求 27 所述的试剂盒,其特征在于,所述 MEFA 含有氨基酸 1689-1735,相对于 HCV-1 编号。

29. 如权利要求 16-20 中任一项所述的试剂盒,其特征在于,所述 MEFA 含有 SEQ ID NO : 4 所示的氨基酸序列。

30. 如权利要求 16-20 中任一项所述的试剂盒,其特征在于,所述 MEFA 含有 SEQ ID NO : 6 所示的氨基酸序列。

31. 含有结合于其上的 HCV 抗原的免疫检测固相支持物和可检测地标记的 HCV MEFA 在制备权利要求 1-15 中任一项所述的检测生物样品中的 HCV 感染用的试剂盒中的用途,其特征在于,所述 HCV 抗原由一种或多种得自 HCV 多蛋白的第一区域的分离的抗原组成,所述标记的 MEFA 含有至少一种得自与所述一种或多种分离的抗原相同的 HCV 多蛋白区域的表位。

32. 含有结合于其上的 HCV 抗原的免疫检测圆相支持物和可检测地标记的分离 HCV 抗原在制备权利要求 16-30 任一项所述的检测生物样品中的 HCV 感染用的试剂盒中的用途,其特征在于,所述 HCV 抗原由一种或多种 MEFA 组成,所述标记的分离 HCV 抗原 HCV 得自存

在于所述一种或多种 MEFA 中的 HCV 多蛋白区域。

HCV 检测

技术领域

[0001] 本发明总的涉及病毒诊断。具体而言,本发明涉及使用了衍生自丙型肝炎病毒多蛋白区域的第一分离抗原和包含得自 HCV 多蛋白的多个表位的多表位融合抗原的抗原/抗体/抗原夹心检测,用于准确诊断丙型肝炎病毒感染,其中,所述多个表位包括一种或多种得自与所述第一抗原相同的多蛋白区域的表位。

[0002] 发明背景

[0003] 丙肝病毒 (HCV) 是胃肠外非甲型非乙型肝炎 (NANBH) 的主要原因,它主要通过输血和性接触传播。该病毒存在于 0.4-2.0% 的供血者中。在约 50% 感染者中产生慢性肝炎,其中约 20% 被感染的个人产生肝硬化,有时导致肝细胞癌。因此,研究和控制该疾病在医学上是重要的。

[0004] Houghten 等首先鉴定并确定 HCV 是 NANBH 的原因。HCV 的病毒基因组序列是已知的,以及获得该序列的方法。参见,如国际出版物 No. W089/04669;W090/11089;和 W090/14436。HCV 具有 9.5kb 的正义单链 RNA 基因组,其是黄病毒科的成员。基于系统发育分析,已鉴定出至少 6 个不同但相关的 HCV 基因型 (Simmonds 等, J. Gen. Virol. (1993) 74 :2391-2399)。该病毒编码一个具有 3000 多个氨基酸残基的多蛋白 (Choo 等, Science (1989) 244 :359-362;Choo 等, Proc. Natl. Acad. Sci. USA (1991) 88 :2451-2455;Han 等, Proc. Natl. Acad. Sci. USA (1991) 88 :1711-1715)。多蛋白在翻译同时或之后加工成结构和非结构 (NS) 蛋白。

[0005] 具体说,如图 1 所示, HCV 基因组编码几种蛋白质。HCV 多蛋白的切割产物的顺序和命名如下: $\text{NH}_2\text{-C-E1-E2-p7-NS2-NS3-NS4a-NS4b-NS5a-NS5b-COOH}$ 。宿主蛋白酶催化多蛋白的最初切割,释放三种结构蛋白: N-末端核衣壳蛋白(称为“核心”)和两种包膜糖蛋白“AE1”(也称为 E)和“AE2”(也称为 E2/NS1),和含有病毒酶的非结构 (NS) 蛋白。NS 区称为 NS2、NS3、NS4 和 NS5。NS2 是具有蛋白水解活性的膜内在蛋白,它与 NS3 组合,切开 NS2-NS3 sissle 键,从而产生 NS3N-末端并释放大量的多蛋白,该多蛋白包含丝氨酸蛋白酶和 RNA 解旋酶活性。NS3 蛋白酶加工剩余的多蛋白。在这些反应中,NS3 释放 NS3 辅因子 (NS4a)、两种蛋白质 (NS4b 和 NS5a),和依赖 RNA 的 RNA 聚合酶 (NS5b)。多蛋白熟化的完成由 NS3-NS4 连接处的自催化裂解引发,由 NS3 丝氨酸蛋白酶催化。

[0006] 已描述了许多用作 HCV 的免疫和诊断试剂,衍生自 HCV 多蛋白的一般和特异性多肽。参见如 Houghton 等,欧洲出版物 No. 318, 216 和 388, 232;Choo 等, Science (1989) 244 : 359-362;Kuo 等, Science (1989) 244 :362-364;Houghton 等, Hepatology (1991) 14 : 381-388;Chien 等, Proc. Natl. Acad. Sci. USA (1992) 89 :10011-10015;Chien 等, J. Gastroent. Hepatol. (1993) 8 :S33-39;Chien 等, 国际出版物 No. W0 93/00365;Chien, D. Y., 国际出版物 No. W094/01778。这些出版物通常对 HCV、HCV 多肽免疫试剂的生产和使用提供了广泛的背景。

[0007] 筛选和鉴定 HCV 载体和 HCV 污染的血液或血液制品的灵敏和专一的方法将对医学提供重要的进步。在约 10% 输血的病人中发生输血后肝炎 (PTH),而 HCV 在这些病例中占

多达 90%。病人护理和预防及 HCV 通过血液或血制品,或通过密切的个人接触传播需要可靠的诊断和预后工具。因此,开发了几种测定方法用于 HCV 感染的血清诊断。参见如 Choo 等, Science(1989)244 :359-362 ;Kuo 等, Science(1989)244 :362-364 ;Choo 等, Br. Med. Bull. (1990)46 :423-441 ;Ebeling 等, Lancet(1990)335 :982-983 ;van der Poel 等, Lancet(1990)335 :558-560 ;van der Poel 等, Lancet(1991)337 :317-319 ;Chien, D. Y., 国际出版物 No. W094/01778 ;Valenzuela 等, 国际出版物 No. W097/44469 ;和 Kashiwakuma 等, 美国专利 No. 5, 871, 904。

[0008] 一些基于血清的测定方法面临的显著问题是,在病毒感染和检测之间有明显的间隔,通常超过 80 天。这种测定间隔对输血受体可能产生很大的危险。为了克服该问题,开发了直接检测病毒 RNA 的基于核酸的测试 (NAT),和检测病毒抗原而不是抗体反应的 HCV 核心抗原试验。参见如 Kashiwakuma 等, 美国专利 No. 5, 871, 904 ;Beld 等, Transfusion(2000), 40 :575-579。

[0009] 然而,仍然需要灵敏、准确的诊断和预后工具,以提供充分的病人护理并防止 HCV 通过血液和血制品,或密切的个人接触传播。

发明概要

[0010] 本发明部分基于以下发现:使用衍生自 HCV 多蛋白的第一抗原与包含得自与该第一抗原相同的 HCV 多蛋白区域的表位的多表位融合抗原的组合能提供敏感且可靠的检测 HCV 的方法。本文所述的检测可检测由六种已知的 HCV 基因型中的任何一种引起的 HCV 感染。在本发明的一个代表性实施例中,所述第一抗原包括 NS3/4a 构象表位,第二抗原是一种多表位融合抗原,包含一种或多种得自 NS3/4a 区域的表位。多表位融合蛋白的使用能在单位面积底物上提供大量的表位,并改善选择性,从而能减少掩蔽问题,提高敏感性,并检测抗体。而且,本文所述的检测可快速进行,并可使用较大的样品体积,而没有背景影响。

[0011] 因此,在一个实施方式中,本发明涉及检测生物样品中的丙型肝炎病毒 (HCV) 的方法。该方法包括:

[0012] (a) 提供免疫检测固相支持物,该支持物含有结合于其上的 HCV 抗原,其中,所述 HCV 抗原由一种或多种得自 HCV 多蛋白的第一区域的分离抗原组成;

[0013] (b) 在使 HCV 抗体(当存在于该生物样品中时)与所述一种或多种 HCV 抗原结合的条件下,使生物样品与该固相支持物混合;

[0014] (c) 在形成复合物的条件下,将可检测的标记 HCV 多表位融合抗原 (MEFA) 加到步骤 (b) 的固相支持物中,其中,该标记的 MEFA 含有至少一个得自与所述一个或多个分离的抗原相同的 HCV 多蛋白区域的表位,所述 MEFA 与结合的 HCV 抗体结合;

[0015] (d) 检测该 HCV 抗体与得自所述 HCV 多蛋白的第一区域的一种或多种抗原以及所述 MEFA 之间形成的复合物,如果有的话,作为生物样品中 HCV 感染的指标。

[0016] 在另一实施方式中,本发明涉及一种检测生物样品中的 HCV 感染的方法,该方法包括:

[0017] (a) 供免疫检测固相支持物,该支持物包含结合于其上的 HCV 抗原,其中,该 HCV 抗原由一种或多种多表位融合抗原 (MEFA) 组成;

[0018] (b) 在使 HCV 抗体(当存在于该生物样品中时)与所述一种或多种 MEFA 结合的条件

件下,使生物样品与该固相支持物混合;

[0019] (c) 在形成复合物的条件下,将存在于一种或多种 MEFA 中的得自 HCV 多蛋白的区域的可检测的标记的分离 HCV 抗原加到步骤 (b) 的固相支持物中;

[0020] (d) 检测该 HCV 抗体与所述分离的 HCV 抗原以及所述 MEFA 之间形成的复合物,如果有的话,作为生物样品中 HCV 感染的指标。

[0021] 在又一实施方式中,本发明涉及一种检测生物样品中的 HCV 感染的方法,该方法包括:

[0022] (a) 提供免疫检测固相支持物,该支持物包含结合于其上的 HCV 抗原,其中,该 HCV 抗原由一种或多种分离的 HCV NS3/4a 构象表位组成;

[0023] (b) 在使 HCV 抗体(当存在于该生物样品中时)与所述一种或多种 NS3/4a 表位结合的条件下,使生物样品与该固相支持物混合;

[0024] (c) 在形成复合物的条件下,将可检测的标记 HCV 多表位融合抗原(MEFA)加到步骤(b)的固相支持物中,其中,所述标记的 MEFA 含有至少一种得自 HCV NS3/4a 区域的表位,其中,该 MEFA 与该结合的 HCV 抗体结合;

[0025] (d) 检测该 HCV 抗体与该 NS3/4a 构象表位以及所述 MEFA 之间形成的复合物,如果有的话,作为生物样品中 HCV 感染的指标。

[0026] 在另一实施方式中,本发明涉及一种检测生物样品中的 HCV 感染的方法,该方法包括:

[0027] (a) 提供免疫检测固相支持物,该支持物含有结合于其上的 HCV 抗原,其中,该 HCV 抗原由一种或多种多表位融合抗原(MEFA)组成,所述一种或多种 MEFA 包含至少一种得自 HCV NS3/4a 区域的表位;

[0028] (b) 在使 HCV 抗体(当存在于该生物样品中时)与所述一种或多种 MEFA 结合的条件下,使生物样品与该固相支持物混合;

[0029] (c) 在形成复合物的条件下,将可检测的标记 HCV NS3/4a 构象表位加到所述固相支持物中,其中,所述 NS3/4a 构象表位与该结合的 HCV 抗体结合;

[0030] (d) 检测该 HCV 抗体与该 NS3/4a 构象表位以及所述 MEFA 之间形成的复合物,如果有的话,作为生物样品中 HCV 感染的指标。

[0031] 在上述方法中,NS3/4a 构象表位和/或 MEFA 包含得自 HCV 多蛋白的 NS3/4a 蛋白酶区域的表位,和/或得自 HCV 多蛋白的 NS3/4a 解旋酶区域的表位。在具体实施例中,该 NS3/4a 构象表位包含图 3A-3D 描述的氨基酸序列(SEQ IDNO:1 和 2)。

[0032] 在其它实施方式中,MEFA 包含氨基酸 1193-1657(依 HCV-1 序列编号)。在又一实施方式中,MEFA 包含得自 HCV 多蛋白 c33c 区域的表位,如氨基酸 1211-1457,和/或氨基酸 1192-1457,依 HCV-1 编号。

[0033] 在另外的实施方式中,MEFA 包含得自 HCV 多蛋白的 5-1-1 区域的表位,如氨基酸 1689-1735,依 HCV-1 编号。

[0034] 在具体的实施方式中,MEFA 包含图 6A-6F 所述的氨基酸序列(SEQ IDNO:3 和 4)或图 8A-8F 所述的氨基酸序列(SEQ ID NO:5 和 6)。

[0035] 在结合以下详细描述和附图后,本发明的这些和其它方面将变得显而易见。此外,本文给出了各种参考文献,这些文献更详细地描述了某些方法或组合物。

[0036] 附图简述

[0037] 图 1 是 HCV 基因组图,描述了衍生出本检测试剂(蛋白质和抗体)的多蛋白的不同区域。

[0038] 图 2 显示本发明使用 MEFA 12 进行的代表性抗原/抗体/抗原夹心检测的示意图。

[0039] 图 3A-3D(SEQ ID NO :1 和 2) 显示了用于本检测的代表性 NS3/4a 构象抗原的 DNA 和相应的氨基酸序列。在图 3A-3D 中 403 和 404 位的氨基酸代表 HCV-1 的天然氨基酸序列中的 Pro 代替 Thr,和 Ile 代替 Ser。

[0040] 图 4 是 pd. HCV1a. ns3ns4aPI 的构建图。

[0041] 图 5 是 MEFA 12 的示意图。

[0042] 图 6A-6F(SEQ ID NO :3 和 4) 描述了 MEFA 12 的 DNA 和相应的氨基酸序列。

[0043] 图 7 是 MEFA 7.1 的示意图。

[0044] 图 8A-8F(SEQ ID NO :5 和 6) 描述了 MEFA 7.1 的 DNA 和相应的氨基酸序列。

[0045] 图 9A-9C 显示用于本发明免疫检测的代表性 MEFA。图 9A 是 MEFA 3 的示意图。图 9B 是 MEFA 5 的示意图。图 9C 是 MEFA 6 的示意图。

[0046] 发明详述

[0047] 本发明的实施除非另外说明,将使用本领域技术人员已知的化学、生物化学、重组 DNA 技术和免疫学的常规方法。这些技术在文献中有完整的解释。参见,如《基础免疫学》(Fundamental Virology),第二版,第 I 和 II 卷(B. N. Fields 和 D. M. Knipe 编);《实验免疫学手册》(Handbook of Experimental Immunology),第 I-IV 卷(D. M. Weir 和 C. C. Blackwell 编, Blackwell Scientific Publications);T. E. Creighton,《蛋白质:结构和分子特性》(Proteins: Structures and Molecular properties)(W. H. Freeman and Company, 1993);A. L. Lehninger,《生物化学》(Biochemistry)(Worth Publishers, Inc. 最新版);Sambrook 等,《分子克隆:实验室手册》(Molecular Cloning: a Laboratory Manual),第二版,1989;《酶学方法》(Methods in Enzymology)(S. Colowick 和 N. Kaplan 编, Academic Press, Inc.)。

[0048] 必须注意的是,如本说明书和权利要求中使用的,单数形式“一”、“一个”、“该”包括复数参考,除非内容明显说明。因此,例如,“一抗原”包括两个或多个抗原的混合物等。

[0049] 文中使用了下列氨基酸缩写:

[0050] 丙氨酸:Ala(A) 精氨酸:Arg(R)

[0051] 天冬酰胺:Asn(N) 天冬氨酸:Asp(D)

[0052] 半胱氨酸:Cys(C) 谷氨酰胺:Gln(Q)

[0053] 谷氨酸:Glu(E) 甘氨酸:Gly(G)

[0054] 组氨酸:His(H) 异亮氨酸:Ile(I)

[0055] 亮氨酸:Leu(L) 赖氨酸:Lys(K)

[0056] 甲硫氨酸:Met(M) 苯丙氨酸:Phe(F)

[0057] 脯氨酸:Pro(P) 丝氨酸:Ser(S)

[0058] 苏氨酸:Thr(T) 色氨酸:Trp(W)

[0059] 酪氨酸:Tyr(Y) 缬氨酸:Val(V)

[0060] I. 定义

[0061] 在本发明的描述中,使用了下列术语,并如下定义。

[0062] 术语“多肽”和“蛋白质”指氨基酸残基的聚合物,并不限于产物的最小长度。因此,肽、寡肽、二聚物、多聚物等都包括在该定义中。全长的蛋白质及其片段包括在该定义中。该术语还包括多肽的表达后修饰,例如糖基化、乙酰化、磷酸化等。另外,为了本发明的目的,“多肽”指包括天然序列的修饰,例如缺失、添加和取代(通常性质保守),只要蛋白质维持所需活性。这些修饰可以是经过设计的,如通过定点诱变,或可以是偶然的,例如通过产生蛋白质的宿主突变,或由于 PCR 扩增引起的错误。

[0063] HCV 多肽是一种如上所述的衍生自 HCV 多蛋白的多肽。该多肽不需要物理衍生自 HCV,但可以是合成或重组产生的。另外,该多肽可衍生自任何各种 HCV 株和分离物,如但不限于来自 HCV 株 1、2、3、4、5 或 6 的任一分离物。在这些株之间有许多保守和可变的区域,一般当两条序列排列时,衍生自这些区域的表位的氨基酸序列将具有高程度的序列同源性,例如高于 30%,优选高于 40%的氨基酸序列同源性。因此,例如术语“NS3/4a”多肽指任何各种 HCV 株的天然 NS3/4a 和 NS3/4a 类似物、突变蛋白和免疫原性的片段(如下进一步定义)。许多这些病毒株的完整基因型是已知的。参见如美国专利 No. 6, 150, 087 和 GenBank 登录号 AJ238800 和 AJ238799。

[0064] 术语“类似物”和“突变蛋白”指参照分子的生物活性衍生物,或这些衍生物保留所需活性(如在本文所述的测定中的免疫反应性)的片段。通常,术语“类似物”指具有天然多肽序列和结构,以及相对于天然分子的一个或多个氨基酸添加、取代(通常性质保守)和/或缺失的化合物,只要修饰不破坏免疫原性的活性。术语“突变蛋白”指具有一种或多种肽模拟物(“类肽”)的肽,如国际出版物号 W091/04282 中所述的。优选类似物或突变蛋白至少具有与天然分子相同的免疫活性。制备多肽类似物和突变蛋白的方法是本领域已知的,如下进一步所述。

[0065] 特别优选的类似物包括性质上保守的取代,即这些取代发生在与它们的侧链有关的一类氨基酸中。具体而言,氨基酸一般被分成四类:(1) 酸性——天冬氨酸和谷氨酸;(2) 碱性——赖氨酸、精氨酸、组氨酸;(3) 非极性——丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸;(4) 无电荷的极性——甘氨酸、天冬酰胺、谷氨酰胺、半胱氨酸、丝氨酸、苏氨酸、酪氨酸。有时将苯丙氨酸、色氨酸和酪氨酸归为芳族氨基酸。例如,有理由预测:单独用异亮氨酸或缬氨酸取代亮氨酸、用谷氨酸取代天冬氨酸、用丝氨酸取代苏氨酸,或者用结构上相关的氨基酸取代类似的保守的氨基酸,这样的取代将不会对生物活性有重要影响。例如,感兴趣的多肽可包括多达约 5-10 个保守的或不保守的氨基酸取代,甚至多达约 15-25 个保守的或不保守的氨基酸取代,或 2-25 之间任何整数,只要该分子的所需功能仍维持完整。本领域的熟练技术人员可结合本领域熟知的 Hopp/Woods 和 Kyte-Doolittle 曲线图,容易地测定感兴趣的分子中可耐受改变的区域。

[0066] 术语“片段”指仅由完整的全长多肽序列和结构的一部分组成的多肽。该片段可包括该天然多肽的 C 末端缺失和/或 N 末端缺失。特定的 HCV 蛋白质的“免疫原性片段”一般包括至少该全长分子的约 5-10 个连续的氨基酸残基,较佳至少含有该全长分子的约 15-25 个连续的氨基酸残基,最佳至少含有该全长分子的约 20-50 个或以上的连续的氨基酸残基,这些氨基酸残基限定了一个表位;或者含有 5 个氨基酸和全长序列之间的任何整数,只要所述的片段在测定中保留免疫活性。例如,用于 MEFA 的优选的免疫原性片段,包括

但不限于 HCV 核心的片段,包含例如多蛋白的氨基酸 10-45、10-53、67-88 和 120-130,表位 5-1-1(在病毒基因组的 NS3 区域中)和衍生自 HCV 多蛋白的 E1、E2、c33c(NS3)、c100(NS4),NS3/4a 和 NS5 区域的限定的表位,以及从 HCV 多蛋白中鉴定出的任何其它各种表位。见例如 Chien 等,Proc. Natl. Acad. Sci. USA(1992),89:10011-10015;Chien 等, J. Gastroent. Hepatol., (1993),8:S33-39;Chien 等,国际出版物 WO 93/00365;Chien, D. Y., 国际出版物 WO 94/01778;美国专利 No. 6, 150, 087 和 6, 121, 020。

[0067] 术语“表位”在文中指至少约有 3-5 个、较佳地约 5 到 10 或 15 个、但不超过约 1000 个氨基酸(或其中的任何整数)的序列;它定义了一条自身或作为较大序列的一部分,与在对其应答中产生的抗体结合的序列。该片段的长度没有严格的上限,它几乎可包含蛋白质序列的全长,甚至包括含有 HCV 多蛋白的两个或多个表位的融合蛋白。用于本发明的表位并不限于具有衍生了它的母体蛋白质部分的确切序列的多肽。实际上,病毒基因组处于恒定流动的状态,且含有几种在隔离种群间具有相对高度易变性的可变结构域。因而,术语“表位”包括与天然序列相同的序列,以及该天然序列的修饰物,如缺失、添加和取代(通常性质保守)。

[0068] 可使用任何数量的本领域周知的表位作图技术来鉴别包括表位在内的给定多肽的区域。参见,如《分子生物学方法》(Methods in Molecular Biology)中的表位作图程序,第 66 卷(Glenn E. Morris 编辑,1996),Humana Press, Totowa, New Jersey。例如,通过如同时在固相支持物上合成大量的肽,对应于蛋白质分子的部分的肽,在这些肽仍连接于该支持物时使肽与抗体反应,就可测定各线性表位。这些技术在本领域中是已知的,并在美国专利 No. 4, 708, 871;Geysen 等(1984), Proc. Natl. Acad. Sci. USA,81:3998-4002;Geysen 等(1985)Proc. Natl. Acad. Sci. USA 82:178-182;Geysen 等(1986), Molec. Immunol., 23:709-715 中有所描述。用这些技术鉴定了许多 HCV 的表位。参见如 Chien 等,《病毒性肝炎和肝病》(Viral Hepatitis and Liver Disease)(1994),第 320-324 页和下文。类似地,通过使用如 X-射线结晶学和二维核磁共振法测定氨基酸的空间构象容易地鉴定构象表位。参见如“表位作图流程”,上述。还可使用标准抗原性曲线图和亲水性图,如用从 Oxford Molecular Group 获得的 Omega 1.0 版软件程序,也可鉴别蛋白质的抗原区域。此计算机程序采用 Hopp/Woods 方法(Hopp 等, Proc. Natl. Acad. Sci. USA(1981),78:3824-3828)确定抗原性图谱,并用 Kyte-Doolittle 技术(Kyte 等, J. Mol. Biol. (1982),157:105-132)绘制亲水性图。

[0069] 关于各种 HCV 表位的描述,可参见 Chien 等, Proc. Natl. Acad. Sci. USA(1992)89:10011-10015;Chien 等, J. Gastroent. Hepatol. (1993)8:S33-39;Chien 等, 国际出版物 No. WO 93/00365;Chien, D. Y., 国际出版物 No. WO94/01778;和美国专利 6280927 和 6150087。

[0070] 术语“构象表位”在文中指全长蛋白质的一部分,或其类似物或突变蛋白,它具有编码该全长天然蛋白质内的表位的氨基酸序列的结构特征。天然的结构特征包括但不限于糖基化和三维结构。表位限定的序列的长度可广泛改变,因为据信这些表位是通过抗原的三维形状形成的(例如折叠)。因此限定该表位的氨基酸可以数量相对少,但沿着分子的长度(或者甚至在二聚体等的情况下对于不同的分子)广泛地分散,通过折叠形成正确的表位构象。限定表位的残基之间的抗原部分对于表位的构象结构不是关键的。例如,这些间

插序列的缺失或取代可能不影响构象表位,条件是保留对于表位构象关键的序列(例如涉及二硫键、糖基化位点的半胱氨酸等)。

[0071] 可用上述方法容易鉴别 NS3/4a 区域中存在的构象表位。另外,在给定的多肽中构象表位的存在与否可通过用抗体(构象表位的多克隆血清或单克隆抗体)筛选感兴趣的抗原,并将它的反应性与仅保留线性表位(如存在)的抗原的变性形式比较来确定。在用多克隆抗体的筛选中,有利的是首先用变性的抗原吸附多克隆血清,看其是否保留针对感兴趣的抗原的抗体。另外,在 NS3/4a 的情况下,保留天然构象的分子还具有蛋白酶活性,和任选的解旋酶活性。这些活性可用酶测定检测(如下所述)。

[0072] 较佳地,重组产生构象表位,并使它在一种在保留其所需结构特征的条件下(如表位没有变性)能被抽提出来的细胞中表达。这样的细胞包括细菌、酵母、昆虫细胞和哺乳动物细胞。HCV 多蛋白的重组构象表位的表达和分离在如国际出版物 W0 96/04301、W0 94/01778、W0 95/33053、W0 92/08734 中有描述。另外,可能表达抗原并在回收后使蛋白质复性。还理解化学合成也可提供构象抗原模拟表位(mimitopes),它与“天然”抗原的构象表位交叉反应。

[0073] 本文所用的术语“多表位融合抗原”或“MEFA”指一种多肽,其中多个 HCV 抗原是一条单一、连续的氨基酸链的部分,该链在天然中不存在。HCV 抗原可以直接通过肽键彼此连接,或通过插入氨基酸序列分隔。融合抗原还可以含有相对于该 HCV 多蛋白而言外源的序列。另外,存在的 HCV 序列可以来自 HCV 的多基因型和/或分离物。本免疫测定中所用的具体的 MEFA 的例子在例如国际出版物 W001/96875、W001/09609、W097/44469 和美国专利 6514731 和 6428792 中描述。

[0074] “抗体”指通过化学或物理方法与感兴趣的多肽专一性结合的分子。因此,HCV NS3/4a 抗体是一种与 HCVNS3/4a 蛋白的表位特异性结合的分子。本文所用的术语“抗体”包括从多克隆和单克隆制备物获得的抗体,及以下:杂交(嵌合)抗体分子(参见如 Winter 等(1991)Nature 349:293-299;和美国专利 No. 4,816,567);F(ab')₂ 和 F(ab) 片段;Fv 分子(非共价异二聚体,参见如 Inbar 等(1972)Proc. Natl. Acad. Sci. USA 69:2659-2662;和 Ehrlich 等(1980)Biochem 19:4091-4096);单链 Fv 分子(sFv)(参见如 Huston 等(1988)Proc. Natl. Acad. Sci. USA 85:5879-5883);二聚和三聚抗体片段构建物;小体(minibodies)(参见如 Pack 等(1992)Biochem 31:1579-1584;Cumber 等(1992)J Immunology 149B:120-126);人源化抗体分子(参见如 Riechmann 等(1988)Nature 332:323-327;Verhoeyan 等(1988)Science 239:1534-1536;和英国专利申请 No. GB 2,276,169,1994 年 9 月 21 日出版);和从这些分子获得的任何功能性片段,其中这些片段维持亲本抗体分子的免疫结合性质。

[0075] “重组”蛋白质是指保留了所需活性的通过本文所述的重组 DNA 技术制得的蛋白质。通常,克隆感兴趣的基因,然后在转化的生物体中表达(如下所述)。宿主生物体表达外源基因,从而在表达条件下产生该蛋白。

[0076] 当涉及一种多肽时,“分离”意味着所述的分子从发现该分子天然存在的整个生物体中分离和分开,或基本不存在其它相同类型的生物大分子。术语“分离”对于多核苷酸是:一种核酸分子,它完全或部分缺乏通常与其天然结合的序列;或一个序列,因为它天然存在,但具有与其结合的异源序列;或从染色体解离的分子。

[0077] “等效抗原决定簇”指来自 HCV 的不同亚种或株的抗原决定簇,例如 HCV 株 1、2、3 等。更特别地,表位是已知的,如 5-1-1,和病毒株 1、2 和 3 之间的表位有变化。因此,这三种不同病毒株的表位 5-1-1 是等效的抗原决定簇,即使它们的序列不相同,它们也是“拷贝”。等效抗原决定簇的氨基酸序列通常具有高度的序列同源性,例如当两个序列对比时,大于 30% 的氨基酸序列同源性,较佳地大于 40% 的氨基酸序列同源性。

[0078] “同源性”指两个多核苷酸或两条多肽部分之间的相似性百分数。两条 DNA 或两条多肽序列在确定的分子长度内,当序列显示有至少约 50%,优选至少约 75%、更佳至少约 80-85%、尤其佳至少约 90%、最佳至少约 95-98% 序列相似性时,彼此“基本上同源”。如本文所述,基本同源也指显示与特定的 DNA 或多肽序列具有完全同一性的序列。

[0079] 通常,“同一性”指两条多核苷酸或多肽序列上准确的核苷酸对核苷酸或者氨基酸对氨基酸对应。通过排列两个分子的序列直接比较它们的序列信息,计算两条排列的序列间匹配的准确数量,将其除以较短序列的长度,然后乘以 100,从而可得到同一性百分数。

[0080] 在相似性和同一性分析中可辅助使用易于获得的计算机程序,如 ALIGN、Dayhoff、M. O. (Atlas of Protein Sequence and Structure、M. O. Dayhoff 编辑, 5Suppl., 3: 353-358, National Biomedical Research Foundation, Washington, DC), 它适用于 Smith 和 Waterman 分析肽用的局部同源性算法 (Advances in Appl. Math., 2: 482-489, 1981)。可从 Wisconsin Sequence Analysis Package (第 8 版, 从 Genetics Computer Group, Madison, WI 获得) 获得测定核苷酸序列相似性和同一性的程序, 例如, BESTFIT、FASTA 和 GAP 程序, 这些程序也依赖于 Smith 和 Waterman 算法。使用制造者建议的和上述 Wisconsin Sequence Analysis Package 所述的默认参数可容易地使用这些程序。例如, 可使用 Smith 和 Waterman 的同源性算法的默认计分表和 6 个核苷酸位置的间隔罚分 (gap penalty) 测定的核苷酸序列与参比序列的相似性百分数。

[0081] 本发明建立相似性百分数的另一方法是使用版权属于爱丁堡大学、由 John F. Collins 和 Shane S. Sturrok 开发、由 IntelliGenetics, Inc. (Mountain View, CA) 发行的 MPSRCH 程序包。Smith-Waterman 算法可在这套程序包中使用, 其中, 在计分表中使用默认参数 (例如, 间隔开放罚分 = 12, 间隔延伸罚分 = 1, 间隔 = 6)。从这批数据产生的“匹配”值反映出“序列相似性”。计算序列间的同一性百分数或相似性百分数的其它合适的程序在本领域中一般都是已知的, 例如, 另一种排列程序是 BLAST, 使用默认参数。例如, 可使用下述默认参数的 BLASTN 和 BLASTP: 基因编码 = 标准; 过滤 = 无; 链 = 两者; 截留 = 60; 期望值 = 10; 矩阵 = BLOSUM62; 描述 = 50 个序列; 排序 = HIGH SCORE; 数据库 = 无冗余, GenBank+EMBL+DDBJ+PDB+GenBank CDS 翻译 +Swiss 蛋白 +Spupdate+PIR。在 <http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST> 网址上可查到这些程序的详细描述。

[0082] 或者, 在同源区域之间形成稳定的双链体的条件下进行多核苷酸杂交, 接着用单链特异性核酸酶消化, 测定消化的片段的大小, 从而测出同源性。在如 (对具体的体系所定义的) 严格条件下进行的 Southern 杂交试验中, 可鉴别基本同源的 DNA 序列。确定适当的杂交条件在本领域熟练技术人员所掌握的知识之内。例如, 参见 Sambrook 等, 同上; DNA 克隆, 同上; 核酸杂交, 同上。

[0083] “一般固相支持物”指一个单一固相基质, 免疫测定所用的 HCV 多肽与其共价或通过非共价方式, 例如疏水性吸附结合。

[0084] “免疫反应性”指所关心的抗原与 HCV 感染的个体的生物样品中存在的抗-HCV 抗体特异地反应。

[0085] “免疫复合物”指当抗体和抗原上的表位结合时形成的组合。

[0086] 本文所用的“生物样品”指从受试者中分离的组织或液体的样品,包括(但不限于)血液、血浆、血清、粪便物、尿液、骨髓、胆汁、脊髓液、淋巴液、皮肤样品、皮肤分泌物、呼吸道分泌物、肠道分泌物和泌尿生殖道分泌物、泪液、唾液、乳液、血细胞、器官、活组织检查和体外细胞培养组分的样品,包括但不限于细胞和组织在培养基中生长获得的条件培养基,例如重组细胞和细胞成分。

[0087] 本文所用的术语“标记”和“可检测标记”指能检测的分子,包括但不限于反射性同位素、荧光素、化学发光物、生色团、酶、酶底物、酶辅因子、酶抑制剂、生色团、染料、金属离子、金属溶胶、配体(例如生物素、链霉亲和素和半抗原)等。术语“荧光素”指能在可检测范围内显示荧光的物质或其部分。可在本发明中使用的标记的特定例子包括但不限于,辣根过氧化物酶(HRP)、荧光素、FITC、罗丹明、丹磺酰、伞形酮、二甲基吡啶鎓酯(DMAE)、德克萨斯红、鲁米诺、NADPH 和 α - β -半乳糖苷酶。

[0088] II. 本发明的实施方式

[0089] 在详细描述本发明之前,应理解本发明并非局限于特定的制剂或工艺参数,它们当然是可以变化的。还应理解本文所用的术语仅用于说明本发明具体实施例的目的,而并非对其限定。

[0090] 虽然大量与本文所述类似或相等的组合物和方法可用于本发明的实践中,但本文描述了优选的材料和方法。

[0091] 如上所述,本发明基于以下发现:新颖的抗原/抗体/抗原夹心检测方法可用于准确检测 HCV 感染。这些方法可快速、有效地实施,并消除使用大样品体积可能产生的背景影响。这些方法较佳使用存在于 HCV 血清转化的早期阶段期间的高度免疫原性 HCV 抗原,从而增加检测准确度,并减少假结果的产生率。

[0092] 具体而言,本文所述的免疫测定使用两种基本的 HCV 抗原,一种存在于固相中,另一种存在于液相中。两种抗原都能结合感染个体的生物样品中存在的 HCV 抗体。用于本检测中的一种抗原是得自 HCV 多蛋白的区域的分离的抗原。所使用的第二抗原是含有得自相同或不同 HCV 基因型和分离物的各种 HCV 多肽的多表位融合抗原(“MEFA”)。该 MEFA 包括至少一种或多种衍生自与分离的 HCV 抗原相同的 HCV 多蛋白区域的表位。

[0093] 在特别优选的实施例中,用于本检测的抗原之一是衍生自 HCV 多蛋白的 NS3/4a 区域的高度免疫原性构象表位。在这些实施例中,所使用的第二抗原是包含一种或多种得自 NS3/4a 区域(线性或构象的)的表位的 MEFA,如下文所述。因此,MEFA 可包括衍生自一种或多种 HCV 分离物的 NS3/4a 区域的多个优势免疫表位。如果在多表位融合中使用多个 NS3/4a 表位,这些表位可以是相同或不同的表位。或者,融合抗原可包括一种或多种衍生自 NS3/4a 区域的表位,以及得自其它 HCV 区域如(无限制的)HCV 核心、E1、E2、P7、NS4b、NS5a 和 NS5b 序列的主要线性表位。

[0094] 可使用下文所述的任何一种检测方式,如(但不限于)使用结合了分离的 HCV 抗原(如 NS3/4a 构象表位)或 MEFA 的固相支持物的检测形式,以简单的检测来方便地实施这些方法。因此,MEFA 可以液相或固相的形式提供。如果提供在溶液中,分离的 HCV 抗原

则存在于固相中。

[0095] 例如,在本发明的一个代表性方法中,在固相支持物上实施检测,该固相支持物已结合有一种或多种包含一种或多种衍生自 HCV 多蛋白的 NS3/4a 区域的构象表位。在此实施例中,MEFA 存在于液相中。在另一实施例中,在构象支持物上实施检测,该固相支持物已结合有一种或多种 MEFA。在此实施例中,含有构象 NS3/4a 表位的多肽存在在液相中。因此,如果构象 NS3/4a 表位存在于固相支持物中,MEFA 则存在在液相中,或者相反。

[0096] 为了进一步理解本发明,以下将就用于本发明方法的各种 HCV 多肽抗原和 MEFA 以及蛋白质的生产和使用蛋白质的方法提供更详细的讨论。

[0097] HCV 抗原和 MEFA

[0098] HCV 株的基因组含有长约 9000-12000 个核苷酸的单一开放读框,它被转录成多蛋白。如图 1 和表 1 所示,HCV 多蛋白在切割后产生至少 10 种不同的产物,其顺序为 NH₂-核心-E1-E2-p7-NS2-NS3-NS4a-NS4b-NS5a-NS5b-COOH。核心多肽的位置为 1-191,相对于 HCV-1 编号 (HCV-1 的基因组可参见,Choo 等人,(1991)Proc. Natl. Acad. Sci. USA 88 : 2451-2455)。进一步加工该多肽产生约 1-173 个氨基酸的 HCV 多肽。包膜多肽 E1 和 E2 分别出现在 192-383 和 384-746 位。在约 747-809 位点发现 P7 区域。NS2 是具有蛋白水解活性的膜内在蛋白,且约位于多蛋白的 810-1026。NS2 与 NS3 (位于 1027-1657),切割 NS2-NS3 的 sissle 键,从而产生 NS3N- 末端并释放大多的多蛋白,具有丝氨酸蛋白酶和 RNA 解旋酶活性。在约 1027-1207 位发现的 NS 蛋白酶用于加工余下的多蛋白。在约 1193-1657 位发现解旋酶活性。NS3 释放出 NS3 辅因子 (NS4a,位置为约 1658-1711)、两种蛋白质 (NS4b 的位置为约 1712-1972,和 NS5a 的位置为约 1973-2420) 和 RNA- 依赖性 RNA 聚合酶 (NS5b 的位置为约 2421-3011)。多蛋白熟化的完成由 NS3-NS4 连接处的自催化裂解引发,由 NS3 丝氨酸蛋白酶催化。

[0099]

表 1	
区域	大约的边界*
C(核心)	1-191
E1	192-383
E2	384-746
P7	747-809
NS2	810-1026
NS3	1027-1657
NS4a	1658-1711
NS4b	1712-1972
NS5a	1973-2420
NS5b	2421-3011

[0100] * 相对于 HCV-1 编号。参见 Choo 等人 (1991)Proc. Natl. Acad. Sci. USA88 : 2451-2455。

[0101] 本发明方法的一个组成是得自上述 HCV 多蛋白任何各个区域中的任一个的分离

的抗原。已测定了多种 HCV 株和分离物的核酸和氨基酸序列,包括上述各种区域的核酸和氨基酸序列。例如,Kubo 等人 (1989),*Japan. Nucl. Acids. Res.*,17 :10367-10372 ;Takeuchi 等人 (1990),*Gene*,91 :287-291 ;Takeuchi 等人 (1990),*J. Gen. Virol.*,71 :3027-3033 以及 Takeuchi 等人 (1990),*Nucl. Acids Res.*,18 :4626 中描述了分离物 HCV J1.1。两个独立的分离物 HCV-J 和 BK 的完整的编码序列分别由 Kato 等人 (1990),*Proc. Natl. Acad. Sci. USA*,87 :9524-9528) 和 Takamizawa 等人 (1991),*J. Virol.*,65 :1105-1113) 描述。

[0102] 描述 HCV-1 分离物的出版物包括 Choo 等人 (1990),*Brit. Med. Bull.*,46 :423-441 ;Choo 等人 (1991),*Proc. Natl. Acad. Sci. USA*,88 :2451-2455 和 Han 等人 (1991),*Proc. Natl. Acad. Sci. USA*,88 :1711-1715。在 Okamoto 等人 (1991),*Japan J. Exp. Med.*,60 :167-177 中描述了 HCV 分离物 HC-J1 和 HC-J4。Weiner 等人 (1991),*Virol.*,180 :842-848 中描述了 HCV 分离物 HCT 18-、HCT 23、Th、HCT 27、EC1 和 EC10。在 Enomoto 等人 (1990),*Biochem. Biophys. Res. Commun.*,170 :1021-1025 中描述了 HCV 分离物 Pt-1、HCV-K1 和 HCV-K2。在 Tsukiyama-Kohara 等人 (1991),*Virus Genes*,5 :243-254 中描述了 HCV 分离物 A、C、D 和 E。

[0103] 因此,例如,可从任何一种这些 HCV 分离物的核心区域衍生得到分离的 HCV 抗原。这个区域出现在该 HCV 多蛋白的第 1-191 位氨基酸 (相对于 HCV-1 编号)。该全长蛋白质、其片段 (如氨基酸 1-150,如氨基酸 1-130、1-120,例如氨基酸 1-121、1-122、1-123 等) 或含有该全长蛋白质的表位的较小片段都可用于本发明方法,所述表位如在第 10-53 个氨基酸、第 10-45 个氨基酸、第 67-88 个氨基酸、第 120-130 个氨基酸之间发现的那些表位,或者在以下文献中鉴别的任何核心表位:如 Houghton 等人,美国专利第 5350671 ;Chien 等人,*Proc. Natl. Acad. Sci. USA* (1992),89 :10011-10015 ;Chien 等人,*J. Gastroent. Hepatol.* (1993),8 :S33-39 ;Chien 等人,国际出版物第 W0 93/00365 号 ;Chien, D. Y., 国际出版物第 W0 94/01778 ;和美国专利第 6280927 号和 6150087 号。此外,可使用由多蛋白的核心区域中的移码产生的蛋白质 (如国际出版物 W099/63941 中所述)。

[0104] 类似地,可将得自 HCV E1 和 / 或 E2 区域的多肽用作本发明方法中的分离的 HCV 抗原。E2 存在于多个物种中 (Spaete 等, *Virol.* (1992) 188 :819-830 ;Selby 等, *J. Virol.* (1996) 70 :5177-5182 ;Grakoui 等, *J. Virol.* (1993) 67 :1385-1395 ;Tomei 等, *J. Virol.* (1993) 67 :4017-4026), 裁剪和蛋白水解可出现在 E2 多肽的 NB 和 C 末端。因此,用于本文的 E2 多肽可含有 HCV 多蛋白的氨基酸 405-661,如 400、401、402、……到 661。如 383 或 384-661、383 或 384-715、383 或 384-746、383 或 384-749 或 383 或 384-809、或 383 或 384 到 661-809 之间的任一 C 末端,相对于全长 HCV-1 多蛋白编号。类似地,用于本文的 E1 多肽可含有 HCV 多蛋白的氨基酸 192-326、192-330、192-360、192-363、192-383、或 192 到 326-393 之间的任一 C 末端。

[0105] 含有表位的 E1 和 / 或 E2 的免疫原性片段也用于本发明方法中。例如,E1 多肽片段可包括从 E1 多肽约 5 个氨基酸到近全长分子,如 6、10、25、50、75、100、125、150、175、185 个或更多氨基酸或所示数字间任何整数。类似地,E2 多肽片段可包含 E2 多肽的 6、10、25、50、75、100、150、200、250、300 或 350 个氨基酸,或所述数字之间的任何整数。

[0106] 例如,衍生自如 E2 的超变区 (例如跨越氨基酸 384-410 或 390-410 的区域) 的表位可包括在这些融合体中。掺入到 E2 序列中的特别有效的 E2 表位是含有衍生自此区域的

共有序列的表位,所述共有序列如 Gly-Ser-Ala-Ala-Arg-Thr-Thr-Ser-Gly-Phe-Val-Ser-Leu-Phe-Ala-Pro-Gly-Ala-Lys-Gln-Asn,代表 HCV 类型 1 基因组氨基酸 390-410 的共有序列。E1 和 E2 的其它表位是已知的,并描述于如 Chien 等,国际公开号 WO 93/00365。

[0107] 此外,E1 和 / 或 E2 多肽可缺乏所有或部分跨膜结构域。关于 E1,一般终止于约氨基酸 370 位和更高(以 HCV-1E1 编号为基础)的多肽由 ER 保留并因此不分泌入培养基。关于 E2,终止于约氨基酸 731 位和更高(也以 HCV-1E2 序列编号为基础)的多肽由 ER 保持并因此不分泌。(参见例如出版于 1996 年 2 月 15 日的国际出版物 WO 96/04301)。应该指出这些氨基酸位置不是绝对的,且可一定程度变化。因此,本发明考虑使用保留跨膜结合结构域的 E1 和 / 或 E2 多肽,以及缺乏所有或部分跨膜结合结构域的多肽,包括终止于约氨基酸 369 位和更低的 E1 多肽和终止于约氨基酸 730 位和更低的 E1 多肽。此外,C-末端截短可延伸超出跨膜结构域到 N-末端。因此,例如发生在位置低于如 360 的 E1 截短和发生在低于如 715 位的 E2 截短也可包括在本发明中。必要的是截短的 E1 和 E2 多肽保持用于所需目的的功能。然而,特别优选的截短 E1 构建物是那些延伸不超过约氨基酸 300 的构建物。最优选的是在 360 位终止。优选的截短的 E2 构建物是其 C 末端截短延伸不超过氨基酸 715 位的构建物。特别优选的是,E2 截短是在氨基酸 715-730 中的任一氨基酸(如 725)之后截短。

[0108] 此外,可将得自 NS3、NS4、NS5a 或 NS5b 区域的表位用作分离的 HCV 抗原。特别优选的是使用得自 NS3/4a 区域的表位。已描述了 HCV 多蛋白的 NS3/4a 区域,且该蛋白的氨基酸序列和整体结构已公开于:如 Yao 等,Structure(1999 年 11 月)7:1353-1363;Sali 等,Biochem. (1998) 37:3392-3401;和 Bartenschlager,R.,J. Viral Hepat. (1999) 6:165-181。同样,可参见 Dasmahapatra 等的美国专利 No. 5,843,752。本发明的免疫测定可使用至少一种衍生自 NS3/4a 区域的构象表位,其以天然的 HCV 粒子或其传染性产物中所发现的构象存在,这可通过由 NS3/4a 基因产物通常显示的蛋白酶和任选地解旋酶活性的保存和 / 或抗原与受 HCV 感染的对象的生物样品中的抗体的免疫反应性以及抗原变性后表位的免疫反应性验证。例如,通过加热、改变 pH 至极酸性或碱性或通过添加已知的有机变性剂(如二硫苏糖醇(DTT))或适当的去污剂破坏构象表位。参见如《蛋白质纯化方法,一种实用的方法》(Protein Purification Methods, a Practical approach)(E. L. V. Harris 和 S. Angal 编辑,IRL Press),并将变性的产物与未进行上述处理的产物相比。

[0109] 可用本领域已知的标准酶试验确定蛋白酶和解旋酶活性。例如,可用本领域已知的方法确定蛋白酶活性。参见如 Takeshita 等,Anal. Biochem. (1997) 247:242-246;Kakiuchi 等,J. Biochem. (1997) 122:749-755;Sali 等,Biochemistry(1998) 37:3392-3401;Cho 等,J. Virol. Meth(1998) 72:109-115;Cerretani 等,Anal. Biochem. (1999) 266:192-197;Zhang 等,Anal. Biochem. (1999) 270:268-275;Kakiuchi 等,J. Virol. Meth(1999) 80:77-84;Fowler 等,J. Biomol. Screen. (2000) 5:153-158;和 Kim 等,Anal. Biochem. (2000) 284:42-48。以下实施例阐述了一种特别方便的测试蛋白酶活性的测定。

[0110] 类似地,解旋酶活性测定是本领域熟知的,且 NS3/4a 表位的解旋酶活性可用如下方法确定,例如 ELISA 分析(如 Hsu 等,Biochem. Biophys. Res. Commun. (1998) 253:594-599 中所述);闪烁亲近测定系统(如 Kyono 等,Anal. Biochem. (1998) 257:120-126

中所述);高通量筛选测定法(如 Hicham 等, *Antiviral Res.* (2000) 46 :181-193 和 Kwong 等, *Methods Mol. Med.* (2000) 24 :97-116 中所述);以及本领域已知的其它测定法。例如,参见 Khu 等, *J. Virol.* (2001) 75 :205-214;Utama 等, *Virology* (2000) 273 :316-324;Paolini 等, *J. Gen. Virol.* (2000) 81 :1335-1345;Preugschat 等, *Biochemistry* (2000) 39 :5174-5183;Preugschat 等, *Methods Mol. Med.* (1998) 19 :353-364;和 Hesson 等, *Biochemistry* (2000) 39 :2619-2625。

[0111] 如果使用构象 NS3/4a 表位,抗原的长度应足以维持一种免疫反应性构象表位。通常,含有所用抗原的多肽几乎是全长的,然而也可以截短该多肽,以用于例如增加溶解性或改善分泌。一般而言,在 NS3/4a 中发现的构象表位在细胞中作为重组多肽表达,且该多肽提供所需形式的表位(如下详述)。

[0112] 图 3A-3D(SEQ ID NO :1 和 2) 显示了 NS3/4a 多肽的代表性氨基酸序列。图 3A-3D 的 2-686 位所示的氨基酸序列相应于 HCV-1 的 1027-1711 位的氨基酸。1 位显示了编码 Met 的起始密码子(ATG)。另外,在 HCV-1 的 1428 位通常出现的 Thr(图 3 的氨基酸 403 位)突变为 Pro,在 HCV-1 的 1429 位通常出现的 Ser(图 3 的氨基酸 404 位)突变为 Ile。然而,无论是天然序列(有或无 N-末端 Met)、所示的类似物(有或无 N-末端 Met)或其它类似物和片段都可用于本发明测定,只要该表位是采用保留或恢复其天然构象(如保留蛋白酶活性和任选地解旋酶活性)的方法产生的。Dasmahapatra 等,美国专利 No. 5,843,752 和 Zhang 等,美国专利 No. 5,990,276,他们都描述了 NS3/4a 的类似物。

[0113] 在约 1027-1207 位发现 NS3/4a 的 NS3 蛋白酶,相对于 HCV-1 编号,图 3 上为 2-182 位。NS3 蛋白酶的结构和活性位点是已知的,参见如 De Francesco 等, *Antivir. Ther.* (1998) 3 :99-109;Koch 等, *Biochemistry* (2001) 40 :631-640。通常为耐受的天然序列上的变化将是分子活性位点以外的那些变化。具体说,需要维持图 3 的氨基酸 1-或 2-155(有少量或仅有保守性替代)。155 位以外的氨基酸可以耐受更大的变化。另外,如果使用 NS3/4a 序列的片段,这些片段通常至少包括氨基酸 1-或 2-155,较佳地包括氨基酸 1-或 2-175,且最佳地包括氨基酸 1-或 2-182(有或无 N-末端 Met)。发现解旋酶结构域约在 HCV-1 的 1193-1657 位(图 3 的 207-632 位)。因此,如果需要解旋酶活性,就要维持分子的该部分有少量或仅有保守性变化。本领域技术人员根据 NS3/4a 的已知结构易确定耐受变化的其它区域。

[0114] 含有衍生自 NS3/4a 的表位的许多抗原是已知的,包括但不限于衍生自 c33c、c200、c100 和 5-1-1 区域、以及含有 NS3 表位的融合蛋白(如 c25)的抗原。

[0115] 关于得自其它 HCV 区域的这些和其它各种 HCV 表位的描述,可参见如 Houghton 等人,美国专利第 5350671;Chien 等人, *Proc. Natl. Acad. Sci. USA* (1992), 89 :10011-10015;Chien 等人, *J. Gastroent. Hepatol.* (1993), 8 :S33-39;Chien 等人,国际出版物 WO 93/00365 号;Chien, D. Y., 国际出版物 WO 94/01778;和美国专利 6280927、6150087。

[0116] 本文所述的免疫检测也使用多表位融合抗原(称为“MEFA”),如国际出版物 W001/96875、W001/09609、W097/44469 和美国专利 6514731、6428792 中所述的。用于本发明检测的 MEFA 包含多个衍生自图 1 和表 1 所述的各种病毒区域的任一区域的表位,如上所述。

[0117] 多个 HCV 抗原是单一、连续的氨基酸链的一部分,该链在自然中不出现。因此,表

位的线性顺序与它们在基因组中出现的线性顺序不同。用于本文的 MEFA 的序列的线性顺序较佳以最优的抗原性排列。优选的是,这些表位得自一种以上的 HCV 株,从而可在单一检测中提供检测多种 HCV 菌株的增强的能力。因此,用于本文的 MEFA 可含有衍生自上述多蛋白的各种免疫原性区域。

[0118] 如上所述,在特别优选的实施例中,将 NS3/4a 表位用作分离的 HCV 抗原。在此实施例中,MEFA 包括至少一种或多种衍生自 NS3/4a 区域的表位(线性的或构象的)。因此,MEFA 可包含衍生自多种一种或多种 HCV 分离物的 NS3/4a 区域的多个免疫优势表位。如果在多表位融合体中使用多个 NS3/4a 表位,则这些表位可以是相同或不同的。或者,融合抗原可包含一种或多种衍生自 NS3/4a 区域的表位,以及得自其它 HCV 区域(如不限于 HCV 核心、E1、E2、P7、NS4b、NS5a 和 NS5b 序列)的主要线性表位。

[0119] 含有衍生自 NS3/4a 区域的表位的多肽包括但不限于含有所有或部分 NS3、NS4a 和 NS3/4a 区域的多肽。得自这些区域的许多表位是已知的,包括但不限于衍生自 c33c、c200 和 c100 区域以及含有 NS3 表位的融合蛋白(如 c25)的抗原。这些和其它 NS3 表位可用于本发明的检测中,并且是本领域已知的,在以下的文献中有描述:如 Houghton 等人,美国专利第 5350671;Chien 等人,Proc. Natl. Acad. Sci. USA (1992), 89:10011-10015;Chien 等人, J. Gastroent. Hepatol. (1993), 8:S33-39;Chien 等人,国际出版物 W0 93/00365 号;Chien, D. Y., 国际出版物第 W0 94/01778;和美国专利 6280927、6150087。

[0120] 此外,已知为 5-1-1 的抗原决定簇部分存在于 NS4a 区域中(可参见图 1),它特别可用于本发明检测所使用的 MEFA 中。此抗原决定簇以三种不同的形式在三种不同的 HCV 病毒株中出现。因此,在本发明的一个优选实施例中,5-1-1 的所有三种形式出现在用于本发明免疫检测的多表位融合抗原中。

[0121] 用于 MEFA 的其它 HCV 表位包括任何上述各种表位,如衍生自 E2 的高变区衍生的表位,如跨越氨基酸 384-410 或 390-410 的区域或得自此区域的共有序列(如上述, Gly-Ser-Ala-Ala-Arg-Thr-Thr-Ser-Gly-Phe-Val-Ser-Leu-Phe-Ala-Pro-Gly-Ala-Lys-Gln-Asn, SEQ ID NO:7),其代表 1 型 HCV 基因组的氨基酸 390-410 的共有序列。本发明 MEFA 中存在的代表性 E2 表位包含跨越氨基酸 390-444 的杂交表位。这种杂交 E2 表位可包含代表氨基酸 390-410 的共有序列,该序列融合于 HCV E2 的氨基酸 411-444 的天然氨基酸序列。

[0122] 如上所述,可从各种 HCV 株衍生抗原。已知 HCV 的多种病毒株,且在融合蛋白中可用从任何这些株衍生的表位。已知任何特定种类的生物体的个体间互不相同,而且特定生物体如病毒可以有大量不同的株。例如,如上所述,HCV 包括至少 6 种基因型。这些基因型各包括等效的抗原决定簇。更具体地说,各株包括大量抗原决定簇,它们存在于该病毒的所有株中,但各病毒株之间略有差异。类似地,也可能存在来自不同 HCV 株的核心区域的等效抗原决定簇。通常,等效抗原决定簇的氨基酸序列具有高度同源性,序列对比时其同源性程度通常为 30%或更高,较佳地为 40%或更高。本发明的多拷贝表位还可包含同一表位的精确拷贝的多拷贝。

[0123] 用于本发明检测的代表性 MEFA 显示在图 5、7 和 9A-9C 中,并在国际出版物 W001/96875、W001/09609、W097/44469 和美国专利 6514731 和 6428792 中描述。用于本文的代表性 MEFA 包括命名为 MEFA 3、MEFA 5、MEFA 6、MEFA 7.1、MEFA 12、MEFA 13 和 MEFA 13.1 的 MEFA。应理解的是,这些 MEFA 仅仅是代表性的,衍生自 HCV 基因组的其它表位也可

用于本发明的检测,并可掺入到这些和其它 MEFA 中。

[0124] 图 6A 到 6F 显示了 MEFA12 的 DNA 序列及相应的氨基酸序列。图 5 显示了 MEFA 12 的结构通式,其如下:hSOD-E1(1 型)-E2HVR 共有序列(1a 型)-E2HVR 共有序列(1 和 2 型)-c33c 短(1 型)-5-1-1(1 型)-5-1-1(3 型)-5-1-1(2 型)-c100(1 型)-NS5(1 型)-NS5(1 型)-核心(1+2 型)-核心(1+2 型)。这种多拷贝表位包括以下氨基酸序列,相对于 HCV-1 编号(以下所述的氨基酸编号是按照 Choo 等人(1991)Proc. Natl. Acad. Sci. USA 88: 2451-2455 中的编号标志进行的,其中氨基酸 #1 是由核心区域的编码序列编码的第一个甲硫氨酸):超氧化物歧化酶的氨基酸 1-69(截短的 SOD,用于提高蛋白质的重组表达);来自 E1 区域多蛋白的氨基酸 303-320;多蛋白的氨基酸 390-410,代表 HCV-1a E2 超变区的共有序列;来自区域 E2 多蛋白的氨基酸 384-414,代表 HCV-1 和 HCV-2 的 E2 超变区的共有序列;HCV-1 多蛋白的氨基酸 1211-1457,其确定解旋酶;来自 5-1-1,氨基酸 1689-1735 的表位的三种拷贝,一种来自 HCV-1,一种来自 HCV-3,另一种来自 HCV-2,这些拷贝是 HCV 的三种不同病毒株的等效抗原性决定簇;HCV-1 的 HCV 多肽 C100,多蛋白的氨基酸 1901-1936;来自 HCV-1 的 NS5 区域的表位的两种精确拷贝,各具有 HCV 多蛋白的氨基酸 2278-2313;和来自核心区域三个表位的两种拷贝,两种来自 HCV-1 和一种来自 HCV-2,这些拷贝是由 HCV-1 的氨基酸 9-53 和 64-88 以及 HCV-2 的氨基酸 67-84 代表的等效抗原决定簇。

[0125] 表 2 显示了参考本文图 6A-6F(SEQ ID NO:3 和 4)MEFA12 中的各种表位的氨基酸位置。表中的编号以 HCV-1 为基础。见 Choo 等,Proc. Natl. Acad. Sci. USA 88:2451-2455。MEFA 13 和 13.1 还共有 MEFA12 上面特定的通式,具有分别在表 3 和表 4 指明的修改。

[0126] 表 2MEFA12

[0127]

MEFAaa#	5' 末端位点	表位	hcvaan#	病毒株
1-69	NcoI	截短的 hSOD		
72-89	MluI	E1	303-320	1
92-112	HindIII	E2HVR1a 共有序列	390-410	1
113-143		E2HVR1+2 共有序列	384-414	1,2
146-392	SpeI	C33C 短	1211-1457	1
395-441	SphI	5-1-1	1689-1735	1
444-490	NruI	5-1-1	1689-1735	3
493-539	ClaI	5-1-1	1689-1735	2
542-577	AvaI	C100	1901-1936	1

MEFAaa#	5' 末端位点	表位	hcvaa#	病毒株
580-615	XbaI	NS5	2278-2313	1
618-653	BglIII	NS5	2278-2313	1
654-741	NcoI	核心 表位	9-53, R47L 64-88 67-84	1 1 2
742-829	BaII	核心 表位	9-53, R47L 64-88 67-84	1 1 2

[0128] 表 3MEFA 13

[0129]

MEFAaa#	5' 末端位点	表位	hcvaa#	病毒株
1-156	NcoI	突变的 hSOD (aa70-72, ALA)		
161-178	MluI	E1	303-320	1
181-201	HindIII	E2HVR1a 共有序列	390-410	1
202-232		E2HVR1+2 共有序列	384-414	1, 2
235-451		C33C 短	1211-1457	1
454-500	HindIII	5-1-1PI 突变 [*]	1689-1735	1
503-549	NruI	5-1-1PI 突变 [*]	1689-1735	3
552-598	ClaI	5-1-1PI 突变 [*]	1689-1735	2
601-636	AvaI	C100	1901-1936	1
639-674	XbaI	NS5	2278-2313	1
677-712	BglIII	NS5	2278-2313	1

MEFAaa#	5' 末端位点	表位	hcvaa#	病毒株
713-800		核心 表位	9-53, R47L 64-88 67-84	1 1 2
801-888		核心 表位	9-53, R47L 64-88 67-84	1 1 2

[0130] * 通过消除可能由 NS3/4a 重组蛋白靶向的切割位点 (CS 或 CA) 修饰 5-1-1 表位。序列已变成 PI, 代替 CS 或 CA。

[0131] 表 4MEFA 13.1

[0132]

MEFAaa#	5' 末端位点	表位	hcvaa#	病毒株
1-86	NcoI	突变的 hSOD (aa70-72, ALA)		
89-106	MluI	E1	303-320	1
109-129	HindIII	E2HVR1a 共有序列	390-410	1
130-160		E2HVR1+2 共有序列	384-414	1, 2
163-379		C33C 短	1211-1457	1
382-428	HindIII	5-1-1PI 突变*	1689-1735	1
431-477	NruI	5-1-1PI 突变*	1689-1735	3
480-526	ClaI	5-1-1PI 突变*	1689-1735	2
529-564	AvaI	C100	1901-1936	1
567-602	XbaI	NS5	2278-2313	1
605-640	BglIII	NS5	2278-2313	1

MEFAaa#	5' 末端位点	表位	hcvaa#	病毒株
641-728		核心 表位	9-53, R47L 64-88 67-84	1 1 2
729-816		核心 表位	9-53, R47L 64-88 67-84	1 1 2

[0133] * 通过消除可能由 NS3/4a 重组蛋白靶向的切割位点 (CS 或 CA) 修饰 5-1-1 表位。序列已变成 PI, 代替 CS 或 CA。

[0134] 图 8A 到 8F (SEQ ID NO :5 和 6) 显示了另一代表性多表位融合抗原 MEFA7.1 的 DNA 序列及相应的氨基酸序列。图 7 显示了 MEFA7.1 的结构通式, 其如下: hSOD-E1 (1 型)-E2HVR 共有序列 (1a 型)-E2HVR 共有序列 (1 和 2 型)-解旋酶 (1 型)-5-1-1 (1 型)-5-1-1 (3 型)-5-1-1 (2 型)-c100 (1 型)-NS5 (1 型)-NS5 (1 型)-核心 (1+2 型)-核心 (1+2 型)。这种多拷贝表位包括以下氨基酸序列, 相对于 HCV-1 编号 (以下所述的氨基酸编号是按照 Choo 等 (1991) Proc. Natl. Acad. Sci. USA 88 :2451-2455 中提供的编号标志, 其中氨基酸 #1 是由核心区域的编码序列编码的第一个甲硫氨酸): 超氧化物歧化酶的氨基酸 1-156 (SOD, 用于提高蛋白质的重组表达); E1 区域多蛋白的氨基酸 303-320; 多蛋白的氨基酸 390-410, 代表 HCV-1a E2 高变区的共有序列; 区域 E2 多蛋白的氨基酸 384-414, 代表 HCV-1 和 HCV-2 的 E2 高变区的共有序列; HCV-1 多蛋白的氨基酸 193-1658, 其确定解旋酶; 5-1-1, 氨基酸 1689-1735 的表位的三种拷贝, 一种来自 HCV-1, 一种来自 HCV-3, 另一种来自 HCV-2, 这些拷贝是 HCV 的三种不同病毒株的等效抗原决定簇; HCV-1 的 HCV 多肽 C100, 多蛋白的氨基酸 1901-1936; HCV-1 的 NS5 区域的表位的两种精确拷贝, 各具有 HCV 多蛋白的氨基酸 2278-2313; 和核心区域的表位的两种拷贝, 一种来自 HCV-1 和一种来自 HCV-2, 这些拷贝是由 HCV-1 的氨基酸 9-32、39-42 和 64-88 以及 HCV-2 的氨基酸 67-84 代表的等效抗原决定簇。

[0135] 表 5 显示了参考本文图 8A-8F (SEQ ID NO :5 和 6) 的各种表位的氨基酸位置。

表 5 MEFA 7.1

mefa aa#	5'末端位点	表位	hcv aa#	菌株
1-156	<i>Nco</i> 1	hSOD		
159-176	<i>Eco</i> R1	E1	303-320	1
179-199	<i>Hind</i> 111	E2 HVR1a 共有序列	390-410	1
200-230		E2 HVR1+2 共有序列	384-414	1+2
231-696	<i>Sal</i> 1	解旋酶	1193-1658	1
699-745	<i>Sph</i> 1	5-1-1	1689-1735	1
748-794	<i>Nru</i> 1	5-1-1	1689-1735	3
797-843	<i>Cl</i> a1	5-1-1	1689-1735	2
846-881	<i>Ava</i> 1	C100	1901-1936	1
884-919	<i>Xba</i> 1	NS5	2278-2313	1
922-957	<i>Bgl</i> 11	NS5	2278-2313	1
958-1028	<i>Nco</i> 1	核心 表位	9-32, 39-42 64-88 67-84	1 1 2
1029-1099	<i>Bal</i> 1	核心 表位	9-32, 39-42, 64-88 67-84	1 1 2

[0136]

[0137] 在一种检测模式中,如下文进一步所述的,将样品与固相支持物混合。该固相支持物包含分离的 HCV 抗原(如一种或多种 NS3/4a 构象表位)或上述 MEFA,该 MEFA 包含一种或多种衍生自与 HCV 抗原相同的多蛋白区域(如 NS3/4a 区域)的表位。如上所述,已知许多包含这些表位的抗原,包括但不限于衍生自 c33c 和 c100 区域以及含有 NS3 表位的融合蛋白(如 c25)的抗原,以及称为 5-1-1 部分存在于 NS4a 区域内的抗原决定簇(可参见图 1)。这些和其它 NS3/4a 表位可用于本发明检测,并且是本领域已知的,如在以下文献中有描述:Houghton 等人,美国专利第 5350671;Chien 等人,Proc. Natl. Acad. Sci. USA(1992), 89:10011-10015;Chien 等人,J. Gastroent. Hepatol. (1993), 8:S33-39;Chien 等人,国际出版物 WO 93/00365 号;Chien, D. Y., 国际出版物第 WO 94/01778;和美国专利 6280927、6150087。

[0138] 如果样品受到 HCV 感染,针对固相支持物上存在的表位的 HCV 抗体将与该固相支持物组分结合。将也与该得自生物样品中的俘获的 HCV 抗体反应的可检测的标记抗原加入液相中。例如,如果结合于该固相支持物的抗原是 NS3/4a 构象表位,则用于液相中的可检测的标记抗原是包含 NS3/4a 表位的 MEFA。如果结合于固相支持物的抗原是包含 NS3/4a 表位的 MEFA,则用于液相的该可检测的标记抗原含有构象 NS3/4a 表位。

[0139] 图2描述了本发明的代表性检测。如该图所示,固相支持物包含构象 NS3/4a 表位。将生物样品加到该固相支持物中。针对存在于该样品中的 NS3/4a 表位的 HCV 抗体将与该固相支持物上的该 NS3/4a 构象表位结合。然后加入辣根过氧化物酶 (HRP) 标记的含有样品抗体与之结合的表位的 MEFA 12。MEFA 12 与也与 NS3/4a 构象表位结合的抗体结合。洗掉未结合的成分,标记物的检测表明存在 HCV 感染。

[0140] 由于使用了两种与样品抗体结合的抗原而使得可使用较大体积的样品,上述抗原 / 抗体 / 抗原夹心检测是特别有利的。此外,该检测可快速完成。

[0141] 用于 HCV 免疫检测的抗原的生产

[0142] 如上所述,通常用重组方法制备本发明的分子。因此,可以用标准分子生物学方法制备用于本发明的编码 HCV 抗原的多核苷酸。例如,用重组方法可以获得编码上述分子的多核苷酸序列,例如通过从表达基因的细胞筛选 cDNA 和基因组文库,或通过从已知的包括该基因的载体衍生该基因。另外,用本领域所述的方法,如 Houghton 等的美国专利 No. 5,350,671 所述的方法,从病毒核酸分子直接分离出所需的基因。也可合成而非克隆制备感兴趣的基因。可用适当的特定序列的密码子设计该分子。然后将用标准方法制备的重叠的寡核苷酸装配完整的序列,并装配入完整的编码序列中。参见如 Edge (1981) Nature 292 :756 ;Nambair 等 (1984) Science 223 :1299 ;和 Jay 等 (1984) J. Biol. Chem. 259 :6311。

[0143] 因此,从携带所需序列的载体可获得具体的核苷酸序列,或用本领域已知的各种寡核苷酸合成方法,如定位诱变和聚合酶链式反应 (PCR),完全或部分合成。参见如 Sambrook,同上。具体说,获得编码所需序列的核苷酸序列的方法是用常规的自动多核苷酸合成仪制备的退火补集的重叠的合成的寡核苷酸,然后用适当的 DNA 连接酶连接,并通过 PVR 扩增连接的核苷酸序列。参见如 Jayaraman 等 (1991) Proc. Natl. Acad. Sci. USA 88 :4084-4088。另外,在本发明中也可使用寡核苷酸定向的合成 (Jones 等, (1986) Nature 54 :75-82)、先有核苷酸区域的寡核苷酸定向诱变 (Riechmann 等, (1988) Nature 332 :323-327 和 Verhoeyen 等 (1988) Science 239 :1534-1536) 和用 T4DNA 聚合酶进行的酶促补平带缺口的寡核苷酸 (Queen 等 (1989) Proc. Natl. Acad. Sci. USA 86 :10029-10033),以提供抗原结合能力变化或提高的和 / 或免疫原性降低的分子。

[0144] 一旦制备或分离了编码序列,就可将这些序列克隆入任何合适的载体或复制子中。对本领域技术人员而言,各种克隆载体是已知的,且适当的克隆载体的筛选只是选择问题。合适的载体包括 (但不限于):质粒、噬菌体、转座子、粘粒、染色体或当与适当的控制元件结合时能复制的病毒。

[0145] 然后将克隆序列置于合适的控制元件的控制下,这取决于用于表达的系统。因此,可以将编码序列置于启动子、核糖体结合位点 (用于细菌表达) 和任选地操纵子的控制下,从而由合适的转化体将感兴趣的 DNA 序列转录到 RNA 中。该编码序列可以包含或不包含信号肽或前导序列 (随后可由宿主在翻译后加工除去)。参见如美国专利 4,431,739、4,425,437、4,338,397。

[0146] 除了控制序列外,可以添加调节序列,从而可以相对于宿主细胞的生长调节序列的表达。调节序列是本领域技术人员已知的,其实例包括那些能导致应答化学或物理刺激 (包括调节化合物的存在) 启动或关闭基因表达的调节序列。载体中还可存在其它类型的调节元件。例如,可以使用增强子元件以增加构建物的表达水平。实例包括 SV40 早期基因

增强子 (Dijkema 等 (1985)EMBO J. 4 :761) ;从 Rous 肉瘤病毒的长末端重复序列 (LTR) 衍生的增强子 / 启动子 (Gorman 等 (1982)Proc. Natl. Acad. Sci. USA 79 :6777) ;和从人 CMV 衍生的元件 (Boshart 等, (1985)Cell 41 :521), 如 CMV 内含子 A 序列中包括的元件 (美国专利 No. 5, 688, 688)。表达盒还可包括在合适的宿主细胞中自主复制的复制起点、一种或多种可选择的标记物、一个或多个限制酶切位点、高拷贝数的潜势 (potential) 和强启动子。

[0147] 构建表达载体, 使具体的编码序列位于该具有适当调节序列的载体中, 与控制序列有关的编码序列的位置和取向使编码序列在控制序列的“控制”下转录 (即, 结合于控制序列中 DNA 分子的 RNA 聚合酶转录该编码序列)。可能需要对编码感兴趣分子的序列进行修饰来实现此目的。例如, 在一些情况中可能需要修饰该序列, 从而使其连接于适合取向的控制序列, 即维持读框。在插入到载体之前, 控制序列和其它调节序列可连接于编码序列。或者, 可以将编码序列直接克隆入已包含控制序列和合适的限制酶切位点的表达载体中。

[0148] 如上所述, 还可能需要制备感兴趣的抗原的突变体或类似物。尤其对 NS3/4a 而言。如此制备的方法在如下文献中有描述, 如 Dasmahapatra 等的美国专利 No. 5, 843, 752 和 Zhang 等的美国专利 No. 5, 990, 276。通过缺失一部分编码感兴趣的多肽的序列、插入序列、和 / 或替代该序列内的一个或多个核苷酸, 可以制备用于分析的这种和其它 HCV 蛋白的突变体或类似物。修饰核苷酸序列的方法, 如定向诱变等是本领域技术人员所熟知的。参见如 Sambrook 等, 上述 ;Kunkel, T. A. (1985)Proc. Natl. Acad. Sci. USA (1985) 82 : 448 ;Geisselsoder 等 (1987)BioiTechniques 5 :786 ;Zoller 和 Smith (1983) MethodsEnzymol. 100 :468 ;Dalbie-McFarland 等 (1982)Proc. Natl. Acad. Sci USA 79 : 6409。

[0149] 这些分子可以在各种系统中表达, 包括本领域熟知的昆虫、哺乳动物、细菌、病毒和酵母表达系统。

[0150] 例如, 昆虫细胞表达系统如杆状病毒系统是本领域技术人员已知的, 描述于如 Summers 和 Smith, Texas Aguricultural Experiment Station Bulletin No. 1555 (1987)。用于杆状病毒 / 昆虫细胞表达系统的材料和方法都是可以试剂盒形式购得的, 尤其是 Invitrogen, San Diego CA (“MaxBac” 试剂盒)。类似地, 细菌和哺乳动物细胞表达系统也是本领域熟知的, 其描述见如 Sambrook 等, 同上。酵母表达系统也是本领域熟知的, 其描述见如《酵母遗传工程》(Yeast Genetic Engineering) (Barr 等编辑, 1989) Butterworths, London。

[0151] 许多用于上述系统的合适的宿主细胞也是已知的。例如, 哺乳动物细胞系是本领域已知的, 其包括可从美国典型培养物保藏中心 (ATCC) 获得的无限增殖的细胞系, 如 (但并非限于) 中国仓鼠卵巢细胞 (CHO)、HeLa 细胞、幼仓鼠肾 (BHK) 细胞、猴肾细胞 (COS)、人胚肾细胞、人肝细胞癌细胞 (如 Hep G2)、Madin-Darby 牛肾 (“MDBK”) 细胞等。类似地, 在本发明的表达构建物中也可使用细菌宿主, 如大肠杆菌 (E. coli)、枯草杆菌 (Bacillus subtilis) 和链球菌 (Streptococcus spp.)。在本发明中还可使用酵母宿主, 特别包括酿酒酵母 (Saccharomyces cerevisiae)、麦芽糖假丝酵母 (Candida maltosa)、白色念珠菌 (Candida albicans)、多形汉逊酵母 (Hansenula polymorpha)、脆壁克鲁维酵母 (Kluyveromyces fragilis)、乳酸克鲁维酵母 (Kluyveromyces lactis)、季也蒙氏毕赤酵母 (Pichia guilliermondii)、巴斯德毕赤酵母 (Pichia pastoris)、粟酒裂殖酵

母 (*Schizosaccharomyces pombe*) 和解脂耶罗威亚酵母 (*Yarrowia lipolytica*)。可与杆状病毒表达载体一起使用的昆虫细胞特别包括埃及伊蚊 (*Aedes aegypti*)、苜蓿银纹夜蛾 (*Autographa californica*)、家蚕 (*Bombyx mori*)、黑腹果蝇 (*Drosophila melanogaster*)、草地夜蛾 (*Spodoptera frugiperda*) 和粉纹夜蛾 (*Trichoplusia ni*)。

[0152] 用本领域熟知的各种基因送递的方法, 可将包含感兴趣的核苷酸序列的核酸分子稳定地整合入宿主细胞基因组中或在合适的宿主细胞中的稳定的附加型元件上维持。参见如美国专利 No. 5, 399, 346。

[0153] 根据所选用的表达系统和宿主, 在表达蛋白质的条件下, 培养用如上所述的表达载体转化的宿主细胞制备该分子。然后从宿主细胞分离出表达的蛋白质并纯化。如果表达系统将蛋白质分泌到生长培养基中, 则直接从培养基纯化产物。如果不是分泌的, 则从细胞裂解液分离。合适的培养条件和回收方法的选择都是本领域技术人员能力之内的。

[0154] 已描述了各种 HCV 抗原的重组制备。参见如 Houghton 等, 美国专利 No. 5, 350, 671; Chien 等, *J. Gastroent. Hepatol.* (1993) 8 :S33-39; Chien 等, 国际出版物 No. WO 93/00365; Chien, D. Y., 国际出版物 No. WO 94/01778。

[0155] 免疫诊断试验

[0156] 一旦产生, 将上述 HCV 抗原置于合适的固相支持物上, 用于本发明的免疫试验。为了本发明的目的, 固相支持物可以是任何物质, 它是不溶性的基质, 并可具有粗糙或半粗糙的表面。示范性的固相支持物包括 (但并非限于) 基质如硝化纤维素 (如以膜或微量滴定孔形式); 聚氯乙烯 (如板或微量滴定孔); 聚苯乙烯乳胶 (如珠或微量滴定板); 聚偏 1,1-二氟乙烯; 重氮化纸; 尼龙膜; 活化珠、磁力感应珠等。具体的支持物包括板、小球、皿、毛细管、空心纤维、针、钉、固体纤维、纤维素珠、微孔玻璃珠、硅胶、聚苯乙烯珠 (任选地与二乙烯苯交联)、嫁接的共聚珠、聚丙烯酰胺珠、乳胶珠、二甲基丙烯酰胺珠 (任选地与 N-N' -二-烯丙基乙二胺交联) 和用疏水聚合物包膜的玻璃珠。

[0157] 如果需要, 易将待加到固相支持物上的分子功能化, 产生苯乙烯或丙烯酸部分, 能让该分子掺入聚苯乙烯、聚丙烯酸酯或其它聚合物如聚酰亚胺、聚丙烯酰胺、聚乙烯、聚乙炔、聚二乙炔、聚苯撑-乙烯、多肽、多糖、聚砜、聚吡咯、聚咪唑、聚噻吩、聚醚、环氧树脂、石英玻璃、硅胶、硅氧烷、多磷酸盐、水凝胶、琼脂糖、纤维素等。

[0158] 在一实例中, 在使分子充分地固定于支持物的合适的结合条件下, 首先将固相支持物与分离的 HCV 抗原或 MEFA (在此称为“固相成分”) 反应。有时, 可以通过首先将抗原与具有较好固相结合特性的蛋白质偶联, 从而提高支持物的固定化。适合的偶联蛋白包括 (但并非限于) 大分子如血清白蛋白, 包括: 牛血清白蛋白 (BSA)、匙孔蛾血蓝蛋白、免疫球蛋白分子、甲状腺球蛋白、卵白蛋白和本领域技术人员已知的其它蛋白质。其它可用于将分子结合于支持物的试剂包括多糖、聚乳酸、聚羟基乙酸、多聚氨基酸、氨基酸共聚物等。将这些分子与抗原偶联的分子和方法是本领域技术人员所熟知的。参见, 如 Brinkley, M. A. (1992) *Bioconjugate Chem.* 3 :2-13; Hashida 等人 (1984) *J. Appl. Biochem.* 6 :56-63; 和 Anjaneyulu 和 Staros (1987) *International J. of Peptide and Protein Res.* 30 : 117-124。

[0159] 在固相支持物与固相成分反应后, 冲洗除去支持物上未固定的固相成分, 然后在适合的结合条件下, 将与支持物结合的成分与怀疑含有 HCV 抗体 (本文将其称为“配体分

子”)的生物样品接触。冲洗除去任何未结合的配体分子后,在合适的结合条件下加入第二 HCV 抗原(分离的 HCV 抗原或 MEFA,取决于结合于固相支持物上的是哪种抗原)。此第二抗原在此称为“液相成分”。加入的抗原带有如上所述的可检测标记,并与已与结合于支持物上的抗原反应的配体分子结合。因此,配体分子与固相成分以及液相成分结合。冲洗除去未结合的配体分子和液相成分。由此,标记物的存在表明生物样品中存在 HCV 抗原。

[0160] 更具体的,可用 ELISA 法,其中微量滴定板的孔用固相成分包涂。然后在包涂的孔中加入含有或怀疑含有配体分子的生物样品。一段足够使配体分子与固定的固相成分结合的培育时间后,洗涤平板以除去未结合的部分,加入可检测标记的液相成分。这些分子能与任何捕获的样品抗体反应,洗涤平板并用本领域熟知的方法检测标记物的存在。

[0161] 上述测定试剂,包括具有结合抗原的免疫测定固相支持物和与捕获样品反应的抗原,可以具有合适的说明书和其它必需试剂的试剂盒提供,以进行如上所述的免疫测定。视所用的特定免疫测定而定,试剂盒还可含有合适的标记和其它包装的试剂和材料(即洗涤缓冲液等)。标准免疫测定,例如上述的测定可用这些试剂盒进行。

[0162] III. 实验

[0163] 下面是实施本发明所需的特定实施例的例子。这些实施例仅为了说明的目的,而不是要以任何形式限制本发明的范围。

[0164] 已做出努力确保使用的数据(例如量和温度等)的准确性,但当然应允许一些实验错误和偏差。

[0165] 实施例 1

[0166] 具有 Thr 到 Pro 和 Ser 到 Ile 取代的 NS3/4a 构象表位的制备

[0167] 如下获得 NS3/4a 的构象表位。该表位具有图 3A-3D(SEQ ID NO:1 和 2)所述的序列,在 403 位(HCV-1 全长序列的氨基酸 1428)和 404 位(HCV-1 全长序列的氨基酸 1429)与天然序列不同。具体说,在天然序列 1428 位上通常出现的 Thr 已突变成 Pro,在天然序列 1429 位上出现的 Ser 已突变为 Ile。

[0168] 具体说,所用的酵母表达载体是 pBS24.1。此酵母表达载体含有 2 μ 用于在酵母中自主复制序列和反向重复序列(IR),确保转录终止的 α -因子终止子和用于选择的酵母 leu2-d 和 URA3。还存在 ColE1 的复制起点和 β -内酰胺酶基因,用于大肠杆菌中的增殖和选择(Pichuanes 等,1996,“酵母中异源基因产物的表达”,在“蛋白质工程:设计和生产指南”,第 5 章,J.L.Cleland 和 C.Craik 编辑,Wiley-Liss,Inc.,纽约,N.Y.,第 129-161 页。

[0169] 如下制备质粒 pd.hcv1a.ns3ns4aPI,其编码用于本发明免疫测定的代表性 NS3/4a 表位。使用两步流程。首先,将以下的 DNA 片段连接在一起:(a)合成的寡核苷酸,其提供 5' HindIII 克隆位点,随后接序列 ACAAACAAA(SEQ IDNO:8),起始密码子 ATG,和 HCV1a 的密码子,从氨基酸 1027 开始继续到氨基酸 1046 的 BglII 位点;(b)来自 pAcHLTns3ns4aPI 的 683bp 的 BglII-ClaI 限制性片段(编码氨基酸 1046-1274);和(c)HindIII 和 ClaI 消化、去磷酸化和凝胶纯化的 pSP72 载体(Promega,Madison,WI,GenBank/EMBL 登录号 X65332)。质粒 pAcHLTns3ns4aPI 衍生自 pAcHLT,从 BD Pharmingen(San Diego,CA)购得的一种杆状病毒表达载体。特别是制备了 pAcHLT EcoRI-PstI 载体和如下片段:EcoRI-AlwNI,935bp,对应于 HCV-1 基因组的氨基酸 1027-1036,AlwNI-SacII,247bp,对应于 HCV-1 基因组的氨基酸 1336-1419;HinfI-BglII,175bp,对应于 HCV-1 基因组的氨基酸 1449-1509;BglII-PstI,

619bp, 对应于 HCV-1 基因组的氨基酸 1510-1711, 加上转录终止密码子。SacII-HinfI 合成产生的 91bp 片段 (对应于 HCV-1 基因组的氨基酸 1420-1448, 含有 PI 突变 (Thr-1428 突变成 Pro, Ser-1429 突变成 Ile)) 与上述的 175bp HinfI-BglI 片段和 619bp 的 BglI-PstI 片段连接, 并亚克隆入用 SacII 和 PstI 消化的 pGEM-5Zf(+) 载体。pGEM-5Zf(+) 是市售的大肠杆菌载体 (Promega, Madison, WI, GenBank/EMBL 登录号 X65308)。在转化感受态 HB101 细胞, 单独克隆的小筛选分析并证实序列后, 凝胶纯化 pGEM5. PI 克隆 2 的 885bp 的 SacII-PstI 片段。该片段与上述的 EcoRI-AlwNI 935bp 片段、AlwNI-SacII 247bp 片段和 pAcHLT EcoRI-PstI 载体连接。得到的构建物命名为 pAcHLTns3ns4aPI。

[0170] 上述连接混合物转化入 HB 101- 感受态细胞, 并铺在含有 100 微克 / 毫升氨苄青霉素的 Luria 琼脂平板上。单独克隆的少量制备物分析鉴定了推定的阳性, 扩增了其中两个。用 Qiagen Maxiprep 试剂盒制备 pSP72 1aHC, 克隆 #1 和 #2 的质粒 DNA, 并测序。

[0171] 接着, 连接下列片段:(a) 来自 pSP721aHC#1 的 761bp HindIII-ClaI 片段 (pSP72. 1aHC 是通过连接下列片段产生的: 用 HindIII 和 ClaI 消化的 pSP72, 提供 5' HindIII 克隆位点的寡核苷酸, 接着是序列 ACAAACAAA (SEQ ID NO:8), 起始密码子 ATG 和 HCV1a 的密码子, 从氨基酸 1027 开始, 直到氨基酸 1046 处的 BglIII 位点, 和来自 pAcHLTns3ns4aPI 的 683bp BglIII-ClaI 限制性片段 (编码氨基酸 1046-1274)); (b) 酵母杂交启动子 ADH2/GAPDH 的 1353bp 的 BamHI-HindIII 片段; (c) 来自 pAcHLTns3ns4aPI 的 1320bp ClaI-SalI 片段 (编码 HCV1a 氨基酸 1046-1711, 其中 Thr1428 突变成 Pro, Ser1429 突变成 Ile); 和 (d) 已用 BamHI 和 SacI 消化, 去磷酸化和凝胶纯化的 pBS23. 1 酵母表达载体。该连接混合物转化入感受态 HB101 细胞, 并铺在含有 100 微克 / 毫升氨苄青霉素的 Luria 琼脂平板上。单独克隆的小量制备物分析鉴定到具有预期的 3446bp 的 BamHI-SalI 插入的克隆, 它带有 ADH2/GAPDH 启动子, 起始密码子 ATG 和氨基酸 1027-1711 的 HCV1a NS3/4a (显示图 3A-3D 的氨基酸 1-686), Thr1428 (图 3A-3D 的氨基酸位置 403) 突变成 Pro, Ser1429 (图 3A-3D 的氨基酸位置 404) 突变成 Ile。构建物称为 pd. HCV1a. ns3ns4aPI (见图 4)。

[0172] 用 pd. HCV1a. ns3ns4aPI 转化酿酒酵母株 AD3, 在培养基中的葡萄糖耗尽后检查单个转化子的表达。在酵母中高水平表达重组蛋白质, 如考马斯蓝染色检测到的, 并使用针对 NS3 的解旋酶结构域的多克隆抗体, 通过免疫印迹分析证实。

[0173] 实施例 2

[0174] NS3/4a 构象表位的纯化

[0175] 如下纯化 NS3/4a 构象表位。如上所述收集表达 NS3/4a 表位的酿酒酵母细胞。将细胞悬浮在裂解缓冲液 (50mM Tris pH8.0, 150mM NaCl, 1mM EDTA, 1mM PMSF, 0.1 μ M 抑胃酶肽, 1 μ M 亮抑酶肽), 并以细胞: 缓冲液: 0.5 毫米玻璃珠 1: 1: 1 的比例, 在 Dyno-Mill (Wab Willy A. Bachofen, Basel, Switzerland) 或等效装置中裂解。裂解液于 30100xg、4°C 离心 30 分钟, 并在洗涤缓冲液 (6ml/g 起始细胞沉淀重量) 中加入含有不溶性细胞组分的沉淀, 室温振摇 15 分钟。洗涤缓冲液含有 50mM NaPO₄ pH8.0、0.3M NaCl、5mM β -巯基乙醇、10% 甘油、0.05% 辛基葡糖苷、1mM EDTA、1mM PMSF、0.1 μ M 抑胃酶肽、1 μ M 亮抑酶肽。30100xg、4°C 离心 30 分钟除去细胞碎片。丢弃上清液并保留沉淀。

[0176] 如下从沉淀中抽提蛋白质。加入 6ml/g 抽提缓冲液并在室温振摇 15 分钟。抽提

缓冲液含有 50mM Tris pH8.0、1M NaCl、5mM β -巯基乙醇、10%甘油、1mMEDTA、1mM PMSF、0.1 μ M 抑胃酶肽、1 μ M 亮抑酶肽。这在 30100xg、4℃离心 30 分钟。保留上清液并用下列配方加入达到 17.5%的硫酸铵：上清液的体积 (ml) 乘以 x%硫酸铵 / (1-x%硫酸铵) = 加到上清液中的 4.1M 饱和硫酸铵的毫升数。一边在冰上搅拌，一边滴加硫酸铵，溶液在冰上搅拌 10 分钟。17700xg 4℃离心溶液 30 分钟，保留沉淀，储藏在 2℃ -8℃达 48 小时。

[0177] 重新悬浮沉淀，并 4℃如下过 Poly U 柱 (Poly U Sepharose 4B, AmershamPharmacia)。沉淀重新悬浮在 6ml Poly U 平衡缓冲液 / 克沉淀重量中。平衡缓冲液含有 25mM HEPES pH8.0、200mM NaCl、5mM DTT (新鲜加入)、10%甘油、1.2%辛基葡糖苷。4℃振摇溶液 15 分钟，31000xg、4℃离心 30 分钟。

[0178] 制备 Poly U 柱 (1ml 树脂 / 克起始沉淀重量)。线性流速为 60cm/hr，填充流速为 133% 60cm/hr。用平衡缓冲液平衡柱，并在平衡好的柱上加入重悬浮硫酸铵沉淀的上清液。用平衡缓冲液将柱洗到基线，在一步洗脱中用下列 Poly U 洗脱缓冲液洗脱蛋白质：25mM HEPES pH8.0、1M NaCl、5mM DTT (新鲜加入)、10%甘油、1.2%辛基葡糖苷。柱洗脱物在 SDS-PAGE (考马斯蓝染色) 上展开，冷冻等份并储藏在 -80℃。用蛋白质印迹，使用针对 NS3 蛋白酶结构域的多克隆抗体和针对 5-1-1 表位 (HCV 4a) 的单克隆抗体证实了 NS3/4a 表位的存在。

[0179] 另外，在纯化过程中如下监测蛋白酶的酶活性。用 90 微升反应缓冲液 (25mMTris, pH7.5, 0.15M NaCl, 0.5mM EDTA, 10%甘油, 0.05 正十二烷基 B-D- 麦芽糖苷, 5mM DTT) 稀释 NS4A 肽 (KKGSVVIVGRIVLSGKPAIIIPKK, SEQ IDNO :9) 和含有 NS3/4a 构象表位的样品，并在室温下混合 30 分钟。将 90 微升混合物加到微量滴定板 (Costar, Inc., Corning, NY) 中，加入 10 微升 HCV 底物 (AnaSpec, Inc., San Jose CA)。混合平板并在 Fluostar 平板阅读仪上阅读。结果表示为相对荧光单位 (RFU) / 分钟。

[0180] 用这些方法，1M NaCl 抽提的产物含有 3.7RFU/ 分钟的活性，硫酸铵沉淀物具有 7.5RFU/ 分钟的活性，Poly U 纯化的产物具有 18.5RFU/ 分钟的活性。

[0181] 实施例 3

[0182] NS3/4a 构象表位对变性 NS3/4a 的免疫反应性

[0183] 将如上所述产生的 NS3/4a 构象表位的免疫反应性与在 NS3/4a 构象表位制备物中加入 SDS 到最终浓度为 2%的变性的 NS3/4a 进行比较。将变性的 NS3/4a 和构象 NS3/4a 涂覆到微量滴定板上 (如上所述)。也将 c200 抗原 (Hepatology (1992) 15 :19-25, 在 ORTHO HCV 3.0 版 ELISA 测试系统, Ortho-Clinical Diagnostics, Raritan, New Jersey 中可得) 涂覆到微量滴定板上。用 c200 作为比较，假定它是非构象的，由于在其配方中存在还原剂 (DTT) 和去污剂 (SDS)。

[0184] 针对两个早期 HCV 血清转化组 PHV904 和 PHV914 (市售人血样，购自 Boston Biomeida, Inc., West Bridgewater, MA) 测试了免疫反应性。结果如表 6 所示。数据提示 NS3/4a (和 c200) 变性或线性化的形式不与 NS3/4a 构象表位一样早地检测到早期血清转化组。

[0185] 表 6

[0186]

	NS3/4a 对变性的 NS3/4a							
	*储存液 NS3/4a 中掺入 2%SDS							
		NS3/4a	dNS3/4a*	c200		NS3/4a	dNS3/4a*	c200
		OD	OD	OD		s/co	s/co	s/co
HCV	PHV 904-1	0.012	0.012	0.009		0.02	0.02	0.01
血清转化	PHV 904-2	0.011	0.009	0.008		0.02	0.01	0.01
	PHV 904-3	1.124	0.071	0.045		1.80	0.11	0.07
	PHV 904-4	2.401	0.273	0.129		3.85	0.44	0.21
	PHV 904-5	3.022	0.793	0.347		4.85	1.28	0.57
	PHV 904-6	2.711	1.472	0.774		4.35	2.37	1.28
	PHV 904-7	3.294	1.860	0.943		5.28	2.99	1.55
	PHV 914-1	0.006	0.004	0.001		0.01	0.01	0.00
	PHV 914-2	0.005	0.004	0.002		0.01	0.01	0.00
	PHV 914-3	0.098	0.003	0.001		0.16	0.00	0.00
	PHV 914-4	1.118	0.006	0.004		1.79	0.01	0.01
	PHV 914-5	2.035	0.044	0.022		3.26	0.07	0.04
	PHV 914-6	2.092	0.074	0.025		3.35	0.12	0.04
	PHV 914-7	2.519	0.281	0.132		4.04	0.45	0.22
	PHV 914-8	2.746	0.907	0.500		4.40	1.46	0.82
	PHV 914-9	3.084	1.730	0.931		4.94	2.78	1.53
HCV 3.0	阴性对照	0.023	0.024	0.008				
对照	阴性对照	0.027	0.024	0.007				
	阴性对照	0.021	0.017	0.005				
	平均	0.024	0.022	0.007				
	截留	0.624	0.622	0.607				
	阳性对照	1.239	0.903	0.575		1.99	1.45	0.95
	阳性对照	1.445	0.916	0.614		2.32	1.47	1.01

[0187] 还使用标准程序制备的针对 NS3/4a 的单克隆抗体测试了构象表位的免疫反应性。然后在针对 NS3/4a 和变性的 NS3/4a 和 c200 抗原的 ELISA 测定中测试这些单克隆抗体。数据显示抗 -NS3/4a 单克隆抗体与 NS3/4a 和变性的 NS3/4a 以与表 7 显示的血清转化组相似的方式反应。该结果还提供了 NS3/4a 天然是构象表位的进一步证据, 因为可制备反应性与早期 c33c 血清转化组相似的单克隆抗体。

[0188]

表 7				
		平板		
		NS3/4a	dNS3/4a	c200
单克隆		OD	OD	OD
4B9/E3	1:100	1.820	0.616	0.369
	1:1000	1.397	0.380	0.246
	1:10000	0.864	0.173	0.070
	1:20000	0.607	0.116	0.085
5B7/D7	1:100	2.885	0.898	0.436
	1:1000	2.866	0.541	0.267
	1:10000	1.672	0.215	0.086
	1:20000	1.053	0.124	0.059
1A8/H2	1:100	1.020	0.169	0.080
	1:1000	0.921	0.101	0.043
	1:10000	0.653	0.037	0.013
	1:20000	0.337	0.027	0.011

[0189] 实施例 4

[0190] 用 HCV 抗原涂覆固相支持物

[0191] 如下将 HCV NS3/4a 构象表位或 MEFA 抗原涂覆到平板上。用 0.22 μ 过滤装置过滤 HCV 涂覆缓冲液 (50mM NaPO₄ pH 7.0, 2mM EDTA 和 0.1% 氯乙酰胺)。然后将以下试剂顺序添加到 HCV 涂覆缓冲液中, 且每次添加后都要搅拌: 来自 10mg/ml 溶液的 2 μ g/ml 巯基修饰的 BSA (Bayer Corp, Pentex, Kankakee, Illinois); 来自 1M 溶液的 5mM DTT (Sigma, St. Louis, MO); 0.45 μ g/ml NS3/4a (蛋白质浓度为 0.3mg/ml); 0.375 μ g/ml MEFA 7.1 (蛋白质浓度为 1mg/ml)。室温搅拌最终溶液 15 分钟。

[0192] 将 200 μ l 上述溶液加到 Costar 高度结合的平底板 (Corning Inc., Corning, New York) 的各孔中, 并将平板在保湿室中培育过夜。然后用洗涤缓冲液 (1 $\times$ PBS, 0.1% TWEEN-20) 洗涤平板, 轻敲干燥, 并加入 285 μ l Ortho Post-涂覆缓冲液 (1 $\times$ PBS pH 7.4, 1% BSA, 3% 蔗糖)。培育平板至少 1 小时, 轻敲并在 2-8 $^{\circ}$ C 干燥过夜。将平板装入有干燥剂的袋内备用。

[0193] 实施例 5

[0194] 免疫测定

[0195] 为了测试本发明免疫测定测试 HCV 感染的能力, 用可购得的人血液样品 (受 HCV 感染的) 的板。这些板是从 Boston Biomedica, Inc., West Bridgewater, MA (BBI); Bioclinical Partners, Franklin, MA (BCP) 和 North American Biologics, Inc., Boca Raton, FL (NABI) 购得。

[0196] 如下进行测定。在涂覆的平板中加入 200 μ l 样品稀释缓冲液 (1g/l 酪蛋白,

100mg/1 重组人 SOD, 1g/1 氯乙酰胺, 10g/1 BSA, 500mg/1 酵母提取物, 0.366g/1 EDTA, 1.162g/1 KPO_4 , 5ml/1 Tween-20, 29.22g/1 NaCl, 1.627g/1 NaPO_4 , 1% SDS)。然后加入 20 μl 样品。将其在 37℃ 培育 30-60 分钟。用洗涤缓冲液 (1xPBS、pH 7.4, 0.1% Tween-20) 洗涤平板。加入 200 μl 在 ORTHO HCV 3.0ELISA 测试系统中用 Enhanced SAvE 批量偶联稀释液 (Ortho-Clinical Diagnostics, Raritan, New Jersey) 以 1 : 22,000 稀释的标记的液相成分 (HRP 标记的 MEFA 或 HRP 标记的 NS3/4a 构象表位, 取决于结合于固相支持物上的抗原), 并在 37℃ 培育 30-60 分钟。如上所述洗涤, 并添加 200 μl 底物溶液 (10PD 片剂/10ml)。OPD 片剂含有辣根过氧化物酶反应显色所需的邻苯二胺二盐酸盐和过氧化氢。将在黑暗中室温培育 30 分钟。加入 50 μl 4N H_2SO_4 终止反应, 并在 492nm 阅读平板, 相对于 690nm 处的吸光度作为对照。

[0197] 因此, 公开了新颖的 HCV 检测试验。根据上述内容, 应理解虽然本文为了说明的目的描述了本发明的具体实施例, 但是可作出各种改变, 而不背离本文的精神和范围。

[0198] 序列表

[0199] <110> 希龙公司

[0200] <120> HCV 检测

[0201] <130> 2300-19199.40

[0202] <150> 60/409, 515

[0203] <151> 2002-09-09

[0204] <160> 9

[0205] <170> PatentIn version 3.2

[0206] <210> 1

[0207] <211> 2058

[0208] <212> DNA

[0209] <213> 人工

[0210] <220>

[0211] <223> NS3/41 构象表位 DNA 序列

[0212] <400> 1

[0213] atg gcg ccc atc acg gcg tac gcc cag cag aca agg ggc ctc cta ggg 48

[0214] Met Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly

[0215] 1 5 10 15

[0216] tgc ata atc acc agc cta act ggc cgg gac aaa aac caa gtg gag ggt 96

[0217] Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly

[0218] 20 25 30

[0219] gag gtc cag att gtg tca act gct gcc caa acc ttc ctg gca acg tgc 144

[0220] Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys

[0221] 35 40 45

[0222] atc aat ggg gtg tgc tgg act gtc tac cac ggg gcc gga acg agg acc 192

[0223] Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr

[0224] 50 55 60

[0225]	atc gcg tca ccc aag ggt cct gtc atc cag atg tat acc aat gta gac	240
[0226]	Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp	
[0227]	65 70 75 80	
[0228]	caa gac ctt gtg ggc tgg ccc gct ccg caa ggt agc cga tca ttg aca	288
[0229]	Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg Ser Leu Thr	
[0230]	85 90 95	
[0231]	ccc tgc act tgc ggc tcc tcg gac ctt tac ctg gtc acg agg cac gcc	336
[0232]	Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala	
[0233]	100 105 110	
[0234]	gat gtc att ccc gtg cgc cgg cgg ggt gat agc agg ggc agc ctg ctg	384
[0235]	Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu	
[0236]	115 120 125	
[0237]	tcg ccc cgg ccc att tcc tac ttg aaa ggc tcc tcg ggg ggt ccg ctg	432
[0238]	Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu	
[0239]	130 135 140	
[0240]	ttg tgc ccc gcg ggg cac gcc gtg ggc ata ttt agg gcc gcg gtg tgc	480
[0241]	Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys	
[0242]	145 150 155 160	
[0243]	acc cgt gga gtg gct aag gcg gtg gac ttt atc cct gtg gag aac cta	528
[0244]	Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu	
[0245]	165 170 175	
[0246]	gag aca acc atg agg tcc ccg gtg ttc acg gat aac tcc tct cca cca	576
[0247]	Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro	
[0248]	180 185 190	
[0249]	gta gtg ccc cag agc ttc cag gtg gct cac ctc cat gct ccc aca ggc	624
[0250]	Val Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly	
[0251]	195 200 205	
[0252]	agc ggc aaa agc acc aag gtc ccg gct gca tat gca gct cag ggc tat	672
[0253]	Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr	
[0254]	210 215 220	
[0255]	aag gtg cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt	720
[0256]	Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly	
[0257]	225 230 235 240	
[0258]	gct tac atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg	768
[0259]	Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly	
[0260]	245 250 255	
[0261]	gtg aga aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc	816
[0262]	Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly	
[0263]	260 265 270	

[0264] aag ttc ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata 864
 [0265] Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile
 [0266] 275 280 285
 [0267] att tgt gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att 912
 [0268] Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile
 [0269] 290 295 300
 [0270] ggc act gtc ctt gac caa gca gag act gcg ggg gcg aga ctg gtt gtg 960
 [0271] Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val
 [0272] 305 310 315 320
 [0273] ctc gcc acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac 1008
 [0274] Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn
 [0275] 325 330 335
 [0276] atc gag gag gtt gct ctg tcc acc acc gga gag atc cct ttt tac ggc 1056
 [0277] Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly
 [0278] 340 345 350
 [0279] aag gct atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc 1104
 [0280] Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe
 [0281] 355 360 365
 [0282] tgt cat tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca 1152
 [0283] Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala
 [0284] 370 375 380
 [0285] ttg ggc atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc 1200
 [0286] Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val
 [0287] 385 390 395 400
 [0288] atc ccg ccc atc ggc gat gtt gtc gtc gtg gca acc gat gcc ctc atg 1248
 [0289] Ile Pro Pro Ile Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
 [0290] 405 410 415
 [0291] acc ggc tat acc ggc gac ttc gac tcg gtg ata gac tgc aat acg tgt 1296
 [0292] Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys
 [0293] 420 425 430
 [0294] gtc acc cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag 1344
 [0295] Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu
 [0296] 435 440 445
 [0297] aca atc acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc 1392
 [0298] Thr Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly
 [0299] 450 455 460
 [0300] agg act ggc agg ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg 1440
 [0301] Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly
 [0302] 465 470 475 480

[0303] gag cgc ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat 1488
 [0304] Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr
 [0305] 485 490 495
 [0306] gac gca ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt 1536
 [0307] Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val
 [0308] 500 505 510
 [0309] agg cta cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac 1584
 [0310] Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
 [0311] 515 520 525
 [0312] cat ctt gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat 1632
 [0313] His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp
 [0314] 530 535 540
 [0315] gcc cac ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac 1680
 [0316] Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr
 [0317] 545 550 555 560
 [0318] ctg gta gcg tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc 1728
 [0319] Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro
 [0320] 565 570 575
 [0321] cca tcg tgg gac cag atg tgg aag tgt ttg att cgc ctc aag ccc acc 1776
 [0322] Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr
 [0323] 580 585 590
 [0324] ctc cat ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat 1824
 [0325] Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn
 [0326] 595 600 605
 [0327] gaa atc acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg 1872
 [0328] Glu Ile Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met
 [0329] 610 615 620
 [0330] tcg gcc gac ctg gag gtc gtc acg agc acc tgg gtg ctc gtt ggc ggc 1920
 [0331] Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly
 [0332] 625 630 635 640
 [0333] gtc ctg gct gct ttg gcc gcg tat tgc ctg tca aca ggc tgc gtg gtc 1968
 [0334] Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val
 [0335] 645 650 655
 [0336] ata gtg ggc agg gtc gtc ttg tcc ggg aag ccg gca atc ata cct gac 2016
 [0337] Ile Val Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp
 [0338] 660 665 670
 [0339] agg gaa gtc ctc tac cga gag ttc gat gag atg gaa gag tgc
 2058
 [0340] Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys

[0341]	675	680	685
[0342]	<210>2		
[0343]	<211>686		
[0344]	<212>PRT		
[0345]	<213> 人工		
[0346]	<220>		
[0347]	<223>NS3/41 构象表位氨基酸序列		
[0348]	<400>2		
[0349]	Met Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly		
[0350]	1	5	10 15
[0351]	Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu Gly		
[0352]	20	25	30
[0353]	Glu Val Gln Ile Val Ser Thr Ala Ala Gln Thr Phe Leu Ala Thr Cys		
[0354]	35	40	45
[0355]	Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg Thr		
[0356]	50	55	60
[0357]	Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val Asp		
[0358]	65	70	75 80
[0359]	Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg Ser Leu Thr		
[0360]	85	90	95
[0361]	Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His Ala		
[0362]	100	105	110
[0363]	Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu Leu		
[0364]	115	120	125
[0365]	Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu		
[0366]	130	135	140
[0367]	Leu Cys Pro Ala Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys		
[0368]	145	150	155 160
[0369]	Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn Leu		
[0370]	165	170	175
[0371]	Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro		
[0372]	180	185	190
[0373]	Val Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly		
[0374]	195	200	205
[0375]	Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr		
[0376]	210	215	220
[0377]	Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly		
[0378]	225	230	235 240
[0379]	Ala Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly		

[0380]		245		250		255
[0381]	Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly					
[0382]		260		265		270
[0383]	Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile					
[0384]		275		280		285
[0385]	Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile					
[0386]		290		295		300
[0387]	Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val					
[0388]	305		310		315	320
[0389]	Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn					
[0390]		325		330		335
[0391]	Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly					
[0392]		340		345		350
[0393]	Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe					
[0394]		355		360		365
[0395]	Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala					
[0396]		370		375		380
[0397]	Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val					
[0398]	385		390		395	400
[0399]	Ile Pro Pro Ile Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met					
[0400]		405		410		415
[0401]	Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys					
[0402]		420		425		430
[0403]	Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu					
[0404]		435		440		445
[0405]	Thr Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly					
[0406]		450		455		460
[0407]	Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly					
[0408]	465		470		475	480
[0409]	Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr					
[0410]		485		490		495
[0411]	Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val					
[0412]		500		505		510
[0413]	Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp					
[0414]		515		520		525
[0415]	His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp					
[0416]		530		535		540
[0417]	Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr					
[0418]	545		550		555	560

[0419]	Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro	
[0420]	565 570 575	
[0421]	Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr	
[0422]	580 585 590	
[0423]	Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn	
[0424]	595 600 605	
[0425]	Glu Ile Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met	
[0426]	610 615 620	
[0427]	Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly	
[0428]	625 630 635 640	
[0429]	Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val Val	
[0430]	645 650 655	
[0431]	Ile Val Gly Arg Val Val Leu Ser Gly Lys Pro Ala Ile Ile Pro Asp	
[0432]	660 665 670	
[0433]	Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu Cys	
[0434]	675 680 685	
[0435]	<210>3	
[0436]	<211>2499	
[0437]	<212>DNA	
[0438]	<213> 人工	
[0439]	<220>	
[0440]	<223>MEFA 12DNA 序列	
[0441]	<400>3	
[0442]	atg gct aca aag gct gtt tgt gtt ttg aag ggt gac ggc cca gtt caa	48
[0443]	Met Ala Thr Lys Ala Val Cys Val Leu Lys Gly Asp Gly Pro Val Gln	
[0444]	1 5 10 15	
[0445]	ggt att att aac ttc gag cag aag gaa agt aat gga cca gtg aag gtg	96
[0446]	Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val	
[0447]	20 25 30	
[0448]	tgg gga agc att aaa gga ctg act gaa ggc ctg cat gga ttc cat gtt	144
[0449]	Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val	
[0450]	35 40 45	
[0451]	cat gag ttt gga gat aat aca gca ggc tgt acc agt gca ggt cct cac	192
[0452]	His Glu Phe Gly Asp Asn Thr Ala Gly Cys Thr Ser Ala Gly Pro His	
[0453]	50 55 60	
[0454]	ttt aat cct cta tcc acg cgt ggt tgc aat tgc tct atc tat ccc ggc	240
[0455]	Phe Asn Pro Leu Ser Thr Arg Gly Cys Asn Cys Ser Ile Tyr Pro Gly	
[0456]	65 70 75 80	
[0457]	cat ata acg ggt cac cgc atg gca tgg aag ctt ggt tcc gcc gcc aga	288

[0458]	His Ile Thr Gly His Arg Met Ala Trp Lys Leu Gly Ser Ala Ala Arg	
[0459]	85 90 95	
[0460]	act acc tcg ggc ttt gtc tcc ttg ttc gcc cca ggt gcc aaa caa aac	336
[0461]	Thr Thr Ser Gly Phe Val Ser Leu Phe Ala Pro Gly Ala Lys Gln Asn	
[0462]	100 105 110	
[0463]	gaa act cac gtc acg gga ggc gca gcc gcc cga act acg tct ggg ttg	384
[0464]	Glu Thr His Val Thr Gly Gly Ala Ala Ala Arg Thr Thr Ser Gly Leu	
[0465]	115 120 125	
[0466]	acc tct ttg ttc tcc cca ggt gcc agc caa aac att caa ttg att act	432
[0467]	Thr Ser Leu phe Ser Pro Gly Ala Ser Gln Asn Ile Gln Leu Ile Thr	
[0468]	130 135 140	
[0469]	agt acg gat aac tcc tct cca cca gta gtg ccc cag agc ttc cag gtg	480
[0470]	Ser Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val	
[0471]	145 150 155 160	
[0472]	gct cac ctc cat gct ccc aca ggc agc ggc aaa agc acc aag gtc ccg	528
[0473]	Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro	
[0474]	165 170 175	
[0475]	gct gca tat gca gct cag ggc tat aag gtg cta gta ctc aac ccc tct	576
[0476]	Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser	
[0477]	180 185 190	
[0478]	gtt gct gca aca ctg ggc ttt ggt gct tac atg tcc aag gct cat ggg	624
[0479]	Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly	
[0480]	195 200 205	
[0481]	atc gat cct aac atc agg acc ggg gtg aga aca att acc act ggc agc	672
[0482]	Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser	
[0483]	210 215 220	
[0484]	ccc atc acg tac tcc acc tac ggc aag ttc ctt gcc gac ggc ggg tgc	720
[0485]	Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys	
[0486]	225 230 235 240	
[0487]	tcg ggg ggc gct tat gac ata ata att tgt gac gag tgc cac tcc acg	768
[0488]	Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr	
[0489]	245 250 255	
[0490]	gat gcc aca tcc atc ttg ggc atc ggc act gtc ctt gac caa gca gag	816
[0491]	Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu	
[0492]	260 265 270	
[0493]	act gcg ggg gcg aga ctg gtt gtg ctc gcc acc gcc acc cct ccg ggc	864
[0494]	Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly	
[0495]	275 280 285	
[0496]	tcc gtc act gtg ccc cat ccc aac atc gag gag gtt gct ctg tcc acc	912

[0497]	Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr	
[0498]	290 295 300	
[0499]	acc gga gag atc cct ttt tac ggc aag gct atc ccc ctc gaa gta atc	960
[0500]	Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile	
[0501]	305 310 315 320	
[0502]	aag ggg ggg aga cat ctc atc ttc tgt cat tca aag aag aag tgc gac	1008
[0503]	Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp	
[0504]	325 330 335	
[0505]	gaa ctc gcc gca aag ctg gtc gca ttg ggc atc aat gcc gtg gcc tac	1056
[0506]	Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala Tyr	
[0507]	340 345 350	
[0508]	tac cgc ggt ctt gac gtg tcc gtc atc ccg acc agc ggc gat gtt gtc	1104
[0509]	Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val	
[0510]	355 360 365	
[0511]	gtc gtg gca acc gat gcc ctc atg acc ggc tat acc ggc gac ttc gac	1152
[0512]	Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp	
[0513]	370 375 380	
[0514]	tcg gtg ata gac tgc aat acg tgt gca tgc tcc ggg aag ccg gca atc	1200
[0515]	Ser Val Ile Asp Cys Asn Thr Cys Ala Cys Ser Gly Lys Pro Ala Ile	
[0516]	385 390 395 400	
[0517]	ata cct gac agg gaa gtc ctc tac cga gag ttc gat gag atg gaa gag	1248
[0518]	Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu	
[0519]	405 410 415	
[0520]	tgc tct cag cac tta ccg tac atc gag caa ggg atg atg ctc gcc gag	1296
[0521]	Cys Ser Gln His Leu Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu	
[0522]	420 425 430	
[0523]	cag ttc aag cag aag gcc ctc ggc ctc tcg cga ggg ggc aag ccg gca	1344
[0524]	Gln Phe Lys Gln Lys Ala Leu Gly Leu Ser Arg Gly Gly Lys Pro Ala	
[0525]	435 440 445	
[0526]	atc gtt cca gac aaa gag gtg ttg tat caa caa tac gat gag atg gaa	1392
[0527]	Ile Val Pro Asp Lys Glu Val Leu Tyr Gln Gln Tyr Asp Glu Met Glu	
[0528]	450 455 460	
[0529]	gag tgc tca caa gct gcc cca tat atc gaa caa gct cag gta ata gct	1440
[0530]	Glu Cys Ser Gln Ala Ala Pro Tyr Ile Glu Gln Ala Gln Val Ile Ala	
[0531]	465 470 475 480	
[0532]	cac cag ttc aag gaa aaa gtc ctt gga ttg atc gat aat gat caa gtg	1488
[0533]	His Gln Phe Lys Glu Lys Val Leu Gly Leu Ile Asp Asn Asp Gln Val	
[0534]	485 490 495	
[0535]	gtt gtg act cct gac aaa gaa atc tta tat gag gcc ttt gat gag atg	1536

[0536]	Val Val Thr Pro Asp Lys Glu Ile Leu Tyr Glu Ala Phe Asp Glu Met	
[0537]	500 505 510	
[0538]	gaa gaa tgc gcc tcc aaa gcc gcc ctc att gag gaa ggg cag cgg atg	1584
[0539]	Glu Glu Cys Ala Ser Lys Ala Ala Leu Ile Glu Glu Gly Gln Arg Met	
[0540]	515 520 525	
[0541]	gcg gag atg ctc aag tct aag ata caa ggc ctc ctc ggg ata ctg cgc	1632
[0542]	Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gly Ile Leu Arg	
[0543]	530 535 540	
[0544]	cgg cac gtt ggt cct ggc gag ggg gca gtg cag tgg atg aac cgg ctg	1680
[0545]	Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu	
[0546]	545 550 555 560	
[0547]	ata gcc ttc gcc tcc aga ggg aac cat gtt tcc ccc acg cac tac gtt	1728
[0548]	Ile Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val	
[0549]	565 570 575	
[0550]	ccg tct aga tcc cgg aga ttc gcc cag gcc ctg ccc gtt tgg gcg cgg	1776
[0551]	Pro Ser Arg Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg	
[0552]	580 585 590	
[0553]	ccg gac tat aac ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac	1824
[0554]	Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr	
[0555]	595 600 605	
[0556]	gaa cca cct gtg gtc cac ggc aga tct tct cgg aga ttc gcc cag gcc	1872
[0557]	Glu Pro Pro Val Val His Gly Arg Ser Ser Arg Arg Phe Ala Gln Ala	
[0558]	610 615 620	
[0559]	ctg ccc gtt tgg gcg cgg ccg gac tat aac ccc ccg cta gtg gag acg	1920
[0560]	Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr	
[0561]	625 630 635 640	
[0562]	tgg aaa aag ccc gac tac gaa cca cct gtg gtc cat ggc aga aag acc	1968
[0563]	Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His Gly Arg Lys Thr	
[0564]	645 650 655	
[0565]	aaa cgt aac acc aac cgg cgg ccg cag gac gtc aag ttc ccg ggt ggc	2016
[0566]	Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly	
[0567]	660 665 670	
[0568]	ggt cag atc gtt ggt gga gtt tac ttg ttg ccg cgc agg ggc cct aga	2064
[0569]	Gly Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg	
[0570]	675 680 685	
[0571]	ttg ggt gtg ctc gcg acg aga aag act tcc cct atc ccc aag gct cgt	2112
[0572]	Leu Gly Val Leu Ala Thr Arg Lys Thr Ser Pro Ile Pro Lys Ala Arg	
[0573]	690 695 700	
[0574]	cgg ccc gag ggc agg acc tgg gct cag ccc ggt tac cct tgg ccc ctc	2160

[0575] Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu
 [0576] 705 710 715 720
 [0577] tat ggc aat aag gac aga cgg tct aca ggt aag tcc tgg ggt aag cca 2208
 [0578] Tyr Gly Asn Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro
 [0579] 725 730 735
 [0580] ggg tac cct tgg cca aga aag acc aaa cgt aac acc aac cgg cgg ccg 2256
 [0581] Gly Tyr Pro Trp Pro Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro
 [0582] 740 745 750
 [0583] cag gac gtc aag ttc ccg ggt ggc ggt cag atc gtt ggt gga gtt tac 2304
 [0584] Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr
 [0585] 755 760 765
 [0586] ttg ttg ccg cgc agg ggc cct aga ttg ggt gtg ctc gcg acg aga aag 2352
 [0587] Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Leu Ala Thr Arg Lys
 [0588] 770 775 780
 [0589] act tcc cct atc ccc aag gct cgt cgg ccc gag ggc agg acc tgg gct 2400
 [0590] Thr Ser Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala
 [0591] 785 790 795 800
 [0592] cag ccc ggt tac cct tgg ccc ctc tat ggc aat aag gac aga cgg tct 2448
 [0593] Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Lys Asp Arg Arg Ser
 [0594] 805 810 815
 [0595] aca ggt aag tcc tgg ggt aag cca ggg tac cct tgg ccc taatgagtcg ac 2499
 [0596] Thr Gly Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro
 [0597] 820 825
 [0598] <210>4
 [0599] <211>829
 [0600] <212>PRT
 [0601] <213> 人工
 [0602] <220>
 [0603] <223>MEFA 12 氨基酸序列
 [0604] <400>4
 [0605] Met Ala Thr Lys Ala Val Cys Val Leu Lys Gly Asp Gly Pro Val Gln
 [0606] 1 5 10 15
 [0607] Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
 [0608] 20 25 30
 [0609] Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
 [0610] 35 40 45
 [0611] His Glu Phe Gly Asp Asn Thr Ala Gly Cys Thr Ser Ala Gly Pro His
 [0612] 50 55 60
 [0613] Phe Asn Pro Leu Ser Thr Arg Gly Cys Asn Cys Ser Ile Tyr Pro Gly

[0614]	65	70	75	80
[0615]	His Ile Thr Gly His Arg Met Ala Trp Lys Leu Gly Ser Ala Ala Arg			
[0616]		85	90	95
[0617]	Thr Thr Ser Gly Phe Val Ser Leu Phe Ala Pro Gly Ala Lys Gln Asn			
[0618]		100	105	110
[0619]	Glu Thr His Val Thr Gly Gly Ala Ala Ala Arg Thr Thr Ser Gly Leu			
[0620]		115	120	125
[0621]	Thr Ser Leu Phe Ser Pro Gly Ala Ser Gln Asn Ile Gln Leu Ile Thr			
[0622]		130	135	140
[0623]	Ser Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val			
[0624]	145	150	155	160
[0625]	Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro			
[0626]		165	170	175
[0627]	Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser			
[0628]		180	185	190
[0629]	Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly			
[0630]		195	200	205
[0631]	Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser			
[0632]		210	215	220
[0633]	Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys			
[0634]	225	230	235	240
[0635]	Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr			
[0636]		245	250	255
[0637]	Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu			
[0638]		260	265	270
[0639]	Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro Gly			
[0640]		275	280	285
[0641]	Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr			
[0642]		290	295	300
[0643]	Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile			
[0644]	305	310	315	320
[0645]	Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp			
[0646]		325	330	335
[0647]	Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala Tyr			
[0648]		340	345	350
[0649]	Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val			
[0650]		355	360	365
[0651]	Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp			
[0652]		370	375	380

[0653]	Ser Val Ile Asp Cys Asn Thr Cys Ala Cys Ser Gly Lys Pro Ala Ile			
[0654]	385	390	395	400
[0655]	Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu			
[0656]		405	410	415
[0657]	Cys Ser Gln His Leu Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu			
[0658]		420	425	430
[0659]	Gln Phe Lys Gln Lys Ala Leu Gly Leu Ser Arg Gly Gly Lys Pro Ala			
[0660]		435	440	445
[0661]	Ile Val Pro Asp Lys Glu Val Leu Tyr Gln Gln Tyr Asp Glu Met Glu			
[0662]		450	455	460
[0663]	Glu Cys Ser Gln Ala Ala Pro Tyr Ile Glu Gln Ala Gln Val Ile Ala			
[0664]	465	470	475	480
[0665]	His Gln Phe Lys Glu Lys Val Leu Gly Leu Ile Asp Asn Asp Gln Val			
[0666]		485	490	495
[0667]	Val Val Thr Pro Asp Lys Glu Ile Leu Tyr Glu Ala Phe Asp Glu Met			
[0668]		500	505	510
[0669]	Glu Glu Cys Ala Ser Lys Ala Ala Leu Ile Glu Glu Gly Gln Arg Met			
[0670]		515	520	525
[0671]	Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gly Ile Leu Arg			
[0672]		530	535	540
[0673]	Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu			
[0674]	545	550	555	560
[0675]	Ile Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val			
[0676]		565	570	575
[0677]	Pro Ser Arg Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg			
[0678]		580	585	590
[0679]	Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr			
[0680]		595	600	605
[0681]	Glu Pro Pro Val Val His Gly Arg Ser Ser Arg Arg Phe Ala Gln Ala			
[0682]		610	615	620
[0683]	Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr			
[0684]	625	630	635	640
[0685]	Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His Gly Arg Lys Thr			
[0686]		645	650	655
[0687]	Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly			
[0688]		660	665	670
[0689]	Gly Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg			
[0690]		675	680	685
[0691]	Leu Gly Val Leu Ala Thr Arg Lys Thr Ser Pro Ile Pro Lys Ala Arg			

[0692]	690	695	700
[0693]	Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu		
[0694]	705	710	715
[0695]	Tyr Gly Asn Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro		
[0696]	725	730	735
[0697]	Gly Tyr Pro Trp Pro Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro		
[0698]	740	745	750
[0699]	Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr		
[0700]	755	760	765
[0701]	Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Leu Ala Thr Arg Lys		
[0702]	770	775	780
[0703]	Thr Ser Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala		
[0704]	785	790	795
[0705]	Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Lys Asp Arg Arg Ser		
[0706]	805	810	815
[0707]	Thr Gly Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro		
[0708]	820	825	
[0709]	<210>5		
[0710]	<211>3297		
[0711]	<212>DNA		
[0712]	<213> 人工		
[0713]	<220>		
[0714]	<223>MEFA7. 1DNA 序列		
[0715]	<400>5		
[0716]	atg gct aca aag gct gtttgt gtt ttg aag ggt gac ggc ccagtt caa	48	
[0717]	Met Ala Thr Lys Ala Val Cys Val Leu Lys Gly Asp Gly Pro Val Gln		
[0718]	1 5 10 15		
[0719]	ggt att att aac ttc gag cag aag gaa agt aat gga cca gtg aag gtg	96	
[0720]	Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val		
[0721]	20 25 30		
[0722]	tgg gga agc att aaa gga ctg act gaa ggc ctg cat gga ttc cat gtt	144	
[0723]	Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val		
[0724]	35 40 45		
[0725]	cat gag ttt gga gat aat aca gca ggc tgt acc agt gca ggt cct cac	192	
[0726]	His Glu Phe Gly Asp Asn Thr Ala Gly Cys Thr Ser Ala Gly Pro His		
[0727]	50 55 60		
[0728]	ttt aat cct cta tcc aga aaa cac ggt ggg cca aag gat gaa gag agg	240	
[0729]	Phe Asn Pro Leu Ser Arg Lys His Gly Gly Pro Lys Asp Glu Glu Arg		
[0730]	65 70 75 80		

[0731]	cat gtt gga gac ttg ggc aat gtg act gct gac aaa gat ggt gtg gcc	288
[0732]	His Val Gly Asp Leu Gly Asn Val Thr Ala Asp Lys Asp Gly Val Ala	
[0733]	85 90 95	
[0734]	gat gtg tct att gaa gat tct gtg atc tca ctc tca gga gac cat tgc	336
[0735]	Asp Val Ser Ile Glu Asp Ser Val Ile Ser Leu Ser Gly Asp His Cys	
[0736]	100 105 110	
[0737]	atc att ggc cgc aca ctg gtg gtc cat gaa aaa gca gat gac ttg ggc	384
[0738]	Ile Ile Gly Arg Thr Leu Val Val His Glu Lys Ala Asp Asp Leu Gly	
[0739]	115 120 125	
[0740]	aaa ggt gga aat gaa gaa agt aca aag aca gga aac gct gga agt cgt	432
[0741]	Lys Gly Gly Asn Glu Glu Ser Thr Lys Thr Gly Asn Ala Gly Ser Arg	
[0742]	130 135 140	
[0743]	ttg gct tgt ggt gta att ggg atc gcc cag aat ttg aat tct ggt tgc	480
[0744]	Leu Ala Cys Gly Val Ile Gly Ile Ala Gln Asn Leu Asn Ser Gly Cys	
[0745]	145 150 155 160	
[0746]	aat tgc tct atc tat ccc ggc cat ata acg ggt cac cgc atg gca tgg	528
[0747]	Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp	
[0748]	165 170 175	
[0749]	aag ctt ggt tcc gcc gcc aga act acc tcg ggc ttt gtc tcc ttg ttc	576
[0750]	Lys Leu Gly Ser Ala Ala Arg Thr Thr Ser Gly Phe Val Ser Leu Phe	
[0751]	180 185 190	
[0752]	gcc cca ggt gcc aaa caa aac gaa act cac gtc acg gga ggc gca gcc	624
[0753]	Ala Pro Gly Ala Lys Gln Asn Glu Thr His Val Thr Gly Gly Ala Ala	
[0754]	195 200 205	
[0755]	gcc cga act acg tct ggg ttg acc tct ttg ttc tcc cca ggt gcc agc	672
[0756]	Ala Arg Thr Thr Ser Gly Leu Thr Ser Leu Phe Ser Pro GlyAla Ser	
[0757]	210 215 220	
[0758]	caa aac att caa ttg att gtc gac ttt atc cct gtg gag aac cta gag	720
[0759]	Gln Asn Ile Gln Leu Ile Val Asp Phe Ile Pro Val Glu Asn Leu Glu	
[0760]	225 230 235 240	
[0761]	aca acc atg cga tct ccg gtg ttc acg gat aac tcc tct cca cca gta	768
[0762]	Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Val	
[0763]	245 250 255	
[0764]	gtg ccc cag agc ttc cag gtg gct cac ctc cat gct ccc aca ggc agc	816
[0765]	Val Pro Gln Ser Phe Gln Val Ala His Leu Hi s Ala Pro Thr Gly Ser	
[0766]	260 265 270	
[0767]	ggc aaa agc acc aag gtc ccg gct gca tat gca gct cag ggc tat aag	864
[0768]	Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys	
[0769]	275 280 285	

[0770]	gtg cta gta ctc aac ccc tct gtt gct gca aca ctg ggc ttt ggt gct	912
[0771]	Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala	
[0772]	290 295 300	
[0773]	tac atg tcc aag gct cat ggg atc gat cct aac atc agg acc ggg gtg	960
[0774]	Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val	
[0775]	305 310 315 320	
[0776]	aga aca att acc act ggc agc ccc atc acg tac tcc acc tac ggc aag	1008
[0777]	Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys	
[0778]	325 330 335	
[0779]	ttc ctt gcc gac ggc ggg tgc tcg ggg ggc gct tat gac ata ata att	1056
[0780]	Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile	
[0781]	340 345 350	
[0782]	tgt gac gag tgc cac tcc acg gat gcc aca tcc atc ttg ggc att ggc	1104
[0783]	Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly	
[0784]	355 360 365	
[0785]	act gtc ctt gac caa gca gag act gcg ggg gcg aga ctg gtt gtg ctc	1152
[0786]	Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu	
[0787]	370 375 380	
[0788]	gcc acc gcc acc cct ccg ggc tcc gtc act gtg ccc cat ccc aac atc	1200
[0789]	Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile	
[0790]	385 390 395 400	
[0791]	gag gag gtt gct ctg tcc acc acc gga gag atc cct ttt tac ggc aag	1248
[0792]	Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys	
[0793]	405 410 415	
[0794]	gct atc ccc ctc gaa gta atc aag ggg ggg aga cat ctc atc ttc tgt	1296
[0795]	Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys	
[0796]	420 425 430	
[0797]	cat tca aag aag aag tgc gac gaa ctc gcc gca aag ctg gtc gca ttg	1344
[0798]	His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu	
[0799]	435 440 445	
[0800]	ggc atc aat gcc gtg gcc tac tac cgc ggt ctt gac gtg tcc gtc atc	1392
[0801]	Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile	
[0802]	450 455 460	
[0803]	ccg acc agc ggc gat gtt gtc gtc gtg gca acc gat gcc ctc atg acc	1440
[0804]	Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met Thr	
[0805]	465 470 475 480	
[0806]	ggc tat acc ggc gac ttc gac tcg gtg ata gac tgc aat acg tgt gtc	1488
[0807]	Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val	
[0808]	485 490 495	

[0809] acc cag aca gtc gat ttc agc ctt gac cct acc ttc acc att gag aca 1536
 [0810] Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr
 [0811] 500 505 510
 [0812] atc acg ctc ccc caa gat gct gtc tcc cgc act caa cgt cgg ggc agg 1584
 [0813] Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg
 [0814] 515 520 525
 [0815] act ggc agg ggg aag cca ggc atc tac aga ttt gtg gca ccg ggg gag 1632
 [0816] Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu
 [0817] 530 535 540
 [0818] cgc ccc tcc ggc atg ttc gac tcg tcc gtc ctc tgt gag tgc tat gac 1680
 [0819] Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp
 [0820] 545 550 555 560
 [0821] gca ggc tgt gct tgg tat gag ctc acg ccc gcc gag act aca gtt agg 1728
 [0822] Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg
 [0823] 565 570 575
 [0824] cta cga gcg tac atg aac acc ccg ggg ctt ccc gtg tgc cag gac cat 1776
 [0825] Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His
 [0826] 580 585 590
 [0827] ctt gaa ttt tgg gag ggc gtc ttt aca ggc ctc act cat ata gat gcc 1824
 [0828] Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala
 [0829] 595 600 605
 [0830] cac ttt cta tcc cag aca aag cag agt ggg gag aac ctt cct tac ctg 1872
 [0831] His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu
 [0832] 610 615 620
 [0833] gta gcg tac caa gcc acc gtg tgc gct agg gct caa gcc cct ccc cca 1920
 [0834] Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro
 [0835] 625 630 635 640
 [0836] tcg tgg gac cag atg tgg aag tgt ttg att cgc ctc aag ccc acc ctc 1968
 [0837] Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu
 [0838] 645 650 655
 [0839] cat ggg cca aca ccc ctg cta tac aga ctg ggc gct gtt cag aat gaa 2016
 [0840] His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu
 [0841] 660 665 670
 [0842] atc acc ctg acg cac cca gtc acc aaa tac atc atg aca tgc atg tcg 2064
 [0843] Ile Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser
 [0844] 675 680 685
 [0845] gcc gac ctg gag gtc gtc acg agc gca tgc tcc ggg aag ccg gca atc 2112
 [0846] Ala Asp Leu Glu Val Val Thr Ser Ala Cys Ser Gly Lys Pro Ala Ile
 [0847] 690 695 700

[0848] ata cct gac agg gaa gtc ctc tac cga gag ttc gat gag atg gaa gag 2160
 [0849] Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu
 [0850] 705 710 715 720
 [0851] tgc tct cag cac tta ccg tac atc gag caa ggg atg atg ctc gcc gag 2208
 [0852] Cys Ser Gln His Leu Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu
 [0853] 725 730 735
 [0854] cag ttc aag cag aag gcc ctc ggc ctc tcg cga ggg ggc aag ccg gca 2256
 [0855] Gln Phe Lys Gln Lys Ala Leu Gly Leu Ser Arg Gly Gly Lys Pro Ala
 [0856] 740 745 750
 [0857] atc gtt cca gac aaa gag gtg ttg tat caa caa tac gat gag atg gaa 2304
 [0858] Ile Val Pro Asp Lys Glu Val Leu Tyr Gln Gln Tyr Asp Glu Met Glu
 [0859] 755 760 765
 [0860] gag tgc tca caa gct gcc cca tat atc gaa caa gct cag gta ata gct 2352
 [0861] Glu Cys Ser Gln Ala Ala Pro Tyr Ile Glu Gln Ala Gln Val Ile Ala
 [0862] 770 775 780
 [0863] cac cag ttc aag gaa aaa gtc ctt gga ttg atc gat aat gat caa gtg 2400
 [0864] His Gln Phe Lys Glu Lys Val Leu Gly Leu Ile Asp Asn Asp Gln Val
 [0865] 785 790 795 800
 [0866] gtt gtg act cct gac aaa gaa atc tta tat gag gcc ttt gat gag atg 2448
 [0867] Val Val Thr Pro Asp Lys Glu Ile Leu Tyr Glu Ala Phe Asp Glu Met
 [0868] 805 810 815
 [0869] gaa gaa tgc gcc tcc aaa gcc gcc ctc att gag gaa ggg cag cgg atg 2496
 [0870] Glu Glu Cys Ala Ser Lys Ala Ala Leu Ile Glu Glu Gly Gln Arg Met
 [0871] 820 825 830
 [0872] gcg gag atg ctc aag tct aag ata caa ggc ctc ctc ggg ata ctg cgc 2544
 [0873] Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gly Ile Leu Arg
 [0874] 835 840 845
 [0875] cgg cac gtt ggt cct ggc gag ggg gca gtg cag tgg atg aac cgg ctg 2592
 [0876] Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu
 [0877] 850 855 860
 [0878] ata gcc ttc gcc tcc aga ggg aac cat gtt tcc ccc acg cac tac gtt 2640
 [0879] Ile Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val
 [0880] 865 870 875 880
 [0881] ccg tct aga tcc cgg aga ttc gcc cag gcc ctg ccc gtt tgg gcg cgg 2688
 [0882] Pro Ser Arg Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg
 [0883] 885 890 895
 [0884] ccg gac tat aac ccc ccg cta gtg gag acg tgg aaa aag ccc gac tac 2736
 [0885] Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr
 [0886] 900 905 910

[0887] gaa cca cct gtg gtc cac ggc aga tct tct cgg aga ttc gcc cag gcc 2784
 [0888] Glu Pro Pro Val Val His Gly Arg Ser Ser Arg Arg Phe Ala Gln Ala
 [0889] 915 920 925
 [0890] ctg ccc gtt tgg gcg cgg ccg gac tat aac ccc ccg cta gtg gag acg 2832
 [0891] Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr
 [0892] 930 935 940
 [0893] tgg aaa aag ccc gac tac gaa cca cct gtg gtc cat ggc aga aag acc 2880
 [0894] Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His Gly Arg Lys Thr
 [0895] 945 950 955 960
 [0896] aaa cgt aac acc aac cgg cgg ccg cag gac gtc aag ttc ccg ggt ggc 2928
 [0897] Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly
 [0898] 965 970 975
 [0899] ggt cag atc gtt ggt cgc agg ggc cct cct atc ccc aag gct cgt cgg 2976
 [0900] Gly Gln Ile Val Gly Arg Arg Gly Pro Pro Ile Pro Lys Ala Arg Arg
 [0901] 980 985 990
 [0902] ccc gag ggc agg acc tgg gct cag ccc ggt tac cct tgg ccc ctc tat 3024
 [0903] Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr
 [0904] 995 1000 1005
 [0905] ggc aat aag gac aga cgg tct aca ggt aag tcc tgg ggt aag cca ggg 3072
 [0906] Gly Asn Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
 [0907] 1010 1015 1020
 [0908] tac cct tgg cca aga aag acc aaa cgt aac acc aac cga cgg ccg cag 3120
 [0909] Tyr Pro Trp Pro Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln
 [0910] 1025 1030 1035 1040
 [0911] gac gtc aag ttc ccg ggt ggc ggt cag atc gtt ggt cgc agg ggc cct 3168
 [0912] Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Arg Arg Gly Pro
 [0913] 1045 1050 1055
 [0914] cct atc ccc aag gct cgt cgg ccc gag ggc agg acc tgg gct cag ccc 3216
 [0915] Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro
 [0916] 1060 1065 1070
 [0917] ggt tac cct tgg ccc ctc tat ggc aat aag gac aga cgg tct acc ggt 3264
 [0918] Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Lys Asp Arg Arg Ser Thr Gly
 [0919] 1075 1080 1085
 [0920] aag tcc tgg ggt aag cca ggg tat cct tgg ccc 3297
 [0921] Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro
 [0922] 1090 1095
 [0923] <210>6
 [0924] <211>1099
 [0925] <212>PRT

[0926] <213> 人工
 [0927] <220>
 [0928] <223>MEFA 7.1 氨基酸序列
 [0929] <400>6
 [0930] Met Ala Thr Lys Ala Val Cys Val Leu Lys Gly Asp Gly Pro Val Gln
 [0931] 1 5 10 15
 [0932] Gly Ile Ile Asn Phe Glu Gln Lys Glu Ser Asn Gly Pro Val Lys Val
 [0933] 20 25 30
 [0934] Trp Gly Ser Ile Lys Gly Leu Thr Glu Gly Leu His Gly Phe His Val
 [0935] 35 40 45
 [0936] His Glu Phe Gly Asp Asn Thr Ala Gly Cys Thr Ser Ala Gly Pro His
 [0937] 50 55 60
 [0938] Phe Asn Pro Leu Ser Arg Lys His Gly Gly Pro Lys Asp Glu Glu Arg
 [0939] 65 70 75 80
 [0940] His Val Gly Asp Leu Gly Asn Val Thr Ala Asp Lys Asp Gly Val Ala
 [0941] 85 90 95
 [0942] Asp Val Ser Ile Glu Asp Ser Val Ile Ser Leu Ser Gly Asp His Cys
 [0943] 100 105 110
 [0944] Ile Ile Gly Arg Thr Leu Val Val His Glu Lys Ala Asp Asp Leu Gly
 [0945] 115 120 125
 [0946] Lys Gly Gly Asn Glu Glu Ser Thr Lys Thr Gly Asn Ala Gly Ser Arg
 [0947] 130 135 140
 [0948] Leu Ala Cys Gly Val Ile Gly Ile Ala Gln Asn Leu Asn Ser Gly Cys
 [0949] 145 150 155 160
 [0950] Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp
 [0951] 165 170 175
 [0952] Lys Leu Gly Ser Ala Ala Arg Thr Thr Ser Gly Phe Val Ser Leu Phe
 [0953] 180 185 190
 [0954] Ala Pro Gly Ala Lys Gln Asn Glu Thr His Val Thr Gly Gly Ala Ala
 [0955] 195 200 205
 [0956] Ala Arg Thr Thr Ser Gly Leu Thr Ser Leu Phe Ser Pro Gly Ala Ser
 [0957] 210 215 220
 [0958] Gln Asn Ile Gln Leu Ile Val Asp Phe Ile Pro Val Glu Asn Leu Glu
 [0959] 225 230 235 240
 [0960] Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro Pro Val
 [0961] 245 250 255
 [0962] Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr Gly Ser
 [0963] 260 265 270
 [0964] Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys

[0965]	275	280	285
[0966]	Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly Ala		
[0967]	2902	95	300
[0968]	Tyr Met Ser Lys Ala His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val		
[0969]	305	310	315
[0970]	Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys		
[0971]	325	330	335
[0972]	Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile		
[0973]	340	345	350
[0974]	Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly Ile Gly		
[0975]	355	360	365
[0976]	Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val Val Leu		
[0977]	370	375	380
[0978]	Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Pro His Pro Asn Ile		
[0979]	385	390	395
[0980]	Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr Gly Lys		
[0981]	405	410	415
[0982]	Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe Cys		
[0983]	420	425	430
[0984]	His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala Leu		
[0985]	435	440	445
[0986]	Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile		
[0987]	450	455	460
[0988]	Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met Thr		
[0989]	465	470	475
[0990]	Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys Val		
[0991]	485	490	495
[0992]	Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr		
[0993]	500	505	510
[0994]	Ile Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg Gly Arg		
[0995]	515	520	525
[0996]	Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro Gly Glu		
[0997]	530	535	540
[0998]	Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp		
[0999]	545	550	555
[1000]	Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr Val Arg		
[1001]	565	570	575
[1002]	Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln Asp His		
[1003]	580	585	590

[1004]	Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile Asp Ala
[1005]	595 600 605
[1006]	His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Leu Pro Tyr Leu
[1007]	610 615 620
[1008]	Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro
[1009]	625 630 635 640
[1010]	Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu
[1011]	645 650 655
[1012]	His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu
[1013]	660 665 670
[1014]	Ile Thr Leu Thr His Pro Val Thr Lys Tyr Ile Met Thr Cys Met Ser
[1015]	675 680 685
[1016]	Ala Asp Leu Glu Val Val Thr Ser Ala Cys Ser Gly Lys Pro Ala Ile
[1017]	690 695 700
[1018]	Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu
[1019]	705 710 715 720
[1020]	Cys Ser Gln His Leu Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu
[1021]	725 730 735
[1022]	Gln Phe Lys Gln Lys Ala Leu Gly Leu Ser Arg Gly Gly Lys Pro Ala
[1023]	740 745 750
[1024]	Ile Val Pro Asp Lys Glu Val Leu Tyr Gln Gln Tyr Asp Glu Met Glu
[1025]	755 760 765
[1026]	Glu Cys Ser Gln Ala Ala Pro Tyr Ile Glu Gln Ala Gln Val Ile Ala
[1027]	770 775 780
[1028]	His Gln Phe Lys Glu Lys Val Leu Gly Leu Ile Asp Asn Asp Gln Val
[1029]	785 790 795 800
[1030]	Val Val Thr Pro Asp Lys Glu Ile Leu Tyr Glu Ala Phe Asp Glu Met
[1031]	805 810 815
[1032]	Glu Glu Cys Ala Ser Lys Ala Ala Leu Ile Glu Glu Gly Gln Arg Met
[1033]	820 825 830
[1034]	Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gly Ile Leu Arg
[1035]	835 840 845
[1036]	Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu
[1037]	850 855 860
[1038]	Ile Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val
[1039]	865 870 875 880
[1040]	Pro Ser Arg Ser Arg Arg Phe Ala Gln Ala Leu Pro Val Trp Ala Arg
[1041]	885 890 895
[1042]	Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro Asp Tyr

[1043]	900	905	910
[1044]	Glu Pro Pro Val Val His Gly Arg Ser Ser Arg Arg Phe Ala Gln Ala		
[1045]	915	920	925
[1046]	Leu Pro Val Trp Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr		
[1047]	930	935	940
[1048]	Trp Lys Lys Pro Asp Tyr Glu Pro Pro Val Val His Gly Arg Lys Thr		
[1049]	945	950	955
[1050]	Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly		
[1051]	965	970	975
[1052]	Gly Gln Ile Val Gly Arg Arg Gly Pro Pro Ile Pro Lys Ala Arg Arg		
[1053]	980	985	990
[1054]	Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr		
[1055]	995	1000	1005
[1056]	Gly Asn Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly		
[1057]	1010	1015	1020
[1058]	Tyr Pro Trp Pro Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln		
[1059]	1025	1030	1035
[1060]	Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Arg Arg Gly Pro		
[1061]	1045	1050	1055
[1062]	Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro		
[1063]	1060	1065	1070
[1064]	Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Lys Asp Arg Arg Ser Thr Gly		
[1065]	1075	1080	1085
[1066]	Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro		
[1067]	1090	1095	
[1068]	<210>7		
[1069]	<211>21		
[1070]	<212>PRT		
[1071]	<213> 人工		
[1072]	<220>		
[1073]	<223> 共有序列		
[1074]	<400>7		
[1075]	Gly Ser Ala Ala Arg Thr Thr Ser Gly Phe Val Ser Leu Phe Ala Pro		
[1076]	1	5	10
[1077]	Gly Ala Lys Gln Asn		15
[1078]	20		
[1079]	<210>8		
[1080]	<211>10		
[1081]	<212>DNA		

[1082] <213> 人工

[1083] <220>

[1084] <223>HindI II 位点后的序列

[1085] <400>8

[1086] acaaaacaaa 10

[1087] <210>9

[1088] <211>23

[1089] <212>PRT

[1090] <213> 人工

[1091] <220>

[1092] <223>NS4A 肽

[1093] <400>9

[1094] Lys Lys Gly Ser Val Val Ile Val Gly Arg Ile Val Leu Ser Gly Lys

[1095] 1 5 10 15

[1096] Pro Ala Ile Ile Pro Lys Lys

[1097] 20

HCV基因组和重组蛋白

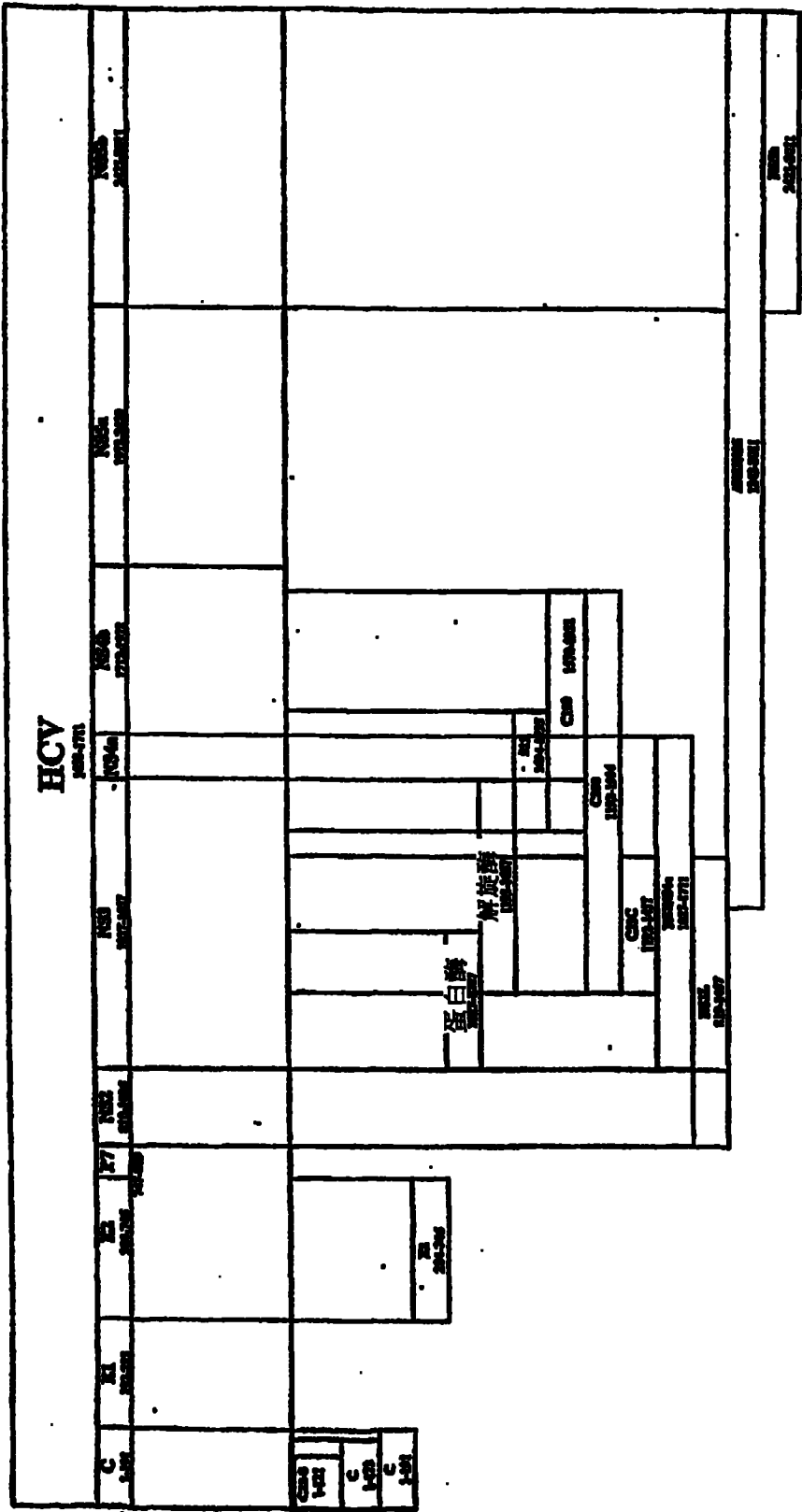


图 1

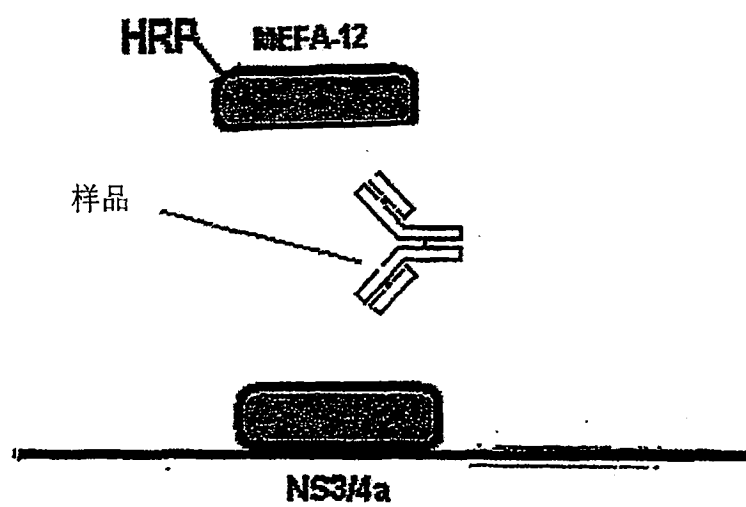


图 2

```

      1                               10
      M A P I T A Y A Q Q
      ATG GCG CCC ATC ACG GCG TAC GCC CAG CAG

      20
      T R G L L G C I I T S L T G R
      ACA AGG GGC CTC CTA GGG TGC ATA ATC ACC AGC CTA ACT GGC CCG

      30                               40
      D K N Q V E G E V Q I V S T A
      GAC AAA AAC CAA GTG GAG GGT GAG GTC CAG ATT GTG TCA ACT GCT

      50
      A Q T F L A T C I N G V C W T
      GCC CAA ACC TTC CTG GCA ACG TGC ATC AAT GGG GTG TGC TGG ACT

      60                               70
      V Y H G A G T R T I A S P K G
      GTC TAC CAC GGG GCC GGA ACG AGG ACC ATC GCG TCA CCC AAG GGT

      80
      P V I Q N Y T N V D Q D L V G
      CCT GTC ATC CAG ATG TAT ACC AAT GTA GAC CAA GAC CTT GTG GGC

      90                               100
      W P A P Q G S R S L T P C T C
      TGG CCC GCT CCG CAA GGT AGC CGA TCA TTG ACA CCC TGC ACT TGC

      110
      G S S D L Y L V T R H A D V I
      GGC TCC TCG GAC CTT TAC CTG GTC ACG AGG CAC GCC GAT GTC ATT

      120                               130
      P V R R R G D S R G S L L S P
      CCC GTG CCG CCG CCG GGT GAT AGC AGG GGC AGC CTG CTG TCG CCC

      140
      R P I S Y L K G S S G G P L L
      CCG CCC ATT TCC TAC TTG AAA GGC TCC TCG GGG GGT CCG CTG TTG

      150                               160
      C P A G H A V G I F R A A V C
      TGC CCC GCG GGG CAC GCC GTG GGC ATA TTT AGG GCC GCG GTG TGC

      170
      T R G V A K A V D F I P V E N
      ACC CGT GGA GTG GCT AAG GCG GTG GAC TTT ATC CCT GTG GAG AAC

      180                               190
      L E T T M R S P V F T D N S S
      CTA GAG ACA ACC ATG AGG TCC CCG GTG TTC ACG GAT AAC TCC TCT

```

图 3A

```

                200
P P V V P Q S F Q V A H L H A
CCA CCA GTA GTG CCC CAG AGC TTC CAG GTG GCT CAC CTC CAT GCT

                210
P T G S G K S T K V P A A Y A
CCC ACA GGC AGC GGC AAA AGC ACC AAG GTC CCG GCT GCA TAT GCA

                220
A Q G Y K V L V L N P S V A A
GCT CAG GGC TAT AAG GTG CTA GTA CTC AAC CCC TCT GTT GCT GCA

                230
T L G F G A Y M S K A H G I D
ACA CTG GGC TTT GGT GCT TAC ATG TCC AAG GCT CAT GGG ATC GAT

                240
P N I R T G V R T I T T G S P
CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC CCC

                250
I T Y S T Y G K F L A D G G C
ATC ACG TAC TCC ACC TAC GGC AAG TTC CTT GCC GAC GGC GGG TGC

                260
S G G A Y D I I I C D E C H S
TCG GGG GGC GCT TAT GAC ATA ATA ATT TGT GAC GAG TGC CAC TCC

                270
T D A T S I L G I G T V L D Q
ACG GAT GCC ACA TCC ATC TTG GGC ATT GGC ACT GTC CTT GAC CAA

                280
A E T A G A R L V V L A T A T
GCA GAG ACT GCG GGG GCG AGA CTG GTT GTG CTC GCC ACC GCC ACC

                290
P P G S V T V P H P N I E E V
CCT CCG GGC TCC GTC ACT GTG CCC CAT CCC AAC ATC GAG GAG GTT

                300
A L S T T G E I P F Y G K A I
GCT CTG TCC ACC ACC GGA GAG ATC CCT TTT TAC GGC AAG GCT ATC

                310
P L E V I K G G R H L I F C H
CCC CTC GAA GTA ATC AAG GGG GGG AGA CAT CTC ATC TTC TGT CAT

                320
S K K K C D B L A A K L V A L
TCA AAG AAG AAG TGC GAC GAA CTC GCC GCA AAG CTG GTC GCA TTG

```

图 3B

```

      390
G   I   N   A   V   A   Y   Y   R   G   L   D   V   S   V
GGC ATC AAT GCC GTG GCC TAC TAC CGC GGT CTT GAC GTG TCC GTC

      400
      410
I   P   P   I   G   D   V   V   V   V   A   T   D   A   L
ATC CCG CCC ATC GGC GAT GTT GTC GTC GTG GCA ACC GAT GCC CTC

      420
M   T   G   Y   T   G   D   F   D   S   V   I   D   C   N
ATG ACC GGC TAT ACG GGC GAC TTC GAC TCG GTG ATA GAC TGC AAT

      430
      440
T   C   V   T   Q   T   V   D   F   S   L   D   P   T   F
ACG TGT GTC ACC CAG ACA GTC GAT TTC AGC CTT GAC CCT ACC TTC

      450
T   I   E   T   I   T   L   P   Q   D   A   V   S   R   T
ACC ATT GAG ACA ATC ACG CTC CCC CAA GAT GCT GTC TCC CGC ACT

      460
      470
Q   R   R   G   R   T   G   R   G   K   P   G   I   Y   R
CAA CGT CGG GGC AGG ACT GGC AGG GGG AAG CCA GGC ATC TAC AGA

      480
F   V   A   P   G   E   R   P   S   G   M   F   D   S   S
TTT GTG GCA CCG GGG GAG CGC CCC TCC GGC ATG TTC GAC TCG TCC

      490
      500
V   L   C   E   C   Y   D   A   G   C   A   W   Y   E   L
GTC CTC TGT GAG TGC TAT GAC GCA GGC TGT GCT TGG TAT GAG CTC

      510
T   P   A   E   T   T   V   E   L   R   A   Y   M   N   T
ACG CCC GCC GAG ACT ACA GTT AGG CTA CGA GCG TAC ATG AAC ACC

      520
      530
P   G   L   P   V   C   Q   D   H   L   E   F   W   E   G
CCG GGG CTT CCC GTG TGC CAG GAC CAT CTT GAA TTT TGG GAG GGC

      540
V   F   T   G   L   T   H   I   D   A   H   F   L   S   Q
GTC TTT ACA GGC CTC ACT CAT ATA GAT GCC CAC TTT CTA TCC CAG

      550
      560
T   K   Q   E   G   E   N   L   P   Y   L   V   A   Y   Q
ACA AAG CAG AGT GGG GAG AAC CTT CCT TAC CTG GTA GCG TAC CAA

      570
A   T   V   C   A   R   A   Q   A   P   P   P   S   W   D
GCC ACC GTG TGC GCT AGG GCT CAA GCC CCT CCC CCA TCG TGG GAC

      580

```

图 3C

590
Q M W K C L I R L K P T L H G
CAG ATG TGG AAG TGT TTG ATT CGC CTC AAG CCC ACC CTC CAT GGG

600
P T P L L Y R L G A V Q N E I
CCA ACA CCC CTG CTA TAC AGA CTG GGC GCT GTT CAG AAT GAA ATC

610

620
T L T H P V T K Y I M T C M S
ACC CTG ACG CAC CCA GTC ACC AAA TAC ATC ATG ACA TGC ATG TCG

630
A D L E V V T S T W V L V G G
GCC GAC CTG GAG GTC GTC ACG AGC ACC TGG GTG CTC GTT GGC GGC

640

650
V L A A L A A Y C L S T G C V
GTC CTG GCT GCT TTG GCC GCG TAT TGC CTG TCA ACA GGC TGC GTG

660
V I V G R V V L S G K P A I I
GTC ATA GTG GGC AGG GTC GTC TTG TCC GGG AAG CCG GCA ATC ATA

670

680
P D R E V L Y R E P D E M E E
CCT GAC AGG GAA GTC CTC TAC CGA GAG TTC GAT GAG ATG GAA GAG

686
C
TGC

图 3D

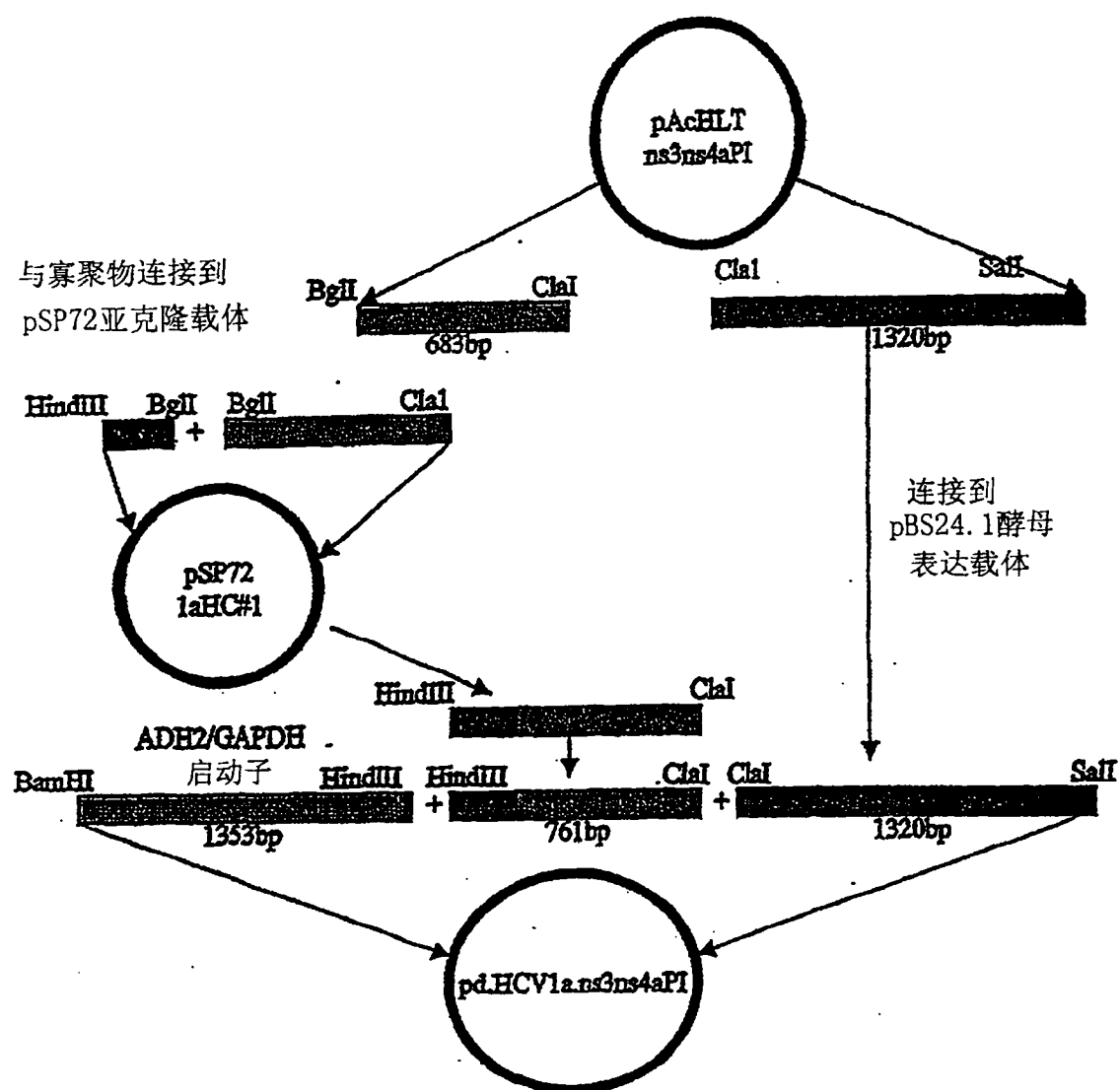


图 4

MEFA12抗原构建物

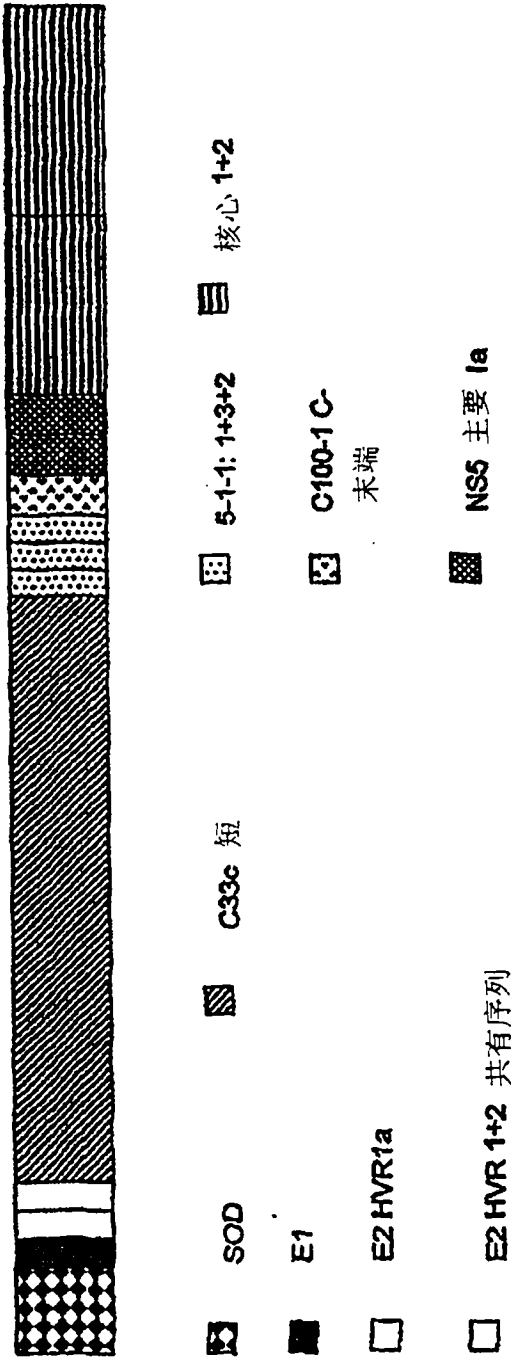


图 5

```

      1                               10
      M   A   T   K   A   V   C   V   L   K   G   D   G   P   V
      ATG GCT ACA AAG GCT GTT TGT GTT TTG AAG GGT GAC GGC CCA GTT 45

      20                               30
      Q   G   I   I   N   F   E   Q   K   E   S   N   G   P   V
      CAA GGT ATT ATT AAC TTC GAG CAG AAG GAA AGT AAT GCA CCA GTG 90

      40
      K   V   W   G   S   I   K   G   L   T   E   G   L   H   G
      AAG GTG TGG GGA AGC ATT AAA GGA CTG ACT GAA GGC CTG CAT GGA 135

      50                               60
      F   H   V   H   E   F   G   D   N   T   A   G   C   T   S
      TTC CAT GTT CAT GAG TTT GGA GAT AAT ACA GCA GGC TGT ACC AGT 180

      70
      A   G   P   H   F   N   P   L   S   T   R   G   C   N   C
      GCA GGT CCT CAC TTT AAT CCT CTA TCC ACG CGT GGT TGC AAT TGC 225

      80                               90
      S   I   Y   P   G   H   I   T   G   H   R   M   A   W   K
      TCT ATC TAT CCC GGC CAT ATA ACG GGT CAC CGC ATG GCA TGG AAG 270

      100
      L   G   S   A   A   R   T   T   S   G   F   V   S   L   F
      CTT GGT TCC GCC GCC AGA ACT ACC TCG GGC TTT GTC TCC TTG TTC 315

      110                               120
      A   P   G   A   K   Q   N   E   T   H   V   T   G   G   A
      GCC CCA GGT GCC AAA CAA AAC GAA ACT CAC GTC ACG GGA GGC GCA 360

      130
      A   A   R   T   T   S   G   L   T   S   L   F   S   P   G
      GCC GCC CGA ACT ACG TCT GGG TTG ACC TCT TTG TTC TCC CCA GGT 405

```

图 6A

```

      140                                150
A   S   Q   N   I   Q   L   I   T   S   T   D   N   S   S
GCC AGC CAA AAC ATT CAA TTG ATT ACT AGT ACG GAT AAC TCC TCT 450

      160
P   P   V   V   P   Q   S   F   Q   V   A   H   L   H   A
CCA CCA GTA GTG CCC CAG AGC TTC CAG GTG GCT CAC CTC CAT GCT 495

      170                                180
P   T   G   S   G   K   S   T   K   V   P   A   A   Y   A
CCC ACA GGC AGC GGC AAA AGC ACC AAG GTC CCG GCT GCA TAT GCA 540

      190
A   Q   G   Y   K   V   L   V   L   N   P   S   V   A   A
GCT CAG GGC TAT AAG GTG CTA GTA CTC AAC CCC TCT GTT GCT GCA 585

      200                                210
T   L   G   F   G   A   Y   M   S   K   A   H   G   I   D
ACA CTG GGC TTT GGT GCT TAC ATG TCC AAG GCT CAT GGG ATC GAT 630

      220
P   N   I   R   T   G   V   R   T   I   T   T   G   S   P
CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC CCC 675

      230                                240
I   T   Y   S   T   Y   G   K   F   L   A   D   G   G   C
ATC ACG TAC TCC ACC TAC GGC AAG TTC CTT GCC GAC GGC GGG TGC 720

      250
S   G   G   A   Y   D   I   I   I   C   D   E   C   H   S
TCG GGG GGC GCT TAT GAC ATA ATA ATT TGT GAC GAG TGC CAC TCC 765

      260                                270
T   D   A   T   S   I   L   G   I   G   T   V   L   D   Q
ACG GAT GCC ACA TCC ATC TTG GGC ATC GGC ACT GTC CTT GAC CAA 810

      280
A   E   T   A   G   A   R   L   V   V   L   A   T   A   T
GCA GAG ACT GCG GGG GCG AGA CTG GTT GTG CTC GCC ACC GCC ACC 855

      290                                300
P   P   G   S   V   T   V   P   H   P   N   I   E   E   V
CCT CCG GGC TCC GTC ACT GTG CCC CAT CCC AAC ATC GAG GAG GTT 900

```

图 6B

```

          310
A  L  S  T  T  G  E  I  P  F  Y  G  K  A  I
GCT CTG TCC ACC ACC GGA GAG ATC CCT TTT TAC GGC AAG GCT ATC 945

          320
P  L  E  V  I  K  G  G  R  H  L  I  F  C  H
CCC CTC GAA GTA ATC AAG GGG GGG AGA CAT CTC ATC TTC TGT CAT 990

          330
          340
S  K  K  K  C  D  E  L  A  A  K  L  V  A  L
TCA AAG AAG AAG TGC GAC GAA CTC GCC GCA AAG CTG GTC GCA TTG 1035

          350
          360
G  I  N  A  V  A  Y  Y  R  G  L  D  V  S  V
GGC ATC AAT GCC GTG GCC TAC TAC CGC GGT CTT GAC GTG TCC GTC 1080

          370
I  P  T  S  G  D  V  V  V  V  A  T  D  A  L
ATC CCG ACC AGC GGC GAT GTT GTC GTC GTG GCA ACC GAT GCC CTC 1125

          380
          390
M  T  G  Y  T  G  D  F  D  S  V  I  D  C  N
ATG ACC GGC TAT ACC GGC GAC TTC GAC TCG GTG ATA GAC TGC AAT 1170

          400
T  C  A  C  S  G  K  P  A  I  I  P  D  R  E
ACG TGT GCA TGC TCC GGG AAG CCG GCA ATC ATA CCT GAC AGG GAA 1215

          410
          420
V  L  Y  R  E  F  D  E  M  E  E  C  S  Q  H
GTC CTC TAC CGA GAG TTC GAT GAG ATG GAA GAG TGC TCT CAG CAC 1260

          430
L  P  Y  I  E  Q  G  M  M  L  A  E  Q  F  K
TTA CCG TAC ATC GAG CAA GGG ATG ATG CTC GCC GAG CAG TTC AAG 1305

          440
          450
Q  K  A  L  G  L  S  R  G  G  K  P  A  I  V
CAG AAG GCC CTC GGC CTC TCG CGA GGG GGC AAG CCG GCA ATC GTT 1350

          460
P  D  K  E  V  L  Y  Q  Q  Y  D  E  M  E  E
CCA GAC AAA GAG GTG TTG TAT CAA CAA TAC GAT GAG ATG GAA GAG 1395

```

图 6C

```

      470                                     480
C   S   Q   A   A   P   Y   I   E   Q   A   Q   V   I   A
TGC TCA CAA GCT GCC CCA TAT ATC GAA CAA GCT CAG GTA ATA GCT 1440

      490
H   Q   F   K   E   K   V   L   G   L   I   D   N   D   Q
CAC CAG TTC AAG GAA AAA GTC CTT GGA TTG ATC GAT AAT GAT CAA 1485

      500                                     510
V   V   V   T   P   D   K   E   I   L   Y   E   A   F   D
GTG GTT GTG ACT CCT GAC AAA GAA ATC TTA TAT GAG GCC TTT GAT 1530

      520
E   M   E   E   C   A   S   K   A   A   L   I   E   E   G
GAG ATG GAA GAA TGC GCC TCC AAA GCC GCC CTC ATT GAG GAA GGG 1575

      530                                     540
Q   R   M   A   E   M   L   K   S   K   I   Q   G   L   L
CAG CGG ATG GCG GAG ATG CTC AAG TCT AAG ATA CAA GGC CTC CTC 1620

      550
G   I   L   R   R   H   V   G   P   G   E   G   A   V   Q
GGG ATA CTG CGC CGG CAC GTT GGT CCT GGC GAG GGG GCA GTG CAG 1665

      560                                     570
W   M   N   R   L   I   A   F   A   S   R   G   N   H   V
TGG ATG AAC CGG CTG ATA GCC TTC GCC TCC AGA GGG AAC CAT GTT 1710

      580
S   P   T   H   Y   V   P   S   R   S   R   R   F   A   Q
TCC CCC ACG CAC TAC GTT CCG TCT AGA TCC CGG AGA TTC GCC CAG 1755

      590                                     600
A   L   P   V   W   A   R   P   D   Y   N   P   P   L   V
GCC CTG CCC GTT TGG GCG CGG CCG GAC TAT AAC CCC CCG CTA GTG 1800

      610
E   T   W   K   K   P   D   Y   E   P   P   V   V   H   G
GAG ACG TGG AAA AAG CCC GAC TAC GAA CCA CCT GTG GTC CAC GGC 1845

      620                                     630
R   S   S   R   R   F   A   Q   A   L   P   V   W   A   R
AGA TCT TCT CGG AGA TTC GCC CAG GCC CTG CCC GTT TGG GCG CGG 1890

```

图 6D

640
 P D Y N P P L V E T W K K P D
 CCG GAC TAT AAC CCC CCG CTA GTG GAG ACG TGG AAA AAG CCC GAC 1935

650
 Y E P P V V H G R K T K R N T
 TAC GAA CCA CCT GTG GTC CAT GGC AGA AAG ACC AAA CGT AAC ACC 1980

670
 N R R P Q D V K F P G G G Q I
 AAC CGG CGG CCG CAG GAC GTC AAG TTC CCG GGT GGC GGT CAG ATC 2025

680
 V G G V Y L L P R R G P R L G
 GTT GGT GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG GGT 2070

700
 V L A T R K T S P I P K A R R
 GTG CTC GCG ACG AGA AAG ACT TCC CCT ATC CCC AAG GCT CGT CGG 2115

710
 P E G R T W A Q P G Y P W P L
 CCC GAG GGC AGG ACC TGG GCT CAG CCC GGT TAC CCT TGG CCC CTC 2160

730
 Y G N K D R R S T G K S W G K
 TAT GGC AAT AAG GAC AGA CGG TCT ACA GGT AAG TCC TGG GGT AAG 2205

740
 P G Y P W P R K T K R N T N R
 CCA GGG TAC CCT TGG CCA AGA AAG ACC AAA CGT AAC ACC AAC CGG 2250

760
 R P Q D V K F P G G G Q I V G
 CCG CCG CAG GAC GTC AAG TTC CCG GGT GGC GGT CAG ATC GTT GGT 2295

770
 G V Y L L P R R G P R L G V L
 GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG GGT GTG CTC 2340

790
 A T R K T S P I P K A R R P E
 GCG ACG AGA AAG ACT TCC CCT ATC CCC AAG GCT CGT CGG CCC GAG 2385

图 6E

800 810
G R T W A Q P G Y P W P L Y G
GGC AGG ACC TGG GCT CAG CCC GGT TAC CCT TGG CCC CTC TAT GGC 2430

820
N K D R R S T G K S W G K P G
AAT AAG GAC AGA CGG TCT ACA GGT AAG TCC TGG GGT AAG CCA GGG 2475

829
Y P W P OC
TAC CCT TGG CCC TAA TGAGTCGAC

图 6F

MEFA12抗原构建物

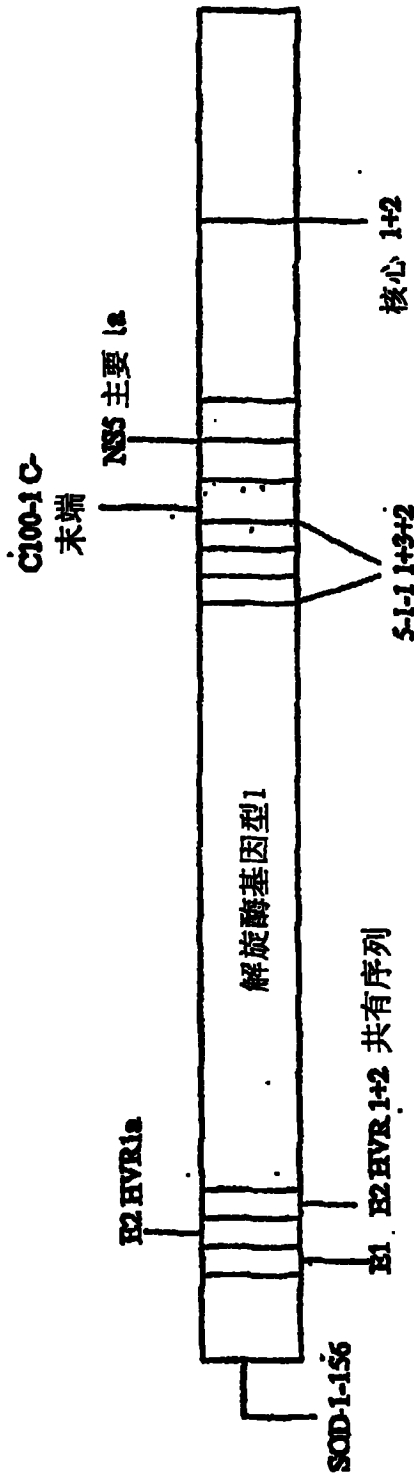


图 7

```

      1                               10
M A T K A V C V L K G D G P V
ATG GCT ACA AAG GCT GTT TGT GTT TTG AAG GGT GAC GGC CCA GTT

      20                               30
Q G I I N F E Q K E S N G P V
CAA GGT ATT ATT AAC TTC GAG CAG AAG GAA AGT AAT GGA CCA GTG

      40
K V W G S I K G L T E G L H G
AAG GTG TGG GGA AGC ATT AAA GGA CTG ACT GAA GGC CTG CAT GGA

      50                               60
F H V H E F G D N T A G C T S
TTC CAT GTT CAT GAG TTT GGA GAT AAT ACA GCA GGC TGT ACC AGT

      70
A G P H F N P L S R K H G G P
GCA GGT CCT CAC TTT AAT CCT CTA TCC AGA AAA CAC GGT GGG CCA

      80                               90
K D E E R H V G D L G N V T A
AAG GAT GAA GAG AGG CAT GTT GGA GAC TTG GGC AAT GTG ACT GCT

      100
D K D G V A D V S I E D S V I
GAC AAA GAT GGT GTG GCC GAT GTG TCT ATT GAA GAT TCT GTG ATC

      110                               120
S L S G D H C I I G R T L V V
TCA CTC TCA GGA GAC CAT TGC ATC ATT GGC GGC ACA CTG GTG GTC

      130
H E K A D D L G K G G N E E S
CAT GAA AAA GCA GAT GAC TTG GGC AAA GGT GGA AAT GAA GAA AGT

      140                               150
T K T G N A G S R L A C G V I
ACA AAG ACA GGA AAC GCT GGA AGT CGT TTG GCT TGT GGT GTA ATT

      160
G I A Q N L N S G C N C S I Y
GGG ATC GCC CAG AAT TTG AAT TCT GGT TGC AAT TGC TCT ATC TAT

      170                               180
P G H I T G H R M A W K L G S
CCC GGC CAT ATA ACG GGT CAC CGC ATG GCA TGG AAG CTT GGT TCC

      190
A A R T T S G F V S L F A P G
GCC GCC AGA ACT ACC TCG GGC TTT GTC TCC TTG TTC GCC CCA GGT

```

图 8A

```

      200
A K Q N E T H V T G G A A A
GCC AAA CAA AAC GAA ACT CAC GTC ACG GGA GGC GCA GCC GCC CGA

      210

T T S G L T S L F S P G A S Q
ACT ACG TCT GGG TTG ACC TCT TTG TTC TCC CCA GGT GCC AGC CAA

      220

N I Q L I V D F I P V E N L E
AAC ATT CAA TTG ATT GTC GAC TTT ATC CCT GTG GAG AAC CTA GAG

      230

T T M R S P V F T D N S S P P
ACA ACC ATG CGA TCT CCG GTG TTC ACG GAT AAC TCC TCT CCA CCA

      240

V V P Q S F Q V A H L H A P T
GTA GTG CCC CAG AGC TTC CAG GTG GCT CAC CTC CAT GCT CCC ACA

      250

G S G K S T K V P A A Y A A Q
GGC AGC GGC AAA AGC ACC AAG GTC CCG GCT GCA TAT GCA GCT CAG

      260

G Y K V L V L N P S V A A T L
GGC TAT AAG GTG CTA GTA CTC AAC CCC TCT GTT GCT GCA ACA CTG

      270

G F G A Y M S K A H G I D P N
GGC TTT GGT GCT TAC ATG TCC AAG GCT CAT GGG ATC GAT CCT AAC

      280

I R T G V R T I T T G S P I T
ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC CCC ATC ACG

      290

Y S T Y G K F L A D G G C S G
TAC TCC ACC TAC GGC AAG TTC CTT GCC GAC GGC GGG TGC TCG GGG

      300

G A Y D I I C D E C H S T D
GGC GCT TAT GAC ATA ATA ATT TGT GAC GAG TGC CAC TCC ACG GAT

      310

A T S I L G I G T V L D Q A H
GCC ACA TCC ATC TTG GGC ATT GGC ACT GTC CTT GAC CAA GCA GAG

      320

T A G A R L V V L A T A T P P
ACT GCG GGG GCG AGA CTG GTT GTG CTC GCC ACC GCC ACC CCT CCG

      330

G S V T V P H P N I H E V A L
GGC TCC GTC ACT GTG CCC CAT CCC AAC ATC GAG GAG GTT GCT CTG

      340

      350

      360

      370

      380

      390

      400

      410
      420

```

图 8B

```

      S T T G E I P F Y G K A I P L
TCC ACC ACC GGA GAG ATC CCT TTT TAC GGC AAG GCT ATC CCC CTC

      430
      H V I K G G R H L I F C H S K
GAA GTA ATC AAG GGG GGG AGA CAT CTC ATC TTC TGT CAT TCA AAG

      440
      K K C D E L A A K L V A L G I
AAG AAG TGC GAC GAA CTC GCC GCA AAG CTG GTC GCA TTG GGC ATC

      450
      N A V A Y Y R G L D V S V I P
AAT GCC GTG GCC TAC TAC CGC GGT CTT GAC GTG TCC GTC ATC CCG

      460
      T S G D V V V V A T D A L M T
ACC AGC GGC GAT GTT GTC GTC GTG GCA ACC GAT GCC CTC ATG ACC

      470
      G Y T G D F D S V I D C N T C
GGC TAT ACC GGC GAC TTC GAC TCG GTG ATA GAC TGC AAT ACG TGT

      480
      V T Q T V D F S L D P T F T I
GTC ACC CAG ACA GTC GAT TTC AGC CTT GAC CCT ACC TTC ACC ATT

      490
      H T I T L P Q D A V S R T Q R
GAG ACA ATC ACG CTC CCC CAA GAT GCT GTC TCC CGC ACT CAA CGT

      500
      R G R T G R G K P G I Y R F V
CGG GGC AGG ACT GGC AGG GGG AAG CCA GGC ATC TAC AGA TTT GTG

      510
      A P G E R P S G N F D S S V L
GCA CCG GGG GAG CGC CCC TCC GGC ATG TTC GAC TCG TCC GTC CTC

      520
      C E C Y D A G C A W Y E L T P
TGT GAG TGC TAT GAC GCA GGC TGT GCT TGG TAT GAG CTC ACG CCC

      530
      A B T T V R L R A Y M N T P G
CCC GAG ACT ACA GTT AGG CTA CGA GCG TAC ATG AAC ACC CCG GGG

      540
      L P V C Q D H L E F W E G V F
CTT CCC GTG TGC CAG GAC CAT CTT GAA TTT TGG GAG GGC GTC TTT

      550
      T G L T H I D A H F L S Q T K
ACA GGC CTC ACT CAT ATA GAT GCC CAC TTT CTA TCC CAG ACA AAG

      560
      Q S G E M L P Y L V A Y Q A T

```

图 8C

```

CAG AGT GGG GAG AAC CTT CCT TAC CTG GTA GCG TAC CAA GCC ACC
                                     640
V   C   A   R   A   Q   A   P   P   P   S   W   D   Q   M
GTG TGC GCT AGG GCT CAA GCC CCT CCC CCA TCG TGG GAC CAG ATG
                                     650
W   K   C   L   I   R   L   K   P   T   L   H   G   P   T
TGG AAG TGT TTG ATT CGC CTC AAG CCC ACC CTC CAT GGG CCA ACA
                                     660
P   L   L   Y   R   L   G   A   V   Q   N   E   I   T   L
CCC CTG CTA TAC AGA CTG GGC GCT GTT CAG AAT GAA ATC ACC CTG
                                     670
T   H   P   V   T   K   Y   I   M   T   C   M   S   A   D
ACG CAC CCA GTC ACC AAA TAC ATC ATG ACA TGC ATG TCG GCC GAC
                                     680
L   H   V   V   T   S   A   C   S   G   K   P   A   I   I
CTG GAG GTC GTC ACG AGC GCA TGC TCC GGG AAG CCG GCA ATC ATA
                                     690
P   D   R   E   V   L   Y   R   E   F   D   E   M   E   E
CCT GAC AGG GAA GTC CTC TAC CGA GAG TTC GAT GAG ATG GAA GAG
                                     700
C   S   Q   H   L   P   Y   I   E   Q   G   M   M   L   A
TGC TCT CAG CAC TTA CCG TAC ATC GAG CAA GGG ATG ATG CTC GCC
                                     710
E   Q   F   K   Q   K   A   L   G   L   S   R   G   G   K
GAG CAG TTC AAG CAG AAG GCC CTC GGC CTC TCG CGA GGG GGC AAG
                                     720
P   A   I   V   P   D   K   E   V   L   Y   Q   Q   Y   D
CCG GCA ATC GTT CCA GAC AAA GAG GTG TTG TAT CAA CAA TAC GAT
                                     730
E   M   E   E   C   S   Q   A   A   P   Y   I   E   Q   A
GAG ATG GAA GAG TGC TCA CAA GCT GCC CCA TAT ATC GAA CAA GCT
                                     740
Q   V   I   A   H   Q   P   K   E   K   V   L   G   L   I
CAG GTA ATA GCT CAC CAG TTC AAG GAA AAA GTC CTT GGA TTG ATC
                                     750
D   N   D   Q   V   V   V   T   P   D   K   H   I   L   Y
GAT AAT GAT CAA GTG GTT GTG ACT CCT GAC AAA GAA ATC TTA TAT
                                     760
E   A   F   D   E   M   E   E   C   A   S   K   A   A   L
GAG GCC TTT GAT GAG ATG GAA GAA TGC GCC TCC AAA GCC GCC CTC
                                     770
I   E   E   G   Q   R   M   A   E   M   L   K   S   K   I
ATT GAG GAA GGG CAG CGG ATG GCG GAG ATG CTC AAG TCT AAG ATA
                                     780

```

图 8D

```

      850
Q  G  L  L  G  I  L  R  R  H  V  G  F  G  H
CAA GGC CTC CTC GGG ATA CTG CGC CGG CAC GTT GGT CCT GGC GAG

      860
G  A  V  Q  W  M  N  R  L  I  A  F  A  S  R
GGG GCA GTG CAG TGG ATG AAC CGG CTG ATA GCC TTC GCC TCC AGA

      870
G  N  H  V  S  P  T  H  Y  V  P  S  R  S  R
GGG AAC CAT GTT TCC CCC ACG CAC TAC GTT CCG TCT AGA TCC CGG

      880
R  F  A  Q  A  L  P  V  W  A  R  P  D  Y  N
AGA TTC GCC CAG GCC CTG CCC GTT TGG GCG CGG CCG GAC TAT AAC

      890
P  P  L  V  E  T  W  K  K  P  D  Y  E  P  P
CCC CCG CTA GTG GAG ACG TGG AAA AAG CCC GAC TAC GAA CCA CCT

      900
V  V  H  G  R  S  S  R  R  F  A  Q  A  L  P
GTG GTC CAC GGC AGA TCT TCT CGG AGA TTC GCC CAG GCC CTG CCC

      910
V  W  A  R  P  D  Y  N  P  P  L  V  E  T  W
GTT TGG GCG CGG CCG GAC TAT AAC CCC CCG CTA GTG GAG ACG TGG

      920
K  K  P  D  Y  E  P  P  V  V  H  G  R  K  T
AAA AAG CCC GAC TAC GAA CCA CCT GTG GTC CAT GGC AGA AAG ACC

      930
K  R  N  T  N  R  R  P  Q  D  V  K  F  P  G
AAA CGT AAC ACC AAC CGG CGG CCG CAG GAC GTC AAG TTC CCG GGT

      940
G  G  Q  I  V  G  R  R  G  P  P  I  P  K  A
GGC GGT CAG ATC GTT GGT CGC AGG GGC CCT CCT ATC CCC AAG GCT

      950
R  R  P  E  G  R  T  W  A  Q  P  G  Y  P  W
CGT CCG CCC GAG GGC AGG ACC TGG GCT CAG CCC GGT TAC CCT TGG

      960
P  L  Y  G  N  K  D  R  R  S  T  G  K  S  W
CCC CTC TAT GGC AAT AAG GAC AGA CCG TCT ACA GGT AAG TCC TGG

      970
G  K  P  G  Y  P  W  P  R  K  T  K  R  N  T
GGT AAG CCA GGG TAC CCT TGG CCA AGA AAG ACC AAA CGT AAC ACC

      980
N  R  R  P  Q  D  V  K  F  P  G  G  G  Q  I
AAC CGA CCG CCG CAG GAC GTC AAG TTC CCG GGT GGC GGT CAG ATC

      990

```

图 8E

1060
V G R R G P P I P K A R R P E
GTT GGT CGC AGG GGC CCT CCT ATC CCC AAG GCT CGT CGG CCC GAG

1070
G R T W A Q P G Y P W P L Y G
GGC AGG ACC TGG GCT CAG CCC GGT TAC CCT TGG CCC CTC TAT GGC

1080
N K D R R S T G K S W G K P G
AAT AAG GAC AGA CGG TCT ACC GGT AAG TCC TGG GGT AAG CCA GGG

1090
Y P W P
TAT CCT TGG CCC

图 8F

MEFA-3 抗原

hSOD-(1-154)	核心	核心	c33c	5-1-1 1型	5-1-1 3型	5-1-1 2型	C-100	C-100	NS5	NS5
氨基酸	10 53	10 53	1192 53	1694 1457	1694 1735	1694 1735	1901 1940	1901 1940	2278 2310	2278 2310

MEFA-5 抗原

hSOD-(1-154)	核心	核心	E1	E2	c33c	5-1-1 1型	5-1-1 3型	5-1-1 2型	C-100	NS5
氨基酸	10 53	10 53	303 320	405 444	1192 1457	1689 1735	1689 1735	1689 1735	1901 1940	2278 2313

MEFA-6 抗原

hSOD-(1-154)	E1	E2	c33c	5-1-1 1型	5-1-1 3型	5-1-1 2型	C-100	NS5	核心	核心
氨基酸	303 320	405 444	1192 1457	1689 1735	1689 1735	1689 1735	1901 1940	2278 2313	10 53	10 53

图 9