[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
C07K 14/18
A61K 39/29 G01N 33/576
C12Q 1/70 C12N 15/40

[12] 发明专利申请公开说明书

[21] 申请号 02800609.7

[43] 公开日 2004 年 11 月 17 日

[11] 公开号 CN 1547588A

[22] 申请日 2002.1.11 [21] 申请号 02800609.7
[30] 优先权
[32] 2001. 1.11 [33] US [31] 60/260,669
[32] 2001. 8.30 [33] US [31] 60/315,768
[86] 国际申请 PCT/EP2002/000219 2002.1.11
[87] 国际公布 WO2002/055548 英 2002.7.18
[85] 进入国家阶段日期 2002.11.11
[71] 申请人 基因创新有限公司
地址 比利时根特
[72] 发明人 G·梅尔藤斯 F·波斯曼
E·德普

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 张广育 孟凡宏

权利要求书 4 页 说明书 89 页 序列表 101 页
附图 65 页

[54] 发明名称 用于诊断和治疗用途的纯化的丙型肝炎病毒外被蛋白

[57] 摘要

本发明涉及一种用于纯化选自 E1 和/或 E2 和/或 E1/E2 的 HCV 单个的或特异性寡聚外被蛋白的方法，其特征在于，在为了分离所述重组表达的蛋白而裂解转化过的宿主细胞时，用一种二硫键裂解剂进行二硫键裂解或还原的步骤。 本发明还涉及一种通过该方法所分离的组合物。 本发明还涉及所述组合物的诊断和治疗用途。 另外，本发明涉及将 HCV E1 蛋白和肽用于预测和监测 HCV 治疗的临床效果和/或临床结果的用途。

ISSN1008-4274

1. 一种治疗用疫苗组合物，含有治疗有效量的:

含有选自 E1 蛋白和 E2 蛋白的至少一种纯化的重组 HCV 单个的或特异性寡聚的重组外被蛋白；并选择性地含有可以药用的佐剂的组合物。

2. 如权利要求 1 的组合物，其中，所述重组 HCV 外被蛋白是由重组哺乳动物细胞产生的。

3. 如权利要求 1 的组合物，其中，所述重组 HCV 外被蛋白是由重组酵母细胞生产的。

4. 一种治疗用疫苗组合物，含有治疗有效量的含有至少一种下列 E1 和 E2 肽的组合物:

跨越核心/E1 V1 区的 181-200 号氨基酸的 E1-31（SEQ ID NO: 56），

跨越 E1 区的 193-212 号氨基酸的 E1-33（SEQ ID NO: 57），

跨越 E1 V2 区的 205-224 号氨基酸（表位 B）的 E1-35（SEQ ID NO: 58），

跨越 E1 V2 区的 208-227 号氨基酸（表位 B）的 E1-35A（SEQ ID NO: 59），

跨越 E1 区 V1、C1、和 V2 区的 192-228 号氨基酸（含有表位 B）的 1bE1（SEQ ID NO: 53），

跨越 E1 区 301-320 号氨基酸的 E1-51（SEQ ID NO: 66），

跨越 E1 C4 区的 313-332 号氨基酸（表位 A）的 E1-53（SEQ ID NO: 67），

跨越 E1 区的 325-344 号氨基酸的 E1-55（SEQ ID NO: 68），

跨越 E2 区的 397-418 号氨基酸（表位 A）的 Env67 或 E2-67（SEQ ID NO: 72），

跨越 E2 区的 409-428 号氨基酸（表位 A）的 Env69 或 E2-69（SEQ ID NO: 73），

跨越 E2 区的 583-602 号位置（表位 A）的 Env23 或 E2-23（SEQ ID NO: 86），

跨越 E2 区的 595-614 号位置（表位 E）的 Env25 或 E2-25（SEQ ID NO: 87），

跨越E2区的607-626号位置(表位E)的Env27或E2-27(SEQ ID NO: 88),

跨越E2区的547-586号位置(表位D)的Env17B或E2-17B(SEQ ID NO: 83),

跨越E2区的523-542号位置(表位C)的Env13B或E2-13B(SEQ ID NO: 82),

跨越E1区的192-211号位置的IGP1626 (SEQ ID NO: 112),

跨越E1区的204-223号位置的IGP1627(SEQ ID NO: 113),

跨越E1区的192-211号位置的IGP1626(SEQ ID NO: 112),

跨越E1区的204-223号位置的IGP1627 (SEQ ID NO: 113),

跨越E1区的216-235号位置的IGP1628(SEQ ID NO: 114),

跨越E1区的228-247号位置的IGP1629 (SEQ ID NO: 115),

跨越E1区的240-259号位置的IGP1630 (SEQ ID NO: 116),

跨越E1区的252-271号位置的IGP1631 (SEQ ID NO: 117),

跨越E1区的264-283号位置的IGP1632 (SEQ ID NO: 118),

跨越E1区的276-295号位置的IGP1633 (SEQ ID NO: 119),

跨越E1区的288-307号位置的IGP1634(SEQ ID NO: 120),

跨越E1区的300-319号位置的IGP1635 (SEQ ID NO: 121),

和

跨越E1区的312-331号位置的IGP1636 (SEQ ID NO: 122)。

5. 一种治疗感染了HCV的哺乳动物的方法,包括给予有效量的如权利要求1-4中任意一项的组合物,并选择性地给予可以药用的佐剂。

6. 如权利要求5的方法,其中,所述哺乳动物是人。

7. 一种含有选自E1蛋白和E2蛋白的至少一种纯化的重组HCV重组外被蛋白,并选择性地含有一种佐剂的组合物。

8. 一种含有至少一种下列E1和E2肽的组合物:

跨越核心/E1 V1区的181-200号氨基酸的E1-31(SEQ ID NO: 56),

跨越E1区的193-212号氨基酸的E1-33(SEQ ID NO: 57),

跨越E1 V2区的205-224号氨基酸(表位B)的E1-35(SEQ ID NO: 58),

跨越E1 V2区的208-227号氨基酸(表位B)的E1-35A(SEQ ID

NO: 59)，

跨越E1区V1、C1、和V2区的192-228号氨基酸（含有表位B）的1bE1（SEQ ID NO: 53），

跨越E1区301-320号氨基酸的E1-51（SEQ ID NO: 66），

跨越E1 C4区的313-332号氨基酸(表位A)的E1-53(SEQ ID NO: 67)，

跨越E1区的325-344号氨基酸的E1-55（SEQ ID NO: 68），

跨越E2区的397-418号氨基酸(表位A)的Env67或E2-67(SEQ ID NO: 72)，

跨越E2区的409-428号氨基酸(表位A)的Env69或E2-69(SEQ ID NO: 73)，

跨越E2区的583-602号位置(表位A)的Env23或E2-23(SEQ ID NO: 86)，

跨越E2区的595-614号位置(表位E)的Env25或E2-25(SEQ ID NO: 87)，

跨越E2区的607-626号位置(表位E)的Env27或E2-27(SEQ ID NO: 88)，

跨越E2区的547-586号位置(表位D)的Env17B或E2-17B(SEQ ID NO: 83)，

跨越E2区的523-542号位置(表位C)的Env13B或E2-13B(SEQ ID NO: 82)，

跨越E1区的192-211号位置的IGP1626 (SEQ ID NO: 112)，

跨越E1区的204-223号位置的IGP1627(SEQ ID NO: 113)，

跨越E1区的192-211号位置的IGP1626(SEQ ID NO: 112)，

跨越E1区的204-223号位置的IGP1627 (SEQ ID NO: 113)，

跨越E1区的216-235号位置的IGP1628(SEQ ID NO: 114)，

跨越E1区的228-247号位置的IGP1629 (SEQ ID NO: 115)，

跨越E1区的240-259号位置的IGP1630 (SEQ ID NO: 116)，

跨越E1区的252-271号位置的IGP1631 (SEQ ID NO: 117)，

跨越E1区的264-283号位置的IGP1632 (SEQ ID NO: 118)，

跨越E1区的276-295号位置的IGP1633 (SEQ ID NO: 119)，

跨越E1区的288-307号位置的IGP1634(SEQ ID NO: 120)，

跨越E1区的300-319号位置的IGP1635（SEQ ID NO: 121），和

跨越E1区的312-331号位置的IGP1636（SEQ ID NO: 122）。

9. 一种用于诱导HCV特异性抗体的治疗用组合物，含有治疗有效量的含有由纯化的重组HCV单个的或特异性寡聚的重组E1或E2蛋白形成的E1/E2复合物；并选择性地含有可以药用的佐剂的组合物。

10. 如权利要求9的组合物，其中，所述重组HCV外被蛋白是由重组哺乳动物细胞生产的。

11. 如权利要求9的组合物，其中，所述重组HCV外被蛋白是由重组酵母细胞生产的。

12. 一种治疗感染了HCV的哺乳动物的方法，包括给予有效量的如权利要求9-11中任意一项的组合物，并选择性地给予可以药用的佐剂。

13. 如权利要求12的方法，其中，所述哺乳动物是人。

14. 一种用于诱导HCV特异性抗体的药用组合物，含有治疗有效量的含有选自E1蛋白和E2蛋白的至少一种纯化的重组HCV单个的或特异性寡聚的重组外被蛋白；并选择性地含有可以药用的佐剂的组合物。

用于诊断和治疗用途的纯化的丙型肝炎病毒外被蛋白

发明领域

本发明涉及重组蛋白表达、重组蛋白纯化、合成肽、HCV 感染诊断、对 HCV 感染进行预性防治疗的一般领域，并且涉及患有慢性肝炎个体的治疗的临床效果的预测/监测治疗效果，或预测/监测天然疾病。

更具体地讲，本发明涉及纯化丙型肝炎病毒外被蛋白的纯化方法，按照本发明所披露的方法所纯化的 HCV 外被蛋白在诊断、预防或治疗中的用途，以及单个的或特异性寡聚 E1 和/或 E2 和/或 E1/E2 外被蛋白在监测疾病、和/或诊断疾病、和/或治疗疾病的测定方法中的用途。本发明还涉及 E1 和/或 E2 外被蛋白表位，以及它的单克隆抗体及其在诊断、预防或治疗中的用途。

发明背景

按照本发明所披露的方法，从细胞裂解物中纯化的 E2 蛋白能够与大约 95%的患者的血清起反应。这种反应性类似于由 CHO 细胞中分泌的 E2 所获得的反应性（Spaete 等，1992）。不过，细胞内表达形式的 E2 可能更接近于天然病毒外被蛋白，因为它含有高含量的甘露糖碳水化合物基序，而由 CHO 细胞中分泌的 E2 蛋白，由半乳糖和唾液酸糖成分进行过进一步的修饰。当 E2 的氨基末端部分是在杆状病毒系统中表达时，只能在来自若干患者组的大约 13-21%的血清中检测到（Inoue 等。1992）。在用大肠杆菌表达 E2 之后，HCV 血清的反应性甚至更低，在 14（Yokosuka 等，1992）-17%（Mita 等，1992）范围内。使用本发明的纯化的、疫苗表达的重组 E1 蛋白，大约有 75%的 HCV 血清（以及 95%的慢性患者）是抗 E1 阳性的，与 Kohara 等（1992）和 Hsu 等（1993）的结果形成了鲜明的对比。Kohara 等使用痘苗病毒表达的 E1 蛋白，并且在 7-23%的患者体内检测到抗 E1 抗体，而 Hsu 等使用杆状病毒表达的 E1 检测了 14/50（28%）的血清。

以上结果表明，为了获得所述外被蛋白与人类患者血清的高的反应性，不仅需要好的表达系统，而且还需要好的纯化方法。这一目的可以通过使用本发明的正确的表达系统和/或纯化方法而实现，该方法能

确保保持所述蛋白的天然折叠，并且本发明的纯化方法可以确保消除污染性蛋白，它能保持 HCV 外被蛋白的构象，并因此保持其反应性。诊断性筛选测定所需要的纯化 HCV 外被蛋白的量为每年数克。对于疫苗用途来说，可能需要更大量的外被蛋白。因此，可以利用所述痘苗病毒系统筛选最佳的表达构建体，并且用于有限的扩大表达，并且，在用若干酵母菌株表达时，可以实现大规模表达和纯化含有高甘露糖碳水化合物的单个的或特异性的寡聚外被蛋白。例如，对于乙型肝炎来说，与酵母产生的乙型肝炎疫苗相比，用哺乳动物细胞生产 H BsAg 的成本更高。

本发明的目的

本发明的一个目的是提供一种纯化重组表达的 E1 和/或 E2 和/或 E1/E2 蛋白的新方法，以便能够以不含杂质的单个的或特异性寡聚重组蛋白形式而不是凝聚体的形式，将所述重组蛋白直接用于诊断和疫苗接种目的。

本发明的另一个目的是提供含有纯化的（单个的或特异性寡聚的）重组 E1 和/或 E2 和/或 E1/E2 糖蛋白的组合物，所述糖蛋白含有来自 HCV 的 E1 和/或 E2 和/或 E1/E2 结构域的构象表位。

本发明的另一个目的是提供用于重组表达 E1 和/或 E2 和/或 E1/E2 蛋白的新的重组载体构建体，以及用所述载体构建体转化过的宿主细胞。

本发明的另一个目的是提供一种用于生产并纯化重组 HCV E1 和/或 E2 和/或 E1/E2 蛋白的方法。

本发明的另一个目的是提供本发明的重组 HCV E1 和/或 E2 和/或 E1/E2 蛋白的诊断和免疫原用途，以及提供将含有本发明的任意一种重组 HCV E1 和/或 E2 和/或 E1/E2 蛋白的用于诊断目的的试剂盒、疫苗或治疗剂。

本发明的另一个目的是提供将 E1 和/或 E2 和/或 E1/E2 蛋白或其适当的部分用于监测/预测对患有 HCV 感染的患者进行治疗（例如，用干扰素治疗）的反应的新的用途。

本发明的另一个目的是提供将本发明的重组 E1 和/或 E2 和/或 E1/E2 蛋白用于 HCV 筛选和证实抗体测定的用途。

本发明的另一个目的是提供可用于诊断 HCV 感染并用于制备抗体的 E1 和/或 E2 肽。所述肽还可用于分离人单克隆抗体。

本发明的另一个目的是提供单克隆抗体，尤其是人单克隆抗体或人源化的小鼠单克隆抗体，这种抗体能与肽或重组蛋白中的构象表位所包含的 E1 和/或 E2 表位特异性地起反应。

本发明的另一个目的是提供将抗 E1 或抗 E2 单克隆抗体用于 HCV 抗原检测或用于慢性 HCV 感染治疗的可能的用途。

本发明的再一个目的是提供用于监测/预测对患有 HCV 感染的患者进行治疗（例如，用干扰素治疗）的反应或监测/预测所述疾病的后果的试剂盒。

本发明的所有目的都可以通过下面所提供的实施方案而实现。

定义

下面的定义用于解释在本发明中所使用的不同的术语和表达形式。

术语‘丙型肝炎病毒单个外被蛋白’表示含有一种氨基酸序列（和/或氨基酸类似物）的多肽或其类似物（例如，模拟表位），该序列确定了 E1 或 E2 区的至少一个 HCV 表位。所述单个外被蛋白在其广义的含义上可以是单体或同型寡聚物形式的、重组表达的外被蛋白。通常，限定所述表位的序列相当于 HCV 的 E1 或 E2 区的氨基酸序列（相同或者通过用天然氨基酸残基类似物取代而不会破坏该表位）。一般，表位限定序列的长度为 3 个或 3 个以上氨基酸，更优选 5 个或 5 个以上氨基酸，更优选 8 个或 8 个以上氨基酸，更优选 10 个或 10 个以上氨基酸。对于构象表位来说，所述表位限定序列的长度可以有较大的改变，因为确信这种表位是由抗原的三维形状形成的（例如，折叠）。因此，限定所述表位的氨基酸的数量可以比较少，而且广泛分布在该分子的长度上，通过折叠形成正确的表位构象。位于限定表位的残基之间的抗原部分，对于该表位的构象的结构来说可能并不重要。例如，只要保留了对表位构象重要的序列（例如，与二硫键结合有关的半胱氨酸，糖基化位点等），缺失或取代这些间插序列可能不会影响该构象表位。构象表位还可以通过同型寡聚物或异源寡聚物的亚基的两个或两个以上关键区而形成。

本发明的 HCV 抗原包括来自 HCV 的 E1 和/或 E2（外被蛋白）结构域的构象表位，被认为对应于病毒外被蛋白的 E1 结构域，目前估计跨越 HCV 多蛋白的 192-383 号氨基酸（Hijikada 等，1991）。在哺乳动物系统中进行表达（糖基化的）时，通过 SDS-PAGE 测定相信具有大约 35kDa 的分子量。据信以前被称为 NS1 的 E2 蛋白跨越 HCV 多蛋白的 384-809 或 384-746 号氨基酸（Grakoui 等，1993），并且它也是一种外被蛋白，在疫苗系统中表达时（糖基化的）据信它具有大约 72kDa 的表观凝胶分子量。相信这些蛋白的终点是大致的（例如，E2 的羧基末端可能位于 730-820 号氨基酸片段中的某一位置上，例如，以 730、735、740、742、744、745 号氨基酸为终点，优选以 746、747、748、750、760、770、780、790、800、809、810、820 号氨基酸为终点）。所述 E2 蛋白还可以与 E1、P7（747-809 号氨基酸）、NS2（810-1026 号氨基酸）、NS4A（1658-1711 号氨基酸）或 NS4B（1712-1972 号氨基酸）一起表达。与上述其他 HCV 蛋白一起表达，对于获得正确的蛋白折叠来说是重要的。

还应当理解的是，在本发明的实施例部分所使用的分离物不是要用来限定本发明范围的，并且来自 1、2、3、4、5、6、7、8、9、10 型或任何其他新基因型的 HCV 的任何 HCV 分离物都是用于实施本发明的 E1 和/或 E2 序列的合适来源。

用于本发明的 E1 和 E2 抗原可以是完整长度的病毒蛋白，其基本为全长的形式，或其功能性片段（例如，没有缺少形成或保留表位所必须的序列的片段）。另外，本发明的 HCV 抗原还可以包括不会妨碍或抑制形成感兴趣的构象表位的其他序列。通过用一种抗体（所述构象表位的多克隆或单克隆血清）筛选所述感兴趣的抗原，并且将它的反应性与仅仅保留了线性表位（如果有的话）的所述抗原的变性形式的反应性进行比较，可以方便地确定构象表位的存在与否。在所述用多克隆抗体进行的筛选中，优选首先用变性的抗原吸附所述多克隆血清，并观察是否能保留感兴趣抗原的抗体。

本发明的 HCV 抗原可以通过能提供感兴趣的表位的任何重组方法制备。例如，对于天然 HCV 抗原来说，在哺乳动物或昆虫细胞中进行重组细胞内表达是提供具有“天然”构象的糖基化 E1 和/或 E2 抗原的优选方法。酵母细胞和突变型酵母菌株（例如，mnn9 突变体（Kniskern

等，1994）或通过矾酸盐抗性筛选所产生的糖基化变体（Ballou 等，1991））很适合生产分泌的高甘露糖型糖类；而由哺乳动物细胞分泌的蛋白可能含有包括半乳糖或唾液酸在内的修饰，这种修饰对于某些诊断或疫苗用途来说可能是不需要的。不过，对于某些用途来说是可行的并且是足够的，因为对蛋白来说，已知可以在其他重组宿主（如大肠杆菌）中表达所述抗原，并且在回收之后对该蛋白进行复性。

术语‘融合多肽’是指这样一种多肽，其中，HCV 抗原是单个的连续氨基酸链的一部分，该氨基酸链不是天然存在的。所述 HCV 抗原可以通过肽键直接彼此连接，或者通过间插氨基酸序列分开。所述融合多肽还可以含有对于 HCV 来说是外源的氨基酸序列。

术语‘固相’是指一种固体物质，单个的 HCV 抗原或由 HCV 抗原组成的融合多肽通过共价结合或诸如疏水性吸附的非共价性结合形式与它连接。

术语‘生物学样品’是指哺乳动物个体（例如，类人猿、人）的体液或组织，它通常含有由所述个体所产生的抗体，尤其是抗 HCV 的抗体。所述体液或组织还可以含有 HCV 抗原。所述成分为本领域所公知，并且包括，但不限于血液、血浆、血清、尿液、脑脊液、淋巴液、呼吸道、肠道或泌尿生殖道的分泌物、眼泪、唾液、乳汁、白细胞和骨髓。身体成分包括生物学液体。术语‘生物学液体’是指从一种生物体内获得的体液。某些生物学体液被用作其他制品的来源，如凝血因子（例如，因子 VIII；C）、血清白蛋白、生长激素等。在这种情况下，重要的是生物学液体的来源没有受到诸如 HCV 的病毒的污染。

术语‘免疫学反应性’表示有关抗原能够与存在于来自 HCV 感染过的个体的身体成分中的抗 HCV 抗体特异性地反应。

术语‘免疫复合物’是指当抗体与抗原上的表位结合以后所形成的组合物。

本文所说的‘E1’是指在 HCV 多蛋白的前 400 个氨基酸内表达的蛋白或多肽，有时候是指 E、ENV 或 S 蛋白。其天然形式是一种 35kDa 的糖蛋白，发现它与膜紧密结合。在大多数天然 HCV 菌株中，E1 蛋白是在病毒多蛋白中紧随 C（核心）蛋白之后编码的。E1 蛋白从所述完整长度多蛋白的大约 192 号氨基酸延伸至 383 号氨基酸。

本文所使用的术语‘E1’还包括能与天然 E1 发生免疫交叉反应的

类似的或截短的形式，并且包括 1、2、3、4、5、6、7、8、9、10 基因型或任何其他新鉴定的 HCV 类型或亚型的 E1 蛋白。

本文所说的 'E2' 是指在 HCV 多蛋白的前 900 个氨基酸内表达的蛋白或多肽，有时候是指 NS1 蛋白。其天然形式是一种 72kDa 的糖蛋白，发现它与膜紧密结合。在大多数天然 HCV 菌株中，E2 蛋白是在所述病毒多蛋白中紧随 E1 蛋白之后编码的。E2 蛋白分布在大约 384 氨基酸至 746 号氨基酸之间，另一种形式的 E2 延伸到 809 号氨基酸。本文所使用的术语 'E2' 还表示能与天然 E2 发生免疫交叉反应的类似的或截短的形式。例如，Kato 等（1992）业已报导了在 383 和 384 号密码子之间插入多个密码子，以及缺失 384-387 号氨基酸。

本文所使用的 'E1/E2' 表示含有至少一个 E1 成分和至少一个 E2 成分的外被蛋白的寡聚物形式。

术语 '特异性寡聚' E1 和/或 E2 和/或 E1/E2 外被蛋白表示重组表达的 E1 和/或 E2 外被蛋白的非凝聚的所有可能的寡聚形式。E1 和/或 E2 特异性寡聚外被蛋白还称作同型寡聚 E1 或 E2 外被蛋白（参见下文）。

术语 '单个的或特异性寡聚' E1 和/或 E2 和/或 E1/E2 外被蛋白表示单个的单体 E1 和/或 E2 蛋白（在该术语的严格含义上是单个的）以及特异性寡聚 E1 和/或 E2 和/或 E1/E2 重组表达的蛋白。本发明的所述单个的或特异性寡聚外被蛋白还可以通过以下公式定义：$(E1)_x(E2)_y$，其中，x 可以是 0-100 之间的数，而 y 可以是 0-100 之间的数，前提是 x 和 y 不能同时为 0。当 x=1，而 y=0 时，所述外被蛋白包括单体 E1。

本文所使用的术语 '同型寡聚物' 是指含有一个以上 E1 或 E2 单体的 E1 和/或 E2 的复合物，例如，E1/ E1 二聚体、E1/ E1/E1 三聚体、E1/ E1/E1/E1 四聚体和 E2/ E2 二聚体、E2/ E2/E2 三聚体、E2/ E2/E2/E2 四聚体、E1 五聚体和六聚体、E2 五聚体和六聚体、或 E1 或 E2 的任何更高数目的同型寡聚物，都是该定义范围内的 '同型寡聚物'。所述寡聚物可以含有来自不同类型或亚型丙型肝炎病毒的 E1 或 E2 的一个、二个或数个不同的单体，例如，包括在本申请人的国际专利申请公开文件 WO94/25601 和欧洲专利申请号 94870166.9 中所披露的寡聚物，所述混合的寡聚物仍然属于本发明范围内的同型寡聚物，并且可以对 HCV 进行更通用的诊断、预防或治疗。

本文中用于蛋白上的术语‘纯化的’是指一种组合物，其中，所需要的蛋白至少占该组合物中总蛋白成分的35%。所述需要的蛋白优选占总蛋白成分的至少40%，更优选至少约50%，更优选至少约60%，更优选至少约70%，更优选至少约80%，更优选至少约90%，最优选至少约95%。所述组合物还可以含有不会影响测定本文所使用的纯度百分比的其他化合物，如碳水化合物、盐类、脂类和溶剂等。‘分离的’HCV蛋白是指纯度至少为35%的HCV蛋白组合物。

术语‘基本上纯化的蛋白’是指纯化过的蛋白，以便可将其用于体外诊断方法并且用作治疗化合物。这种蛋白基本上不含细胞蛋白、载体衍生的蛋白或其他HCV病毒成分。通常，这些蛋白被纯化到均质，至少80%的纯度，优选90%，更优选95%，更优选97%，更优选98%，更优选99%，更优选99.5%，最优选纯化到通过诸如SDS-PAGE和银染色的常规方法检测不到污染蛋白的程度。

在本发明中所使用的术语‘重组表达’是指本发明的蛋白是通过重组表达方法在下面所披露的原核生物或低等或高等真核生物中生产的。

术语‘低等真核生物’是指诸如酵母和真菌之类的宿主细胞。低等真核生物通常是（但不一定是）单细胞。优选的低等真核生物有酵母，特别是糖酵母属、裂殖酵母属、克鲁维酵母属、毕赤酵母属（例如，巴斯德毕赤酵母）、汉逊酵母属（例如，多型汉逊酵母）、Yarowia、许盛酵母属、裂殖酵母属和接合酵母属等属内的种。酿酒酵母、卡尔酵母和乳克鲁维酵母是最常用的酵母宿主，并且是常见的真菌宿主。

术语‘原核生物’是指诸如大肠杆菌、乳杆菌、乳球菌、沙门氏菌、链球菌、枯草芽孢杆菌或链霉菌等宿主。这些宿主同样可用于本发明中。

术语‘高等真核生物’是指来自诸如哺乳动物、爬行动物和昆虫之类的高等动物的宿主细胞。本发明优选的高等真核生物宿主细胞来自中国仓鼠（例如，CHO）、猴（例如，COS和Vero细胞）、幼仓鼠肾（BHK）、猪肾（PK15）、兔肾13细胞（RK13）、人骨肉瘤细胞系143B、人细胞系HeLa，以及诸如Hep G2的人肝癌细胞系和昆虫细胞系（例如，草地贪夜蛾）。所述宿主细胞可以以悬浮液或烧瓶培养物、组织培养物和器官培养物之类的形式提供。另外，所宿主细胞还可以是转

基因动物。

术语‘多肽’是指氨基酸聚合物，而且不表示特定长度的产物；因此，肽、寡肽和蛋白都包括在多肽的定义范围内。该术语也不表示或排除对所述多肽的表达后修饰，例如，糖基化、乙酰化和磷酸化等。例如，该定义包括含有一个或多个氨基酸类似物的多肽（例如，包括非天然氨基酸，PNA 等）、具有取代过的键的多肽，以及本领域已知的其他修饰，包括天然发生的和非天然发生的形式。

术语‘重组多核苷酸或核酸’表示来自基因组、cDNA、半合成或合成来源的多核苷酸或核酸，由于其来源或操作：（1）它不与在天然状态下与它结合的多肽的全部或一部分结合，（2）与天然状态下不与它连接的多核苷酸连接，或（3）在天然状态下不存在。

术语‘重组宿主细胞’、‘宿主细胞’、‘细胞’、‘细胞系’、‘细胞培养物’以及其他作为单细胞实体培养的表示微生物或高等真核细胞系的诸如此类的术语，是指可以是或者曾经被用作重组载体或其他转移多核苷酸的受体的细胞，并且包括转染过的原始细胞的后代。可以理解的是，由于天然、意外或人为突变，所述单个亲本细胞的后代不一定在形态上或者在基因组或总 DNA 互补上与原始亲本完全相同。

术语‘复制子’是任何遗传因子，例如，质粒、染色体、病毒、粘粒等，它在细胞内的表现就如同自主性核苷酸复制单元；即能够在它自己的控制之下复制。

术语‘载体’是一种另外包括提供所需要的开放读框复制和/或表达的序列的复制子。

术语‘控制序列’表示实现与它连接的编码序列表达所必需的多核苷酸序列。所述控制序列的性质根据宿主生物而有所不同；在原核生物中，所述控制序列通常包括启动子、核糖体结合位点和终止子；在真核生物中，所述控制序列通常包括启动子、终止子以及在某些场合下包括增强子。术语‘控制序列’意在最低限度包括表达时所需要存在的所有成分，并且还可以包括其他成分，这些成分的存在是有利的，例如，能控制分泌的前导序列。

术语‘启动子’是由共有序列所组成的核苷酸序列，它使得 RNA 聚合酶能够以这样的方式与 DNA 模板结合：在相邻的结构基因的正常

转录起始位点上开始产生 mRNA。

短语‘可操作地连接’表示一种并列，其中，有关成分的相互关系使得它们能够以预期的方式起起用。控制序列‘可操作地连接’于编码序列是这样一种连接，该编码序列的表达是在与该控制序列相容的条件下实现的。

‘开放读框’（ORF）是一段多核苷酸序列，它编码一种多肽，并且不含终止密码子；该片段可以是编码序列的一部分或者是整个编码序列。

‘编码序列’是在适当的调控序列控制下能转录成 mRNA 和/或翻译成多肽的多核苷酸序列。所述编码序列的边界是由位于 5’末端的翻译起始密码子和位于 3’末端的翻译终止密码子确定的。编码序列可以包括，但不限于 mRNA、DNA（包括 cDNA）以及重组多核苷酸序列。

在本文中，‘表位’或‘抗原决定族’表示具有免疫活性的氨基酸序列。通常，一个表位由至少 3-4 个氨基酸，更常见的是至少 5 或 6 个氨基酸组成，有时候所述表位由约 7-8 个，或者甚至是大约 10 个氨基酸组成。在本文中，特定多肽的表位表示具有与所述特定多肽的表位相同的氨基酸序列的表位及其免疫学等同物。所述等同物还包括菌株、亚型（=基因型）或类型（组）特异性变体，例如，属于 1a、1b、1c、1d、1e、1f、2a、2b、2c、2d、2e、2f、2g、2h、2i、3a、3b、3c、3d、3e、3f、3g、4a、4b、4c、4d、4e、4f、4g、4h、4i、4j、4k、41、5a、5b、6a、6b、6c、7a、7b、7c、8a、8b、9a、9b、10a 或任何其他新确定的 HCV（亚）型的目前已知的序列或菌株。可以理解的是，构成所述表位的氨基酸不一定是线性序列的一部分，而且可以由任何数量的氨基酸间隔开，因此形成一种构象表位。

术语‘免疫原性’表示一种物质在有或没有佐剂的情况下，单独或者与一种载体结合时，能导致体液和/或细胞应答的能力。‘中和’表示能部分或完全抑制一种感染剂的感染性的免疫应答。‘疫苗’是能够部分或全面诱导针对 HCV 的保护作用的免疫原性组合物。疫苗还可用于治疗个体，在这种情况下，它又被称为治疗性疫苗。

术语‘治疗性’是指一种能治疗 HCV 感染的组合物。

术语‘有效量’表示足以在给予它的个体体内诱导免疫应答或者在其预期的系统（例如，免疫测定）中产生可检测的免疫反应的具有表

位的多肽的量。所述有效量优选足以实现上文所定义的治疗。所需要的确切的用量会根据用途有所不同。例如，对于疫苗用途或对于制备多克隆抗血清/抗体的用途来说，所述有效量可能根据物种、年龄、个体的一般状况、要治疗的状况的严重程度、以及所选择的特定多肽及其给药方式等而有所不同。另外，相信所述有效量可以在相对较大的非关键范围内。仅仅通过常规实验，就可以方便地确定合适的有效量。用于预防 HCV 疾病的 E1 和/或 E2 和/或 E1/E2 单个的或特异性寡聚外被蛋白的优选范围是 0.01-100 微克/剂，优选 0.1-50 微克/剂。为了获得足够的免疫应答，以及随后预防 HCV 病，每个个体可能需要若干个剂量。

发明详述

更具体地讲，本发明涉及一种用于分离或纯化选自 E1 和/或 E2 和/或 E1/E2 的重组 HCV 单个的或特异性寡聚外被蛋白的方法，其特征在于在裂解转化过的宿主细胞以便分离重组表达的蛋白时，用一种二硫键裂解剂完成一个二硫键裂解或还原步骤。

本发明的所述‘单个的或特异性寡聚’ 外被蛋白的实质是，它们不含污染性蛋白，并且，它们不通过二硫键与污染物连接。

本发明的蛋白是在低等或高等真核细胞或原核细胞中重组表达的。本发明的重组蛋白优选是糖基化的，并且可以含有高甘露糖型、杂合型或复杂的糖基化。所述蛋白优选是由在实施例部分详细披露的哺乳动物细胞系表达的，或者是诸如同样在实施例部分详细披露的突变型酵母菌株等酵母中表达的。

本发明的蛋白可以是分泌的或者在细胞的成分，如 ER 或高尔基体内表达。不过，本发明的蛋白优选具有高甘露糖型糖基化，并且保留在哺乳动物细胞的 ER 或高尔基体中，或者保留在酵母细胞中或者从酵母细胞中分泌出来，优选是由诸如 mnn9 突变体的酵母突变菌株（Kniskern 等，1994），或通过矾酸盐抗性筛选过的突变体（Ballou 等，1991）分泌。

在表达 HCV 外被蛋白时，本发明人证实半胱氨酸的某些游离的巯基没有参与细胞内或细胞间二硫桥，与宿主或表达系统产生的（例如，痘苗）蛋白或其他 HCV 外被蛋白（单个的或寡聚物的）的半胱氨酸起

反应，并形成一种特殊的分子间桥。这会导致 HCV 外被蛋白与污染蛋白一起形成‘凝聚物’。在 WO92/08734 中还证实了‘凝聚物’是在纯化之后获得的，不过，该文献没有披露与哪一种蛋白相互作用有关。在专利申请 WO92/08734 中由痘苗病毒系统表达的重组 E1/ E2 蛋白是部分纯化的凝聚物，并且发现它的纯度仅仅为 70%，使得这种纯化的凝聚物不能用于诊断、预防或治疗目的。

因此，本发明的一个主要目的是从污染的蛋白中分离单个的或特异性寡聚 HCV 外被蛋白，并将纯化的蛋白（＞95%的纯度）用于诊断、预防和治疗目的。为达到以上目的，本发明人业已能够提供在二硫桥和非共价蛋白-蛋白相互作用基础上形成了凝聚的蛋白复合物（凝聚物）的证据。因此，本发明提供了一种用于在特定情况下选择性地裂解二硫键并且用于从能极大干扰诊断、预防和治疗用途的污染蛋白中分离蛋白的方法。为了防止再次形成二硫键，可以封闭（可逆的或不可逆的）游离的巯基基团，或者对它进行氧化并与其他外被蛋白寡聚化（参见同型寡聚物的定义）。可以理解的是，所述蛋白寡聚物基本上不同于在 WO92/08734 和 WO94/01778 中所披露的凝聚物，因为它的污染蛋白的含量是检测不到的。

所述二硫键的裂解可以通过以下方法实现:

（1）通过半胱磺酸进行的过甲酸氧化，在这种情况下，半胱氨酸残基被修饰成半胱磺酸（Moore 等，1963）。

（2）例如，通过亚硫酸盐（SO^{2-}_3）与诸如铜离子的适当的氧化剂一起进行亚硫酸分解（R-S-S-R-2R-SO_3），在这种情况下，半胱氨酸被修饰成 S-磺基-半胱氨酸（Bailey 和 Cole，1959）。

（3）由诸如二硫苏糖醇（DDT）、β-巯基乙醇、半胱氨酸、谷胱甘肽红、ε-巯基乙胺或巯基乙酸的硫醇进行的还原，其中，DTT 和β-巯基乙醇是常用的（Cleland，1964），它是本发明的优选方法，因为该方法可以在水环境中进行，并且因为半胱氨酸保持未被修饰。

（4）通过膦（例如，Bu3P）进行还原（Ruegg 和 Rudinger，1977）。

因此，上述所有化合物都被认为是本发明用于裂解二硫键的制剂或手段。

本发明的所述二硫键裂解（或还原）步骤，优选是一个部分二硫键裂解（还原）步骤（在部分裂解或还原条件下进行）。

本发明的优选的二硫键裂解或还原制剂是二硫苏糖醇（DTT），部分还原是通过使用低浓度的所述还原剂而实现的，即，对于 DTT 来说，例如，在大约 0.1-大约 50mM 的浓度范围内，优选大约 0.1-大约 20mM，优选大约 0.5-大约 10mM，优选高于 1mM，高于 2mM 或高于 5mM，更优选大约 1.5mM，大约 2.0mM，大约 2.5mM，大约 5mM 或大约 7.5mM。

所述二硫键裂解步骤还可以在存在诸如 Decy1PEG、EMPIGEN-BB、NP-40、草酸钠、Triton X-100 的能裂解所表达的蛋白的合适的洗涤剂（作为裂解二硫键的例子或与裂解剂组合）的条件下进行。

所述还原或裂解步骤（优选部分还原或裂解步骤）优选在有（含）洗涤剂的条件下进行。本发明优选的洗涤剂是 Empigen-BB。诸如 Empigen-BB 的洗涤剂的用量优选在 1-10%的范围内，优选高于 3%，更优选大约 3.5%。

用于实现二硫键裂解的特别优选的方法采用了上述常规二硫键裂解制剂和洗涤剂（同样如上文所述）的组合。正如在实施例部分所提到的，低浓度 DTT（1.5-7.5mM）和大约 3.5%Empigen-BB 的特定组合，被证明是用于纯化重组表达的 E1 和 E2 蛋白的还原剂和洗涤剂的特别优选的组合。在进行凝胶过滤层析时，发现所述部分还原会导致产生可能的二聚体 E1 蛋白，并且从导致免疫测定时产生假反应性的污染蛋白中分离这种 E1 蛋白。

不过，同样可以理解的是，使得半胱氨酸更容易接近的本领域已知的任何还原剂与本领域已知的任何洗涤剂或其他方式的任何其他组合，同样属于本发明的范围，只要所述组合能够达到与本发明所披露的优选组合相同的二硫键裂解目的就行。

除了还原二硫键之外，本发明的二硫键裂解方式还可以包括本领域已知的任何二硫键交换制剂（有机类或蛋白类竞争性制剂，例如，参见 Creighton，1988），这种制剂能够导致发生以下反应：

$$R1S\text{-}SR2+R3SH \rightarrow R1S\text{-}SR3+R2SH$$

*R1、R2：蛋白凝聚物的化合物

R3SH：竞争性制剂（有机类、蛋白类）

术语‘二硫桥交换制剂’被理解为包括二硫键再生成剂以及二硫键

抑制剂。

本发明还涉及用于纯化或分离上述 HCV 单个或特异性寡聚外被蛋白的方法，包括利用本领域已知的任何 SH 基团抑制剂或结合制剂，如选自下面所列举的制剂：

谷胱甘肽

5，5'-二硫代双-（2-硝基苯甲酸）或双-（3-羧基-4-硝基苯基）-二硫化物（DTNB 或 Ellman's 制剂）（Elmann，1959）

N-乙基马来酰亚胺（NEM；Benesch 等，1956）

N-（4-二甲基氨基-3，5-二硝基苯基）马来酰亚胺或 Tuppy's 马来酰亚胺，它能为蛋白提供颜色

P-氯代苯甲酸汞（Grassetti 等，1969）

4-乙烯吡啶（Friedman 和 Krull，1969），可以在通过酸水解反应之后释放出来

乙腈，可以在通过酸水解之后释放出来（Weil 和 Seibles，1961）

NEM-生物素（例如，购自 Sigma B1267）

2．2'-二硫代吡啶（Grassetti 和 Murray，1967）

4．4'-二硫代吡啶（Grassetti 和 Murray，1967）

6．6'-二硫代二烟酸（DTDAN；Brown 和 Cunigham，1970）

2．2'-二硫代双-（5'-硝基吡啶）（DTNP；US3597160）或其他二硫代双（杂环衍生物）化合物（Grassetti 和 Murray，1969）

综合所引用的文献可以发现，通常用于巯基基团的不同的制剂会与同一种蛋白或酶分子的不同数量的硫醇基团起反应。人们可能会认为所述硫醇基团在反应性方面的变化可能是由于该基团的空间环境所导致的，如所述分子的形状和原子周围的基团及其电荷，以及所述制剂分子或离子的大小、形状和电荷。通常，足够浓度的诸如十二烷基硫酸钠、尿素或盐酸胍的变性剂的存在，会导致蛋白分子具有足够的非折叠，使得所有的制剂都能同样接触硫醇基团。通过改变变性剂的浓度，可以控制非折叠的程度，并且通过这种方法可以发现具有不同程度反应性的硫醇基团。尽管到目前为止，所报导的大部分研究都是用 p-氯代苯甲酸汞、N-乙基马来酰亚胺和 DTNB 进行的，其他最近开发出的制剂可能会被证明同样有用。由于它们具有不同的结构，事实上，很有可能它们会对硫醇基团的空间环境的改变产生不同的反应。

另外，用于抑制游离的SH基团氧化，并因此抑制在重组表达和纯化E1和E2（外被）蛋白时形成大的分子间凝聚物的诸如低的pH（优选低于pH6）的条件也属于本发明的范围。

本发明的优选的SH基团封闭制剂是N-乙基马来酰亚胺（NEM）。所述SH基团封闭制剂可以在重组宿主细胞裂解期间，并且在上述部分还原过程之后或者在用于裂解二硫桥的任何其他过程之后使用。所述SH基团封闭制剂还可以用能够提供可检测的标记的任何基团和/或有助于将所述重组蛋白固定在固体底物上的任何基团，例如，生物素化的NEM进行修饰。

Darbre（1987），Means和Feeney（1971）以及Wong（1993）还披露了用于裂解半胱氨酸桥和封闭游离的半胱氨酸的方法。

如上文所述的本发明的纯化单个的或特异性寡聚重组蛋白E1和/或E2和/或E1/E2蛋白的方法的特征还在于包括以下步骤:

-优选在有诸如N-乙基马来酰亚胺（NEM）的SH基团封闭制剂，以及可能存在的合适的洗涤剂，优选Empigen-BB的条件下，裂解重组E1和/或E2和/或E1/E2表达宿主细胞，

-通过亲和纯化回收所述HCV外被蛋白，例如，通过凝集素层析纯化，如lentil凝集素层析、或用抗E1和/或抗E2特异性单克隆抗体进行的免疫亲和层析，然后

-用诸如DTT的二硫键裂解制剂还原或裂解二硫键，同样优选在有诸如NEM或生物素-NEM的SH基团封闭制剂的情况下进行，以及

-通过诸如凝胶过滤（大小排阻层析或分子筛）的方法回收所述还原的HCV E1和/或E2和/或E1/E2外被蛋白。随后还可以进行额外的Ni^{2+}-IMAC层析以及脱盐步骤。

应当理解的是，上述回收步骤还可以用本领域技术人员所公知的任何其他合适的技术完成。

优选的凝集素层析系统包括雪花莲凝集素（GAN）-层析，或兵豆凝集素（LCA）（lentil）凝集素层析，正如在实施例部分所披露的。其他有用的凝集素包括那些能识别高甘露糖型糖类的凝集素，如黄水仙凝集素（NPA）、豌豆凝集素（PSA）或熊葱凝集素（AUA）。

优选将上述方法用于纯化按上述方法在细胞内产生的单个的或特异性寡聚HCV外被蛋白。

对于分泌型 E1 或 E2 或 E1/E2 寡聚物来说，诸如蓖麻凝集素 I(RCA I）的凝集素结合复合糖是优选的凝集素。

更具体地讲，本发明涉及选自 E1 和/或 E2 和/或 E1/E2 的基本上纯化的重组单个的或特异性寡聚 HCV 外被蛋白，其特征是通过上述方法分离或纯化的。

更具体地讲，本发明涉及纯化或分离由诸如痘苗的重组哺乳动物细胞表达的重组单个的或特异性寡聚 HCV 外被蛋白。

本发明还涉及纯化或分离由重组酵母细胞表达的重组外被蛋白。

本发明同样涉及纯化或分离由重组细菌（原核生物）细胞表达的重组外被蛋白。

本发明还涉及一种重组载体，该载体包括一个载体序列，一个合适的原核生物、真核生物或病毒或合成的启动子序列，随后是一个可以表达本发明的单个的或特异性寡聚 E1 和/或 E2 和/或 E1/E2 的核苷酸序列。

具体地讲，本发明涉及一种重组载体，该载体包括一个载体序列，一个合适的原核生物、真核生物或病毒或合成的启动子序列，随后是一个可以表达本发明的单个的 E1 或 E1 的核苷酸序列。

具体地讲，本发明涉及一种重组载体，该载体包括一个载体序列，一个合适的原核生物、真核生物或病毒或合成的启动子序列，随后是一个可以表达本发明的单个的 E1 或 E2 的核苷酸序列。

插入所述载体序列中的编码需要的 E1 和/或 E2 序列的 HCV cDNA 片段，可以与一个信号序列连接。所述信号序列可以是来自非 HCV 来源的信号序列，例如，用于在哺乳动物细胞中表达的 IgG 或组织纤溶酶原激活物（tpa）前导序列，或用于在酵母细胞中表达的α交配因子序列，不过，本发明特别优选的构建体含有出现在 HCV 基因组上位于 E1 和 E2 蛋白的相应的起始位点前面的信号序列。插入所述载体序列中的编码需要的 E1 和/或 E2 序列的 HCV cDNA 片段还可以包括缺失，例如，在实施例部分所披露的疏水性结构域的缺失，或 E2 高可变区 I 的缺失。

具体地讲，本发明的重组载体包括具有编码所述多蛋白的 HCV cDNA 片段的核酸，该片段始于 HCV 多蛋白的 1 号和 192 号氨基酸位点之间，并且止于 250-400 号位点之间的部分，更优选止于 250-341 号

位点之间的部分，更优选止于 290-341 号位点之间的部分，以便表达 HCV 单个的 E1 蛋白。最优选的是，本发明的重组载体包括具有编码部分 HCV 多蛋白的 HCV cDNA 片段的重组核酸，该片段始于 117-192 位置之间的部分，并且止于 263-326 号位置之间的任意位点，以便表达 HCV 单个的 E1 蛋白。缺失了第一个疏水性结构域（264-293 号位置加或减 8 个氨基酸）的形式，或添加了 5'末端 ATG 密码子和 3'末端终止密码子的形式，或添加了因子 Xa 裂解位点和/或 3-10 个，优选 6 个组氨酸密码子的形式同样属于本发明的范围。

更具体地，本发明的重组载体包括具有编码所述多蛋白的 HCV cDNA 片段的核酸，该片段始于 290-409 号氨基酸位点之间的部分，并且止于 HCV 多蛋白的 600-820 号位点之间的部分，更优选始于 322-406 号位点之间的部分，更优选始于 347-406 号位点之间的部分，更优选始于 364-406 号位点之间的部分，以便表达 HCV 单个的 E2 蛋白。最优选的是本发明的重组载体包括具有编码所述多蛋白的 HCV cDNA 片段的重组核酸，该片段始于 290-406 号位置之间的部分，并且止于 623、650、673、710、715、720、746 或 809 号位置中的任意一个位置，以便表达 HCV 单个的 E2 蛋白。添加了 5'末端 ATG 和 3'末端终止密码子的形式，或添加了因子 Xa 裂解位点和/或 3-10 个，优选 6 个组氨酸密码子的形式同样属于本发明的范围。

有多种载体可用于实现本发明的 HCV 单个的或特异性的寡聚外被蛋白的重组表达。诸如酵母或糖基化突变菌株的低等真核生物通常是用质粒转化的，或者是用重组病毒转化的。所述载体可以在所述宿主内独立复制，或者也可以整合到宿主细胞基因组中。

高等真核生物可以用载体转化，或者可以用重组病毒转染，例如用重组痘苗病毒转染。用于将外源 DNA 插入痘苗病毒中的方法和载体为本领域所熟知，并且使用诸如同源重组的方法。在本领域可以获得哺乳动物表达所需要的多种病毒启动子序列，可能需要的终止子序列和聚腺苷酸添加序列，可能的增强子序列以及可能的扩增序列。痘苗是特别优选的，因为痘苗能抑制宿主细胞蛋白的表达。痘苗同样是非常优选的，因为它能在用活的重组痘苗病毒免疫的细胞或个体中表达 HCV 的 E1 和 E2 蛋白。对于人类免疫接种来说，禽痘病毒和 Ankara 修饰病毒（AMV）是特别优选的载体。

另外，已知源于杆状病毒苜蓿银纹夜蛾核型多角体病毒（AcNPV）的昆虫表达转移载体是辅助依赖型病毒表达载体。源于该系统的表达载体通常使用强的病毒多角体蛋白基因启动子驱动异源基因的表达。本领域技术人员可以获得用于将异源 DNA 导入杆状病毒的理想位点以便进行杆状病毒表达的不同的载体和方法。由昆虫细胞识别的用于翻译后修饰的不同的信号在本领域中同样是公知的。

本发明还包括用于生产纯化的重组单个的或特异性寡聚 HCV E1 或 E2 或 E1/E2 蛋白的方法，其中，参与凝聚物形成的半胱氨酸残基在核酸序列水平上被其他残基所取代，以便抑制凝聚物的形成。由携带所述突变的 E1 和/或 E2 蛋白编码核酸的重组载体所表达的所述重组蛋白同样属于本发明的范围。

本发明还涉及重组 E1 和/或 E2 和/或 E1/E2 蛋白，其特征在于业已去掉了它们的至少一个糖基化位点，并因此被称为糖基化变体。正如在实施例部分所述，为了诊断（筛选、证实、预测等）和抑制有关患者的 HCV 诊断和预防有关患者的 HCV 疾病，可能需要不同的糖基化变体。例如，业已发现缺乏 GLY4 的 E2 蛋白糖基化变体在诊断中能改善某些血清的反应性。所述糖基化变体优选是按照本发明所披露的方法纯化的。本发明还涉及具有编码诸如 E1 和/或 E2 和/或 E1/E2 糖基化变体的核酸插入片段的重组载体，以及用所述重组载体转化过的宿主细胞。

本发明还涉及包括同样构成本发明的一部分的多核苷酸的重组载体。本发明特别涉及如 SEQ ID NO: 3、5、7、9、11、13、21、23、25、27、29、31、35、37、39、41、43、45、47 和 49 所示出的重组核酸或其部分。

本发明还涉及用上述重组载体转化过的宿主细胞，其中，所述载体包括编码上述 HCV E1 和/或 E2 和/或 E1/E2 蛋白的核酸序列，另外还包括可操作地连接在所述 HCV E1 和/或 E2 和/或 E1/E2 序列上的，并且能够调控所述 HCV E1 和/或 E2 和/或 E1/E2 蛋白表达的调控序列。

真核宿主包括在定义部分所披露的低等和高等真核宿主。低等真核宿主包括本领域所熟知的酵母细胞。高等真核宿主主要包括本领域所公知的哺乳动物细胞系，并且包括可以从 ATCC 获得的很多永生化细胞系，包括 HeLa 细胞、中国仓鼠卵巢（CHO）细胞、幼仓鼠肾（BHK）细

胞、PK15、PK13和多种其他细胞系。

本发明特别涉及由含有上述重组载体的上述宿主细胞所表达的重组E1和/或E2和/或E1/E2蛋白。所述重组蛋白特别是用本发明的方法纯化过的。

上述用于分离或纯化HCV外被蛋白的优选方法的特征还在于至少包括以下步骤:

-在合适的培养基中生长用本发明的重组载体或能表达E1和/或E2和/或E1/E2 HCV外被蛋白的重组载体转化过的上述宿主细胞,

-在合适的条件下使所述载体序列表达,以及

-优选在有诸如N-乙基马来酰亚胺(NEM)的SH基团封闭制剂,以及可能存在的合适的洗涤剂,优选Empigen-BB的条件下裂解所述转化过的宿主细胞,

-通过亲和层析回收所述HCV外被蛋白,例如,通过凝集素亲和层析或使用抗E1和/或抗E2特异性单克隆抗体的免疫亲和层析,所述凝集素优选是lentil-凝集素或GNA,然后

-用诸如DTT的二硫键裂解制剂温育以上步骤的洗脱物,优选随后用诸如NEM或生物素NEM的SH基团封闭制剂进行温育,

-通过诸如凝胶过滤的方法分离HCV单个的或特异性寡聚E1和/或E2和/或E1/E2蛋白,随后还可以进行额外的Ni^{2+}-IMAC层析以及脱盐步骤。

正如在实施例部分所披露的,上述过程的结果是,以从含有载体来源的成分和/或凝胶过滤柱或IMAC柱的空隙体积中的细胞成分的大的凝聚物中差异洗脱的形式生产出了E1和/或E2和/或E1/E2蛋白。由于存在宿主和/或表达系统产生的蛋白,所述二硫键裂解步骤,优选还消除了假的反应性。在裂解细胞期间NEM和合适的洗涤剂的存在已经部分或者甚至是完全抑制了HCV外被蛋白和污染物之间的凝聚。

正如Janknecht等1991和Hochuli等1998所披露的,优选将Ni^{2+}-IMAC层析和随后的脱盐步骤用于具有$(His)_6$的构建体。

本发明还涉及用于在诸如小鼠和大鼠的小动物中生产单克隆抗体的方法,以及利用本发明的HCV单个的或特异性寡聚外被蛋白筛选并分离能识别抗HCV抗体的人B-细胞的方法。

本发明还涉及一种含有在表3中所列举的至少一种下列E1肽的组

合物:

跨越核心/E1 V1 区的 181-200 号氨基酸的 E1-31 (SEQ ID NO: 56),

跨越 E1 区的 193-212 号氨基酸的 E1-33 (SEQ ID NO: 57),

跨越 E1 V2 区的 205-224 号氨基酸(表位 B)的 E1-35(SEQ ID NO: 58),

跨越 E1 V2 区的 208-227 号氨基酸 (表位 B) 的 E1-35A (SEQ ID NO: 59),

跨越 E1 区的 192-228 号氨基酸(V1、C1、和 V2 区(含有表位 B))的 1bE1 (SEQ ID NO: 53),

跨越 E1 区 301-320 号氨基酸的 E1-51 (SEQ ID NO: 66),

跨越 E1C4 区的 313-332 号氨基酸(表位 A)的 E1-53(SEQ ID NO: 67),

跨越 E1 的 325-344 号氨基酸的 E1-55 (SEQ ID NO: 68),

本发明还涉及含有在表 3 中所列举的至少一种下列 E2 肽的组合物;

跨越 E2 区的 397-418 号氨基酸 (表位 A, 由单克隆抗体 2F10H10 识别, 参见图 19) 的 Env67 或 E2-67 (SEQ ID NO: 72),

跨越 E2 区的 409-428 号氨基酸(表位 A)的 Env69 或 E2-69(SEQ ID NO: 73),

跨越 E2 区的 583-602 号位置(表位 E)的 Env23 或 E2-23(SEQ ID NO: 86),

跨越 E2 区的 595-614 号位置(表位 E)的 Env25 或 E2-25(SEQ ID NO: 87),

跨越 E2 区的 607-626 号位置(表位 E)的 Env27 或 E2-27(SEQ ID NO: 88),

跨越 E2 区的 547-586 号位置(表位 D)的 Env17B 或 E2-17B(SEQ ID NO: 83),

跨越 E2 区的 523-542 号位置 (表位 C; 由单克隆抗体 16A6E7 识别, 参见图 19) 的 Env13B 或 E2-13B (SEQ ID NO: 82)。

本发明还涉及含有至少一种下列 E2 构象表位的组合物。

由单克隆抗体 15C8C1、12D11F1 和 8G10D1H9 单克隆抗体识别的表

位 F，

由单克隆抗体 9G3E6 识别的表位 G，

由单克隆抗体 10D3C4 和 4H6B2 识别的表位 H（或 C），

由单克隆抗体 17F2C2 识别的表位 I。

本发明还涉及通过用一种肽或蛋白组合物免疫所产生的 E1 或 E2 特异性抗体，所述抗体能够与上述任意一种多肽或肽特异性地反应，并且所述抗体优选是单克隆抗体。

本发明还涉及通过本领域公知的方法从质粒或噬菌体中可变链文库中或者从人 B 细胞群体中筛选的 E1 或 E2 特异性抗体，所述抗体能够与上述任意一种多肽或肽起反应，并且所述抗体优选是单克隆抗体。

在一方面，本发明的 E1 或 E2 特异性单克隆抗体可以通过用本发明的 HCV 多肽或肽免疫过的动物，特别是小鼠或大鼠脾细胞，通过常规方法形成的任意一种杂交瘤产生，另一方面，可以由骨髓瘤细胞系的细胞产生，可以根据杂交瘤产生能识别起初用于免疾所述动物的多肽的单克隆抗体的能力进行选择。

本发明所涉及到的抗体可以通过酶促、荧光、或放射性类型的合适的标记物进行标记。

本发明的该优选实施方案的单克隆抗体，可以是通过重组 DNA 技术制备的人源化形式的小鼠单克隆抗体，脱离了部分来自编码 H 和 L 链的 cDNA 或基因组克隆的编码 H 和 L 链的小鼠和/或人基因组 DNA 序列。

另外，本发明的该优选实施方案的单克隆抗体可以是人单克隆抗体。本发明的该实施方案的抗体还可以来自受 HCV 感染的或对 HCV 有免疫力的患者的人外周血淋巴细胞。

例如，所述人单克隆抗体是通过使人外周血淋巴细胞（PBL）重新集中于重度合并性免疫缺陷（SCID）小鼠而制备的（最新的综述参见 Duchosal 等，1992）。

本发明还涉及利用本发明的蛋白或肽通过所有成分克隆的方法筛选重组抗体（Persson 等，1991）。

针对来自特定基因型的肽或单个的或特异性寡聚的外被蛋白的抗体可以被用作药物，更具体地讲，用于检测 HCV 基因型的免疫测定中

（用于检测 HCV E1 或 E2 抗原的存在），用于预测/监测 HCV 疾病，或者用作治疗剂。

另外，本发明还涉及利用上述任意一种 E1 或 E2 特异性单克隆抗体制备用于检测生物学样品中 E1 或 E2 抗原存在的免疫测定试剂盒，用于制备预测/监测 HCV 疾病的试剂盒或用于制备 HCV 药物。

本发明还涉及一种用于体外诊断或检测生物学样品中 HCV 抗原存在的方法，至少包括以下步骤：

（i）在能够形成免疫复合物的合适条件下，让所述生物学样品与上述任意一种 E1 和/或 E2 特异性单克隆抗体接触，优选固定化形式的单克隆抗体，

（ii）清除未结合的成分，

（iii）在合适的条件下，温育由异源抗体形成的免疫复合物，它能特异性地结合存在于要分析的样品中的抗体，在所述异源抗体上结合了一种可检测的标记，

（iv）通过肉眼或机械方法（例如，通过密度测定、荧光测定、比色测定）检测所述免疫复合物的存在。

本发明还涉及用于体外诊断存在于生物学样品中的 HCV 抗原的试剂盒，包括：

-至少一种上述单克隆抗体，所述单克隆抗体优选固定在一种固体基质上，

-生产使得所述抗体和存在于所述生物学样品中的 HCV 抗原之间的结合反应能够进行的缓冲液所需要的缓冲物或成分，

-一种用于检测在上述结合反应中所形成的免疫复合物的装置，

-还可能包括一个用于根据所观察到的结合形式推测存在于所述样品中的 HCV 抗原的自动化扫描和翻译装置。

本发明还涉及被用作药物的含有用本发明的方法纯化过的 E1 和/或 E2 和/或 E1/E2 重组 HCV 蛋白的组合物，或含有上述至少一种肽的组合物。

更具体地讲，本发明涉及含有上述至少一种外被肽的组合物或上述重组外被蛋白组合物，将该组合物用作疫苗对哺乳动物，优选人类进行 HCV 免疫，包括给予足够量的所述组合物，还可以同时给予可以药用的佐剂，以便产生免疫应答。

更具体地讲，本发明还涉及将上述任意一种组合物用于制备上述疫苗。

另外，本发明涉及一种用于对哺乳动物，优选人类进行 HCV 免疫的疫苗，它含有来自上述 E1 和/或 E2 区的 HCV 单个的或特异性寡聚蛋白或肽。

可以按照本领域公知的方法制备免疫原性组合物。本发明的组合物含有免疫原量的上述重组 E1 和/或 E2 和/或 E1/E2 单个的或特异性寡聚蛋白或上述 E1 或 E2 肽，通常与一种可以药用的载体组合，优选还含有一种佐剂。

预计本发明的 E1 和/或 E2 和/或 E1/E2 单个的或特异性寡聚外被蛋白能够提供特别有用的疫苗抗原，因为形成抗 E1 或 E2 的抗体比形成抗其他外被蛋白的抗体更理想，并且由于 E2 蛋白在 HCV 类型之间有交叉反应，而且 E1 蛋白是类型特异性的。含有来自若干种基因型的 1 型 E2 蛋白和 E1 蛋白的混合物可能是特别有利的。含有摩尔数过量的 E1 与 E2 或 E2 与 E1 的混合物可能同样是特别有用的。可以给动物给予免疫原性组合物，以便诱导抗体的产生，从而提供抗体的来源或在所述动物体内诱导保护性免疫性。

可以药用的载体包括任何载体，它本身不会诱导对接受该组合物的个体有害的抗体的产生。合适的载体通常是大的、缓慢代谢的大分子，如蛋白、多糖、聚乳酸、聚乙醇酸、聚氨基酸、氨基酸共聚物；以及无活性的病毒颗粒。所述载体为本领域普通技术人员所熟知。

能增强所述组合物效果的优选佐剂包括，但不限于：披露于 US4606918 中的氢氧化铝（明矾）、N-乙酰-肉桂酰-L-苏氨酰-D-异谷氨酰胺（thr-MDP），N-乙酰-正肉桂酰-L-丙氨酰-D-异谷氨酰胺（nor-MDP）、N-乙酰肉桂酰-L-丙氨酰-D-异谷氨酰胺-L-丙氨酸-2-（1'-2'-二棕榈酰-sn-甘油-3-羟基磷氧基）-乙胺（MTP-PE）和 RIBI，在 2%角鲨烯/Tween80 乳液中含有三种从细菌中提取的成分-单磷脂 A、海藻糖二霉菌酸酯、以及细胞壁骨架（MPL+TDM+CWS）。MPL、TDM 或 CWS 这三种成分中的任意一种都可以单独使用或者两两组合使用。另外，还可以使用诸如 Stimulon（Cambridge Bioscience，Worcester，MA）或 SAF-1（Syntex）的佐剂。另外，可将完全弗氏佐剂（CFA）和不完全弗氏佐剂（IFA）用于非人用途和研究目的。

所述免疫原性组合物通常含有可以药用的载体，如水、盐水、甘油、乙醇等。另外，在所述载体中可以包括辅助性物质，如湿润剂或乳化剂，pH 缓冲物质和防腐剂等。

通常，所述免疫原性组合物是作为可以注射的液体溶液或悬浮液形式制备的；还可以制备适合在注射之前溶解在或悬浮在液体载体中的固体形式。所述制剂还可以乳化在或胶囊化在脂质体中，以便增强佐剂的效果。E1 和 E2 蛋白还可以与诸如 Quil A 的皂苷一起结合在免疫刺激复合物中（ISCOMS）。

被用作疫苗的免疫原性组合物含有‘足够量的’或‘免疫有效量的’本发明的外被蛋白，以及根据需要含有上述任何其他成分。‘免疫有效量的’表示作为单个剂量或作为一系列剂量的一部分给予这一用量就能实现上述有效治疗。所述用量根据接受治疗的个体的健康和体能状况、接受治疗的个体的分类类型（例如，非人灵长类、灵长类等）、所述个体的免疫系统合成抗体的能力、需要保护的程度、疫苗的配方、治疗医师对医学状况的评估、感染 HCV 的菌株、以及其他相关因素而有所不同。预计，所述用量落在一个较宽的范围内，该范围可以通过常规实验确定。通常，所述用量在 0.01-1000 微克/剂之间，更优选 0.1-100 微克/剂。

所述单个的或特异性寡聚外被蛋白还可以用作疫苗载体，以与乙型肝炎表面抗原相同的方式呈递同源的（例如，来自核心、NS2、NS3、NS4 或 NS5 区的 T 细胞表位或 B 细胞表位）或异源的（非 HCV）半抗原（参见欧洲专利申请 174444）。在这一用途中，外被蛋白提供了一种能够刺激对与所述凝聚物结合的半抗原或抗原的免疫应答的免疫原性载体。所述抗原可以通过常规化学方法连接，或者可以克隆到编码 E1 和/或 E2 的基因的相当于该蛋白的亲水性区的位点上。所述亲水区包括 V1 区（包括 191-202 号氨基酸）、V2 区（包括 213-223 号氨基酸）、V3 区（包括 230-242 号氨基酸）、V4 区（包括 230-242 号氨基酸）、V5 区（包括 294-303 号氨基酸）和 V6 区（包括 329-336 号氨基酸）。另一种可用于插入半抗原的位点是疏水区（包括 264-293 号氨基酸）。在本发明中证实，该区可以缺失，而不会影响缺失了的 E1 蛋白与血清的反应性。因此，可以将半抗原插在所述缺失位点上。

所述免疫原性组合物通常是通过注射方式经肠胃外途径给予的，例

如，皮下或肌内注射。适合其他给予方法的其他制剂包括口服制剂和栓剂。剂量治疗可以是单个剂量方案或多剂量方案。所述疫苗可以与其他免疫调控剂组合使用。

本发明还涉及一种用于体外检测存在于生物学样品中的 HCV 抗体的含有上述肽或多肽的组合物。

本发明还涉及将上述组合物用于制备用来检测存在于生物学样品中的 HCV 抗体的免疫测定试剂盒的用途。

本发明还涉及一种用于体外诊断存在于生物学样品中的 HCV 抗体的方法，该方法至少包括以下步骤:

（i）在能够形成免疫复合物的合适条件下，让所述生物学样品与含有上述任意一种上述外被肽或蛋白的组合物接触，优选固定化形式的组合物，其中，所述肽或蛋白可以是通过链霉亲和素或亲和素复合物与一种固体基质共价结合的生物素化的肽或蛋白，

（ii）清除未结合的成分，

（iii）在合适的条件下，温育由异源抗体形成的免疫复合物，所述异源抗体上结合了可检测的标记，

（iv）通过肉眼或机械方法（例如，通过密度测定、荧光测定、比色测定）检测所述免疫复合物的存在。

另外，本发明还涉及竞争性免疫测定制剂，其中，将上文所披露的重组生产的单个的或特异性寡聚 E1 和/或 E2 和/或 E1/E2 蛋白与 E1 和/或 E2 蛋白组合使用，以便竞争存在于生物学样品中的 HCV 抗体。

本发明还涉及一种用于确定 HCV 抗体在生物学样品中的存在的试剂盒，包括:

-上述至少一种肽或蛋白组合物，有可能与来自 HCV 或其他类型 HCV 的其他的多肽或肽组合，所述肽或蛋白优选固定在一种固体基质上，更优选固定在同一块 ELISA 平板上的不同的微孔上，更优选固定在同一个膜片上，

-生产使得所述多肽或肽与所述存在于所述生物学样品中的抗 HCV 的抗体之间的结合反应能够进行的缓冲液所需要的缓冲物或成分，

-用于检测在上述结合反应中所形成的免疫复合物的装置，

-还可能包括一个用于根据所观察到的结合形式推测存在于所述样品中的 HCV 基因型的自动化扫描和翻译装置。

本发明的免疫测定方法使用来自 E1 和/或 E2 结构域的单个的或特异性寡聚抗原，该抗原保留了由存在于来自受 HCV 感染的个体的血清中的抗体识别的线性（对于肽来说）和构象表位（单个的或特异性寡聚蛋白）。例如，使用单个的或特异性寡聚抗原、二聚体抗原、以及单个的或特异性寡聚抗原的组合也属于本发明的范围。本发明的 HCV E1 和 E2 抗原实际上可应用于采用已知的抗原检测抗体的任何测定方案中。当然，应当避免或改进使 HCV 构象表位变性的方案。上述所有测定的共同特征是，在所述抗原能够结合存在于所述成分中的任何所述抗体的条件下，让所述抗原与被怀疑含有 HCV 抗体的身体成分接触。所述条件通常是使用过量抗原时的生理学温度、pH 和离子强度。在将所述抗原与所述样品一起温育之后，检测包括所述抗原的免疫复合物。

免疫测定的设计可以有很大的不同，并且许多方案为本领域所公知。例如可以使用固体支持物或免疫沉淀的方案。大多数测定包括使用标记过的抗体或多肽；所述标记可以是，例如，酶促、荧光、化学发光、放射性、或染料分子。用于放大来自所述免疫复合物的信号的测定同样是公知的；其例子有利用生物素和亲和素或链霉亲和素的测定，以及未标记的和介导的免疫测定，如 ELISA 测定。

所述免疫测定可以是,但不限于标准类型或竞争类型的异源或同源方案。在异源方案中，所述多肽通常与一种固体基质或支持物结合，以便有利于在温育之后将所述样品与所述多肽分离。可以使用的固体支持物的例子有硝酸纤维素(例如，薄膜形式或微量滴定孔形式)，聚氯乙烯（例如，片材或微量滴定孔），聚丙乙烯乳胶（例如，小球或微量滴定板），聚偏氟乙烯（被称为 Immunolon™），重氮化纸、尼龙膜、活性的小球、以及蛋白 A 小球。例如，在异源方案中可以使用 Dynatech Immunolon™1 或 Immunolon™2 微量滴定板或 0.25 英寸聚苯乙烯小球（精密塑料球）。在与测试样品分离之后，并且在检测结合的抗体之前，通常要洗涤含有所述抗原性多肽的固体支持物。标准方案和竞争性方案都是本领域所公知的。

在同源方案中，所述测试样品在溶液中与抗原的组合一起温育。例如，可以在能使所形式的任何抗原-抗体复合物沉淀的条件下温育。用于这种测定的标准方案和竞争方案为本领域所公知。

在标准方案中，直接监测抗体-抗原复合物中 HCV 抗体的量。这一目的可以通过测定能识别抗 HCV 抗体上的一个表位的标记过的抗异源（例如，抗人）抗体是否会因为复合物的形成而结合来实现。在竞争方案中，通过监测结合所述复合物中已知量的标记过的抗体（或其他竞争配体）的竞争效应推测样品中 HCV 抗体的量。

根据所述方案，通过多种已知方法中的任意一种检测所形成的含有抗 HCV 抗体的复合物（或者对于竞争测定来说，检测竞争抗体的量）。例如，所述复合物中的未标记过的 HCV 抗体可以利用与一种标记物（例如，酶标记物）复合的抗异源 IgG 的缀合物检测。

在免疫沉淀或凝聚测定方案中，HCV 抗原和抗体之间的反应形成了一个从溶液或悬浮液中沉淀出来的网络，并形成一个肉眼可见沉淀层或膜。如果在测试样品中不存在抗 HCV 抗体，就不会形成可见的沉淀。

目前有三种特殊类型的颗粒凝聚（PA）测定。这些测定被用于检测涂在一种支持物上的各种抗原的抗体。上述一种类型的测定是利用红细胞（RBCs）进行的凝血测定，所述红细胞通过在 RBC 上被动吸附抗原（或抗体）而被敏化。通过添加特殊抗原，存在于身体成分中的抗体（如果有的话）会导致用纯化的抗原包衣的 RBCs 发生凝聚。

为了消除血凝聚测定中潜在的非特异性反应，可以用两种人工载体取代 PA 中的 RBC。最常见的是乳胶颗粒。不过，也可以使用凝聚颗粒。使用上述任意一种载体的测定都是基于用纯化抗原包衣的颗粒的被动凝聚。

包括构象表位的本发明的 HCV 单个的或特异性寡聚 E1 和/或 E2 和/或 E1/E2 抗原通常以试剂盒进行包装，用于上述免疫测定。该试剂盒通常包括独立的容器，分别装有天然 HCV 抗原、对照抗体制剂（阳性和阴性）标记过的抗体（如果该测定方案需要这种抗体），以及信号发生试剂（例如，酶底物）（如果所述标记不能直接产生一种信号）。所述天然 HCV 抗原可能已经与一种固体基质，或者与用于将它结合在所述基质上的制剂分开。在该试剂盒中通常还包括实施该测定的说明书（例如，文字、磁带、CD-ROM 等）。

采用天然 HCV 抗原的免疫测定可用于筛选血液，以便制备缺乏潜在的传染性 HCV 的血液供应物。用于制备血液供应物的方法包括以下步骤。让本发明的 HCV E1 和/或 E2 蛋白与来自供血个体的身体成分，

优选血液或血液成分起反应，以便发生 HCV 抗体（如果有的话）和 HCV 抗原之间的免疫反应。检测是否因为上述反应形成了 HCV 抗体-HCV 抗原复合物。作为血液供应物的血液来自不存在天然 HCV 抗原 E1 或 E2 抗体的供体。

如果出现了对 HCV 抗原的阳性反应，优选重复所述免疫测定，以便降低假阳性的可能性。例如，在大规模筛选血液以便生产血液制品（例如，输血用血、血浆、因子 VIII、免疫球蛋白等）时，‘筛选’测试通常以牺牲特异性为代价来提高灵敏度（以便确保没有污染的传染物，即提高了假阳性的比例）。因此，通常仅仅是推迟进一步检测‘反复反应’的供体，即在对供体样品进行两轮或两轮以上的免疫测定时都是阳性。不过，为了证实 HCV 阳性，通常要进行‘验证’实验，以牺牲灵敏度为代价来提高特异性（以便确保不会确认假阳性样品）。因此，本发明所披露的用于纯化 E1 和 E2 的方法非常有利于将单个的或特异性寡聚外被蛋白包括在 HCV 诊断测定中。

所选择的固相可以包括聚合物小球或玻璃小球、硝酸纤维素、微颗粒、反应平板的微孔、试管和磁性小球。信号发生化合物可以包括酶、发光化合物、生色剂、放射性元素和化学发光化合物。酶的例子包括碱性磷酸酶、辣根过氧化物酶和 β-半乳糖苷酶。增强化合物的例子包括生物素、抗生物素和亲和素。增强化合物结合成分 的例子包括生物素、抗生物素和亲和素。为了抑制类风湿因子样物质的作用，要用足以抑制类风湿因子样物质的作用的条件对测试样品进行处理。所述条件包括让所述测试样品与一定量的抗人 IgG 接触，以便形成一种混合物，并且在处理形成基础上不含类风湿因子样物质的反应混合物的条件下温育所述混合物一段时间。

本发明还涉及将 E1 蛋白或其部分，更具体说是上述 HCV 单个的或特异性寡聚 E1 蛋白用于体外监测 HCV 疾病或预测患有 HCV 感染的患者对治疗（例如用干扰素治疗）的反应的用途，包括:

-在能够形成免疫复合物的条件下将来自感染了丙型肝炎的患者的生物学样品与 E1 蛋白或其合适的部分一起温育。

-除去未结合的成分，

-计算存在于所述样品中的抗 E1 效价（例如，在（干扰素）治疗开始和/或期间测定）。

-监测 HCV 疾病的天然过程，或者根据在治疗开始和/或治疗期间出现在所述样品中的抗 E1 效价的量预测所述患者对治疗的反应。

与最初的抗 E1 效价相比，表现出效价降低了 2、3、4、5、7、10、15 或优选降低了 20 倍以上的患者被认为是 HCV 治疗，尤其是干扰素治疗的长期的、持续的反应者。在实施例部分说明了抗 E1 测定一般是非常有利于预测对 IFN 治疗的长期反应，或者对丙型肝炎病毒疾病治疗的长期反应。

更具体地讲，在表 3 中所列举的以下 E1 肽可用于体外监测 HCV 疾病或预测对患有 HCV 感染的患者进行干扰素治疗的反应。

跨越核心/E1 V1 区的 181-200 号氨基酸的 E1-31（SEQ ID NO: 56），

跨越 E1 区的 193-212 号氨基酸的 E1-33（SEQ ID NO: 57），

跨越 E1 V2 区的 205-224 号氨基酸（表位 B）的 E1-35（SEQ ID NO: 58），

跨越 E1 V2 区的 208-227 号氨基酸（表位 B）的 E1-35A（SEQ ID NO: 59），

跨越 E1 区的 192-228 号氨基酸（V1、C1 和 V2 区（含有表位 B））的 1bE1（SEQ ID NO: 53），

跨越 E1 区的 301-320 号氨基酸的 E1-51（SEQ ID NO: 66），

跨越 E1 C4 区的 313-332 号氨基酸（表位 A）的 E1-53（SEQ ID NO: 67），

跨越 E1 区的 325-344 号氨基酸的 E1-55（SEQ ID NO: 68）。

可以理解的是，上述肽的更小的片段也属于本发明的范围。通过化学合成可以方便地制备所述较小的片段，并且可以检测其用于在上文和在实施例部分所述的测定中使用的能力。

本发明还涉及用于监测 HCV 疾病或预测患有 HCV 感染的患者对治疗（例如用干扰素治疗）的反应的试剂盒，包括:

-至少一种 E1 蛋白或 E1 肽，更具体地讲，上述 E1 蛋白或 E1 肽，

-用于生产所述蛋白或肽和存在于生物学样品中的抗 E1 抗体之间发生结合反应的缓冲液所需要的缓冲物或成分，

-用于检测在上述结合反应中所形成的免疫复合物的装置，

-可能还包括用于了解在治疗进行期间抗 E1 效价降低的自动化扫

描和翻译装置。

可以理解的是，本发明的E2蛋白和肽也可用于在一定程度上监测/预测HCV治疗，正如上面针对E1蛋白或肽所说明，因为与其他HCV抗原的抗体相比，抗E2水平也会降低。不过，应当理解的是，还可以确定E2区中的某些表位，这些表位可能适用于监测/预测HCV疾病的实验。

本发明还涉及血清分类测定，用于检测存在于生物学样品中的HCV的一种或多种血清学类型，特别是用于检测要中和在一个测定方案中检测的不同类型的HCV抗体，至少包括以下步骤：

（i）在能够形成免疫复合物的合适条件下，让需要分析是否存在一种或多种血清学类型的HCV抗体的生物学样品与上述至少一种E1和/或E2和/或E1/E2蛋白组合物或至少一种E1和/或E2肽组合物接触，优选固定化形式的组合物，

（ii）清除未结合的成分，

（iii）在合适的条件下，将与异源抗体形成的免疫复合物与所述可检测标记结合的异源抗体一起温育，

（iv）通过肉眼或机械方法（例如，通过密度测定、荧光测定、比色测定）检测所述免疫复合物的存在，并且根据所观察到的结合形式了解一种或多种HCV血清学类型的存在。

应当理解的是，用于该方法中的蛋白或肽的组合物是重组表达的类型特异性外被蛋白或类型特异性肽。

本发明还涉及一种用于对存在于生物学样品中的一种或多种血清类型的HCV进行血清分类，尤其是用于检测所述血清类型的HCV的抗体的试剂盒，包括：

-至少一种上述E1和/或E2和/或E1/E2蛋白或E1或E2蛋白，

-用于生产使得所述蛋白或肽和存在于所述生物学样品中的抗E1抗体发生结合反应的缓冲液所需要的缓冲物或成分，

-用于检测在上述结合反应中所形成的免疫复合物的装置，

-还可能包括一个用于根据所观察到的结合形式检测一种或多种血清类型的存在的自动化扫描和翻译装置。

本发明还涉及将上述肽或蛋白组合物固定在一种固体基质上，并且用于反相杂交测定，优选用于作为对照系固定在膜片的固体支持物

上，用于按照上述方法测定 HCV 的存在或基因型。与来自其他 HCV 多蛋白区的类型特异性抗原的组合也属于本发明的范围。

本发明提供了一种用于纯化选自通过重组方法生产的 E1 和/或 E2 和/或 E1/E2 蛋白的重组 HCV 单个的或特异性寡聚外被蛋白的方法，包括让所述蛋白与一种二硫键裂解或还原制剂接触，本发明的该方法的接触可以在部分裂解或还原的条件下进行。所述二硫链裂解制剂优选是二硫苏糖醇（DTT），其浓度范围优选为 0.1-50mM，优选 0.1-20mM，更优选 0.5-10mM。另外，所述二硫键裂解制剂还可以诸如 Empigen-BB（它是含有 N-十二烷基-N,N-二甲基甘油作为主要成分的混合物）的洗涤剂，其浓度优选为 1-10%，更优选 3.5%。还可以使用洗涤剂，二硫键裂解制剂和/或还原制剂的混合物。在一种实施方案中，利用诸如 N-乙基马来酰亚胺（NEM）或其衍生物的 SH 基团封闭制剂抑制二硫键的形成。在一种优选实施方案中，所述二硫键再形成是通过使用低的 pH 条件抑制的。

本发明还提供了一种本文所披露的方法，该方法还包括以下步骤：在选择性地存在诸如 N-乙基马来酰亚胺（NEM）的 SH 封闭基团的条件下，裂解重组 E1 和/或 E2 和/或 E1/E2 表达宿主细胞；通过亲和纯化回收所述 HCV 外被蛋白，如通过凝集素层析，如 lentil 凝集素层析，或通过使用抗 E1 和/或抗 E2 特异性单克隆抗体的免疫亲和；同样优选在存在于 NEM 或生物素-NEM 的 SH 封闭制剂的条件下，用诸如 DTT 的二硫键裂解制剂还原或裂解二硫键；乙基，通过凝胶过滤回收还原的 E1 和/或 E2 和/或 E1/E2 外被蛋白，并且选择性地进行额外的 Ni^{2+}-IMAC 层析和脱盐步骤。

本发明提供了一种含有选自 E1 和/或 E2 和/或 E1/E2 的基本上分离的和/或纯化的，和/或分离的和/或纯化的重组 HCV 单个的或特异性寡聚外被蛋白的组合物，所述蛋白是用本文所披露的方法分离的。在一种优选实施方案中，本发明的重组 HCV 外被蛋白是在哺乳动物细胞中表达的，如疫苗、重组酵母细胞。

本发明提供了一种重组载体，它含有可操作地结合在一起的载体序列，原核、真核或病毒启动子序列，以及可以表达单个的或特异性寡聚 E1 和/或 E2 和/或 E1/E2 蛋白的核苷酸序列。在一种实施方案中，所述载体的核苷酸序列编码单个的 HCV E1 蛋白，该蛋白始于 1-192 号

氨基酸之间的部分，并且止于 250-400 号氨基酸之间的部分，特别是止于 250-341 号位置之间的部分，更优选止于 290-341 号位置之间的部分。在另一种实施方案中，所述载体的核苷酸序列编码单个的 HCV E1 蛋白，该蛋白始于 117-192 号氨基酸之间的部分，并且止于 263-400 号氨基酸之间的部分，特别是止于 250-326 号位置之间的部分。在另一种实施方案中，所述载体的核苷酸序列编码单个的 HCV E1 蛋白，该蛋白缺失了 293 号位置之间的第一个疏水性结构域，加或减 8 个氨基酸。在另一种实施方案中，所述载体的核苷酸序列编码单个的 HCV E2 蛋白，该蛋白始于 290-406 号氨基酸之间的部分，并且止于 600-820 号氨基酸之间的部分，特别是始于 322-406 号位置之间的部分，更优选始于 347-406 号位置之间的部分，最优选始于 346-406 号位置之间的部分；并且优选止于 623、650、661、673、710、715、720、746 或 809 中任意一个氨基酸位置。在一种实施方案中，本发明的载体含有可操作地连接于所述核苷酸序列上的 5'末端 ATG 密码子和 3'末端终止密码子。在一种实施方案中，所述载体还包括一个因子 Xa 裂解位点和/或添加在所述编码区 3'末端的 3-10，优选 6 个组氨酸密码子。本发明的载体选择性地包括一个核苷酸序列，其中，业已在核酸水平上去掉了存在于 E1 或 E2 蛋白上的至少一个糖基化位点。

本发明提供了含有 SEQ ID NO: 3、5、7、9、11、13、21、23、25、27、29、31、35、37、39、41、43、45、47 和 49 中的任意一种或其部分的核酸。本发明的载体优选含有一种核苷酸序列，该序列含有包括 SEQ ID NO: 3、5、7、9、11、13、21、23、25、27、29、31、35、37、39、41、43、45、47 和 49 中的任意一种或其部分的核酸。

本发明的组合物还含有重组 HCV 外被蛋白，该蛋白是由上述载体表达或者是该载体的表达产物。

本发明提供了一种上述至少一种重组载体转化过的宿主细胞，其中，所述载体含有本文所披露的编码 HCV E1 和/或 E2 和/或 E1/E2 蛋白的核苷酸序列，此外还包括能够在该宿主细胞中起作用并且能够调控 HCV E1 和/或 E2 和/或 E1/E2 蛋白表达的调控序列。另外，本发明提供了一种由本发明的宿主细胞表达的重组 E1 和/或 E2 和/或 E1/E2 蛋白。

本发明还提供了一种本文所披露的方法，该方法包括以下步骤：在

合适的培养基中生长用本文所披露的重组载体转化过的本文所披露的宿主细胞；在合适的条件下使本文所披露的载体核苷酸序列表达；优选在存在诸如 N-乙基马来酰亚胺（NEM）的 SH 基团封闭制剂的条件下裂解转化过的宿主细胞；通过亲和层析回收所述 HCV 外被蛋白，如通过凝集素层析或使用抗 E1 和/或抗 E2 特异性单克隆抗体进行免疫亲和层析，所述凝聚素优选是 lentil 凝聚素，然后将上一步骤所获得的洗脱物与 DTT 的二硫键裂解剂一起温育，同样优选在存在诸如 NEM 或生物素 NEM 的 SH 基团封闭制剂的条件下进行；以及，通过凝胶过滤分离 HCV 单个的或特异性寡聚 E1 和/或 E2 和/或 E1/E2 蛋白，还可以通过额外的 Ni^{2+}-IMAC 层析和脱盐步骤分离。

本发明提供了一种含有至少一种下列 E1 和/或 E2 肽的组合物：

跨越核心/E1 V1 区的 181-200 号氨基酸的 E1-31（SEQ ID NO: 56），

跨越 E1 区的 193-212 号氨基酸的 E1-33（SEQ ID NO: 57），

跨越 E1 V2 区的 205-224 号氨基酸（表位 B）的 E1-35（SEQ ID NO: 58），

跨越 E1 V2 区的 208-227 号氨基酸（表位 B）的 E1-35A（SEQ ID NO: 59），

跨越 E1 区的 192-228 号氨基酸（V1、C1 和 V2 区（含有表位 B））的 1bE1（SEQ ID NO: 53），

跨越 E1 区的 301-320 号氨基酸的 E1-51（SEQ ID NO: 66），

跨越 E1 C4 区的 313-332 号氨基酸（表位 A）的 E1-53（SEQ ID NO: 67），

跨越 E1 区的 325-344 号氨基酸的 E1-55（SEQ ID NO: 68）。

跨越 E2 区的 397-418 号氨基酸（表位 A）的 Env67 或 E2-67（SEQ ID NO: 72），

跨越 E2 区的 409-428 号氨基酸（表位 A）的 Env69 或 E2-69（SEQ ID NO: 73），

跨越 E2 区的 583-602 号位置（表位 E）的 Env23 或 E2-23（SEQ ID NO: 86），

跨越 E2 区的 595-614 号位置（表位 E）的 Env25 或 E2-25（SEQ ID NO: 87），

跨越 E2 区的 607-626 号位置(表位 E)的 Env27 或 E2-27(SEQ ID NO: 88),

跨越 E2 区的 547-586 号位置(表位 D)的 Env17B 或 E2-17B(SEQ ID NO: 83),

跨越 E2 区的 523-542 号位置(表位 C)的 Env13B 或 E2-13B(SEQ ID NO: 82),

本发明提供了一种含有至少一种下列 E2 构象表位的组合物:

由单克隆抗体 15C8C1、12D11F1 和 8G10D1H9 识别的表位 F

由单克隆抗体 9G3E6 识别的表位 G,

由单克隆抗体 10D3C4 和 4H6B2 识别的表位 H(或 C),

由单克隆抗体 17F2C2 识别的表位 I。

本发明提供了用上述组合物免疫接种所产生的 E1 和/或 E2 特异性单克隆抗体,例如,本发明的抗体可以被用作药物,使用在免疫测定试剂盒中,用于检测 HCV E1 或 E2 抗原的存在,用于预测/监测 HCV 疾病或用于 HCV 治疗。本发明提供了将本文所披露的 E1 和/或 E2 特异性单克隆抗体用于制备用来检测 HCV E1 或 E2 抗原的免疫测定试剂盒,用于制备用来预测/监测 HCV 疾病的试剂盒或用于制备 HCV 药物。

本发明提供了一种用于体外诊断存在于生物学样品中的 HCV 抗原的方法,至少包括以下步骤:

(i)在能够形成免疫复合物的合适条件下,让所述生物学样品与本文所披露的 E1 和/或 E2 特异性单克隆抗体接触,优选固定化形式的单克隆抗体,

(ii)清除未结合的成分,

(iii)在合适的条件下,温育与异源抗体形成的免疫复合物,所述异源抗体与一种可检测的标记结合,

(iv)通过肉眼或机械方法检测所述免疫复合物的存在。

本发明提供了一种用于确定存在于生物学样品中的 HCV 抗原的存在的试剂盒,该试剂盒至少包括以下部分:至少一种本文所披露的 E1 和/或 E2 特异性单克隆抗体,优选是固定在固体基质上的形式,生产使得所述抗体和存在于生物学样品中的 HCV 抗原之间能发生结合反应的缓冲液所需要的缓冲物或成分,以及选择性地包括用于检测在上述结合反应中所形成的免疫复合物的装置。

本发明的化合物能够以药物形式提供。

本发明提供了一种组合物，如本文所述，该组合物被用作疫苗，对哺乳动物，优选人类进行 HCV 免疫，所述免疫包括给予有效量的所述组合物，并选择性地同时给予可以药用的佐剂，以便产生免疫反应。

本发明提供了一种利用本文所披露的组合物制备用于对哺乳动物，优选人类进行 HCV 免疫的疫苗的方法，包括给予有效量的所述组合物，并选择性地给予可以药用的佐剂，以便产生免疫反应。

本发明提供了一种对哺乳动物，优选人类进行 HCV 免疫的疫苗组合物，该组合物包含有效量的含有本文所披露的 E1 和/或 E2 的组合物，并选择性地含有可以药用的佐剂。

本发明的组合物可以体外检测存在于生物学样品中的 HCV 抗体的形式提供。本发明还提供了一种制备用于检测存在于生物学样品中的 HCV 抗体的免疫测定试剂盒的方法，以及利用本发明的试剂盒检测存在于生物学样品中的 HCV 抗体的方法，以便诊断存在于生物学样品中的 HCV 抗体。本发明的上述方法至少包括以下步骤：

（i）在能够与所述生物学样品中的 HCV 抗体形成免疫复合物的合适条件下，让所述生物学样品与本文所披露的组合物接触，优选固定化形式的组合物，

（ii）清除未结合的成分，

（iii）在合适的条件下，温育与异源抗体形成的免疫复合物，所述异源抗体与一种可检测的标记结合，

（iv）通过肉眼或机械方法检测所述免疫复合物的存在。

本发明提供了一种用于确定存在于生物学样品中的 HCV 抗体的存在的试剂盒，包括：至少一种本文所披露的肽或蛋白组合物，优选是固定在固体基质上的固定化形式的；生产使得所述蛋白或肽和抗存在于生物学样品中的 HCV 的抗体之间能发生结合反应的缓冲液所需要的缓冲物或成分；以及选择性地包括用于检测在上述结合反应中所形成的免疫复合物的装置。

本发明提供了一种体外监测 HCV 疾病或诊断患有 HCV 感染的患者对治疗、优选用干扰素治疗的反应的方法，该方法包括：在能形成免疫复合物的条件下将来自受到 HCV 感染的患者的生物学样品与 E1 蛋白或其合适的部分一起温育；除去未结合的成分。计算在治疗开始时和

治疗期间存在于所述样品中的抗 E1 效价；监测 HCV 疾病的自然进程，或根据在治疗开始时和/或在治疗期间存在于所述样品中的抗 E1 效价的量诊断患者对治疗的反应。

本发明提供了一种用于监测 HCV 疾病或预测患有 HCV 感染的患者对治疗，特别是用干扰素治疗的反应的试剂盒，其中，所述试剂盒包括:

-至少一种 E1 蛋白或 E1 肽，特别是本文所披露的 E1 蛋白或 E1 肽 ；

-用于生产使得所述蛋白或肽与所述存在于所述生物学样品中的抗 E1 抗体之间的能发生结合反应的缓冲液所需要的缓冲物或成分;

-并且选择性地包括用于检测在上述结合反应中所形成的免疫复合物的装置;

-还选择性地包括用于了解在治疗过程中抗 E1 效价降低的自动化扫描和翻译装置。

本发明还提供了一种血清分类测定,用于检测存在于生物学样品中的一种或多种血清类型的 HCV,特别是用于检测需要组合在一个测定方案中检测的不同类型的 HCV 抗体，至少包括以下步骤:

（i）在能够形成免疫复合物的合适条件下，让需要分析一种或多种血清类型的 HCV 抗体的存在的所述生物学样品与至少一种本文所披露的 E1 和/或 E2 和/或 E1/E2 蛋白组合物或至少一种本文所排版的 E1 和/或 E2 肽组合物接触，优选固定化形式的组合物；（ii）清除未结合的成分，（iii）在合适的条件下，温育与异源抗体形成的免疫复合物，所述异源抗体与一种可检测的标记结合；以及选择性地（iv）通过肉眼或机械方法（例如，通过密度测定、荧光测定、比色测定）检测所述免疫复合物的存在，并且根据所观察到的结合形式了解一种或多种 HCV 血清类型的存在。

本发明提供了一种用于对存在于生物学样品中的一种或多种血清类型的 HCV 进行血清分类的试剂盒，尤其是用于检测所述血清类型的 HCV 的抗体的试剂盒，包括：至少一种本文所披露的 E1 和/或 E2 和/或 E1/E2 蛋白或本文所披露的 E1 或 E2 肽；用于生产使得所述蛋白或肽与所述存在于所述生物学样品中的抗 E1 抗体之间能够发生结合反应的缓冲液所需要的缓冲物或成分；选择性地包括用于检测在上述结合

反应中所形成的免疫复合物的装置，另外还选择性地包括用于根据所观察到的结合形式检测一种或多种血清类型的存在的自动化扫描和翻译装置。

本发明还提供了本文所披露的肽或蛋白组合物，用于固定在固体基质上，并且用于反相杂交测定中，优选用于固定在诸如膜片的固体支持物上作为对照系，用于按照本文所披露的方法确定 HCV 的存在或基因型。

本发明还提供了一种治疗用疫苗组合物，该组合物含有治疗有效量的以下成分：

含有选自 E1 蛋白和 E2 蛋白的至少一种纯化的重组 HCV 单个的或特异性寡聚重组外被蛋白的组合物；以及选择性地含有可以药用的佐剂。本发明疫苗的 HCV 外被蛋白可以选择性地由重组哺乳动物细胞或重组酵母细胞生产。本发明提供了一种含有治疗有效量的含有至少一种下列 E1 和 E2 肽的组合物的治疗用疫苗组合物：

跨越核心/E1 V1 区的 181-200 号氨基酸的 E1-31（SEQ ID NO: 56），

跨越 E1 区的 193-212 号氨基酸的 E1-33（SEQ ID NO: 57），

跨越 E1 V2 区的 205-224 号氨基酸（表位 B）的 E1-35（SEQ ID NO: 58），

跨越 E1 V2 区的 208-227 号氨基酸（表位 B）的 E1-35A（SEQ ID NO: 59），

跨越 E1 区的 192-228 号氨基酸（V1、C1 和 V2 区（含有表位 B））的 1bE1（SEQ ID NO: 53），

跨越 E1 区的 301-320 号氨基酸的 E1-51（SEQ ID NO: 66），

跨越 E1 C4 区的 313-332 号氨基酸（表位 A）的 E1-53（SEQ ID NO: 67），

跨越 E1 区的 325-344 号氨基酸的 E1-55（SEQ ID NO: 68）。

跨越 E2 区的 397-418 号氨基酸（表位 A）的 Env67 或 E2-67（SEQ ID NO: 72），

跨越 E2 区的 409-428 号氨基酸（表位 A）的 Env69 或 E2-69（SEQ ID NO: 73），

跨越 E2 区的 583-602 号位置（表位 E）的 Env23 或 E2-23（SEQ ID

NO: 86)，

跨越 E2 区的 595-614 号位置(表位 E)的 Env25 或 E2-25(SEQ ID NO: 87)，

跨越 E2 区的 607-626 号位置(表位 E)的 Env27 或 E2-27(SEQ ID NO: 88)，

跨越 E2 区的 547-586 号位置(表位 D)的 Env17B 或 E2-17B(SEQ ID NO: 83)，和

跨越 E2 区的 523-542 号位置(表位 C)的 Env13B 或 E2-13B(SEQ ID NO: 82)。

本发明提供了一种治疗诸如人类受 HCV 感染的哺乳动物的方法，包括给予有效量的本文所披露的组合物，如上述疫苗，并选择性地给予可以药用的佐剂。在一种实施方案中，本发明的组合物与抗病毒治疗组合给予或在抗病毒治疗时给予，所述抗病毒治疗是在给予本发明的组合物之前或之后或同时进行。

本发明提供了一种含有选自 E1 蛋白和 E2 蛋白至少一种纯化的重组 HCV 重组外被蛋白的组合物，并且选择性地含有一种佐剂。在一种优选实施方案中，所述组合物至少含有一种下面的 E1 和 E2 肽:

跨越核心/E1 V1 区的 181-200 号氨基酸的 E1-31 (SEQ ID NO: 56)，

跨越 E1 区的 193-212 号氨基酸的 E1-33 (SEQ ID NO: 57)，

跨越 E1 V2 区的 205-224 号氨基酸(表位 B)的 E1-35(SEQ ID NO: 58)，

跨越 E1 V2 区的 208-227 号氨基酸 (表位 B) 的 E1-35A (SEQ ID NO: 59)，

跨越 E1 区的 192-228 号氨基酸 (V1、C1 和 V2 区 (含有表位 B)) 的 1bE1 (SEQ ID NO: 53)，

跨越 E1 区的 301-320 号氨基酸的 E1-51 (SEQ ID NO: 66)，

跨越 E1 C4 区的 313-332 号氨基酸(表位 A)的 E1-53(SEQ ID NO: 67)，

跨越 E1 区的 325-344 号氨基酸的 E1-55 (SEQ ID NO: 68)。

跨越 E2 区的 397-418 号氨基酸(表位 A)的 Env67 或 E2-67(SEQ ID NO: 72)，

跨越 E2 区的 409-428 号氨基酸（表位 A）的 Env69 或 E2-69（SEQ ID NO: 73），

跨越 E2 区的 583-602 号位置（表位 E）的 Env23 或 E2-23（SEQ ID NO: 86），

跨越 E2 区的 595-614 号位置（表位 E）的 Env25 或 E2-25（SEQ ID NO: 87），

跨越 E2 区的 607-626 号位置（表位 E）的 Env27 或 E2-27（SEQ ID NO: 88），

跨越 E2 区的 547-586 号位置（表位 D）的 Env17B 或 E2-17B（SEQ ID NO: 83），和

跨越 E2 区的 523-542 号位置（表位 C）的 Env13B 或 E2-13B（SEQ ID NO: 82）。

本发明提供了一种诱导 HCV 特异性抗体的治疗组合物，该组合物含有治疗有效量的含有由纯化的重组 HCV 单个的或特异性寡聚重组 E1 或 E2 蛋白形成的 E1/E2 复合物的组合物；并且选择性地含有可以药用的佐剂。本发明的重组 HCV 外被蛋白可以用重组哺乳动物细胞生产，或者重组 HCV 外被蛋白是用重组酵母细胞生产的。本发明提供了一种治疗诸如人类的受 HCV 感染的哺乳动物的方法，包括给予有效量的本文所披露的组合物，并选择性地给予可以药用的佐剂。本发明提供了一种用于诱导 HCV 特异性抗体的组合物，该组合物含有治疗有效量的含有选自 E1 和 E2 蛋白的至少一种纯化的重组 HCV 单个的或特异性寡聚重组外被蛋白的组合物；并选择性地含有可以药用的佐剂。

附图和表格的说明

图 1：质粒 pgptATA18 的限制图

图 2：质粒 pgsATA18 的限制图

图 3：质粒 pMS66 的限制图

图 4：质粒 pvHCV-11A 的限制图

图 5：对 IFN 治疗无反应者体内的抗 E1 水平

图 6：对 IFN 治疗有反应者体内的抗 E1 水平

图 7：对 IFN 治疗有完全反应的患者体内的抗 E1 水平

图 8：对 IFN 治疗有不完全反应的患者体内的抗 E1 水平

图 9：对 IFN 治疗无反应者体内的抗 E2 水平

图 10：对 IFN 治疗有反应者体内的抗 E2 水平

图 11：对 IFN 治疗有完全反应的患者体内的抗 E2 水平

图 12：对 IFN 治疗有不完全反应的患者体内的抗 E2 水平

图 13：与肽竞争的人抗 E1 反应性

图 14：抗 E1 单克隆抗体与肽反应性的竞争

图 15：对 IFN 治疗无反应者体内的抗 E1（表位 1）水平

图 16：对 IFN 治疗有反应者体内的抗 E1（表位 1）水平

图 17：对 IFN 治疗无反应者体内的抗 E1（表位 2）水平

图 18：对 IFN 治疗有反应者体内的抗 E1（表位 2）水平

图 19：抗 E2 单克隆抗体与肽反应性的竞争

图 20：与肽竞争的人抗 E2 反应性

图 21：本发明的核酸序列。本发明的编码 E1 或 E2 蛋白的核酸序列可以翻译（SEQ ID NO：3-13，21-31，35 和 41-49 是在从 1 号残基开始的读框中翻译的，SEQ ID NO：37-39 是在从 2 号残基开始的读框中翻译的）成序列表中所示出的相应 E1 或 E2 蛋白的氨基酸序列。

图 22：由受到 vvHCV39（1b 型）、vvHCV40（1b 型）、vvHCV62（3a 型）和 vvHCV63（5a 型）感染的细胞裂解物的四种不同的 E1 纯化物的 lentil 凝集素层析洗脱物级份所获得的 ELISA 的结果。

图 23：根据图 22 所示出的值由四种不同的 E1 结构所获得的 lentil 凝集素层析的洗脱曲线。

图 24：由受到 vvHCV39（1b 型）、vvHCV40（1b 型）、vvHCV62（3a 型）和 vvHCV63（5a 型）感染的细胞裂解物的四种不同的 E1 纯化物的凝胶过滤层析级份所获得的 ELISA 的结果。

图 25：由纯化的 1b 型（1）、3a 型（2）和 5a 型（3）E1 蛋白（来自分别用 vvHCV39、vvHCV62 和 vvHCV63 感染过的 RK13 细胞；按照例 5.2-5.3 所述用 lentil 凝集素纯化并还原）和标准物（4）获得的曲线。峰值用‘1’、‘2’和‘3’表示，它表示纯的 E1 蛋白峰（参见图 24，E1 反应性主要存在于 26-30 号级份中）。

图 26：按例 4 所述方法对 SDS-PAGE 进行银染色，E1 HWCV40（1b 型）的粗制裂解物（泳道 1）vvHCV40 的凝胶过滤收集物 1 相当于图 25 中所示的级份 10-17（泳道 2），vvHCV40 的凝胶过滤收集物 2 相当于

图 25 中所示的级份 18-25（泳道 3），和 E1 收集物（级份 26-30）（泳道 4）。

图 27：E1 构建体 39（1b 型）和 62（3a 型）的凝胶过滤级份的链霉亲和素-碱性磷酸酶的印迹。用 NEM-生物素标记蛋白。泳道 1：开始的凝胶过滤构建体 39，泳道 2：级份 26 构建体 39，泳道 3，级份 27 构建体 39，泳道 4，级份 28 构建体 39，泳道 5，级份 29 构建体 39，泳道 6，级份 30 构建体 39，泳道 7，级份 31 构建体 39，泳道 8，分子量标记，泳道 9，开始的凝胶过滤构建体 62，泳道 10：级份 26 构建体 62，泳道 11：级份 27 构建体 62，泳道 12：级份 28 构建体 62，泳道 13：级份 29 构建体 62，泳道 14：级份 30 构建体 62，泳道 15：级份 31 构建体 62。

在与图 26 相同条件下进行的 vvHCV-39（E1s，1b 型）和 vvHCV-62（E1s，3a 型）的凝胶过滤级份的 SDS-PAGE 凝胶的银染色。泳道 1：开始的凝胶过滤构建体 39，泳道 2：级份 26 构建体 39，泳道 3，级份 27 构建体 39，泳道 4，级份 28 构建体 39，泳道 5，级份 29 构建体 39，泳道 6，级份 30 构建体 39，泳道 7，级份 31 构建体 39，泳道 8，分子量标记，泳道 9，开始的凝胶过滤构建体 62，泳道 10：级份 26 构建体 62，泳道 11：级份 27 构建体 62，泳道 12：级份 28 构建体 62，泳道 13：级份 29 构建体 62，泳道 14：级份 30 构建体 62，泳道 15：级份 31 构建体 62。

用抗 E1 小鼠单克隆抗体 5E1A10 进行的 Western 印迹分析，提供了该纯化方法的完整的概述。泳道 1：粗制裂解物，泳道 2：lentil 层析的流通物，泳道 3：在 lentil 层析之后用 Empigen-BB 洗涤，泳道 4：lentil 层析的洗脱物，泳道 5：在浓缩 lentil 洗脱物期间流通物，泳道 6：在大小排阻层析（凝胶过滤）之后的 E1 合并物。

图 30：在来自受 vvHCV44 感染的 RK13 细胞的 E2 蛋白的 lentil 凝集素层析的 OD_{280} 曲线（连续的曲线）。点划线表示通过 ELISA 检测的 E2 反应性（如实施例 6 所述）。

图 31A：在来自受 vvHCV44 感染的 RK13 细胞的 lentil-凝集素凝胶过滤层析 E2 蛋白收集物的 OD_{280} 曲线（连续的曲线），其中，所述 E2 收集物被马上加样到凝胶过滤柱上（非还原条件）。点划线表示通过 ELISA 检测的 E2 反应性（如实施例 6 所述）。

图 31B：在来自受 vvHCV44 感染的 RK13 细胞的 lentil-凝集素凝胶过滤层析 E2 蛋白收集物的 OD_{280} 曲线（连续的曲线），其中，按照实施例 5.3 的方法还原或封闭 E2 收集物（还原条件）。点划线表示通过 ELISA 检测的 E2 反应性（如实施例 6 所述）。

图 32：在凝胶过滤之后，在图 31B 的还原条件下，由 vvHCV44 表达的 E2 蛋白的 Ni^{2+}-IMAC 层析和 ELISA 反应性。

图 33：通过图 32 中所示的 Ni^{2+}-IMAC 层析的 200mM 咪唑洗脱步骤（泳道 2）和 30mM 咪唑洗涤（泳道 1）步骤所回收的 0.5 微克纯化的 E2 蛋白的 SDS-PAGE 的银染色。

图 34：通过图 33 所示的 200mM 咪唑洗脱所回收的纯化的 E2 蛋白的脱盐步骤的 OD 曲线，意在去除咪唑。

图 35A：在治疗期间以及在治疗之后 6-12 个月的时间内通过 LIAscan 方法测定的 NR 和 LTR 的针对不同的 HCV 抗原（核心 1、核心 2、E2HCVR、NS3）的抗体含量。平均值用带有空心方框的曲线表示。

图 35 B：在治疗期间以及在治疗之后 6-12 个月的时间内通过 LIAscan 方法测定的 NR 和 LTR 的针对不同的 HCV 抗原（NS3、NS5、E1 和 E2）的抗体含量。平均值用带有空心方框的曲线表示。

图 36：LTR 和 NR 组中的平均 E1 抗体（E1Ab）和 E2 抗体（E2Ab）含量。

图 37：1b 型和 3a 型非反应者（NR）和长期反应者（LTR）的平均 E1 抗体（E1Ab）水平。

图 38：抗 E2 单克隆抗体的相对作图位置。

图 39：HCV E1 外被蛋白的部分脱糖基化。按照生产商的说明，将 vvHCV10A-感染过的 RK13 细胞的裂解物与不同浓度的糖苷酶一起温育。右侧图片：糖肽酶 F（PNGase F）。左侧图片：内切糖苷酶 H（Endo H）。

图 40：HCV E2 外被蛋白的部分脱糖基化。按照生产商的说明，将 vvHCV64-感染过的（E2）和 vvHCV41-感染过的（E2s）RK13 细胞的裂解物与不同浓度的糖肽酶 F（PNGase F）按照生产商的说明一起温育。

图 41：HCV E1 糖蛋白的体外诱变。诱变序列的作图和新的限制位点的产生。

图 42A：HCV E1 糖蛋白的体外诱变（第一部分）。PCR 的扩增的第

一个步骤。

图 42B: HCV E1 糖蛋白的体外诱变（第二部分）。重叠延伸和嵌套 PCR。

图 43: HCV E1 糖蛋白的体外诱变。在第一个扩增步骤中合成的 PCR 诱变片段（GLY#和 OVR#）的作图。

图 44A: 通过 Western 印迹分析在 HeLa（左侧）和 RK13（右侧）细胞中表达的 E1 糖蛋白变体。泳道 1: 野生型 VV（痘苗病毒），泳道 2: 原始 E1 蛋白（vvHCV-10A），泳道 3: E1 变体 G1y-1（vvHCV-81），泳道 4: E1 变体 G1y-2(vvHCV-82)，泳道 5: E1 变体 G1y-3(vvHCV-83)，泳道 6: E1 变体 G1y-4(vvHCV-84)，泳道 7: E1 变体 G1y-5(vvHCV-85)，泳道 8: E1 变体 G1y-6（vvHCV-86）。

图 44B: 通过 PCR 扩增/限制分析 E1 糖基化变体痘苗病毒。泳道 1: E1（vvHCV-10A）BspEI，泳道 2: E1. GLY-1（vvHCV-81），BspEI，泳道 4: E1（vvHCV-10A），SacI，泳道 5: E1. GLY-2（vvHCV-82），SacI，泳道 7: E1(vvHCV-10A)，SacI，泳道 8: E1. GLY-3(vvHCV-83) SacI，泳道 10: E1(vvHCV-10A)，StuI，泳道 11: E1. GLY-4(vvHCV-84)，StuI，泳道 13: E1(vvHCV-10A)，SmaI，泳道 14: E1. GLY-5(vvHCV-85)，SmaI，泳道 15: E1(vvHCV-10A)，StuI，泳道 17: E1. GLY-6(vvHCV-86)，StuI，泳道 3-6-9-12-15: 低分子量标记，pBluescript II SK+，MspI 。

图 45: 在酿酒酵母中表达的重组 E2 的 SDS 聚丙烯酰胺凝胶电泳。使接种物在亮氨酸选择培养基中生长 72 小时，并且用完全培养基以 1/15 的比例稀释。在 28℃下培养 10 天之后，取培养基样品。将通过 speedvac 浓缩的相当于 200 微升培养上清液加样到所述凝胶上。分析两个独立的转化体。

图 46: 在糖基化缺陷型酿酒酵母变体中表达的重组 E2 的 SDS 聚丙烯酰胺凝胶电泳。将接种物在亮氨酸选择培养基中生长 72 小时，并且用完全培养基以 1/15 的比例稀释。在 28℃下培养 10 天之后，取培养基样品。将通过离子交换层析浓缩的相当于 350 微升培养上清液加样到所述凝胶上。

图 47: 黑猩猩的特征和免疫方案。

图 48: 在经过 3 次免疫接种后的细胞应答。

图 49：在充分进行 E1 免疫接种时细胞应答的演化。

图 50：在经过 NS3 免疫接种时的细胞应答。

表 1：用于扩增构建实施例 1 中所示的不同形式的 E1 蛋白的相应的克隆和引物的特征

表 2：抗 E1 测定的总结

表 3：用于竞争研究的合成肽

表 4：外被抗体水平随时间的变化

表 5：LTR 和 NR 之间的差别

表 6：鼠 E2 单克隆抗体之间的竞争实验

表 7：用于构建 E1 糖基化变体的引物

表 8：通过 ELISA 分析 E1 糖基化变体

表 9：使用过佐剂的 E1 Balb/c 小鼠的特征

表 10：体液应答：不同的 E1 抗体水平所需要的免疫接种次数。

例 1：丙型肝炎病毒 E1 蛋白的克隆和表达

1.构建痘苗病毒重组载体

pgptATA18 痘苗重组质粒是 pATA18 的修饰形式（Stunnenberg 等，1988），其中额外插入了含有受痘苗病毒 13 中间启动子控制的大肠杆菌黄嘌呤鸟嘌呤磷酸核糖转移酶基因的片段（图 1）。质粒 pgsATA18 是通过将在 3 个开放读框内含有终止密码子的具有 SEQ ID NO: 1/94 的寡核苷酸接头插入 PstI 和 HindIII 切断的 pATA18 载体上而构建的。由此产生了一个额外的 PacI 限制位点（图 2）。没有恢复原有的 HindIII 位点。具有 SEQ ID NO: 1/94 的寡核苷酸接头：

```
5'        G GCATGC AAGCTT AATTAATT          3'
3'    ACGTC CGTACG TTCGAA TTAATTAA TCGA     5'
      PstI  SphI  HindIII  Pac I (HindIII)
```

为了有利于通过融合在重组蛋白上的工程产生的组氨酸片段的 Ni^{2+}螯合方法快速而有效地进行纯化，可以将疫苗重组载体 pMS66 设计成能表达具有一个额外的羧基末端组氨酸标记的分泌型蛋白。用以下方式合成含有通过三种限制酶产生平端（SmaI、StuI 和 PmlI/BbrPI）的独特位点的具有 SEQ ID NO: 2/95 的寡核苷酸接头，任何 cDNA 的羧基末端都可以插入具有一个编码蛋白酶因子 Xa 裂解位点的序列的框

内，随后是一个编码 6 个组氨酸和 2 个终止密码子的核苷酸序列（在 3'末端下游还形成了一个 PacI 位点）。将具有 SEQ ID NO: 2/95 的寡核苷酸插入 pgptATA18 的 XmaI 和 PstI 位点之间（图 3）。

具有 SEQ ID NO: 2/95: 的寡核苷酸接头

```
'5' CCGGG GAGGCCTGCACGTGATCGAGGGCAGACACCATCACCACCATCACTAATAGTTAATTAA CTGCA
  3'   .                                                                .
3'     C CTCCGGACGTGCACTAGCTCCCGTCTGTGGTAGTGGTGGTAGTGATTATCAATTAATT G
       ____                                                        ____
       XmaI                                                        PstI
```

业已按照布达佩斯条约将包含在大肠杆菌 MC1061（λ）中的质粒 pgptATA18 保存在 BCCM/LMBP（比利时微生物协调保藏中心/Laboratorium voor Moleculaire Biologie Plasmidencollectie, Universiteit Gent, K.L.Ledeganckstraat35B-9000 Ghent, Belgium），保藏号为 LMBP4486。所述保藏是在 2002 年 1 月 9 日进行的。

实施例 2.构建 HCV 重组质粒

2.1 编码不同形式的 E1 蛋白的构建体

通过以前所披露的 RNA 制备，随后进行逆转录和 PCR(Stuyver 等，1993b）从血清样品中获得聚合酶链式反应（PCR）产物。表 1 表示用于扩增相应的克隆和引物的特征。将所述 PCR 片段克隆到 SmaI 裂解过的 pSP72（Promega）质粒上。选择以下克隆插入痘苗重组载体上：HCCI9A（SEQ ID NO: 3）、HCCI10A（SEQ ID NO: 5）、HCCI11A（SEQ ID NO: 7）、HCCI12A（SEQ ID NO: 9）、HCCI13A（SEQ ID NO: 11）和 HCCI17A(SEQ ID NO: 13)，如图 21 所示。通过用 EcoRI 和 HindIII 限制裂解将含有编码 E1 区的 cDNA 片段从相应的 pSP72 质粒上裂解下来，并插入 EcoRI/HindIII 裂解过的 pgptATA18 疫苗重组载体上（如例 1 所述），位于 11K 痘苗病毒晚期启动子的下游。相应的质粒被命名为 pvSCV-9A，pvSCV-10A，pvSCV-11A，pvSCV-12A，pvSCV-13A 和 pvSCV-17A，在图 7 中示出了其中了 pvSCV-11A。

2.2 疏水区 E1 缺失变体

按以下方法制备含有密码子 Asp264-Val287 缺失的克隆 HCCI37

（790-861 号氨基酸，编码疏水性结构域 I 的片段）：用引物对 HCPr52（SEQ ID NO: 16）/ HCPr107（SEQ ID NO: 19）和 HCPr108（SEQ ID NO: 20）和 HCPr54（SEQ ID NO: 18）由克隆 HCCI10A 制备两种 PCR 片段。所述引物如图 21 所示。在电泳之后从琼脂糖凝胶中纯化以上两种 PCR 片段，并将 1 纳克每一种片段同时用作模板，通过引物 HCPr52（SEQ ID NO: 16）和 HCPr54（SEQ ID NO: 18）进行 PCR。将所得到的片段克隆到 SmaI 裂解过的 pSP72 载体上，并且能很方便地鉴定含有所述缺失的克隆，因为缺失了 24 个密码子（72 个碱基对）。筛选含有克隆 HCCI37（SEQ ID NO: 15）的质粒 pSP72-HCCI37。通过将所述缺失周围的 HCV 序列（通过 XmaI 和 BamHI 从载体 pSP72-HCCI37 裂解的片段）插入疫苗质粒 pvHCV-10A 上的 XmaI 和 BamHI 位点上构建含有缺乏疏水性结构域 I 的完整长度 E1 cDNA 的重组疫苗质粒。所得到的质粒被命名为 pvHCV-37。在验证性测序之后，将含有所述内部缺失的氨基末端片段从载体 pvHCV-37 上分离下来（用 EcoRI 和 BstEII 裂解），并重新插入 EcoRI 和 BstEII 裂解过的 pvHCV-11A 质粒上。预计该构建体能表达缺失了两个疏水性结构域 E1 蛋白，并命名为 pvHCV-38。克隆 HCCI38 的 E1 编码区如 SEQ ID NO: 23 所示。

由于在构建体 pvHCV-38 上不包括完整的位于 E1 羧基末端的亲水区（理论上讲延伸到 337-340 号氨基酸周围），通过 EcoRI/BamHI 裂解从 pvHCV-37 质粒上分离了较大的缺乏疏水结构域 I 的 E1 片段，并将其克隆到 EcoRI/BamHI 裂解过的 pgsATA-18 载体上。所得到的质粒被命名为 pvHCV-39，并且含有克隆 HCCI39（SEQ ID NO: 25）。通过 BamHI（其中的粘性末端用 KlenowDNA 聚合酶 I（Boehringer）补平）随后用 EcoRI（5'末端）消化将相同的片段从 pvHCV-37 载体上裂解下来。将该序列插入 EcoRI 和 BbrPI 裂解过的载体 pMS-66 上。由此得到了包含在质粒 pvHCV-40 上的克隆 HCCI40（SEQ ID NO: 27），它在其羧基末端含有 6 个组氨酸尾巴。

2.3 其他基因型的 E1

克隆 HCCI62（SEQ ID NO: 29）来自 3a 型感染的慢性丙型肝炎患者（在 WO94/25601 中披露为血清 BR36，克隆 BR36-9-13，SEQ ID NO: 18，并且还可以参见 Stuyver 等，1993a），而 HCCI63（SEQ ID NO: 31）来自 5a 型感染的输血后感染肝炎的儿童（在 WO94/25601 中披露

为血清 BE95，克隆 PC-4-1，SEQ ID NO: 45）。

2.4 E2 构建体

利用 Stuyver 等，1993b 所披露的 RNA 制备、逆转录和 PCR 技术通过引物 HCPr109（SEQ ID NO: 33）和 HCPr72（SEQ ID NO: 34）由血清 BE11（基因型 b）获得 HCV E2 PCR 片段 22，并将该片段克隆到 SamI 裂解过的 pSP72 载体上。用 NcoI/A1wNI 或通过 BamHI/A1wNI 裂解克隆 HCCI22A（SEQ ID NO: 35），并将该片段的粘性末端补平（用 K1enowDNA 聚合酶 I(Boehringer)补平 NcoI 和 BamHI 位点，用 T4DNA 聚合酶（Boehringer）补平 A1wNI 位点）。将所述 BamHI/A1wNI cDNA 片段插入疫苗 pgsATA18 载体上，该载体业已通过 EcoRI 和 HindIII 裂解线性化，并且用 K1enowDNA 聚合酶（Boehringer）将其粘性末端补平。所得到的质粒被命名为 pvSCV-41，并且编码从氨基酸 Met347 到 G1n673 的 E2 区，包括 E1 蛋白的 37 个氨基酸（从 Met347 到 G1y383），该片段可能起着信号序列的作用。将相同的 HCV cDNA 插入 EcoRI/BbrPI 裂解过的载体 pMS-66 上，随后用 K1enowDNA 聚合酶将末端补平，所述的质粒被命名为 pvSCV-42，并且同样编码 347-683 号氨基酸。用类似方法将 NcoI/A1wNI 片段插入 pgsATA18（pvSCV-43）或 pMS-66 疫苗载体（pvSCV-44）的相同位点上。pvSCV-4 和 pvSCV-444 编码多蛋白的 346-673 号氨基酸，其中来自 E1 蛋白的天然羧基末端区的 364-383 号氨基酸编码 E2 信号序列，以及成熟 E2 蛋白的 384-673 号氨基酸。

2.5 制备重组 HCV-痘苗病毒

兔肾 RK13 细胞（ATCC CCL37）、人骨肉瘤 143B 胸苷激酶缺陷型（TK-）（ATCC CRL8303）HeLa（ATCC CCL2）和 Heb G2（ATCC HB8065）细胞系是从美国模式培养物保藏所（ATCC ，Rockville，Md，美国）获得的。在 Dulbecco's 改进的 Eagle 培养基(DMEM)上生长所述细胞，用于生长 RK13 和 143B（TK-）的培养基补充了 10%的胎牛血清和 Earle's 盐（EMEM），而用于 Hep G2 细胞生长的培养基补充了葡萄糖（4 克/升）。按以前披露的方法(Panicali&Paoletti，1982; Piccini 等，1987; Mackett 等，1982，1984 和 1986）在 143B 或 RK13 细胞中常规增殖所述痘苗病毒 WR 菌株（Western Reserve，ATCC VR119）。用野生型痘苗病毒以 0.1 的感染复数（m.o.i.）（0.1 噬斑形成单位

（PFU）/细胞）感染 153B 细胞的铺满的单层。2 小时之后，以含有 500 纳克质粒 DNA 的磷酸钙共沉淀形式将所述疫苗重组质粒转染到受感染的细胞中，以便进行同源重组（Graham&van de Eb，1973; Makett 等，1985）。在用选择培养基（含有 25 微克/毫升霉酚酸（MPA）、250 微克/毫升黄嘌呤和 15 微克/毫升次黄嘌呤的 EMEM; Falkern 和 Moss，1988; Janknecht，1991）培养的兔肾 RK13 细胞上选择表达大肠杆菌黄嘌呤-鸟嘌呤转移酶（gpt）蛋白的重组病毒。在选择培养基的 0.9% 琼脂糖叠层下面的新的 RK13 细胞新的单层上纯化单个的重组病毒。筛选胸苷激酶缺陷型（TK-）重组病毒，然后在有 25 微克/毫升 5-溴-2'-脱氧尿苷的条件下在新的人 143B 细胞（TK-）单层上纯化噬斑。通过以 0.05 的 m.o.i.感染人 143B 或兔 RK13 细胞制备纯化的重组 HCV-痘苗病毒原种（Mackett 等，1988）。在经过 MPA 选择之后，通过利用用于克隆相应的 HCV 片段的引物（参见表 1）进行 PCR，用一份细胞裂解物样品（50 微升）证实所述 HCV cDNA 片段插入了重组痘苗病毒上。所述重组疫苗 HCV 病毒按照疫苗重组质粒编码进行命名，例如，重组痘苗病毒 vvHCV-10A 来自野生型 WR 菌株与 pvHCV-10A 的质粒的重组。

实施例 3：用重组痘苗病毒感染细胞

用实施例 2 中所披露的重组 HCV 痘苗病毒以 3 的感染复数感染 RK13 细胞的铺满的单层。为了进行感染，用磷酸缓冲的盐溶液 pH7.4（PBS）洗涤所述细胞单层 2 次，并用 MEM 培养基稀释所述重组痘苗病毒原种。向每 10^6 细胞中添加 200 微升病毒溶液，使得感染复数为 3，并在 24℃下培养 45 分钟。吸出所述病毒溶液，并向每 10^6 细胞中添加 2 毫升完整的生长培养基（参见例 2）。在 37℃下培养所述细胞 24 小时，在此期间进行了 HCV 蛋白的表达。

实施例 4：通过 Western 印迹分析重组蛋白

用 PBS 洗涤感染过的细胞 2 次，直接用裂解缓冲液（50mM Tris-HCl，pH7.5，100mM 氯化钠，1% Triton X-100，5mM 氯化镁，1 微克/毫升抑蛋白酶肽（Sigma，Bornem，比利时））直接裂解或通过在 50mM Tris-HCl pH7.5/10mMEDTA/150mM 氯化钠培养 5 分钟与烧瓶分离，并通过离心收集（1000g，5 分钟）。然后将细胞沉淀物每 10^6 细胞重新

悬浮在200微升裂解缓冲液(50mM Tris-HCl，pH8.0，2mM EDTA，150mM氯化钠，5mM 氯化镁，抑蛋白酶肽，1% Triton X-100）以14000rpm的速度在Eppendorf 离心机中离心5分钟使细胞裂解物澄清，以便除去不溶性的碎片。通过十二烷基硫酸钠-聚丙烯酰胺凝胶电泳（SDS-PAGE）分离20微升裂解物的蛋白。然后用冷却到4℃的Hoefer HIS转移装置将所述蛋白从所述凝胶上转移到硝酸纤维素膜（Amersham）上，在转移缓冲液中（25mM Tris-HCl，pH8.0，192mM 甘氨酸，20%（v/v）甲醇）以100V的稳定电压转移2小时。用Blotto（用PPS制备的5%(w/v)脱脂速溶奶粉; Johnson等，1981）封闭硝酸纤维素膜，并且与用 Blotto/0.1% Tween-20 稀释的一级抗体一起培养。通常将人类阴性对照血清或感染了HCV的患者的血清稀释200倍，并且在室温下与稀释 200 倍的野生型痘苗病毒感染的细胞裂解物一起预培养1小时，以便减弱非特异性结合。在用 Blotto/0.1% Tween-20 洗涤之后，将所述硝酸纤维素膜与用 Blotto/0.1% Tween-20 稀释的碱性磷酸酶底物溶液一起培养。用由PBS配置的0.1% Tween-20洗涤，然后将所述滤膜与碱性磷酸底物溶液（100mM Tris-HCl，pH9.5，100mM氯化钠，5mM氯化镁，0.38微克/毫升淡兰四唑，0.165微克/毫升5-溴-4-氯-3-吲哚磷酸）一起培养。除了电转移步骤之后，所有步骤都是在室温下进行的。

实施例5：纯化重组E1或E2蛋白

5.1 裂解

用磷酸缓冲的盐溶液（PBS）洗涤感染过的RK13细胞（具有E1或E2构建体）2次，并通过在含有10mM EDTA的PBS中培养将其从培养受体上分离下来。分离的细胞用PBS洗涤2次，并为每10^5细胞添加4℃的1毫升裂解缓冲液（50mM Tris-HCl，pH7.5，150mM氯化钠，1% Triton X-100，5mM氯化镁，1微克/毫升抑蛋白酶肽（Sigma，Bornem，比利时），含有 2mM 生物素化的 N-乙基马来酰亚胺（生物素-NEM）（Sigma））。用B型douncer对该裂解物进行匀浆，并在室温下放置0.5小时。将另外5倍体积的含有 10mM N-乙基马来酰亚胺（NEM，Aldrich，Bornem，比利时）的裂解缓冲液添加到所述原始裂解物中，并将该混合物在室温下放置15分钟。在4℃下通过在Beckman JA14

转子中以 14000rpm 的速度（最大速度为 30100g）离心 1 小时将不溶性细胞碎片从该溶液中清除。

5.2 凝集素层析

以 1 毫升/分钟将澄清的细胞裂解物加样到 0.8×10 厘米的 lentil-凝集素 Sepharose 4B 柱（Pharmacia）上，该柱用 5 倍柱体积的裂解缓冲液以 1 毫升/分钟的流速平衡过。用 5-10 倍柱体积的缓冲液 1（0.5M 磷酸钾，pH7.3，500mM 氯化钾，5%甘油，1mM6-NH2-己烷酸，1mM 氯化镁，和 1%圭基 PEG（KWANT，Bedum，荷兰））洗涤所述 lentil-凝集素柱。在某些实验中，随后用 10 倍柱体积的含有 0.5% Empigen-BB（Calbiochem，San Diego，CA，USA）而不是 1%圭基 PEG 的缓冲液 1 洗涤该柱。通过使用洗脱缓冲液（10mM 磷酸钾，pH7.3，5%甘油，1mM 己烷酸，1mM 氯化镁，0.5% Empigen-BB，和 0.5Mα-甲基-甘露糖吡喃糖苷）洗脱结合的材料。分离洗脱的材料，并通过例 6 所述的 ELISA 筛选级份中 E1 或 E2 蛋白的存在。图 22 表示由用 vvHCV39（1b 型）、vvHCV40（1b 型）、vvHCV62（3a 型）和 vvHCV63（5a 型）感染过的细胞裂解物的四种不同的 E1 纯化物的 lentil 凝集素洗脱物级份所获得的 ELISA 的结果。图 23 表示由图 22 所示出的值所获得的曲线。以上结果表明凝集素亲和层析可用于不同类型 HCV 的外被蛋白。

5.3 浓缩和部分还原

通过在 4℃下，在 Beckman JA20 转子中以 5000rpm 的速度离心 3 小时收集 E1 或 E2 阳性级份，在 Centricon30kDa（Amicon）上浓缩。在某些实验中，E1 或 E2 阳性级份是通过氮气蒸发收集并浓缩的。将相当于 3×10^8 细胞的量浓缩到大约 200 微升。为了进行部分还原，将 30%的 Empigen-BB（Calbiochem，San Diego，CA，USA）添加到这 200 微升样品中，使最终浓度为 3.5%，随后添加用水配制的 1M DTT，使最终浓度为 1.5-7.5mM，并在 37℃下培养 30 分钟。随后添加 NEM（用二甲基亚砜配制成 1M 浓度）至 50mM 的最终浓度，并在 37℃下让它们再反应 30 分钟，以便封闭游离的巯基基团。

5.4 凝胶过滤层析

用 3 倍柱体积的 PBS/3% Empigen-BB 平衡 Superdex-200HR10/20 柱（Pharmacia）。将还原的混合物注入 Smart 系统（Pharmacia）的 500 微升样品回路中，并添加 PBS/3% Empigen-BB 缓冲液，进行凝

胶过滤。收集从 V0 至 Vt 的 250 微升的级份。按例 6 所述方法筛选所述级份中 E1 或 E2 蛋白的存在。图 24 表示由受到 vvHCV39（1b 型）、vvHCV40（1b 型）、vvHCV62（3a 型）和 vvHCV63（5a 型）感染的细胞裂解物的四种不同的 E1 纯化物的凝胶过滤层析级份所获得的 ELISA 的结果。图 25 表示由纯化的 1b 型（1）、3a 型（2）和 5a 型（3）E1 蛋白（来自分别用 vvHCV39、vvHCV62 和 vvHCV63 感染过的 RK13 细胞；按照例 5.2-5.3 所述用 lentil 凝集素纯化并还原）和标准物（4）获得的曲线。峰值用 '1'、'2' 和 '3' 表示，它表示纯的 E1 蛋白峰（参见图 24，E1 反应性主要存在于 26-30 号级份中）。这些峰表示大约为 70kDa 的非常接近的分子量，相当于二聚体蛋白。在这三条曲线上的其他峰表示痘苗病毒和/或细胞蛋白，这些蛋白只所有能与 E1 分离，仅仅是因为在例 5.3 中所述的还原步骤，并且因为随后在有适当的洗涤剂的条件下进行的凝胶过滤步骤。如图 26 所示，收集物 1（表示 10-17 号级份）和收集 2（表示 18-25 号级份）含有不存在于 E1 收集物（26-30 号级份）中的污染性蛋白。将 E1 峰级份在 SDS-PAGE 上分离，并按例 4 所述方法吸印。按图 27 所示通过图 27 所示的链霉亲和素-碱性磷酸酶检测用 NEM-生物素标记过的蛋白。不难看出，在凝胶过滤层析之前所存在的 29kDa 和 45kDa 的污染蛋白（泳道 1）仅以非常低的含量出现在 26-30 号级份中。大约 65kDa 的带表示 E1 二聚体形式，这种二聚体形式不能完全裂解成单体 E1 形式。3aE1 蛋白获得了类似结果（泳道 10-15），它在 SDS-PAGE 上表现出更快的迁移速度，因为仅存在 5 个碳水化合物而不是 6 个。图 28 表示以与图 26 相同的条件进行的 SDS-PAGE 凝胶电泳的银染色。在图 29 中提供了该纯化方法的完整的综述。

通过在例 4 所披露的 Western 印迹进一步证实了纯化 E1 蛋白的存在。二聚体 E1 蛋白似乎是非凝聚的，并且不含杂质。用 477Perkins-Elmer 测序仪对按照上述方法从用 vvHCV-40 感染过的细胞中纯化的亚型 1b E1 蛋白的氨基末端进行测序，并证实它的第一个残基是酪氨酸。由此证实了所述 E1 蛋白业已在正确的位置上（A191 和 Y192 之间）通过信号肽酶裂解与它的信号序列分离。这一结果证实了 Hijikata 等（1991）的发现，即成熟的 E1 蛋白的氨基末端始于 192 号氨基酸位置。

5.5 纯化 E2 蛋白

按照例 5.1-5.4 所述方法从用 vvHCV-44 感染过的 RK13 细胞中纯化 E2 蛋白(384-673 号氨基酸)。图 30 表示 lentil 凝集素层析的 OD_{280} 曲线(连续的曲线)。点划线表示通过 ELISA 检测的 E2 反应性(如例 6 所述)。图 31 表示通过 lentil 凝集素 E2 蛋白收集物的凝胶过滤层析所获得的相同的曲线(参见图 30),其中的一部分是按照例 5.3 所述方法还原并封闭的,而另一部分被马上加样到所述柱上。这两部分 E2 收集物在独立的凝胶过滤柱上过滤。可以证实,如果没有进行还原的话,E2 能与污染性蛋白形成共价连接的凝集物。在还原和封闭之后,污染性蛋白的主要部分分离成 V0 级份。与 E2 蛋白一起纯化的其他污染性蛋白不再与 E2 蛋白共价连接,因为这些污染物可以在随后的步骤除去。图 32 表示为了 E2 蛋白纯化而进行的额外的 Ni^{2+}-IMAC 纯化步骤。该亲和素纯化步骤采用了由 vvHCV44 所表达的添加在 E2 蛋白上的 6 个组氨酸残基。污染性蛋白要么可以从所述柱中流过,或者通过 30mM 咪唑洗涤清除。图 33 表示 0.5 微克纯化的 E2 蛋白的银染色的 SDS-PAGE 和 30mM 的咪唑洗涤。通过 200 Mm 的咪唑洗脱步骤可以轻易地回收纯的 E2 蛋白。图 34 表示用于除去咪唑的额外的脱盐步骤,并可以转移到需要的缓冲液中,例如,PBS、碳酸盐缓冲液、盐水。

在例 5.1-5.5 中所述的方法以大约 50000 平方厘米的 RK13 细胞为原材料(用于生产 E1 的细胞用 vvHCV11A(或 vvHCV40)感染过,或为了生产 E2 蛋白所述细胞用 vvHCV41、vvHCV42、vvHCV43 或 vvHCV44 感染过),可以纯化大约 1.3 毫克 E1 蛋白和 0.6 毫克 E2 蛋白。

还应当指出的是,分泌型 E2 蛋白(占细胞内形式的大约 30-40%,60-70%)的特征是形成凝聚物(与预期结果相反)。因此纯化分泌型 E2 存相同的问题。可以按上述方法纯化分泌型 E2。

实施例 6:用于检测抗 E1 或抗 E2 抗体或用于检测 E1 或 E2 蛋白的 ELISA

以每孔 1 倍体积(例如 50 微升或 100 微升或 200 微升)的用 PBS 配制的 5 微克/毫升链霉亲和素溶液(Boehringer Mannheim)对 Maxisorb 微孔平板(Nunc, Roskilde, 丹麦)进行包衣,在 4℃下进行 16 小时或者在 37℃下进行 1 小时。或者,用 1 倍体积的用 50mM 碳

酸钠缓冲液 pH9.6 配制的 5 微克/毫升雪花莲凝集素（GAN）所述孔进行包衣，4℃下进行 16 小时或者在 37℃下进行 1 小时。在用 GAN 进行包衣时，用 400 微升 Innotes HCV Ab III 试剂盒（Innogenetics，Zwijndrecht，比利时）的洗涤溶液将所述平板洗涤 2 次。用 1.5-2 倍体积的封闭溶液（用 PBS 配制的 0.1%酪蛋白和 0.1%氮化钠）封闭未结合的包衣表面，在 37℃下进行 1 小时或者在 4℃下进行 16 小时。吸出封闭溶液。将纯化的 E1 或 E2 稀释到 100-1000 纳克/毫升（浓度是在 A=280nm 的波长下测定的）或者筛选柱级份中的 E1 或 E2（参见例 5），或将未纯化过的细胞裂解物中的 E1 或 E2（例 5.1）用封闭溶液稀释 20 倍，并将 1 倍体积的 E1 或 E2 溶液添加到每一个孔中，并在 37℃下在链霉亲和素或 GAN 包衣的平板上培养 1 小时。用 1 倍体积的 Innotes HCV Ab III 试剂盒（Innogenetics，Zwijndrecht，比利时）洗涤溶液将所述微孔洗涤 3 次。用 Innotes HCV Ab III 试剂盒的样品稀释剂将血清样品稀释 20 倍，或者将单克隆抗 E1 或抗 E2 抗体稀释到 20 纳克/毫升的浓度，并让 1 体积的溶液与 E1 或 E2 蛋白在 37℃下反应 1 小时。用 100 微升 Innotes HCV Ab III 试剂盒（Innogenetics，Zwijndrecht，比利时）的洗涤溶液洗涤所述微孔 5 次。按上述方法检测结合的抗体：在 37℃下用由 1 体积的 Innotes HCV Ab III 试剂盒（Innogenetics，Zwijndrecht，比利时）的缀合稀释剂稀释 1/80000 倍的山羊抗人或抗小鼠 IgG、过氧化物酶结合的二级抗体（DAKO，Glostrup，丹麦）培养每一个孔 1 小时，用 400 微升 Innotes HCV Ab III 试剂盒（Innogenetics，Zwijndrecht，比利时）的洗涤溶液洗涤所述平板 3 次之后，通过添加用 1 体积的 Innotes HCV Ab III 试剂盒（Innogenetics，Zwijndrecht，比利时）底物溶液稀释 100 倍的 Innotes HCV Ab III 试剂盒的底物（Innogenetics，Zwijndrecht，比利时）在 24℃下反应 30 分钟实现显色。

实施例 7：用不同的临床特征随访患者组

7.1 监测抗 E1 和 E2 抗体

业已开发出了用于筛选并证实 HCV 抗体存在的现有的丙型肝炎病毒（HCV）诊断测定方法。所述测定方法似乎不能提供用于监测疾病治疗或预测治疗结果的信息。不过，对于乙型肝炎来说，检测并定量抗

外被蛋白抗体在临床实践上可能更为有用。为了研究利用抗 E1 抗体效价和抗 E2 抗体效价作为预测丙型肝炎疾病的后果的标记，将一系列用 IFN-α治疗的具有长期持续的反应（被定义为在治疗之后至少 1 年时间内，在血液中具有正常的转氨酶水平和阴性 HCV-RNA 测试（在 5'非编码区上进行的 PCR）的患者）与表现出物反应或表现出在治疗结束时具有延迟的生物化学反应的患者进行比较。

将一组 8 个具有长期持续反应的 IFN-α治疗过的患者（LTR，随访 1-3.5 年，3 个 3a 型和 5 个 1b 型）与 9 个对治疗表现出不完全反应的患者（NR，随访 1-4 年，6 个 1b 型和 3 个 3a 型）进行比较。1b 型（vvHCV-39，参见例 2.5）和 3a E1（vvHCV-62，参见例 2.5）蛋白是由痘苗病毒系统表达的（参见例 3 和 4），并纯化至均匀（例 5）。检验来自受 1b 型丙型肝炎病毒感染的患者的样品与纯化的 1b 型 1E 蛋白的反应性。同时通过例 6 所示的 ELISA 检验受 3a 型感染的样品与抗 3aE1 型抗体的反应性。通过 Inno-LiPA 基因型分类测定（Innogenetics，Zwijndrecht，比利时）确定感染不同患者的丙型肝炎病毒的基因型。图 5 表示在干扰素治疗期间观测到的以及在治疗之后的时间内观测到的所述患者的抗 E1 信噪比。LTR 例子中，一致性地表现出抗 E1 水平的迅速下降（在所有三个例子中都是阴性的），而 NR 例子的抗 E1 水平大体上保持稳定。所获得的某些抗 E1 数据在表 2 中是以平均 S/N 比例±SD 的形式表示的（平均抗 E1 效价）。可以根据表 5、6、7 和 8 中所示的信噪比推测出抗 E1 效价。

在治疗结束时，就已经能够观察到两组之间的显著差别。在 LTR 组中，抗 E1 抗体效价下降了 6.9 倍，而在 NR 组中仅下降了 1.5 倍。在随访观察结束时，在具有持续反应的患者体内的抗 E1 效价下降到原来效价的 1/22.5，在 NR 组中甚至还略有增加。因此，基于以上数据，在监测 IFN-α治疗期间抗 E1 抗体含量的降低与对治疗的长期持续性反应相关。所述抗抗 E1 测定对于预测 IFN 治疗的长期反应或者对于总体上治疗丙型肝炎疾病非常有用。

这一发现是出乎预料的，相反，本发明人曾经预计在具有长期反应的患者中，在 IFN 治疗期间抗 E1 抗体含量会提高。对于乙型肝炎来说，由于抗 H BsAg 抗体的血清转变而清除了这种病毒。另外，在许多其他病毒感染中，在抗外被蛋白抗体产生以后，消除了有关病毒。不过，

在本发明的实验中，在对治疗具有长期反应的患者体内抗 E1 抗体明显降低，而在非反应患者体内该抗体的含量大体上保持在同一水平上。尽管以上实验的结果是出乎预料的，不过这种非显而易见的发现可能对于 HCV 感染的临床诊断非常重要并且有用。如图 9、10、11 和 12 所示，在所研究的相同的患者体内抗 E2 含量表现出很大差别，而对于抗 E1 抗体来说没有发现其效价有明显的下降。图 35 提供了该实验研究的完整的概况。

从表 2 可以看出，在治疗开始时，长期反应者体内的抗 E1 效价比对治疗的不完全反应者的抗 E1 效价平均高至少 2 倍。因此，测定治疗开始时抗 E1 抗体的效价，或者将监测在感染过程中对患者进行监测并测定抗 E1 效价。可以作为临床诊断丙型肝炎的有用的标志。

正如在例 7.3 中所述，可能需要使用 E1 或 E2 蛋白的更明确的部分。

7.2 在较大的患者群中分析 E1 和 E2 抗体

以上实验研究使得本发明人得出了这样的结论:在感染被彻底消除以后，抗 HCV 外被蛋白的抗体的变化比抗更常见的研究过的 HCV 抗原的抗体的变化更快，以 E1 抗体的改变最明显。因此，我们在研究中包括了更多的 1b 型和 3a 感染的 LTR 患者，并且在患者群中补充了匹配的一系列 NR，这样两个组就各包括 14 位患者。还分析了某些部分反应者（PR）以及反应延迟的反应者（RR）。

图 36 表示 LTR 和 NR 组中平均 E1 抗体（E1Ab）和 E2 抗体（E2Ab）含量，而表 4 和 5 表示统计学分析。在这一较大的群体中，在 IFN-α 治疗之前较高的 E1 抗体与 LTR 相关（$p < 0.03$）。由于在 3a 型感染的患者中观察到了比 1b 型感染的患者体内的更高的 E1 抗体含量（图 37），也要将基因型考虑在内（表 4）。在 1b 感染型组内，在开始治疗时 LTR 的 E1 抗体含量同样高于 NR 的 E1 抗体含量（$p < 0.05$）；有限数量的 3a 型感染的 NR 无法进行统计学分析。

在对 LTR 进行 1.5 年的抗体含量监测时，只有 E1 抗体与在治疗开始时测定的含量相比被快速清除[治疗结束时 P=0.0058;在治疗之后 6 个月和 12 个月分别为 P=0.0047 和 P=0.0051]。在 1 型或 3 型感染的 LTR 中这种消除依然保持明显（平均 P 值 < 0.05）。以上数据证实了最初的发现，即在发展的早期阶段 E1Ab 含量迅速降低。这一特征似乎

与病毒基因型无关。在 NR、PR 或 RR 中，在整个随访观察期间没有发现测定的任何抗体有所改变。在对治疗有积极反应的患者中，在治疗期间表现出 ALT 水平的正常化和 HCV-RNA 阴性，而在持续的反应者（LTR）和反应延迟的反应者（RR）之间存在显著差别。与 LTR 相反，RR 没有表现出 E1 抗体含量的任何降低，这表明隐藏的 HCV 感染的存在既不能通过 PCR 或用于检测 HCV-RNA 的传统方法证实，又不能通过提高了的 ALT 含量证实。在治疗期间在 RR 组中仍然存在的少量病毒 RNA 似乎能进行抗 E1 B 细胞刺激。因此，抗 E1 监测不仅能够区分 LTR 和 NR，而且还能区分 LTR 和 RR。

7.3 监测 E1 蛋白的特定区域的抗体

尽管鉴定 HCV 抗原的分子生物学方法导致了在病毒诊断发展中的前所未有的突破，对 λgt11 文库进行免疫筛选的方法主要得到了分布在核心和非结构区上的线性表位，而有关外被蛋白区的分析还有待于在哺乳动物细胞中进行 E1/E2 区的克隆和表达。这一方法与很多其他病毒感染明显不同，其中所述外被蛋白区的表位在解释基因组结构很久以前就已经进行了作图。所述表位和相应的抗体通常具有中和活性，这种活性可用于疫苗开发和/或用于开发具有临床或预测价值的诊断测定方法（例如，乙型肝炎表面抗原的抗体）。由于目前还没有可以对丙型肝炎疾病进行临床诊断和预测的 HCV 疫苗或检测方法，用于免疫监督的病毒外被区的鉴定可能会明显导致对 HCV 诊断和预防的新的指导作用。

按照以前披露的方法(EP-A-0489968)，根据 HC-J1 序列(Okamoto 等，1990）合成了若干二十聚体肽（表 3），这些肽各自有 8 个氨基酸相互重叠。除了肽 env35（又被称为 E1 35）之外，这些肽无一能够检测大约 200 个 HCV 患者血清中的抗体。只有两种血清与 env35 肽有微弱的反应。不过，通过例 6 所披露的抗 E1 ELISA，可以发现以下额外的表位：对例 6 所述的抗 E1 ELISA 进行以下改进：在样品稀释剂中将 50 微克/毫升 E1 肽与 1/20 稀释过的人血清混合。图 13 表示在有单个的或 E1 肽混合物的情况下，人血清对重组 E1 （由 vvHCV-40 表达）蛋白的反应性结果。尽管在线性免疫测定方案中，通过涂在小片上的 E1 肽只能检测 2%的血清，不过超过半数的抗 E1 抗体的血清可以在用重组 E1 蛋白测定时通过相同的肽竞争。在注射纯化的 E1 蛋白之后从

Balb/C 小鼠中获得的某些鼠单克隆抗体随后与 E1 竞争对所述单个肽的反应性（图 14）。很明显，含有所述优势表位的 env53 的部分在添加 env53 时能够明显竞争如果种血清与 E1 的反应性，并且还检测到了 env31 区抗体。这一发现是惊人的，因为在将 env53 和 env31 肽直接涂在固相上时没有发现任何反应性。

因此，用申请人以前所披露的技术合成了有关肽（参见 WO93/18054）。合成了以下肽:

肽 env35A-生物素

NH_2-SNSSEAADMIMHTPGCV-GK 生物素（SEQ ID NO: 51）

跨越 HCV 多蛋白的 E1 区的 208-227 号氨基酸

肽生物素-env53（“表位 A”）

生物素-GG-ITGHRMAWDMMMNWSPTTAL-COOH （SEQ ID NO: 52）

跨越 HCV 多蛋白的 E1 区的 313-332 号氨基酸

肽 1bE1（“表位 B”）

H_2N-YEVRNVSGIYHVTNDCSNSSIVYEAADMIMHTPGCGK-生物素（SEQ ID NO: 53）

跨越 HCV 多蛋白的 E1 区的 192-228 号氨基酸

并且与肽 E1a-BB(生物素-GG-TPTVATRDGKLPATQLRRHIDLL,SEQ ID NO: 54）和 E1b-BB（生物素-GG-TPTLAARDASVPTTTIRRHVDLL，SEQ ID NO: 55）的反应性进行比较，后者分别来自 1a 和 1b 基因型序列的相同部分，并且于 1993 年在 Glasgow 召开的第 9 届国际病毒学会议上公开（表位 C）。并且将用表位 A、B 和 C 测试了一组 HCV 血清的反应性，并且将表位 B 与 env35A 进行比较（在 47 个 HCV 阳性血清中，有 8 个在表位 B 上是阳性的，而没有一个能与 env35A 起反应）。按照例 6 所披露的方法直接测定了结合在用链霉亲和素包衣的平板上的生物素化的肽（50 微克/毫升）对表位 A、B 和 C 的反应性。很明显，表位 A 和 B 反应性最强，而表位 C 和 env35A-生物素反应性要低得多。检测了监测过对完整 E1 蛋白的反应性的相同系列的患者（例 7.1）对表位 A、B 和 C 的反应性。观察到了对表位 C 的很弱的反应性，而如图 15、16、17 和 18 所示，表位 A 和 B 能与大部分血清起反应。不过，针对反应性最强的表位（表位 A）的抗体似乎不能预测疾病的缓解，而抗 1bE1 抗体(表位 B)几乎只是在 IFN 治疗开始时存在于长期反应者体内。因此，

抗 1bE1（表位 B）抗体和抗 env35（表位 A）抗体可以提供用于预测丙型肝炎疾病的有用的标志。Env53 表位可用于检测交叉反应抗体（能在主要基因型之间交叉反应的抗体），而 Env53 区的抗体对于在血清或肝脏组织中检测通用 E1 抗原来说是非常有用的。能识别 Env53 区的单克隆抗体能够与随机表位文库起反应。在用单克隆抗体 5E1A10 进行免疫筛选时能起反应的 4 个克隆中存在序列-GWD-。由于它与存在于所有 HCV 变体的 env53 区中的通用 HCV 序列类似，因此认为序列 AWD 含有 env53 交叉反应鼠表位的必需序列。Env31 明显还含有一个可变区，该可变区可能含有一个存在于氨基末端序列-YQVRNSTGL-（SEQ ID NO: 93）上的表位，并且可用于诊断。在表 3 中所示的 Env31 和 E1-31 是肽 1bE1 的一部分。肽 E1-33 和 E1-51 还能在某种程度上与鼠抗体起反应，而肽 E1-55（含有可变区 6（V6）；跨越 329-336 号氨基酸）也能与某些患者血清起反应。

抗 E2 抗体明显具有不同于抗 E1 抗体的模式，尤其是在对治疗具有长期反应的患者中。因此，很明显的是，抗外被抗体的减少不能向使用单个的抗 E1 或抗 E2 蛋白那样使用重组 E1/E2 蛋白进行有效测定。在同时测定两种类型抗体的测定中，抗 E2 反应能明显干扰抗 E1 反应。因此，证实了检测抗外被蛋白抗体对单个的 E1 和 E2 蛋白的能力。

7.4 在 24 个抗 E2MAB 中只有三个能够竞争肽对重组 E2 的反应性。其中的两个与 HVRI 区反应（肽 E2-67 和 E2-69，被称为表位 A），而一个能识别由肽 E2-13B 识别的表位（表位 C）。大部分鼠抗体能识别构象抗 E2 表位（图 19）。通常还能观察到人类对 HVRI（表位 A）的反应，以及较低程度的对 HVRII（表位 B）和第三种线性表位区（由肽 E2-23、E2-25 或 E2-27 竞争，并称为表位 E）和第四种线性表位区（由肽 E2-17B 竞争，表位 D）的反应，不过大部分血清与构象表位起反应（图 20）。可以按照以下方法根据其相对位置对所述构象表位进行分类：通过蛋白 A 亲和层析纯化能识别构象表位的杂交瘤的 15C8C1、12D11F1、9G3E6、8G10D1H9、10D3C4、4H6B2、17F2C2、5H6A7、15B7A2 的上清液中的 IgG 抗体，在硼酸缓冲液中在有生物素的条件下将所得到的 1 毫克/毫升 IgG 生物素化。通过凝胶过滤层析将生物素化的抗体与游离的生物素分离。将合并的生物素化抗体级份稀释 100-10000

倍。在有 100 倍含量的非生物素化竞争抗体存在的条件下通过生物素化的 IgG 检测与固相结合的 E2 蛋白，然后再通过碱性磷酸酶标记的链霉亲和素检测。

在表 6 中给出了的竞争的百分比。根据以上结果，可以描绘出四种构象抗 E2 表位区（表位 F、G、H 和 I）（图 38）。另外，上述 Mabs 可以识别在本研究中所没有出现过的肽的突变型线性表位。Mabs 4H6B2 和 10D3C4 能竞争 16A6E7 的反应性，但是，与 16A6E7 不同，它们不能识别肽 E2-13B。这种 Mabs 能够识别相同的线性表位（表位 C）的变体，或者识别受到了空间位阻或在将 16A6E7 结合到 E2-13B 区上之后的构象表位（表位 H）。

实施例 8: E1 糖基化变体

8.1 前言

由哺乳动物细胞表达的，由 vvHCV10A 编码的 E1 蛋白，和由 vvHCV41-44 编码的 E2 蛋白分别含有 6 个和 11 个碳水化合物部分。这一结果可以通过将 vvHCV10A 感染过的或 vvHCV44 感染过的 RK13 细胞的裂解物与浓度出现降低的糖苷酶（PNGase F 或内切糖苷酶 H（Boehringer Mannheim Piochemica），按照生产商的说明操作）一起培养，以便使裂解物中的蛋白（包括 E1）被部分脱糖基化（分别如图 39 和 40 所示）。

缺乏某些糖基化位点的变体可以选择具有改善了的免疫学反应性的外被蛋白。以 HIV 为例，业已发现缺乏某些特定的糖添加基序的 gp120 蛋白特别适用于诊断或疫苗目的。通过在 A/香港/3/68（H3N2）流感病毒的逃脱变体的血凝集素蛋白上添加新的寡聚糖侧链，可以抑制与中和单克隆抗体的反应性（Skehel 等，1984）。当通过位点特异性诱变将新的糖基化位点导入所述流感血凝集素蛋白中时，会出现明显的抗原性改变，这表明所述糖类起着抗原性调节物的作用（Gallagher 等，1988）。在另一个分析中，来自 Friend 鼠白血病病毒的表面蛋白 gp70 的 8 个碳水化合物添加基序缺失了。尽管 7 种这样的突变都不影响病毒的感染性，但是针对氨基末端的糖基化信号的第四个突变导致出现了非感染性表型（Kayman 等，1991）。另外，本领域众所周知的是添加 N-连接的糖链对于折叠中间物的稳定来说是重要

的，并因此对于有效折叠、抑制错折叠和内质网降解、寡聚化、生物学活性、以及糖蛋白的转运来说也是重要的（有关综述参见 Rose 等，1988; Doms 等，1993; Helenius，1994）。

在对 HCV 基因型的不同的外被蛋白基因序列进行的排比之后，可以了解到并不是HCV亚型1bE1蛋白上的所有6个糖基化位点都是正确折叠和反应性所必需的，因为在某些类型或亚型中缺少了某些糖基化位点。在1b、6a、7、8和9型中所存在的第4个糖基序（在Asn251上）在目前已知的所有其他类型中都缺乏。可以对这种糖添加基序进行诱变，以便产生具有改善了的反应性的 1bE1 蛋白类型。另外，2b型序列在V5区出现了一个额外的糖基化位点（在Asn299上）上。属于基因型2c的分离物I83甚至缺少V1区的第一个糖基序(在Asn上)，而该基序存在于所有其他的分离物上（Stuyver 等，1994）。不过，即使是在完全保守的糖添加基序中，所述糖的存在可能并非是折叠所必需的，不过，它可能是避免免疫监督方面发生作用。因此，对正确折叠（以及反应性）来说并非必需的糖添加基序的确定不是明确的，并且必须分析每一种变体，并检测其反应性。糖基化基序的诱变（NXS或NXT序列）可以通过以下方式实现：诱变编码N、S或T的密码子，使这些密码子编码的氨基酸不同于它原来所编码的N，和/或使它所编码的氨基酸不同于它原来所编码的S或T。另外，可以将X位点诱变成P，因为已知NPS或NPT通常不会受到糖类的修饰。在确定哪一个糖添加基序是折叠和/或反应性所需要的，哪一个是不需要的之后，可以实施所述诱变的组合。

8.2E1蛋白的诱变

所有突变都是用克隆HCCI10A的E1序列（SEQ ID NO: 5）进行的。第一轮PCR反应是用针对位于疫苗11K晚期启动子上游的GPT序列的有义引物‘GPT’（参见表7），和含有需要的碱基变化的反义引物（被命名为GLY#，其中的#表示糖基化位点的编码，参见图41）进行的，以便获得所述诱变。设计了6种GLY#引物（各自对特定的糖基化位点有特异性），以便：

将编码N-糖基化Asn的密码子(AAC或AAT)修饰成Gln密码子(CAA或CAG）。之所以选择谷氨酰胺，是因为它非常类似于天冬酰胺（这两种氨基酸都是中性的，并且含有非极性残基，谷氨酰胺具有较长的侧

链（一个以上的-CH2-基团））。

在糖基化位点下游的若干密码子的一个上导入沉默突变，以便产生新的特殊的或罕见的（例如，E1G1y5 的第二个 SmaI 位点）限制酶位点。在不改变氨基酸序列的前提下，这种突变可用于区分突变型序列和原始 E1 序列（pvHCV-10A）或与其他序列区分（图 41）。这种额外的限制位点还可用于构建新的杂合（2 倍、3 倍等）糖基化变体。

在第一个错配的核苷酸的 5'末端延长 18 个核苷酸，并向 3'末端延长 12-16 个核苷酸。表 7 示出了与 N-连接的糖基化位点的序列的重叠的六种 GLY#引物的序列。

为了进行定向诱变，使用了‘错误启动’或‘重叠延伸’（Horton，1993）。这一概念如图 42 和 43 所示。首先，由靶基因扩增两种独立的片段，以便得到每一种突变位点。由 5'末端获得的 PCR 产物（产物 GLY#）是用 5'有义 GPT 引物（参见表 7）和相应的 3'反义 GLY#引物扩增的。第二个片段（产物 OVR#）是用 3'反义 TK_R 引物和相应的 5'有义引物（OVR#引物，参见表 7，图 43）扩增的。

所述 OVR#引物针对部分 GLY#引物序列。因此，这两组 PCR 产物拥有一个具有相同序列的重叠区。当混合上述中间产物（GLY-1 和 OVR-1，GLY-2 和 OVR-2 等），以较高的温度熔化并退火时，产物 GLY#的上部有义链可以与产物 OVR#的反义链退火（反之亦然），使得这两股彼此起着对方的引物的作用（参见图 42B）。在两轮 PCR 中通过 Taq 聚合酶延伸退火的重叠部分，产生了完整长度的突变型分子 E1G1y#，该分子具有能破坏#号糖基化位点的突变。通过常见的两种内部嵌套引物在第三轮 PCR 中制备了用于克隆的足够量的 E1GLY#产物。这两种新的引物分别与疫苗 11K 启动子的 3'末端（有义 GPT-2 引物）和疫苗胸苷激酶基因座的 5'末端（反义 TK_R-2 引物，参见表 7）重叠。所有 PCR 条件是按照在 Stuyver 等（1993）所披露的条件下进行的。

通过 EcoRI/BamHI 裂解将上述每一种 PCR 产物克隆到 EcoRI/BamHI 裂解过的含有原始 E1 序列的疫苗载体上（pvHCV10-A）。通过 EcoRI/BamHI 裂解分析特所选择克隆的插入片段的长度，以及每一种新的限制位点的存在。通过双链测序证实重叠的突变位点的序列。

8.3 分析 E1 糖基化变体

以例 8.2 中所披露的含有突变型 E1 序列的 6 种质粒为起点，通过与例 2.5 中所披露的野生型痘苗病毒重组制备了重组型痘苗病毒。简单地讲，用 6 种具有突变型 E1 序列的重组痘苗病毒、vvHCV-10A（具有非突变型 E1 序列）和野生型痘苗病毒感染 175 平方厘米的亚铺满的 RK13 细胞的烧瓶。在感染之后 24 小时裂解细胞，并按照例 4 所述方法在 Western 印迹上进行分析（参见图 44A）。所有突变体在 SDS-PAGE 上都表现出比原始 E1 蛋白更快的迁移率（相当于大约为 2-3kDa 的较小的分子量）；证实增加了一个糖基部分。还通过 PCR 和限制酶分析分析了重组病毒，以便证实不同突变体的身份。图 44B 表示所有突变体（如图 41 所示）都含有预期的额外限制位点。将另一部分细胞裂解物通过 ELISA 检测不同突变体的反应性。将所述裂解物稀释 20 倍，并添加到按照例 6 所述方法用凝集素 GNA 包衣的微孔平板上。让被捕获（突变型）糖蛋白与稀释 20 倍的在例 6 中所披露的 24 个 HCV 感染患者的血清起反应。在表 8 中示出了 6 个突变体和 E1 的信噪比（S/N）值（GLY#的 OD/野生型的 OD）。该表还示出了 GLY#的 S/N 值与 E1 蛋白的比例。应当理解的是，将不同突变体的细胞裂解物用于比较同患者血清的反应性的方法，可能会导致出现的结果是由于不同的表达水平所致，而不是由于反应水平所致。所述困难可以通过按例 5 所述方法纯化不同的突变体而得到克隆，并通过检测所有不同 E1 蛋白的相同的量而克服。不过，表 5 所示的结果已经表明，去掉第一个（GLY1）、第三个（GLY3）和第六个（GLY6）糖基化基序，能减弱某些血清的反应性，而去掉第二个和第五个位点则不能。去掉 GLY4 似乎能提高某些血清的反应性。以上结果表明，不同的患者对本发明糖基化突变体的反应不同。因此，所述突变型 E1 蛋白可用于 HCV 疾病的诊断（筛选、证实、预测等）和预防。

实施例 9：在糖基化缺陷型酵母中表达 HCV E2 蛋白

为相当于克隆 HCCL41 的 E2 序列提供 α 交配因子前/原信号序列，将其插入酵母表达载体上，并且用该构建体转化酿酒酵母细胞，将 E2 蛋白分泌到生长培养基中。业已发现，在酿酒酵母菌株中表达所述构建体时，大多数糖基化位点被高甘露糖型糖基化所修饰（图 45）。这会导致过高的不均匀性，并且抑制其反应性，这是疫苗或诊断用途所

不需要的。为了克服这一问题，通过筛选钒酸盐抗性克隆的方法制备了具有修饰过的糖基化途径的酿酒酵母突变体。通过分析糖蛋白转化酶的分子量和不均匀性，分析了所述克隆的修饰过的糖基化途径。这使得我们能够鉴定不同的糖基化缺陷型酿酒酵母突变体。然后在某些特定的突变体中表达 E2 蛋白，并在例 4 所述的 Western 印迹上让这种蛋白与例 7 所述的单克隆抗体起反应（图 46）。

实施例 10：一般用途

本发明的结果表明，为了获得 HCV 外被蛋白与人类患者血清的高的反应性，不仅需要好的表达系统，而且还需要好的纯化方法。这一目的可以通过使用本发明的合适的HCV外被蛋白表达系统和/或纯化方法而实现，所述表达系统能确保保持所述蛋白的天然折叠，而本发明的纯化方法能确保消除污染性蛋白，并保留其构象，并因此保留了 HCV 外被蛋白的反应性。用于诊断性筛选测定的纯化 HCV 外被蛋白的量为每年数克。对于疫苗用途来说，可能需要更大量的外被蛋白。因此，所述痘苗病毒系统可用于筛选最佳的表达构建体，并用于有限的扩大和大规模表达并纯化含有高甘露糖糖类的单个的或特异性寡聚外被蛋白。在用若干酵母菌株表达时可以实现这一目的。例如，对于乙型肝炎来说，用哺乳动物细胞生产 H BsAg 的成本比用酵母产生的乙型肝炎疫苗生产成本高。

本发明所披露的纯化方法总体上还可用于‘病毒外被蛋白’。这些病毒的例子来自黄病毒、新发现的 GB-A、GB-B 和 GB-C 肝炎病毒、瘟病病毒（如牛病毒性腹泻病毒（BVDV）、猪霍乱病毒（HCV）、边界病病毒（BDV）），不过，还在较低的程度上涉及到诸如乙型肝炎病毒的病毒（主要是用于纯化 H BsAg）。

本发明的外被蛋白纯化方法，可用于在详细说明部分所列举的低等或高等真核细胞或原核细胞中细胞内和细胞外表达的蛋白。

实施例 11：证实预防和治疗作用

通过接种 E1 可以减轻慢性感染 HCV 的黑猩猩的肝脏疾病。不过，为了获得明显的免疫反应，需要多次接种。普通技术人员可以理解的是，通过免疫调节可以产生病毒的抗性，这种抗性是由病毒本身产生

或者是由宿主产生的。为了分析在 HCV 上是否存在这种免疫调节，比较了在初次感染的黑猩猩和慢性感染的黑猩猩中对 E1 和 NS3 的免疫反应。由于预计在慢性感染的动物体内会出现较弱的反应，选择这一组动物进行更严格的免疫接种方案，包括以下步骤：采用在小鼠上证实更有效的佐剂诱导细胞反应（表 9），与明矾进行比较。这种佐剂是用于初次感染的动物的佐剂；而用于慢性感染动物的免疫接种方案包括 12 次免疫接种，与之对应的是初次感染的动物有 6 次免疫接种（图 47）。

尽管由于免疫接种的动物数量无法进行统计学分析，不过在人体液反应中可以检测到以下明显的倾向（表 10）：在初次感染的动物的血清转化的接种次数较少；而在初次感染的动物中的免疫反应的幅度明显较大，有 2/3 的感染过的动物没有达到 10 个内部单位的水平，即使是在免疫接种 12 次之后也是如此。

在 3 次免疫接种之后进行的细胞反应分析发现了更大的差别（图 48a-d），包括以下差别：在慢性感染的动物体内几乎缺乏 E1-特异性 T 细胞增殖，而在初次感染的动物组中可以观察到明显的刺激作用；IL-2 测定证实了在慢性携带者的 T 细胞区室中的较低的刺激作用；并且在初次感染的动物中诱导了明显的 Th2（IL-4）反应，与含有明矾佐剂的疫苗预期的结果一样。

这证实了，至少 E1 免疫接种能在初次感染的动物中提供预防性作用，并且表明 E2 和/或 E1 和 E1 和 E2 蛋白和肽的组合可以在初次感染的动物中提供有用的治疗和/预防效果。

正如以下结果所证实的，通过多次免疫接种只能部分克服对诱导针对 HCV E1 抗原的细胞和体液应答的‘损害’：在每一次注射之后观察到了抗体效价的提高，但是在 2/3 的动物中没有达到初次感染的动物的水平；并且 T 细胞增殖反应依然非常（图 49）。不过 ELISPOT 结果证实了 IL-2 略有增加（未发表），IFN-g 没有变化（未发表），并且 IL-4 增加（图 49），这表明，Th2 型应答更容易诱导。与在初次感染的动物中经过 3 次免疫之后所达到的水平相比，发现 IL-4 仍然处于低水平上。

对 NS3 免疫接种进行了十分相似的观察，其中，在慢性黑猩猩上使用了更强的佐剂（RIBI）。与明矾制剂相比，在初次感染的动物上

发现了以下现象：在两组中诱导的抗体效价相当（未发表）；与在初次感染的动物中的应答相比，在慢性动物中细胞分泌和 T 细胞增殖几乎缺乏（图 49a-b）。

目前已经有某些证据表明，在慢性携带者中对 HCV 的免疫应答较低，或者至少不足于消除感染。上述结果支持了这样的假说，即 HCV 慢性携带者的免疫系统可能受到了损害，并且它对 HCV 抗原的应答不如在初次感染的动物中有效。

在 Wiedmann 等所做的研究中（Hepatology2000; 31: 230-234）的研究中，在 HCV 慢性携带者上，对 HBV 的免疫效果较差，这表明这种免疫损伤不局限于 HCV 抗原。DeMaria 等（Hepatology2000; 32: 444-445）证实了以上结果，并且提供了适合 HCV 患者的改进的疫苗用药方案。本文所提供的数据表明，增加免疫接种的次数，确实可以增强体液的应答，但是，细胞（特别是 Th1 细胞）应答是难以诱导的，即使在使用有效的佐剂时也是如此。在抗病毒治疗的时候开始免疫接种是有利的，此时免疫系统更容易作出应答。

表 1: 重组痘苗质粒和病毒

质粒名称	名称	cDNA 亚克隆构建	长度 (nt/aa)	用于插入的载体
pvHCV-13A	E1s	EcoR I - Hind III	472/157	pgptATA-18
pvHCV-12A	E1s	EcoR I - Hind III	472/158	pgptATA-18
pvHCV-9A	E1	EcoR I - Hind III	631/211	pgptATA-18
pvHCV-11A	E1s	EcoR I - Hind III	625/207	pgptATA-18
pvHCV-17A	E1s	EcoR I - Hind III	625/208	pgptATA-18
pvHCV-10A	E1	EcoR I - Hind III	783/262	pgptATA-18
pvHCV-18A	COREs	Acc I (KI) - EcoR I (KI)	403/130	pgptATA-18
pvHCV-34	CORE	Acc I (KI) - Fsp I	595/197	pgptATA-18
pvHCV-33	CORE-E1	Acc I (KI)	1150/380	pgptATA-18
pvHCV-35	CORE-E1b.his	EcoR I - BamH I (KI)	1032/352	pMS-66
pvHCV-36	CORE-E1n.his	EcoR I - Nco I (KI)	1106/376	pMS-66
pvHCV-37	E1Δ	Xma I - BamH I	711/239	pvHCV-10A
pvHCV-38	E1Δs	EcoR I - BstE II	553/183	pvHCV-11A
pvHCV-39	E1Δb	EcoR I - BamH I	960/313	pgsATA-18
pvHCV-40	E1Δb.his	EcoR I - BamH I (KI)	960/323	pMS-66
pvHCV-41	E2bs	BamH I (KI)-AlwN I (T4)	1005/331	pgsATA-18
pvHCV-42	E2bs.his	BamH I (KI)-AlwN I (T4)	1005/341	pMS-66
pvHCV-43	E2ns	Nco I (KI) - AlwN I (T4)	932/314	pgsATA-18
pvHCV-44	E2ns.his	Nco I (KI) - AlwN I (T4)	932/321	pMS-66
pvHCV-62	E1s (type 3a)	EcoR I - Hind III	625/207	pgsATA-18
pvHCV-63	E1s (type 5)	EcoR I - Hind III	625/207	pgsATA-18
pvHCV-64	E2	BamH I - Hind III	1410/463	pgsATA-18
pvHCV-65	E1-E2	BamH I - Hind III	2072/691	pvHCV-10A
pvHCV-66	CORE-E1-E2	BamH I - Hind III	2427/809	pvHCV-33

nt:核苷酸 aa:氨基酸 K1: Klenow DNA Pol 填充 T4: T4 DNA Pol 填充
位置: HCV 多蛋白序列上的氨基酸位置

表 1（续）：重组痘苗质粒和病毒

质粒名称	HCV cDNA 亚克隆			
	名称	构建	长度 (nt/aa)	用于插入的载体
pvHCV-81	E1*-GLY 1	EcoRI - BamH I	783/262	pvHCV-10A
pvHCV-82	E1*-GLY 2	EcoRI - BamH I	783/262	pvHCV-10A
pvHCV-83	E1*-GLY 3	EcoRI - BamH I	783/262	pvHCV-10A
pvHCV-84	E1*-GLY 4	EcoRI - BamH I	783/262	pvHCV-10A
pvHCV-85	E1*-GLY 5	EcoRI - BamH I	783/262	pvHCV-10A
pvHCV-86	E1*-GLY 6	EcoRI - BamH I	783/262	pvHCV-10A

nt: 核苷酸 aa: 氨基酸 Kl: Klenow DNA Pol 填充 T4: T4 DNA Pol 填充

位置: HCV 多蛋白序列上的氨基酸位置

表 2: 抗 E1 测定总结

S/N±SD（平均抗 E1 效价）

	治疗开始	治疗结束	随访
LTR	6.94 ± 2.29 (1:3946)	4.48 ± 2.69 (1:568)	2.99 ± 2.69 (1:175)
NR	5.77 ± 3.77 (1:1607)	5.29 ± 3.99 (1:1060)	6.08 ± 3.73 (1:1978)

LTR: 1 年以上的长期、持续的反应

NR: 无反应、反应延迟或部分反应

表 3

用于竞争研究的合成肽

蛋白	肽	氨基酸序列	位置	SEQ ID NO
E1	E1-31	LLSCLTVPASAYQVRNSTGL	181-200	56
	E1-33	QVRNSTGLYHVTNDCPNSSI	193-212	57
	E1-35	NDCPNSSIVYEAHDAILHTP	205-224	58
	E1-35A	SNSSIVYEAADMIMHTPGCV	208-227	59
	E1-37	HDAILHTPGCVPCVREGNVS	217-236	60
	E1-39	CVREGNVSRCWVAMTPTVAT	229-248	61
	E1-41	AMTPTVATRDGKLPATQLRR	241-260	62
	E1-43	LPATQLRRHIDLLVGSATLC	253-272	63
	E1-45	LVGSATLCSALYVGDLCGSV	265-284	64
	E1-49	QLFTFSPRRHWTTQGCNCSI	289-308	65
	E1-51	TQGCNCSIYPGHITGHRMAW	301-320	66
	E1-53	ITGHRMAWDMMMNWSPTAAL	313-332	67
	E1-55	NWSPTAALVMAQLLRIPQAI	325-344	68
	E1-57	LLRIPQAILDMIAGAHWGVL	337-356	69
	E1-59	AGAHWGVLAGIAYFSMVGNM	349-368	70
	E1-63	VVLLLFAGVDAETIVSGGQA	373-392	71
E2	E2-67	SGLVSLFTPGAKQNIQLINT	397-416	72
	E2-69	QNIQLINTNGSWHINSTALN	409-428	73
	E2-83B	LNCNESLNTGWWLAGLIYQHK	427-446	74
	E2-81B	AGLIYQHKFNSSGCPERLAS	439-458	75
	E2-1B	GCPERLASCRPLTDFDQGWG	451-470	76
	E2-3B	TDFDQGWGPISYANGSGPDQ	463-482	77
	E2-5B	ANGSGPDQRPYCWHYPPKPC	475-494	78
	E2-7B	WHYPPKPCGIVPAKSVCGPV	487-506	79
	E2-9B	AKSVCGPVYCFTPSPVVVGT	499-518	80
	E2-11B	PSPVVVGTTDRSGAPTYSWG	511-530	81
	E2-13B	GAPTYSWGENDTDVFVLNNT	523-542	82
	E2-17B	GNWFGCTWMNSTGFTKVCGA	547-566	83
	E2-19B	GFTKVCGAPPVCIGGAGNNT	559-578	84
	E2-21	IGGAGNNTLHCPTDCFRKHP	571-590	85
	E2-23	TDCFRKHPDATYSRCGSGPW	583-602	86
	E2-25	SRCGSGPWITPRCLVDYPYR	595-614	87
	E2-27	CLVDYPYRLWHYPCTINYTI	607-626	88
	E2-29	PCTINYTIFKIRMYVGGVEH	619-638	89
	E2-31	MYVGGVEHRLEAACNWTPGE	631-650	90
	E2-33	ACNWTPGERCDLEDRDRSEL	643-662	91
	E2-35	EDRDRSELSPLLLTTTQWQV	655-674	92

表4：外被抗体含量随时间的变化（完整的研究，28位患者）

Wilcoxon Signed Rank test (P values)	E1Ab HR	E1Ab HR 全部	E1Ab HR	E1Ab LTR Ib型	E1Ab LTR 3a型	E1Ab LTR 全部	E2Ab HR	E1Ab LTR Ib型	E1Ab LTR 3a型	E1Ab LTR 全部	E1Ab LTR 全部
End of therapy'	0.1167	0.2004	0.285		0.0058"	0.043"	0.0499"	0.0180"	0.0648		
6 months follow up'	0.66		0.7213	0.5938	0.0047"	0.043"	0.063		0.04326	0.0464"	
12 months follow up'	0.7989	0.3185	1		0.0051"	0.0679	0.0277"	0.0669	0.0058"		

'数值是与治疗开始时获得的值相比较

"P值<0.05

表5. LTR和NR的差异（完整的研究）

Mann-Withney U检验（P值）	E1Ab S/N 全部	E1Ab 效价 全部	E1Ab S/N 1b型	3a型	E1Ab S/N 全部	E2Ab S/N
治疗开始	0.0257*	0.05*	0.68		0.1078	
治疗结束	0.1742				0.1295	
6个月随访	1	0.6099	0.425		0.3081	
12个月随访	0.67	0.23	0.4386		0.6629	

*P值<0.05

表6. 鼠E2单克隆抗体之间的竞争实验

竞争剂	生物素化的抗E2 mabs的抗E2反应性的降低（%）									
	17H10F4D10	2F10H10	16A6E7	10D3C4	4H6B2	17C2F2	9G3E8	12D11F1	15C8C1	8G10D1H9
17H10F4D10	-	62	10	ND	11	ND	5	6	30	ND
2F10H10	90	-	1	ND	30	ND	0	4	12	ND
16A6E7	ND	ND	-	ND	ND	ND	ND	ND	ND	ND
10D3C4	11	60	92	-	94	26	28	43	53	30
4H6B2	ND	ND	82	ND	-	ND	ND	ND	ND	ND
17C2F2	2	ND	75	ND	56	-	11	10	0	0
9G3E8	ND	ND	68	ND	11	ND	-	60	76	ND
12D11F1	ND	ND	26	ND	13	ND	ND	-	88	ND
15C8C1	ND	ND	18	ND	10	ND	ND	ND	-	ND
8G10D1H9	2	2	11	ND	15	ND	67	082	81	-
竞争剂对照										
15B7A2	0	0	9	15	10	9	0	0	0	5
5H6A7	0	2	0	12	8	0	0	4	0	0
23C12H9	ND	ND	2	12	ND	4	ND	ND	ND	2

ND = 未测定

表7.引物

SEQ ID NO. 96 GPT	5'-GTTTAACCACTGCATGATG-3'
SEQ ID NO. 97 TKn	5'-GTCCCATCGAGTGCGGCTAC-3'
SEQ ID NO. 98 GLY1	5'-CGTGACATGGTACATTCCGGACACTTGGCGCACTTCATAAGCGGA-3'
SEQ ID NO. 99 GLY2	5'-TGCCTCATACACAATGGAGCTCTGGGACGAGTCGTTCGTGAC-3'
SEQ ID NO. 100 GLY3	5'-TACCCAGCAGCGGGAGCTCTGTTGCTCCCGAACGCAGGGCAC-3'
SEQ ID NO. 101 GLY4	5'-TGTCGTGGTGGGGACGGAGGCCTGCCTAGCTGCGAGCGTGGG-3'
SEQ ID NO. 102 GLY5	5'-CGTTATGTGGCCCGGGTAGATTGAGCACTGGCAGTCCTGCACCGTCTC-3'
SEQ ID NO. 103 GLY6	5'-CAGGGCCGTTGTAGGCCTCCACTGCATCATCATATCCCAAGC-3'
SEQ ID NO. 104 OVR1	5'-CCGGAATGTACCATGTCACGAACGAC-3'
SEQ ID NO. 105 OVR2	5'-GCTCCATTGTGTATGAGGCAGCGG-3'
SEQ ID NO. 106 OVR3	5'-GAGCTCCCGACTGCTGGGTAGCGC-3'
SEQ ID NO. 107 OVR4	5'-CCTCCGTCCCCACCACGACAATACG-3'
SEQ ID NO. 108 OVR5	5'-CTACCCGGGCCACATAACGGGTCACCG-3'
SEQ ID NO. 109 OVR6	5'-GGAGGCCTACAACGGCCCTGGTGG-3'
SEQ ID NO. 110 GPT-2	5'-TTCTATCGATTAAATAGAATTC-3'
SEQ ID NO. 111 TKn-2	5'-GCCATACGCTCACAGCCGATCCC-3'

加下划线的核苷酸表示额外的限制位点

用粗体字表示的核苷酸表示相对于原始HCC110A序列的突变

表8. 通过ELISA分析E1糖基化变体

血清

	1	2	3	4	5	6	7	8	9	10	11	12		
SN GLY1	1.802402	2.120971	1.403871	1.205597	2.120191	2.866913	1.950345	1.850183	1.730193	2.466162	1.220654	1.629403		
SN GLY2	2.400795	1.76818	2.326495	2.639308	2.459019	5.043993	2.146302	1.595477	1.668973	2.402212	1.467582	2.070524		
SN GLY3	1.042718	1.715477	2.261648	2.354748	1.691818	4.033742	1.90092	1.402099	1.602222	2.191658	1.464210	1.721104		
SN GLY4	2.570154	3.824038	3.874605	1.499387	3.15	4.71302	4.198751	3.959542	3.710507	5.170841	4.250784	3.955153		
SN GLY5	2.462051	1.793761	2.409344	2.627368	1.716311	4.984765	2.13912	1.570330	1.708937	3.021607	1.562092	2.07278		
SN GLY6	2.031467	1.495737	2.131613	2.627925	2.494833	4.704027	2.02069	1.496480	1.704978	2.677757	1.529608	1.744221		
SN E1	2.628205	2.227036	2.512792	2.790861	3.131679	4.869128	2.287753	1.954198	1.805566	2.616822	1.55719	2.593886		
	13	14	15	16	17	18	19	20	21	22	23	24	总计 S/N	平均 S/N
SN GLY1	6.886561	3.233604	3.763498	1.985105	2.317721	6.675179	1.93470	2.47171	4.378033	1.168748	2.166889	1.708992	59.88534	2.495223
SN GLY2	7.556882	2.507813	3.621928	3.055849	2.933792	7.05433	2.127712	2.921208	4.660101	1.150701	1.661914	1.632765	69.65243	2.902185
SN GLY3	7.930538	2.783055	3.016099	2.046628	2.516305	5.776357	1.980165	2.657364	4.268033	0.97707	1.336775	1.20376	62.09872	2.587447
SN GLY4	8.176818	6.561122	5.707668	5.664498	5.004813	8.4125	3.813321	3.002535	4.293038	2.393011	3.08713	2.481585	102.6978	4.279076
SN GLY5	8.883408	2.940334	3.125501	3.330912	2.654224	5.424107	2.442804	3.126761	4.84557	1.153656	1.817901	1.636211	69.26511	2.886046
SN GLY6	8.005561	2.499952	2.621704	2.672389	2.363301	6.194107	1.600716	2.665433	2.781003	1.280743	1.475062	1.716423	61.32181	2.555075
SN E1	8.825112	3.183771	3.067265	3.280335	2.980364	7.191984	2.771218	3.678066	6.35443	1.167286	2.083333	1.70252	76.04068	3.169195

血清

	1	2	3	4	5	6	7	8	9	10	11	12		
GLY1/E1	0.637316	0.952374	0.55869	0.431977	0.677036	0.588794	0.852516	0.954901	0.958261	0.94319	0.783682	0.628171		
GLY2/E1	0.848878	0.793901	0.925463	0.94569	0.785233	1.035913	0.93617	0.816436	0.925431	0.94656	0.942455	0.798232		
GLY3/E1	0.660834	0.770298	0.900053	0.84373	0.508312	0.992733	0.859761	0.758418	0.887385	0.837488	0.940294	0.663547		
GLY4/E1	0.911587	1.717097	1.541952	0.537245	1.005882	0.967939	1.635317	2.028172	2.05505	1.976	2.72978	1.524798		
GLY5/E1	0.877607	0.805447	0.959831	0.941408	0.647746	1.019642	0.935031	0.806641	0.946488	1.154762	1.003148	0.799102		
GLY6/E1	0.716296	0.671826	0.648305	0.90578	0.798669	0.982522	0.883204	0.765781	0.944284	1.023288	0.962288	0.672435		
	13	14	15	16	17	18	19	20	21	22	23	24	总计 E1/GLY#	平均 E1/GLY#
GLY1/E1	0.644248	1.015852	1.226986	0.605153	0.777666	0.928144	0.696162	0.672013	0.817759	1.018300	1.030207	0.957620	19.36524	0.806885
GLY2/E1	0.86827	0.800489	1.180833	0.931505	0.984377	1.084289	0.78779	0.794245	0.874061	0.98586	0.797719	0.815998	21.67304	0.903077
GLY3/E1	0.690033	0.887856	0.983319	0.997968	0.843962	0.803029	0.714554	0.695308	0.797215	0.837558	0.641052	0.675314	19.19921	0.799967
GLY4/E1	0.92654	2.060802	1.860833	1.732902	1.680587	0.89102	1.376045	0.810335	0.901773	2.050064	1.787422	1.392178	36.38592	1.51608
GLY5/E1	1.006606	0.923536	1.019006	1.017857	0.690574	0.75419	0.861491	0.850109	0.867612	0.968323	0.872593	0.919042	21.70879	0.904533
GLY6/E1	0.907134	0.705217	0.854737	0.784184	0.79296	0.72221	0.543702	0.724683	0.519396	1.097197	0.70003	0.962919	19.59691	0.816538

表 9. Balb/c 小鼠体内有佐剂的 E1 的特征

	明矾	T细胞佐剂	RBI
抗体效价(平均值 ± SD, n=6)	96000 ± 101000	62000 ± 60000	176000 ± 149000
抗体同种型	IgG1	IgG1/2b	IgG1/2a
脾脏中的T细胞殖[1] (n=3)	11750 (2/3)	48300 (3/3)	26000 (3/3)
淋巴节中的T细胞增殖[2]	无特异性刺激	4000	8000
细胞因子特征(脾)	Il-4	IFN-g/Il-4	IFN-g/Il-4

[1] 在经过 3 次皮下/肌肉免疫接种之后，分别对 3 只随机挑选的小鼠进行分析，结果以用 E1 刺激（1 μg/ml）4 天之后获得的平均比 cpm 形式表示，括号中的数字表示具有上述特殊刺激背景的小鼠数量

[2] 在经过 1 次足底免疫接种之后（n=2），结果以用 E1 刺激（1 μg/ml）5 天之后获得的平均比 cpm 形式表示。

表 10. 体液应答：不同 E1 抗体水平需要的免疫接种次数

动物	状态	血清转变 [1]	>1 U/ml[2]	>10 U/ml
Marcel	慢性	3	4	5
Peggy	慢性	3	5	>12
Femma	慢性	4	5	>12
Yoran	初次感染	3	4	5
Marti	初次感染	2	3	5

[1] 如果在免疫接种之前不存在 E1 抗体，将 ELISA 信号确定为高于截止水平，在其它病例中，高于免疫接种之前效价的 3 个独立的时间点的结果被视为血清转变点。

[2] 单位确定如下：在干扰素治疗之前和用 1b 基因型感染后，在人慢性载体中的 E1 抗体含量对 50%的患者来说为<0.1 U/ml，对 25%的患者来说为 0.1-1 U/ml，其余 25%的患者 >1 U/ml，n=58。

实施例 12：对受到 HCV1b 亚型慢性感染的黑猩猩进行免疫接种

用 E1(192-326 号氨基酸)接种在 13 年以前(在免疫接种之前 5015 天)就已经感染了 HCV1b 亚型菌株的黑猩猩(Phi1)，所述 E1 来自不同的 1b 基因型菌株，在氨基酸水平上具有 95.1%的相同性（还可参见 WO99/67285 的表 2，该文献被全文收作本文的参考），并且是按照 WO99/97285 中的例 1-3 所披露的方法制备的）。按照生产商的方法

（Ribi 公司，Hamilton，MT）对黑猩猩一共进行了 6 次肌内免疫接种，每次 50 微克 E1，所述 E1 溶解在 PBS/0.05%CHAPS，与 RIBI R-730（MPLA+TDM+CWS）混合。以上 6 次免疫接种是用两个系列每个系列 3 次接种的方式进行的，每个系列中的每次接种之间间隔 3 周时间，而 2 个系列之间间隔 6 周时间。从免疫接种之前 150 天开始，在免疫接种期间一直到免疫接种后 1 年（参见下文和 WO99/67285）对所述黑猩猩进行连续监测，监测表示 HCV 诱导的疾病的活性的各种参数。所述参数包括血液化学、ALT、AST、γGT、血液化学、血清中的病毒负荷，肝脏和肝组织中的病毒负荷。另外，在体液和细胞水平监测对所述接种的免疫应答。在此期间，还要监测所述接种的任何负面影响，如行为的变化、临床症状、体重、体温和局部反应（发红、肿胀、硬结），不对这些效果进行检测。

很明显，一旦抗 E1 的抗体水平达到最大值 ALT（特别是γGT，数据未发表）含量就会降低（参见 WO99/67285 的图 8）。一旦所述抗体含量开始下降，ALT 就能相当快地反弹，但是，只要仍然可以检测到抗 E1，γGT 就保持在较低的水平上。

在可检测到抗 E1 的时间内，肝脏内的 E2 抗原降低到几乎检测不到的水平，而 E2 抗原在所述抗体消失之后很快就反弹。随着肝脏中所述核心和 E2 抗原变得无法检测，肝脏的炎症也明显减弱（还可参见 WO99/67285 的表 3）。这是疫苗诱导肝脏损伤减轻的一个主要证据，有可能是通过清除，至少是部分清除其主要靶器官肝脏中的病毒抗原所致。

在整个研究期间，通过 Amplico HCV 监测仪（Roche，Basel，瑞士）测定在血清中的病毒血症水平大体上保持不变。

对体液应答所作的更详细的分析发现，最大终点效价达到了 14.5×10^3（在第 6 次免疫接种之后），并且该效价在免疫接种之后 1 年内降低到无法检测的水平（WO99/67285 的图 8）。WO99/67285 的图 9 表明，由 B 细胞识别的可以由肽模拟的主要表位位于 E2 N-末端区（肽 V1V2 和 V2V3，有关所使用的肽的细节参见 WO99/67285 的表 4）。由于针对重组 E1 的反应性强而且持续时间长，由该附图可以推测由该表位所识别的抗体仅仅代表抗 E1 的所有抗体群体的一部分。其余部分是针对不能由肽模拟的表位的，即不连续的表位。所述表位仅仅出现在

完整的 E1 分子上或者甚至只能出现在颗粒样结构上。所述对 E1 的免疫应答是独特的，至少与在人类慢性 HCV 携带者（授予 Maertens 的 WO96/13590）和黑猩猩（van Doorn 等，1996）上正常出现的情况相比是独特的，它能在它的初次感染过程中产生抗 E1 抗体。在上述患者中，抗 E1 也是部分针对不连续表位的，不过，大部分是针对 C4 表位的（占患者血清的±50%），小部分是针对 V1V2（根据基因型在 2%-70%之间波动），而针对 V2V3 的反应性只能偶然记录到（Maertens 等，1997）。

分析 T 细胞反应性发现，免疫系统的该区室还是由所述疫苗以特殊方式刺激的，由于所述 T 细胞的刺激指数从 1 提高到 2.5，并且在随后的时间内依然会有某些提高（WO99/67285 的图 10）。这种 T 细胞反应性只能出现在对干扰素治疗的长期反应者体内（参见授予 Leroux-Roels 等的 PCT/EP94/03555; Leroux-Roels 等，1996）。

实施例 13：用不同的亚型免疫接种慢性 HCV 携带者

用 HCV1b 基因型的 E1 接种来自 1a 基因型的 HCV 在 10 年以前就已经感染了来自 1a 基因型的 HCV（在免疫接种之前 3809 天）亚型菌株的黑猩猩（Ton），在氨基酸水平上具有 79.3%的相同性（还可参见 WO99/67285 的表 2，该文献被全文收作本文的参考），并且是按照前面实施例中的方法制备的）。按照生产商的方法（Ribi 公司，Hamilton，MT）对黑猩猩一共进行了 6 次肌内免疫接种，每次 50 微克 E1，所述 E1 溶解在 PBS/0.05%CHAPS，与 RIBI R-730（MPLA+TDM+CWS）混合。以上 6 次免疫接种是用两个系列每个系列 3 次接种的方式进行的，每个系列中的每次接种之间间隔 3 周时间，而 2 个系列之间间隔 4 周时间。从免疫接种之前 250 天开始，在免疫接种期间一直到免疫接种后 9 个月（参见下文和 WO99/67285）对所述黑猩猩进行连续监测，监测表示 HCV 诱导的疾病的活性的各种参数。所述参数包括血液化学、ALT、AST、γGT、血清中的病毒负荷，肝脏和肝组织中的病毒负荷。另外，在体液和细胞水平监测对所述接种的免疫应答。在此期间，还要监测所述免疫接种对动物产生的任何负面影响，如行为的变化、临床症状、体重、体温和局部反应（发红、肿胀、硬结），不对这些效果进行检测。

很明显，一旦抗 E1 的抗体含量达到最大值 ALT（特别是 γGT 含量，数据未发表）含量就会降低（参见 WO99/67285 的图 11）。一旦所述抗体含量开始下降，ALT 和 γGT 就会反弹，不过在整个随访观察期间 ALT 和 γGT 都保持在较低的水平上。与免疫之前的含量时间相比（85±11U/L），在免疫接种之后 ALT 的含量甚至会显著降低（62±6U/L）。由于在血清中回收到的组织损伤的标记物较少，以上发现是免疫诱导的肝脏疾病改善的第一证据。

在抗 E1 保持在高于 1.0×10^3 的效价期间，E2 抗原水平变得无法检测，不过，在较低的 E1 抗体水平又再次变得可以检测。随着 HCV 抗原的消失，肝脏的炎症也会从中度慢性活性肝炎向较弱形式的慢性持续性肝炎显著降低（WO99/67285 的表 3）。这是所述疫苗能诱导肝脏损伤减轻的另一个主要证据，有可能是通过清除，至少是部分清除其主要靶器官肝脏中的病毒所致。

在整个研究期间，通过 Amplico HCV 监测仪（Roche，Basel，瑞士）测定在血清中的病毒血症水平大体上保持不变。对体液应答所作的更详细的分析发现，所达到的最大终点效价为 30×10^3（第六次免疫接种之后）。并且该效价在免疫接种之后 9 个月降低到 0.5×10^3（WO99/67285 的图 11）。WO99/67285 的图 12 表明由 B 细胞识别的可以由肽模拟的主要表位，该表位位于 N-末端区（肽 V1V2 和 V2V3，有关所使用的肽的细节参见 WO99/67285 的表 4）。由于针对重组 E1 的反应性强而且持续时间长，根据该附图可以推测由该表位所识别的抗体仅仅代表抗 E1 的所有抗体群体的一部分。其余部分很有可能是针对不能由肽模拟的表位的，即不连续的表位。所述表位有可能仅仅出现在完整的 E1 分子上或者甚至只能出现在颗粒样结构上。所述对 E1 的免疫应答是独特的，至少与具有可检测的 E1 的人类慢性 HCV 携带者上正常出现的免疫应答相比比是独特的。在上述患者中，抗 E1 还是部分不连续的，大部分是针对 C4 表位的（占患者血清的 50%），小部分是针对 V1V2（根据基因型在 2%-70%之间波动），而只能偶然记录到针对 V2V3 的反应性（Maertens 等，1997）。由于该黑猩猩是用 1a 分离物感染的，还要评估所述抗体反应对 E1-1a 抗原的交叉反应性。从 WO99/67285 的图 13 中可以看出，确实产生了所述交叉反应抗体，不过，这种抗体只占整个抗体群体的一部分。引人注目的是，在肝脏中

病毒抗原的重新出现和血清中可检测的抗1aE1抗体的消失之间的相关性。

分析T细胞反应性还发现，免疫系统的该区室也是由所述疫苗以特殊方式刺激的，由于所述T细胞的刺激指数从1提高到2.5，并且在随后的随访观察时间内保持在较高的水平上（WO99/67285的图14）。

实施例14：用E1再加强HCV慢性携带者

由于在例12和13中观察到的E1抗体效价是不稳定的，并且随着时间推移而降低，对于1b感染的黑猩猩来说甚至会降低到无法检测的水平，对有关这种抗体应答是否能通过额外的加强免疫再次增强的问题进行了研究。以3周的间隔时间对这两种黑猩猩再进行3次连续的免疫接种（与RIBI佐剂混合的50微克E1）。从WO99/67285的图8和11中可以看出，一旦肝脏中的病毒抗原再次降低到低于检测极限的水平，就可以群鼠加强抗E1应答。不过，在Ton的黑猩猩的血清中病毒负荷保持稳定（WO99/67285的图11）。在随访观察期间，第一次测出了低于每毫升10^5基因组当量的病毒血症水平。

值得注意的是以下发现，正如与第一个系列的免疫接种情形相同，与用1a亚型HCV菌株感染的黑猩猩（Phi1）相比用1b亚型HCV菌株感染的黑猩猩（Phi1）以较低的抗E1效价作出应答。第一轮最大效价为14.5×10^3对Ton黑猩猩的30×10^3，在额外加强免疫之后，Phi1的效价仅为1.2×10^3，而Ton黑猩猩的效价为40×10^3）。尽管对这两种动物的有利影响似乎是相似的，不过，根据该实验可以得出这样的结论：为了获得更高的效价，用来自另一种亚型或基因型的E1蛋白对慢性携带者进行免疫接种可能是特别有利的，并且通过用所述感染性亚型或基因型诱导可以避免在所述宿主中预先存在的和特异性的免疫抑制作用。另外，在同源组（1b疫苗+1b感染）中所观察到的较低的效价可能表明了有大量抗体与病毒结合。因此，所诱导的抗体可能具有中和能力。

实施例15：验证E1免疫接种在黑猩猩中的预防作用

HCV E1s蛋白（192-326号氨基酸）是用重组痘苗病毒HCV11B在Vero细胞中表达的。该痘苗病毒基本上与vvHCV11A相同（披露于

US6150134 中，该专利的所有内容被收作本文参考）。不过，已经从 RK13 传入了 Vero 细胞。大体上按照 PCT/E99/04342（WO99/67285）的例 9 中所披露的方法纯化所述蛋白（通过 lentil 层析，还原烷基化和大小排阻层析），利用碘代乙酰胺作为半胱氨酸的烷基化试剂。在纯化之后，通过 PCT/E99/04342 中所披露的大小排阻层析方法将 3%的 Empigen-BB 换成 3%的甜菜碱，这一方法可以回收到颗粒形式的 E1。最后将该材料脱盐成含有 0.5%的甜菜碱的 PBS，E1 的浓度为 500 微克/毫升。将这种 E1 与等体积的 Alhydrogel 1.3%（Superfos，丹麦）混合，并最后用 8 倍体积的 0.9%的氯化钠进一步稀释，以便获得浓度为 50 微克 E1/毫升和 0.13% Alhydrogel 的明矾辅佐的 E1。

按照所披露的有关表达和纯化 E1 的方法用重组痘苗病毒 HCV101 在 Vero 中表达并从中纯化 HCV E2ΔHVRI（412-715 号氨基酸），该重组痘苗病毒是按照 PCT/E99/04342 的例 8 中所披露的方法由 pvHCV-101 和野生型痘苗病毒重组而产生的。同样，在将 Empigen-BB 换成甜菜碱之后，E2ΔHVRI 的表现就如同一种颗粒（通过动态光反射测定）。

选择在 HCV-RNA 和 HCV-抗体检测中呈阴性的 5 只黑猩猩。其中的 1 只动物（Huub）不进行免疫接种，2 只动物接受以明矾为佐剂的 50 微克的 E1 的 6 次免疫接种（Marti 和 Yoran），而其余的 2 只动物接受以明矾为佐剂的 50 微克 E2ΔHVRI 的免疫接种(Boost 和 Karlien)。所有免疫接种都是通过肌内方式进行的，间隔时间为 3 周。在每一只动物中评估针对用来免疫接种的抗原的体液和细胞免疫应答，并且在每一只动物中检测两种类型的应答，如表 11 所示。

表 11：在第 6 次免疫接种 2 周之后，通过 ELISA 测定抗体效价。将该样品的一系列稀释液与一种室内标准物进行比较（这种室内标准物被定义为具有 1000mU/ml 的 E1 或抗 E2ΔHVRI 抗体，它是来自根据高抗外被效价筛选的 HCV 慢性携带者的三种血清的混合物）。体现细胞免疫应答的刺激指数是通过在有或没有外被抗原的条件下培养 PBMC 获得的，所述 PBMC 来自第 3 次免疫之后 2 周的动物，并测定在培养 5 天之后为期 18 小时的时间内结合到所述细胞中的含氚胸苷的含量。刺激指数是掺入用外被抗原培养的细胞中的胸苷与掺入不用抗原培养的细胞中的胸苷的比例。大于 3 的刺激指数被认为是阳性信号。

	抗E1反应		抗E2ΔHVRI反应	
	抗体效价	刺激指数	抗体效价	刺激指数
Yoran	14110	10.9		
Marti	5630	14.2		
Joost			3210	8.5
Karlien			1770	11.2

在6次免疫接种的最后一次接种之后3周，用100VID（黑猩猩感染剂量）的1b基因型接种物（J4.91，由J.Bukh博士馈赠，NIH，Bethesda，马里兰州）对包括对照在内的所有动物进行刺激。对于E1来说，疫苗蛋白和J4.91分离物（其序列信息可以从保藏号BAA01583获得）为7%（135个氨基酸中的9个）。而对于E2ΔHVRI来说为11%（304个氨基酸中的32个）；因此，该刺激被认为是异源的，并且体现了真实的活的刺激。

在刺激之后第7天所有黑猩猩都变成HCV-RNA阳性（用HCV监测仪测定，Roche，Basel，瑞士），并且在第35和63天之间检测到了第一个ALT和γGT峰。这证明了所有黑猩猩都出现了急性肝炎。很明显，2只E1免疫接种的动物都解决了它们的感染，而于E2ΔHVRI和对照动物没有。这一结论得到了以下事实的验证：E1免疫接种过的动物在第98天（Yoran）和133天（Marti）丧失了HCV-RNA（用HCV监测仪测定，Roche，Basel，瑞士），并且在每个月进行的检测中直到第273天都保持阴性。所有其他的动物在整个273天的随访观察期间都保持RNA阳性，其ALT和γGT值与E1免疫接种的黑猩猩不同没有恢复正常，而是逐渐提高。

总之，我们业已证实了E1免疫接种通过抑制慢性感染的发展改变了HCV感染的天然历史，这种感染是与HCV相关的主要健康问题。

实施例16：在人体内可以诱导的能够清除黑猩猩体内感染的类似的E1应答

为了在人体内获得E1免疫接种的预防作用，需要在人体内诱导与黑猩猩相应的类似免疫应答。因此，我们对20位男性人类志愿者进行了免疫接种，没有检测到抗E1应答（体液或细胞中），使用三剂20

微克 E1 免疫接种，该剂量是以 0.5 毫升的体积在 0.13%的 Alhydrogl 上配制的。所有免疫接种都是以 3 周的间隔通过肌内注射进行的。正如在表 12、17 中所证实的，在 20 位志愿者中确实产生了针对 E1 的显著的体液和细胞免疫应答，这种应答不会没有严重的负面影响。仅有 1 位志愿者（对象 021）被认为是非应答者，因为在 3 次 E1 免疫接种之后，体液或细胞应答都没有超过截段水平。体液应答低于黑猩猩的结果与这一事实相关：仅以 20 微克的用量进行了 3 次免疫接种，而不是使用 50 微克的用量进行 6 次免疫接种。

表 12：在第 3 次免疫接种 2 周之后，通过 ELISA 测定抗体效价。将该样品的一系列稀释液与一种室内标准物进行比较（这种室内标准物被定义为具有 1000mU/ml 的 E1 或抗 E2ΔHVRI 抗体，它是来自根据高抗外被效价筛选的 HCV 慢性携带者的三种血清的混合物）。体现细胞免疫应答的刺激指数是通过在有或没有外被抗原的条件下培养 PBMC 获得的，所述 PBMC 来自第 3 次免疫之后 2 周的动物，并测定在培养 5 天之后为期 18 小时的时间内结合到所述细胞中的含氚胸苷的含量。刺激指数是掺入用外被抗原培养的细胞中的胸苷与掺入不用抗原培养的细胞中的胸苷的比例。大于 3 的刺激指数被认为是阳性信号。

受试者编号	抗体效价	刺激指数
002	1370	30.9
003	717	13.2
004	800	9.1
007	680	3.8
008	1026	3.9
009	325	4.6
010	898	7.7
011	284	4.1
012	181	3.6
013	<20	3.5
014	49	4.6
015	228	3.8
016	324	4.1
017	<20*	6.2
018	<20	6.7
019	624	3.1
020	84	5.5
021	<20	2.1
022	226	2.7
023	163	7.6

* 该个体在免疫接种后被认为是抗E1阳性，因为在免疫之前的样品和3次免疫接种之后的样品之间出现了ELISA信号的显著增强，不过，效价非常低，无法进行精确测定。

实施例 17：在免疫过的健康志愿者体内加强 E1 应答

在第 26 周(即在第 3 次免疫接种之后 20 周),用在 0.13%Alhydrogl 上配制的体积为 0.5 毫升的 20 微克 E1 对例 16 中的 20 位人类志愿者中的 19 位再进行 1 次加强接种。在该额外的免疫接种之后 2 周，再次测定抗体效价和细胞免疫应答。在 20 周的时间内，所有个体的抗体效价都降低了，不过，通过所述额外的免疫接种可以增强到等于或高于在第 8 周所出现的水平。与第 8 周的效价相比，在该加强免疫接种之后抗体的效价平均增加了 1 倍,而与第 26 周的效价相比增加了 7 倍(表 13）。

表 13：在第 3 次免疫接种之后 2 周（=8 周）和 20 周（26 周）以及在加强免疫接种之后 2 周（=28 周）通过 ELISA 测定抗体效价。将该样品的一系列稀释液与一种室内标准物进行比较（这种室内标准物被定义为具有 1000mU/ml 的 E1 或抗 E2ΔHVRI 抗体，它是来自根据高

抗外被效价筛选的 HCV 慢性携带者的三种血清的混合物）。为了精确比较用第 26 周和 28 周的样品重复进行在第 8 周进行的效价测定，它解释了在例 16 的表 12 中的差别。

受试者编号	抗体效价		
	第8周	第26周	第28周
002	1471	443	3119
003	963	95	2355
004	1006	409	2043
007	630	65	541
008	926	81	819
009	704	77	269
010	1296	657	3773
011	253	65	368
012	254	148	760
013	36	<20	166
014	53	40	123
016	159	45	231
017	109	39	568
018	43	23	50
019	425	157	1894
020	73	33	113
021	25	<20	26
022	280	150	357
024	177	81	184
平均值	467	138	936

很明显，对于大部分个体的 T 细胞应答在经过 20 周的时间之后仍然很高。考虑到对破伤风反应的正常化（它作为以前免疫的接种出现在大多数个体提内），所述刺激指数的几何平均值没有变化。在经过所述额外的加强免疫之后，考虑到所述破伤风反应的正常化，没有发现变化(图 51)。这证实了在 3 次 E1 免疫接种之后诱导了 T 辅助应答，并且表明这 3 次免疫接种已经诱导了非常好的 T 辅助记忆，它在至少 6 个月的时间内不需要进一步的加强免疫。

图 51 的说明：刺激指数（细胞免疫应答）是通过以下方法获得的：在有或没有 3 微克重组 E1 或 2 微克破伤风类毒素的条件下培养来自免疫接种之前（0 周），第 3 次免疫接种之后 2 周（8 周），加强免疫接种之前（26 周）以及加强免疫接种之后 2 周（28 周）的个体的 PBMC

（10^5细胞），并且在培养 5 天之后，测定在为期 18 小时的时间内结合到所述细胞中的含氚的胸苷的量。刺激指数是结合到用外被抗原培养的细胞中的胸苷与结合到不用抗原培养的细胞中的胸苷的比例。在第 1 个测定（A）中，测定 0 周和 8 周的样品，而在第 2 个测定（B）中，测定 26 周和 28 周的样品，其中，对 0 周的样品进行再次分析。结果以所有 20 位（A，实验）或 19 位（B，实验）志愿者的刺激指数的几何平均值形式表示。

除了 Th1 细胞因子干扰素γ之外，还在第 26 周和 28 周采集的并用 E1 再次刺激过的样品的 PBMC 培养物的上清液中测定 Th2 细胞因子白介素-5。从图 52 中可以看出，由 E1 刺激的 PBMC 分泌的主要细胞因子是干扰素γ。另人吃惊的是，用明矾辅佐的 E1 观察到了偏向 Th1 的强的应答，因为已知明矾是 Th2 诱导物。这一结果再次证实了诱导了良好的 T 细胞记忆应答，因为在最后的加强免疫之前（26 周）就观察到了非常强的应答。在另一个实验中我们发现所述干扰素γ分泌是特异性的，使用在第 0 周采集的样品我们没有发现在这些志愿者的 E1 刺激的细胞培养物和未刺激过的细胞培养物之间在干扰素γ的分泌方面有差别。

图 52 的说明：有 3 微克重组 E1（E1）或 2 微克破伤风类毒素（TT）或无抗原（BI）的条件下培养来自加强免疫接种之前（26 周）和加强免疫接种之后 2 周（28 周）的个体的 PBMC（10^5细胞）。通过 ELISA 测定取自 24 小时之后（白介素-5）或 120 小时之后（干扰素γ）采集的上清液中的细胞因子。刺激指数是在用外被抗原培养的细胞上清液中测定的细胞因子与在不用抗原培养的细胞中测定的细胞因子的比例。结果以所有 19 位志愿者分泌的 pg 细胞因子/毫升的几何平均值形式表示。细胞因子量低于检测极限的样品被确定为检测极限值。类似的，具有超出该测定的线性范围的极高浓度的细胞因子的样品被确定为该测定的线性范围的上限。

实施例 18：免疫接种过的健康志愿者对 E1 的细胞应答的精细作图

为了对 E1 特异性反应进行作图，用标准 Fmoc 化合方法合成了一系列二十聚体肽，具有 8 个氨基酸的重叠，并且覆盖整个 E1 序列。对所有的肽进行 C-末端氨基化和 N-末端乙酰化，除 IGP1626 以外，它具

有游离的氨基末端。

IGP1626 YEVRNVSGIYHVTNDCSNSS（192-211 号氨基酸）

IGP1627 TNDCSNSSIVYEAADMIMHT（204-223 号氨基酸）

IGP1628 AADMIMHTPGCVPCVRENNS（216-235 号氨基酸）

IGP1629 PCVRENNSSRCWVALTPTLA（228-247 号氨基酸）

IGP1630 VALTPTLAARNASVPTTTIR（240-259 号氨基酸）

IGP1631 SVPTTTIRRHVDLLVGAAAF（252-271 号氨基酸）

IGP1632 LLVGAAAFCSAMYVGDLCGS（264-283 号氨基酸）

IGP1633 YVGDLCGSVFLVSQLFTISP（276-295 号氨基酸）

IGP1634 SQLFTISPRRHETVQDCNCS（288-307 号氨基酸）

IGP1635 TVQDCNCSIYPGHITGHRMA（300-319 号氨基酸）

IGP1636 HITGHRMAWDMMMNWSPTTA（312-331 号氨基酸）

在分别有 25 微克/毫升（未免疫接种的人体）或 10 微克/毫升（免疫接种过的人体），在第 3 次或加强注射之后采集的样品）肽的情况下培养来自 14 个没有用 E1 免疫接种的不同的健康供体或 10 个用 E1 免疫接种的 10 个供体的 PBMC。从图 53 中可以看出，与未免疫接种过的人体相比，在免疫接种过的人体内肽 IGP1627、1629、1630、1631、1633、1635 和 1635 都能诱导明显较强的反应。使用 3 的刺激指数作为肽 IGP1627、1629、1631 和 1635 的截断值，1635 是最经常被识别的（至少能被测试的半数免疫接种过的人体所识别）。该实验证实，由来自哺乳动物细胞培养物的 E1 诱导的 T 细胞应答是对 E1 特异性的，因为这种应答不能被来自哺乳动物细胞培养物的相同的 E1 再次诱导，不过还可以通过合成肽诱导。另外，该实验证实了 E1 中的大部分免疫原性 T 细胞结构域位于 204-223、228-271、276-295、300-311 号氨基酸之间，尤其是位于 204-223、228-247、252-271 和 300-319 号氨基酸之间。

图 53 的说明：刺激指数（细胞免疫应答）是通过以下方法获得的：在有或没有肽的条件下培养 PBMC（3×10^5 细胞），并测定在培养 5-6 天之后在一段时间内结合到所述细胞中的含氚的胸苷的量。刺激指数是结合到用肽培养的细胞中的胸苷与结合到不用肽培养的细胞中的含氚胸苷的比例。结果以接种过的人体（上图）或未接种过的或对照（下图）的单个值的形式表示。

因此，本发明还提供了以下E1肽、蛋白、含有它们的组合物和试剂盒、编码所述肽和蛋白的核酸序列及其生产方法和用途。正如在本文中针对本发明的其他E1和相关肽所作的一般性说明。

跨越E1区的192-211号位置的IGP1626（SEQ ID NO: 112），
跨越E1区的204-223号位置的IGP1627（SEQ ID NO: 113），
跨越E1区的192-211号位置的IGP1626（SEQ ID NO: 112），
跨越E1区的204-223号位置的IGP1627（SEQ ID NO: 113），
跨越E1区的216-235号位置的IGP1628（SEQ ID NO: 114），
跨越E1区的228-247号位置的IGP1629（SEQ ID NO: 115），
跨越E1区的240-259号位置的IGP1630（SEQ ID NO: 116），
跨越E1区的252-271号位置的IGP1631（SEQ ID NO: 117），
跨越E1区的264-283号位置的IGP1632（SEQ ID NO: 118），
跨越E1区的276-295号位置的IGP1633（SEQ ID NO: 119），
跨越E1区的288-307号位置的IGP1634（SEQ ID NO: 120），
跨越E1区的300-319号位置的IGP1635（SEQ ID NO: 121），
跨越E1区的312-331号位置的IGP1636（SEQ ID NO: 122）。

参考文献:

Bailey, J. and Cole, R. (1959) J. Biol. Chem. 234, 1733-1739.

Ballou, L., Hitzeman, R., Lewis, M. & Ballou, C. (1991) PNAS 88, 3209-3212.

Benesch, R., Benesch, R.E., Gutcho, M. & Lanfer, L. (1956) Science 123, 981.

Cavins, J. & Friedman. (1970) Anal. Biochem. 35, 489.

Cleland, W. (1964) Biochemistry 3, 480

Creighton , E. (1988) BioEssays 8, 57

Darbre, A., John Wiley & Sons Ltd. (1987) Practical Protein Chemistry - A Handbook.

Darbre, A., John Wiley & Sons Ltd. (1987) Practical Proteinchemistry p.69-79.

Doms et al, (1993), Virology 193, 545-562.

Ellman, G. (1959) Arch. Biochem. Biohys. 82, 70.

Falkner, F. & Moss, B. (1988) J. Virol. 62, 1849-1854.

Friedman, M. & Krull. (1969) Biochem. Biophys. Res. Commun. 37, 630.

Gallagher J. (1988) J. Cell Biol. 107, 2059-2073.

Glazer, A., Delange, R., Sigman, D. (1975) North Holland publishing company, Elsevier, Biomedical. Part : Modification of protein (p. 116).

Graham, F. & van der Eb, A. (1973) Virology 52, 456-467.

Grakoui et al. (1993) Journal of Virology 67:1385-1395.

5

Grassetti, D. & Murray, J. (1969) Analyt. Chim. Acta. 46, 139.

Grassetti, D. & Murray, J. (1967) Arch. Biochem Biophys. 119, 41.

Helenius, Mol. Biol. Cell (1994), 5: 253-265.

Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. (1991) Proc. Natl. Acad. Sci. U.S.A. 88(13):5547-51.

Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., Stüber, D. (1988) Biochemistry 88, 8976.

Hsu, H., Donets, M., Greenberg, H. & Feinstone, S. (1993) Hepatology 17:763-771.

Inoue, Y., Suzuki, R., Matsuura, Y., Harada, S., Chiba, J., Watanabe, Y., Saito, I. & Miyamura, T. (1992) J. Gen. Virol. 73:2151-2154.

Janknecht, R., de Martynoff, G. *et al.*, (1991) Proc. Natl. Acad. Sci. USA 88, 8972-8976.

Kayman (1991) J. Virology 65, 5323-5332.

Kato, N., Oostuyama, Y., Tanaka, T., Nakagawa, M., Muraiso, K., Ohkoshi, S., Hijikata, M., Shimitohno, K. (1992) Virus Res. 22:107-123.

Kniskern, P., Hagopian, A., Burke, P., Schultz, L., Montgomery, D., Hurni, W., Yu Ip, C., Schulman, C., Malgetter, R., Wampler, D., Kubek, D., Sitrin, R., West, D., Ellis, R., Miller, W. (1994) Vaccine 12:1021-1025.

Kohara, M., Tsukiyama-Kohara, K., Maki, N., Asano, K., Yoshizawa, K., Miki, K., Tanaka, S., Hattori, N., Matsuura, Y., Saito, I., Miyamura, T. & Nomoto, A. (1992) J. Gen. Virol. 73:2313-2318.

Mackett, M., Smith, G. & Moss, B. (1985) In: 'DNA cloning: a practical approach' (Ed. Glover, D.) IRL Press, Oxford.

Mackett, M., & Smith, G. (1986) J. Gen. Virol. 67, 2067-2082.

Mackett, M., Smith, G. & Moss, B. (1984) J. Virol. 49, 857-864.

Mackett, M., Smith, G. & Moss, B. (1984) Proc. Natl. Acad. Sci. USA 79, 7415-7419.

Means, G. (1971) Holden Day, Inc.

Means, G. & Feeney, R. (1971) Holden Day p.105 & p. 217.

Mita, E., Hayashi, N., Ueda, K., Kasahara, A., Fusamoto, H., Takamizawa, A., Matsubara, K., Okayama, H. & Kamada T. (1992) Biochem. Biophys. Res. Comm. 183:925-930.

Moore, S. (1963) J. Biol. Chem. 238, 235-237.

Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T., Yoshizawa, H., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1990) Jpn. J. Exp. Med. 60:167-177.

Panicali & Paoletti (1982) Proc. Natl. Acad. Sci. USA 79, 4927-4931.

Piccini, A., Perkus, M. & Paoletti, E. (1987) Meth. Enzymol. 153, 545-563.

Rose (1988) Annu. Rev. Cell Biol. 1988, 4: 257-288;

Ruegg, V. and Rudinger, J. (1977) Methods Enzymol. 47, 111-116.

Shan, S. & Wong (1993) CRC-press p. 30-33.

Spaete, R., Alexander, D., Rugroden, M., Choo, Q., Berger, K., Crawford, K., Kuo, C., Leng, S., Lee, C., Ralston, R., et al. (1992) Virology 188(2):819-30.

Skehel, J., (1984) Proc. Natl. Acad. Sci. USA 81, 1179-1783.

Stunnenberg, H., Lange, H., Philipson, L., Miltenburg, R. & van der Vliet, R. (1988) Nucl. Acids Res. 16, 2431-2444.

Stuyver, L., Van Arnhem, W., Wyseur, A., DeLeys, R. & Maertens, G. (1993a) Biochem. Biophys. Res. Commun. 192, 635-641.

Stuyver, L., Rossau, R., Wyseur, A., Duhamel, M., Vanderborght, B., Van Heuverswyn, H., & Maertens, G. (1993b) J. Gen. Virol. 74, 1093-1102.

Stuyver, L., Van Arnhem, W., Wyseur, A., Hernandez, F., Delaporte, E., Maertens, G. (1994), Proc. Natl. Acad. Sci. USA 91:10134-10138.

Weil, L. & Selbler, S. (1961) Arch. Biochem. Biophys. 95, 470.

Yokosuka, O., Ito, Y., Imazeki, F., Ohto, M. & Omata, M. (1992) Biochem. Biophys. Res. Commun. 189:565-571.

Miller P, Yano J, Yano E, Carroll C, Jayaram K, Ts'o P (1979) Biochemistry 18:5134-43.

Nielsen P, Egholm M, Berg R, Buchardt O (1991) Science 254:1497-500.

Nielsen P, Egholm M, Berg R, Buchardt O (1993) Nucleic-Acids-Res. 21:197-200.

Asseline U, Delarue M, Lancelot G, Toulme F, Thuong N (1984) Proc. Natl. Acad. Sci. USA 81:3297-301.

Matsukura M, Shinozuka K, Zon G, Mitsuya H, Reitz M, Cohen J, Broder S (1987) Proc. Natl. Acad. Sci. USA 84:7706-10.

WO 96/04385 (PCT/EP95/03031) - Purified Hepatitis C Virus Envelope Proteins for Diagnostic and Therapeutic Use.

All references cited herein are incorporated in their entirety by reference.

<110> 基因创新有限公司

<120> 用于诊断和治疗用途的纯化的丙型肝炎病毒外被蛋白

<140> PCT/EP02/00219
<160> 122

<170> PatentIn 3.1

<210> 1

<211> 21
<212> DNA
<213> 丙型肝炎病毒

<400> 1

ggcatgcaag cttaattaat t 21

<210> 2

<211> 68
<212> DNA
<213> 丙型肝炎病毒

<400> 2

ccggggaggc ctgcacgtga tcgagggcag acaccatcac caccatcact aatagttaat 60

taactgca 68

<210> 3

<211> 642
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..639

<220>
<221> mat_肽
<222> 1..636

<400> 3

```
atg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct tta ctg tcc tgt        48
Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
  1               5                  10                  15

ctg acc att cca gct tcc gct tat gag gtg cgc aac gtg tcc ggg atg        96
Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Met
             20                  25                  30

tac cat gtc acg aac gac tgc tcc aac tca agc att gtg tat gag gca       144
Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
         35                  40                  45

gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc gtt cgg gag       192
Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
     50                  55                  60

aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg ctc gca gct       240
Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
 65                  70                  75                  80

agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac gtc gat ttg       288
Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu
                 85                  90                  95

ctc gtt ggg gcg gct gct ctc tgt tcc gct atg tac gtg ggg gat ctc       336
Leu Val Gly Ala Ala Ala Leu Cys Ser Ala Met Tyr Val Gly Asp Leu
            100                 105                 110

tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc tcg cct cgc       384
Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile Ser Pro Arg
        115                 120                 125

cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat ccc ggc cac       432
Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
    130                 135                 140
```

```
ata aca ggt cac cgt atg gct tgg gat atg atg atg aac tgg tcg cct      480
Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro
145                 150                 155                 160

aca acg gcc ctg gtg gta tcg cag ctg ctc cgg atc cca caa gct gtc      528
Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val
                165                 170                 175

gtg gac atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg ggc ctc gcc      576
Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala
            180                 185                 190

tac tat tcc atg gtg ggg aac tgg gct aag gtt ttg att gtg atg cta      624
Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Ile Val Met Leu
        195                 200                 205

ctc ttt gct ctc taatag                                              642
Leu Phe Ala Leu
    210
```

<210> 4

<211> 212
<212> PRT
<213> 丙型肝炎病毒

<400> 4

```
Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
  1               5                  10                  15

Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Met
             20                  25                  30

Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
         35                  40                  45

Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
     50                  55                  60

Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
 65                  70                  75                  80
```

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

Leu Val Gly Ala Ala Ala Leu Cys Ser Ala Met Tyr Val Gly Asp Leu
100 105 110

Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile Ser Pro Arg
115 120 125

Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
130 135 140

Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro
145 150 155 160

Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val
165 170 175

Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala
180 185 190

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

Leu Phe Ala Leu
210

<210> 5

<211> 795
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..792

<220>
<221> mat_肽
<222> 1..789
<400> 5

```
atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc        48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
 1               5                  10                  15

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg        96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
            20                  25                  30

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca       144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
        35                  40                  45

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg       192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
    50                  55                  60

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg       240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65                  70                  75                  80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg       288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                85                  90                  95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc       336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg       384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac       432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg       480
Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145                 150                 155                 160

ggg gac ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc       528
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
                165                 170                 175

tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat       576
```

```
Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
            180                 185                 190

ccc ggc cac ata acg ggt cac cgt atg gct tgg gat atg atg atg aac      624
Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
        195                 200                 205

tgg tcg cct aca acg gcc ctg gtg gta tcg cag ctg ctc cgg atc cca      672
Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro
    210                 215                 220

caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg      720
Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala
225                 230                 235                 240

ggt ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag gtt ttg att      768
Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Ile
                245                 250                 255

gtg atg cta ctc ttt gct ccc taatag                                  795
Val Met Leu Leu Phe Ala Pro
            260
<210> 6

<211> 263
<212> PRT
<213> 丙型肝炎病毒


<400> 6

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
  1               5                  10                  15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
             20                  25                  30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
         35                  40                  45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
     50                  55                  60

Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
```

65 70 75 80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
100 105 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
115 120 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
130 135 140

Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145 150 155 160

Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
165 170 175

Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
180 185 190

Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
195 200 205

Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro
210 215 220

Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala
225 230 235 240

Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Ile
245 250 255

Val Met Leu Leu Phe Ala Pro
260

<210> 7

<211> 633
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..630

<220>
<221> mat_肽
<222> 1..627
<400> 7

```
atg ttg ggt aag gtc atc gat acc ctt acg tgc ggc ttc gcc gac ctc        48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
  1               5                  10                  15

atg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggt gct gcc aga        96
Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
             20                  25                  30

gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac ggc gtg aac tat gca       144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
         35                  40                  45

aca ggg aat ttg cct ggt tgc tct ttc tct atc ttc ctc ttg gct tta       192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
     50                  55                  60

ctg tcc tgt ctg acc att cca gct tcc gct tat gag gtg cgc aac gtg       240
Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
 65                  70                  75                  80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg       288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                 85                  90                  95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc       336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg       384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gca gct agg aac gcc agc gtc ccc act acg aca ata cga cgc cac       432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
```

130 135 140

gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg 480
Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145 150 155 160

ggg gat ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc 528
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
165 170 175

tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat 576
Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
180 185 190

ccc ggc cac ata aca ggt cac cgt atg gct tgg gat atg atg atg aac 624
Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
195 200 205

tgg taatag 633
Trp
210

<210> 8

<211> 209
<212> PRT
<213> 丙型肝炎病毒

<400> 8

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

```
Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
 65                  70                  75                  80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                85                  90                  95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145                 150                 155                 160

Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
                165                 170                 175

Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
            180                 185                 190

Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
        195                 200                 205

Trp
```

<210> 9

<211> 483
<212> DNA
<213> 丙型肝炎病毒
<220>
<221> CDS
<222> 1..480

<220>
<221> mat_肽
<222> 1..477
<400> 9

```
atg ccc ggt tgc tct ttc tct atc ttc ctc ttg gcc ctg ctg tcc tgt        48
Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
 1               5                  10                  15

ctg acc ata cca gct tcc gct tat gaa gtg cgc aac gtg tcc ggg gtg        96
Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val
            20                  25                  30

tac cat gtc acg aac gac tgc tcc aac tca agc ata gtg tat gag gca       144
Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
        35                  40                  45

gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc gtt cgg gag       192
Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
    50                  55                  60

ggc aac tcc tcc cgt tgc tgg gtg gcg ctc act ccc acg ctc gcg gcc       240
Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
65                  70                  75                  80

agg aac gcc agc gtc ccc aca acg aca ata cga cgc cac gtc gat ttg       288
Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu
                85                  90                  95

ctc gtt ggg gct gct gct ttc tgt tcc gct atg tac gtg ggg gat ctc       336
Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu
            100                 105                 110

tgc gga tct gtt ttc ctt gtt tcc cag ctg ttc acc ttc tca cct cgc       384
Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg
        115                 120                 125

cgg cat caa aca gta cag gac tgc aac tgc tca atc tat ccc ggc cat       432
Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
    130                 135                 140

gta tca ggt cac cgc atg gct tgg gat atg atg atg aac tgg tcc taatag     483
Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser
145                 150                 155                 160
```

<210> 10

<211> 159
<212> PRT
<213> 丙型肝炎病毒

<400> 10

Met Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
1 5 10 15

Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val
20 25 30

Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
35 40 45

Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
50 55 60

Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
65 70 75 80

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

Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu
100 105 110

Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg
115 120 125

Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
130 135 140

Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser
145 150 155
<210> 11

<211> 480
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..477

<220>
<221> mat_肽
<222> 1..474
<400> 11

```
atg tcc ggt tgc tct ttc tct atc ttc ctc ttg gcc ctg ctg tcc tgt        48
Met Ser Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
 1               5                  10                  15

ctg acc ata cca gct tcc gct tat gaa gtg cgc aac gtg tcc ggg gtg        96
Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val
            20                  25                  30

tac cat gtc acg aac gac tgc tcc aac tca agc ata gtg tat gag gca       144
Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
        35                  40                  45

gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc gtt cgg gag       192
Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
    50                  55                  60

ggc aac tcc tcc cgt tgc tgg gtg gcg ctc act ccc acg ctc gcg gcc       240
Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
65                  70                  75                  80

agg aac gcc agc gtc ccc aca acg aca ata cga cgc cac gtc gat ttg       288
Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu
                85                  90                  95

ctc gtt ggg gct gct gct ttc tgt tcc gct atg tac gtg ggg gat ctc       336
Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu
            100                 105                 110

tgc gga tct gtt ttc ctt gtt tcc cag ctg ttc acc ttc tca cct cgc       384
Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg
        115                 120                 125

cgg cat caa aca gta cag gac tgc aac tgc tca atc tat ccc ggc cat       432
Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
    130                 135                 140

gta tca ggt cac cgc atg gct tgg gat atg atg atg aac tgg taatag       480
Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
145                 150
```

145 150 155

<210> 12

<211> 158
<212> PRT
<213> 丙型肝炎病毒

<400> 12

Met Ser Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys
1 5 10 15

Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Val
20 25 30

Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala
35 40 45

Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val Arg Glu
50 55 60

Gly Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala
65 70 75 80

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

Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu
100 105 110

Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg
115 120 125

Arg His Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His
130 135 140

Val Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
145 150 155

<210> 13

<211> 636
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..633

<220>
<221> mat_肽
<222> 1..630

<400> 13

```
atg ctg ggt aag gcc atc gat acc ctt acg tgc ggc ttc gcc gac ctc        48
Met Leu Gly Lys Ala Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
  1               5                  10                  15
gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg        96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
             20                  25                  30

gcc ctg gcg cat ggc gtc cgg gtt ctg gaa gac ggc gtg aac tat gca       144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
         35                  40                  45

aca ggg aat ttg cct ggt tgc tct ttc tct atc ttc ctc ttg gct tta       192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
     50                  55                  60
ctg tcc tgt cta acc att cca gct tcc gct tac gag gtg cgc aac gtg       240
Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
 65                  70                  75                  80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg       288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                 85                  90                  95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc       336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg       384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gcg gct agg aac gcc agc atc ccc act aca aca ata cga cgc cac       432
Leu Ala Ala Arg Asn Ala Ser Ile Pro Thr Thr Thr Ile Arg Arg His
```

130 135 140

gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg 480
Val Asp Leu Leu Val Gly Ala Ala Phe Cys Ser Ala Met Tyr Val
145 150 155 160

ggg gat ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc 528
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
165 170 175

tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat 576
Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
180 185 190

ccc ggc cac ata acg ggt cac cgt atg gct tgg gat atg atg atg aac 624
Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
195 200 205

tgg tac taatag 640
Trp Tyr
210

<210> 14

<211> 210
<212> PRT
<213> 丙型肝炎病毒

<400> 14

Met Leu Gly Lys Ala Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

Leu Ser Cys Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val

65 70 75 80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
100 105 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
115 120 125

Leu Ala Ala Arg Asn Ala Ser Ile Pro Thr Thr Thr Ile Arg Arg His
130 135 140

Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145 150 155 160

Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
165 170 175

Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
180 185 190

Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
195 200 205

Trp Tyr
210

<210> 15

<211> 26
<212> DNA
<213> 丙型肝炎病毒
<400> 15

atgcccggtt gctctttctc tatctt 26

<210> 16

<211> 26
<212> DNA
<213> 丙型肝炎病毒

<400> 16

atgttgggta aggtcatcga taccct 26

<210> 17

<211> 30
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> misc_特征
<223>/注释="反义"

<400> 17

ctattaggac cagttcatca tcatatccca 30

<210> 18

<211> 27
<212> DNA
<213> 丙型肝炎病毒

<400> 18

ctattaccag ttcatcatca tatccca 27

<210> 19

<211> 36
<212> DNA
<213> 丙型肝炎病毒
<400> 19

atacgacgcc acgtcgattc ccagctgttc accatc 36

<210> 20

<211> 36
<212> DNA
<213> 丙型肝炎病毒
<220>
<221> misc_特征
<223>/注释="反义"

<400> 20

gatggtgaac agctgggaat cgacgtggcg tcgtat 36

<210> 21

<211> 723
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..720

<220>
<221> mat_肽
<222> 1..717
<400> 21

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc 48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg 96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca 144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg 192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg 240

```
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
 65                  70                  75                  80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg      288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                85                  90                  95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc      336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg      384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac      432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140
gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg      480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145                 150                 155                 160

cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt      528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
                165                 170                 175

atg gct tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg      576
Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val
            180                 185                 190

gta tcg cag ctg ctc cgg atc cca caa gct gtc gtg gac atg gtg gcg      624
Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala
        195                 200                 205

ggg gcc cat tgg gga gtc ctg gcg ggt ctc gcc tac tat tcc atg gtg      672
Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val
    210                 215                 220

ggg aac tgg gct aag gtt ttg att gtg atg cta ctc ttt gct ccc taatag    723
Gly Asn Trp Ala Lys Val Leu Ile Val Met Leu Leu Phe Ala Pro
225                 230                 235                 240
<210> 22

<211> 239
```

<212> PRT
<213> 丙型肝炎病毒

<400> 22

```
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
 1               5                  10                  15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
            20                  25                  30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
        35                  40                  45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
    50                  55                  60

Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65                  70                  75                  80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                85                  90                  95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145                 150                 155                 160

Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
                165                 170                 175

Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val
            180                 185                 190

Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala
        195                 200                 205
```

Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val
210 215 220

Gly Asn Trp Ala Lys Val Leu Ile Val Met Leu Leu Phe Ala Pro
225 230 235

<210> 23

<211> 561
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..558

<220>
<221> mat_肽
<222> 1..555
<400> 23

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc 48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg 96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca 144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg 192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg 240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65 70 75 80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg 288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

```
tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc        336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg        384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac        432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg        480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145                 150                 155                 160

cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt        528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
                165                 170                 175

atg gct tgg gat atg atg atg aac tgg taatag                            561
Met Ala Trp Asp Met Met Met Asn Trp
            180                 185
```

<210> 24

<211> 185
<212> PRT
<213> 丙型肝炎病毒

<400> 24

```
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
  1               5                  10                  15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
             20                  25                  30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
         35                  40                  45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
     50                  55                  60
```

Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65 70 75 80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
100 105 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
115 120 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
130 135 140

Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145 150 155 160

Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
165 170 175

Met Ala Trp Asp Met Met Met Asn Trp
180 185

<210> 25

<211> 606
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..603

<220>
<221> mat_肽
<222> 1..600
<400> 25

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc 48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg 96

```
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
            20                  25                  30

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca      144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
        35                  40                  45

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg      192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
    50                  55                  60

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg      240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65                  70                  75                  80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg      288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                85                  90                  95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc      336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg      384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac      432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg      480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145                 150                 155                 160

cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt      528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
                165                 170                 175

atg gct tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg      576
Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val
            180                 185                 190

gta tcg cag ctg ctc cgg atc ctc taatag                              606
```

Val Ser Gln Leu Leu Arg Ile Leu
195 200

<210> 26

<211> 200
<212> PRT
<213> 丙型肝炎病毒

<400> 26

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65 70 75 80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
100 105 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
115 120 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
130 135 140

Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145 150 155 160

Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
165 170 175

Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val
180 185 190

Val Ser Gln Leu Leu Arg Ile Leu
195 200

<210> 27

<211> 636
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..633

<220>
<221> mat_肽
<222> 1..630
<400> 27

atg ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc 48
Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg 96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca 144
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg 192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg 240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65 70 75 80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg 288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc 336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
100 105 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg 384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
115 120 125

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac 432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
130 135 140

gtc gat tcc cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg 480
Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145 150 155 160

cag gac tgc aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt 528
Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
165 170 175

atg gct tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg 576
Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val
180 185 190

gta tcg cag ctg ctc cgg atc gtg atc gag ggc aga cac cat cac cac 624
Val Ser Gln Leu Leu Arg Ile Val Ile Glu Gly Arg His His His His
195 200 205

cat cac taatag 636
His His
210

<210> 28

<211> 210
<212> PRT
<213> 丙型肝炎病毒

<400> 28

Met Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg

20 25 30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
35 40 45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
50 55 60

Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65 70 75 80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
85 90 95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
100 105 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
115 120 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
130 135 140

Val Asp Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val
145 150 155 160

Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg
165 170 175

Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val
180 185 190

Val Ser Gln Leu Leu Arg Ile Val Ile Glu Gly Arg His His His His
195 200 205

His His
210

<210> 29

<211> 630
<212> DNA
<213> 丙型肝炎病毒
<220>
<221> CDS

<222> 1..627

<220>
<221> mat_肽
<222> 1..624
<400> 29

```
atg ggt aag gtc atc gat acc ctt acg tgc gga ttc gcc gat ctc atg        48
Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met
 1               5                  10                  15

ggg tac atc ccg ctc gtc ggc gct ccc gta gga ggc gtc gca aga gcc        96
Gly Tyr Ile Pro Leu Val Gly Ala Pro Val Gly Gly Val Ala Arg Ala
            20                  25                  30

ctt gcg cat ggc gtg agg gcc ctt gaa gac ggg ata aat ttc gca aca       144
Leu Ala His Gly Val Arg Ala Leu Glu Asp Gly Ile Asn Phe Ala Thr
        35                  40                  45

ggg aat ttg ccc ggt tgc tcc ttt tct att ttc ctt ctc gct ctg ttc       192
Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Phe
    50                  55                  60

tct tgc tta att cat cca gca gct agt cta gag tgg cgg aat acg tct       240
Ser Cys Leu Ile His Pro Ala Ala Ser Leu Glu Trp Arg Asn Thr Ser
65                  70                  75                  80

ggc ctc tat gtc ctt acc aac gac tgt tcc aat agc agt att gtg tac       288
Gly Leu Tyr Val Leu Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr
                85                  90                  95

gag gcc gat gac gtt att ctg cac aca ccc ggc tgc ata cct tgt gtc       336
Glu Ala Asp Asp Val Ile Leu His Thr Pro Gly Cys Ile Pro Cys Val
            100                 105                 110

cag gac ggc aat aca tcc acg tgc tgg acc cca gtg aca cct aca gtg       384
Gln Asp Gly Asn Thr Ser Thr Cys Trp Thr Pro Val Thr Pro Thr Val
        115                 120                 125

gca gtc aag tac gtc gga gca acc acc gct tcg ata cgc agt cat gtg       432
Ala Val Lys Tyr Val Gly Ala Thr Thr Ala Ser Ile Arg Ser His Val
    130                 135                 140

gac cta tta gtg ggc gcg gcc acg atg tgc tct gcg ctc tac gtg ggt       480
```

Asp Leu Leu Val Gly Ala Ala Thr Met Cys Ser Ala Leu Tyr Val Gly
145 150 155 160

gac atg tgt ggg gct gtc ttc ctc gtg gga caa gcc ttc acg ttc aga 528
Asp Met Cys Gly Ala Val Phe Leu Val Gly Gln Ala Phe Thr Phe Arg
165 170 175

cct cgt cgc cat caa acg gtc cag acc tgt aac tgc tcg ctg tac cca 576
Pro Arg Arg His Gln Thr Val Gln Thr Cys Asn Cys Ser Leu Tyr Pro
180 185 190

ggc cat ctt tca gga cat cga atg gct tgg gat atg atg atg aac tgg 624
Gly His Leu Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
195 200 205

taatag 634

<210> 30

<211> 208
<212> PRT
<213> 丙型肝炎病毒

<400> 30

Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met
1 5 10 15

Gly Tyr Ile Pro Leu Val Gly Ala Pro Val Gly Gly Val Ala Arg Ala
20 25 30

Leu Ala His Gly Val Arg Ala Leu Glu Asp Gly Ile Asn Phe Ala Thr
35 40 45

Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Phe
50 55 60

Ser Cys Leu Ile His Pro Ala Ala Ser Leu Glu Trp Arg Asn Thr Ser
65 70 75 80

Gly Leu Tyr Val Leu Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr
85 90 95

Glu Ala Asp Asp Val Ile Leu His Thr Pro Gly Cys Ile Pro Cys Val

100 105 110

Gln Asp Gly Asn Thr Ser Thr Cys Trp Thr Pro Val Thr Pro Thr Val
115 120 125

Ala Val Lys Tyr Val Gly Ala Thr Thr Ala Ser Ile Arg Ser His Val
130 135 140

Asp Leu Leu Val Gly Ala Ala Thr Met Cys Ser Ala Leu Tyr Val Gly
145 150 155 160

Asp Met Cys Gly Ala Val Phe Leu Val Gly Gln Ala Phe Thr Phe Arg
165 170 175

Pro Arg Arg His Gln Thr Val Gln Thr Cys Asn Cys Ser Leu Tyr Pro
180 185 190

Gly His Leu Ser Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
195 200 205

<210> 31

<211> 630
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..627

<220>
<221> mat_肽
<222> 1..624
<400> 31

atg ggt aag gtc atc gat acc cta acg tgc gga ttc gcc gat ctc atg 48
Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met
1 5 10 15

ggg tat atc ccg ctc gta ggc ggc ccc att ggg ggc gtc gca agg gct 96
Gly Tyr Ile Pro Leu Val Gly Gly Pro Ile Gly Gly Val Ala Arg Ala
20 25 30

ctc gca cac ggt gtg agg gtc ctt gag gac ggg gta aac tat gca aca 144
Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala Thr

35 40 45

ggg aat tta ccc ggt tgc tct ttc tct atc ttt att ctt gct ctt ctc 192
Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Ile Leu Ala Leu Leu
50 55 60

tcg tgt ctg acc gtt ccg gcc tct gca gtt ccc tac cga aat gcc tct 240
Ser Cys Leu Thr Val Pro Ala Ser Ala Val Pro Tyr Arg Asn Ala Ser
65 70 75 80

ggg att tat cat gtt acc aat gat tgc cca aac tct tcc ata gtc tat 288
Gly Ile Tyr His Val Thr Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr
85 90 95

gag gca gat aac ctg atc cta cac gca cct ggt tgc gtg cct tgt gtc 336
Glu Ala Asp Asn Leu Ile Leu His Ala Pro Gly Cys Val Pro Cys Val
100 105 110

atg aca ggt aat gtg agt aga tgc tgg gtc caa att acc cct aca ctg 384
Met Thr Gly Asn Val Ser Arg Cys Trp Val Gln Ile Thr Pro Thr Leu
115 120 125

tca gcc ccg agc ctc gga gca gtc acg gct cct ctt cgg aga gcc gtt 432
Ser Ala Pro Ser Leu Gly Ala Val Thr Ala Pro Leu Arg Arg Ala Val
130 135 140

gac tac cta gcg gga ggg gct gcc ctc tgc tcc gcg tta tac gta gga 480
Asp Tyr Leu Ala Gly Gly Ala Ala Leu Cys Ser Ala Leu Tyr Val Gly
145 150 155 160

gac gcg tgt ggg gca cta ttc ttg gta ggc caa atg ttc acc tat agg 528
Asp Ala Cys Gly Ala Leu Phe Leu Val Gly Gln Met Phe Thr Tyr Arg
165 170 175

cct cgc cag cac gct acg gtg cag aac tgc aac tgt tcc att tac agt 576
Pro Arg Gln His Ala Thr Val Gln Asn Cys Asn Cys Ser Ile Tyr Ser
180 185 190

ggc cat gtt acc ggc cac cgg atg gca tgg gat atg atg atg aac tgg 624
Gly His Val Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
195 200 205

taatag 630

<210> 32
<211> 208
<212> PRT
<213> 丙型肝炎病毒

<400> 32

Met Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met
1 5 10 15

Gly Tyr Ile Pro Leu Val Gly Gly Pro Ile Gly Gly Val Ala Arg Ala
20 25 30

Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala Thr
35 40 45

Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Ile Leu Ala Leu Leu
50 55 60

Ser Cys Leu Thr Val Pro Ala Ser Ala Val Pro Tyr Arg Asn Ala Ser
65 70 75 80

Gly Ile Tyr His Val Thr Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr
85 90 95

Glu Ala Asp Asn Leu Ile Leu His Ala Pro Gly Cys Val Pro Cys Val
100 105 110

Met Thr Gly Asn Val Ser Arg Cys Trp Val Gln Ile Thr Pro Thr Leu
115 120 125

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

Asp Tyr Leu Ala Gly Gly Ala Ala Leu Cys Ser Ala Leu Tyr Val Gly
145 150 155 160

Asp Ala Cys Gly Ala Leu Phe Leu Val Gly Gln Met Phe Thr Tyr Arg
165 170 175

Pro Arg Gln His Ala Thr Val Gln Asn Cys Asn Cys Ser Ile Tyr Ser
180 185 190

Gly His Val Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
195 200 205

<210> 33

<211> 23
<212> DNA
<213> 丙型肝炎病毒
<400> 33
tgggatatga tgatgaactg gtc 23

<210> 34

<211> 30
<212> DNA
<213> 丙型肝炎病毒
<400> 34

ctattatggt ggtaagccac agagcaggag 30

<210> 35

<211> 1476
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..1473

<220>
<221> mat_肽
<222> 1..1470
<400> 35

tgg gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg gta tcg 48
Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser
1 5 10 15

cag ctg ctc cgg atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc 96
Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala
20 25 30

```
cat tgg gga gtc ctg gcg ggc ctc gcc tac tat tcc atg gtg ggg aac      144
His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn
         35                  40                  45

tgg gct aag gtt ttg gtt gtg atg cta ctc ttt gcc ggc gtc gac ggg      192
Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly
     50                  55                  60

cat acc cgc gtg tca gga ggg gca gca gcc tcc gat acc agg ggc ctt      240
His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu
 65                  70                  75                  80

gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa atc cag ctc gta aac      288
Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn
                 85                  90                  95

acc aac ggc agt tgg cac atc aac agg act gcc ctg aac tgc aac gac      336
Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp
            100                 105                 110

tcc ctc caa aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc      384
Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe
        115                 120                 125

aac tcg tct gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac      432
Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp
    130                 135                 140

aag ttc gct cag ggg tgg ggt ccc ctc act tac act gag cct aac agc      480
Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser
145                 150                 155                 160

tcg gac cag agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt      528
Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly
                165                 170                 175

att gta ccc gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg      576
Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro
            180                 185                 190

agc cct gtt gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat      624
Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr
        195                 200                 205
```

```
aac tgg ggg gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg      672
Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg
    210                 215                 220

ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg      720
Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly
225                 230                 235                 240

ttc acc aag acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc      768
Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly
                245                 250                 255

aac aac acc ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag      816
Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu
            260                 265                 270

gcc acc tac gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt      864
Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys
        275                 280                 285

atg gtt cat tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac      912
Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn
    290                 295                 300

ttc acc atc ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg      960
Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg
305                 310                 315                 320

ttc gaa gcc gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag     1008
Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu
                325                 330                 335

gac agg gat aga tca gag ctt agc ccg ctg ctg ctg tct aca aca gag     1056
Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu
            340                 345                 350

tgg cag ata ctg ccc tgt tcc ttc acc acc ctg ccg gcc cta tcc acc     1104
Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr
        355                 360                 365

ggc ctg atc cac ctc cat cag aac atc gtg gac gtg caa tac ctg tac     1152
Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr
    370                 375                 380
```

```
ggt gta ggg tcg gcg gtt gtc tcc ctt gtc atc aaa tgg gag tat gtc      1200
Gly Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val
385                 390                 395                 400

ctg ttg ctc ttc ctt ctc ctg gca gac gcg cgc atc tgc gcc tgc tta      1248
Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu
                405                 410                 415

tgg atg atg ctg ctg ata gct caa gct gag gcc gcc tta gag aac ctg      1296
Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu
            420                 425                 430

gtg gtc ctc aat gcg gcg gcc gtg gcc ggg gcg cat ggc act ctt tcc      1344
Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser
        435                 440                 445

ttc ctt gtg ttc ttc tgt gct gcc tgg tac atc aag ggc agg ctg gtc      1392
Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val
    450                 455                 460

cct ggt gcg gca tac gcc ttc tat ggc gtg tgg ccg ctg ctc ctg ctt      1440
Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu
465                 470                 475                 480

ctg ctg gcc tta cca cca cga gct tat gcc tagtaa                       1476
Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala
                485                 490
```

<210> 36

<211> 490
<212> PRT
<213> 丙型肝炎病毒

<400> 36

```
Trp Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser
  1               5                  10                  15

Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala
             20                  25                  30

His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn
```

35 40 45

Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly
50 55 60

His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu
65 70 75 80

Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn
85 90 95

Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp
100 105 110

Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe
115 120 125
Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp
130 135 140

Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser
145 150 155 160

Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly
165 170 175

Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro
180 185 190

Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr
195 200 205

Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg
210 215 220

Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly
225 230 235 240

Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly
245 250 255

Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu
260 265 270

Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys

275 280 285

Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn
290 295 300

Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg
305 310 315 320

Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu
325 330 335

Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu
340 345 350

Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr
355 360 365

Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr
370 375 380

Gly Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val
385 390 395 400

Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu
405 410 415

Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu
420 425 430

Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser
435 440 445

Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val
450 455 460

Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu
465 470 475 480

Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala
485 490

<210> 37

<211> 1021
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 2..1018

<220>
<221> mat_肽
<222> 2..1015
<400> 37

```
g atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga          46
  Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly
    1               5                  10                  15

gtc ctg gcg ggc ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag        94
Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys
                20                  25                  30

gtt ttg gtt gtg atg cta ctc ttt gcc ggc gtc gac ggg cat acc cgc       142
Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg
            35                  40                  45

gtg tca gga ggg gca gca gcc tcc gat acc agg ggc ctt gtg tcc ctc       190
Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu
        50                  55                  60

ttt agc ccc ggg tcg gct cag aaa atc cag ctc gta aac acc aac ggc       238
Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly
    65                  70                  75

agt tgg cac atc aac agg act gcc ctg aac tgc aac gac tcc ctc caa       286
Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln
80                  85                  90                  95

aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc aac tcg tct       334
Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser
                100                 105                 110

gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac aag ttc gct       382
Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala
            115                 120                 125
```

```
cag ggg tgg ggt ccc ctc act tac act gag cct aac agc tcg gac cag      430
Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln
        130                 135                 140

agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt att gta ccc      478
Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro
    145                 150                 155

gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg agc cct gtt      526
Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val
160                 165                 170                 175
gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat aac tgg ggg      574
Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly
                180                 185                 190

gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg ccg ccg cga      622
Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg
            195                 200                 205

ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg ttc acc aag      670
Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys
        210                 215                 220

acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc aac aac acc      718
Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr
    225                 230                 235

ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag gcc acc tac      766
Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr
240                 245                 250                 255

gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt atg gtt cat      814
Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His
                260                 265                 270

tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac ttc acc atc      862
Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile
            275                 280                 285

ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg ttc gaa gcc      910
Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala
        290                 295                 300
```

gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag gac agg gat 958
Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp
305 310 315

aga tca gag ctt agc ccg ctg ctg ctg tct aca aca gag tgg cag agt 1006
Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ser
320 325 330 335

ggc aga gct taatta 1021
Gly Arg Ala

<210> 38

<211> 338
<212> PRT
<213> 丙型肝炎病毒

<400> 38

Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val
1 5 10 15

Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val
20 25 30

Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val
35 40 45

Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe
50 55 60

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

Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr
85 90 95

Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly
100 105 110

Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln
115 120 125

```
Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg
    130                 135                 140

Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala
145                 150                 155                 160

Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val
                165                 170                 175

Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala
            180                 185                 190

Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly
        195                 200                 205

Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr
    210                 215                 220

Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu
225                 230                 235                 240

Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala
                245                 250                 255

Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr
            260                 265                 270

Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe
        275                 280                 285

Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala
    290                 295                 300

Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg
305                 310                 315                 320

Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ser Gly
                325                 330                 335

Arg Ala
```

<210> 39

<211> 1034
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 2..1032

<220>
<221> mat_肽
<222> 2..1029
<400> 39

```
g atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga          46
  Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly
    1               5                  10                  15

gtc ctg gcg ggc ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag        94
Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys
                20                  25                  30

gtt ttg gtt gtg atg cta ctc ttt gcc ggc gtc gac ggg cat acc cgc       142
Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg
            35                  40                  45

gtg tca gga ggg gca gca gcc tcc gat acc agg ggc ctt gtg tcc ctc       190
Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu
        50                  55                  60

ttt agc ccc ggg tcg gct cag aaa atc cag ctc gta aac acc aac ggc       238
Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly
    65                  70                  75

agt tgg cac atc aac agg act gcc ctg aac tgc aac gac tcc ctc caa       286
Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln
80                  85                  90                  95

aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc aac tcg tct       334
Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser
                100                 105                 110

gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac aag ttc gct       382
Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala
```

```
            115                 120                 125

cag ggg tgg ggt ccc ctc act tac act gag cct aac agc tcg gac cag      430
Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln
        130                 135                 140

agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt att gta ccc      478
Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro
    145                 150                 155

gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg agc cct gtt      526
Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val
160                 165                 170                 175

gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat aac tgg ggg      574
Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly
                180                 185                 190

gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg ccg ccg cga      622
Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg
            195                 200                 205

ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg ttc acc aag      670
Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys
        210                 215                 220

acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc aac aac acc      718
Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr
    225                 230                 235
ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag gcc acc tac      766
Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr
240                 245                 250                 255

gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt atg gtt cat      814
Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His
                260                 265                 270

tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac ttc acc atc      862
Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile
            275                 280                 285

ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg ttc gaa gcc      910
Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala
        290                 295                 300
```

```
gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag gac agg gat      958
Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp
    305                 310                 315

aga tca gag ctt agc ccg ctg ctg ctg tct aca aca ggt gat cga ggg     1006
Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gly Asp Arg Gly
320                 325                 330                 335

cag aca cca tca cca cca tca cta at ag                               1034
Gln Thr Pro Ser Pro Pro Ser Leu
                340
```

<210> 40

<211> 343
<212> PRT
<213> 丙型肝炎病毒

<400> 40

```
Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val
  1               5                  10                  15

Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val
             20                  25                  30

Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val
         35                  40                  45

Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe
     50                  55                  60

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

Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr
                 85                  90                  95

Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly
            100                 105                 110

Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln
        115                 120                 125
```

```
Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg
    130                 135                 140

Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala
145                 150                 155                 160
Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val
                165                 170                 175

Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala
            180                 185                 190

Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly
        195                 200                 205

Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr
    210                 215                 220

Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu
225                 230                 235                 240

Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala
                245                 250                 255

Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr
            260                 265                 270

Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe
        275                 280                 285

Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala
    290                 295                 300

Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg
305                 310                 315                 320

Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gly Asp Arg Gly Gln
                325                 330                 335

Thr Pro Ser Pro Pro Ser Leu
            340
```

<210> 41

<211> 945

<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..942

<220>
<221> mat_肽
<222> 1..939
<400> 41

```
atg gtg ggg aac tgg gct aag gtt ttg gtt gtg atg cta ctc ttt gcc        48
Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala
  1               5                  10                  15

ggc gtc gac ggg cat acc cgc gtg tca gga ggg gca gca gcc tcc gat        96
Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp
             20                  25                  30

acc agg ggc ctt gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa atc       144
Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile
         35                  40                  45

cag ctc gta aac acc aac ggc agt tgg cac atc aac agg act gcc ctg       192
Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu
     50                  55                  60

aac tgc aac gac tcc ctc caa aca ggg ttc ttt gcc gca cta ttc tac       240
Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr
 65                  70                  75                  80

aaa cac aaa ttc aac tcg tct gga tgc cca gag cgc ttg gcc agc tgt       288
Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys
                 85                  90                  95

cgc tcc atc gac aag ttc gct cag ggg tgg ggt ccc ctc act tac act       336
Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr
            100                 105                 110

gag cct aac agc tcg gac cag agg ccc tac tgc tgg cac tac gcg cct       384
Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro
        115                 120                 125
```

```
cga ccg tgt ggt att gta ccc gcg tct cag gtg tgc ggt cca gtg tat      432
Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr
    130                 135                 140

tgc ttc acc ccg agc cct gtt gtg gtg ggg acg acc gat cgg ttt ggt      480
Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly
145                 150                 155                 160

gtc ccc acg tat aac tgg ggg gcg aac gac tcg gat gtg ctg att ctc      528
Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu
                165                 170                 175

aac aac acg cgg ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg atg      576
Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met
            180                 185                 190

aat ggc act ggg ttc acc aag acg tgt ggg ggc ccc ccg tgc aac atc      624
Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile
        195                 200                 205

ggg ggg gcc ggc aac aac acc ttg acc tgc ccc act gac tgt ttt cgg      672
Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg
    210                 215                 220

aag cac ccc gag gcc acc tac gcc aga tgc ggt tct ggg ccc tgg ctg      720
Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu
225                 230                 235                 240

aca cct agg tgt atg gtt cat tac cca tat agg ctc tgg cac tac ccc      768
Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro
                245                 250                 255

tgc act gtc aac ttc acc atc ttc aag gtt agg atg tac gtg ggg ggc      816
Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly
            260                 265                 270

gtg gag cac agg ttc gaa gcc gca tgc aat tgg act cga gga gag cgt      864
Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg
        275                 280                 285

tgt gac ttg gag gac agg gat aga tca gag ctt agc ccg ctg ctg ctg      912
Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu
    290                 295                 300
tct aca aca gag tgg cag agc tta att aat tag                          945
```

Ser Thr Thr Glu Trp Gln Ser Leu Ile Asn
305 310

<210> 42

<211> 314
<212> PRT
<213> 丙型肝炎病毒

<400> 42

Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala
1 5 10 15

Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp
20 25 30

Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile
35 40 45

Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu
50 55 60

Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr
65 70 75 80

Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys
85 90 95

Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr
100 105 110

Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro
115 120 125

Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr
130 135 140

Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly
145 150 155 160

Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu
165 170 175

Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met
180 185 190

Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile
195 200 205

Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg
210 215 220

Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu
225 230 235 240

Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro
245 250 255

Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly
260 265 270

Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg
275 280 285
Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu
290 295 300

Ser Thr Thr Glu Trp Gln Ser Leu Ile Asn
305 310

<210> 43

<211> 961
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..958

<220>
<221> mat_肽
<222> 1..955
<400> 43

atg gtg ggg aac tgg gct aag gtt ttg gtt gtg atg cta ctc ttt gcc 48
Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala

```
1               5                  10                  15

ggc gtc gac ggg cat acc cgc gtg tca gga ggg gca gca gcc tcc gat      96
Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp
            20                  25                  30

acc agg ggc ctt gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa atc     144
Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile
        35                  40                  45

cag ctc gta aac acc aac ggc agt tgg cac atc aac agg act gcc ctg     192
Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu
    50                  55                  60

aac tgc aac gac tcc ctc caa aca ggg ttc ttt gcc gca cta ttc tac     240
Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr
65                  70                  75                  80

aaa cac aaa ttc aac tcg tct gga tgc cca gag cgc ttg gcc agc tgt     288
Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys
                85                  90                  95

cgc tcc atc gac aag ttc gct cag ggg tgg ggt ccc ctc act tac act     336
Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr
            100                 105                 110

gag cct aac agc tcg gac cag agg ccc tac tgc tgg cac tac gcg cct     384
Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro
        115                 120                 125

cga ccg tgt ggt att gta ccc gcg tct cag gtg tgc ggt cca gtg tat     432
Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr
    130                 135                 140

tgc ttc acc ccg agc cct gtt gtg gtg ggg acg acc gat cgg ttt ggt     480
Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly
145                 150                 155                 160

gtc ccc acg tat aac tgg ggg gcg aac gac tcg gat gtg ctg att ctc     528
Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu
                165                 170                 175

aac aac acg cgg ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg atg     576
Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met
```

```
            180                 185                 190

aat ggc act ggg ttc acc aag acg tgt ggg ggc ccc ccg tgc aac atc      624
Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile
        195                 200                 205

ggg ggg gcc ggc aac aac acc ttg acc tgc ccc act gac tgt ttt cgg      672
Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg
    210                 215                 220

aag cac ccc gag gcc acc tac gcc aga tgc ggt tct ggg ccc tgg ctg      720
Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu
225                 230                 235                 240

aca cct agg tgt atg gtt cat tac cca tat agg ctc tgg cac tac ccc      768
Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro
                245                 250                 255

tgc act gtc aac ttc acc atc ttc aag gtt agg atg tac gtg ggg ggc      816
Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly
            260                 265                 270

gtg gag cac agg ttc gaa gcc gca tgc aat tgg act cga gga gag cgt      864
Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg
        275                 280                 285

tgt gac ttg gag gac agg gat aga tca gag ctt agc ccg ctg ctg ctg      912
Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu
    290                 295                 300

tct aca aca ggt gat cga ggg cag aca cca tca cca cca tca cta a        958
Ser Thr Thr Gly Asp Arg Gly Gln Thr Pro Ser Pro Pro Ser Leu
305                 310                 315

tag                                                                 961
<210> 44

<211> 319
<212> PRT
<213> 丙型肝炎病毒


<400> 44
```

```
Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe Ala
 1               5                  10                  15

Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp
            20                  25                  30

Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile
        35                  40                  45

Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu
    50                  55                  60

Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr
65                  70                  75                  80

Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys
                85                  90                  95

Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr
            100                 105                 110

Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro
        115                 120                 125

Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr
    130                 135                 140

Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly
145                 150                 155                 160

Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu
                165                 170                 175

Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met
            180                 185                 190

Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile
        195                 200                 205

Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg
    210                 215                 220

Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu
225                 230                 235                 240
```

Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro
245 250 255

Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly
260 265 270

Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg
275 280 285

Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu
290 295 300

Ser Thr Thr Gly Asp Arg Gly Gln Thr Pro Ser Pro Pro Ser Leu
305 310 315

<210> 45

<211> 1395
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..1392

<220>
<221> mat_肽
<222> 1..1389
<400> 45
atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg ggc ctc gcc tac tat 48
Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr
1 5 10 15

tcc atg gtg ggg aac tgg gct aag gtt ttg gtt gtg atg cta ctc ttt 96
Ser Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe
20 25 30

gcc ggc gtc gac ggg cat acc cgc gtg tca gga ggg gca gca gcc tcc 144
Ala Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser
35 40 45

gat acc agg ggc ctt gtg tcc ctc ttt agc ccc ggg tcg gct cag aaa 192

```
Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys
     50                  55                  60

atc cag ctc gta aac acc aac ggc agt tgg cac atc aac agg act gcc      240
Ile Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala
 65                  70                  75                  80

ctg aac tgc aac gac tcc ctc caa aca ggg ttc ttt gcc gca cta ttc      288
Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe
                 85                  90                  95

tac aaa cac aaa ttc aac tcg tct gga tgc cca gag cgc ttg gcc agc      336
Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser
            100                 105                 110

tgt cgc tcc atc gac aag ttc gct cag ggg tgg ggt ccc ctc act tac      384
Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr
        115                 120                 125

act gag cct aac agc tcg gac cag agg ccc tac tgc tgg cac tac gcg      432
Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala
    130                 135                 140

cct cga ccg tgt ggt att gta ccc gcg tct cag gtg tgc ggt cca gtg      480
Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val
145                 150                 155                 160

tat tgc ttc acc ccg agc cct gtt gtg gtg ggg acg acc gat cgg ttt      528
Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe
                165                 170                 175

ggt gtc ccc acg tat aac tgg ggg gcg aac gac tcg gat gtg ctg att      576
Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile
            180                 185                 190

ctc aac aac acg cgg ccg ccg cga ggc aac tgg ttc ggc tgt aca tgg      624
Leu Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp
        195                 200                 205

atg aat ggc act ggg ttc acc aag acg tgt ggg ggc ccc ccg tgc aac      672
Met Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn
    210                 215                 220

atc ggg ggg gcc ggc aac aac acc ttg acc tgc ccc act gac tgt ttt      720
```

```
Ile Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe
225                 230                 235                 240

cgg aag cac ccc gag gcc acc tac gcc aga tgc ggt tct ggg ccc tgg      768
Arg Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp
                245                 250                 255

ctg aca cct agg tgt atg gtt cat tac cca tat agg ctc tgg cac tac      816
Leu Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr
            260                 265                 270

ccc tgc act gtc aac ttc acc atc ttc aag gtt agg atg tac gtg ggg      864
Pro Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly
        275                 280                 285

ggc gtg gag cac agg ttc gaa gcc gca tgc aat tgg act cga gga gag      912
Gly Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu
    290                 295                 300

cgt tgt gac ttg gag gac agg gat aga tca gag ctt agc ccg ctg ctg      960
Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu
305                 310                 315                 320

ctg tct aca aca gag tgg cag ata ctg ccc tgt tcc ttc acc acc ctg     1008
Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu
                325                 330                 335

ccg gcc cta tcc acc ggc ctg atc cac ctc cat cag aac atc gtg gac     1056
Pro Ala Leu Ser Thr Gly Leu Ile His Leu His Gln Asn Ile Val Asp
            340                 345                 350

gtg caa tac ctg tac ggt gta ggg tcg gcg gtt gtc tcc ctt gtc atc     1104
Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val Val Ser Leu Val Ile
        355                 360                 365

aaa tgg gag tat gtc ctg ttg ctc ttc ctt ctc ctg gca gac gcg cgc     1152
Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg
    370                 375                 380

atc tgc gcc tgc tta tgg atg atg ctg ctg ata gct caa gct gag gcc     1200
Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala
385                 390                 395                 400

gcc tta gag aac ctg gtg gtc ctc aat gcg gcg gcc gtg gcc ggg gcg     1248
```

```
Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala
                405                 410                 415

cat ggc act ctt tcc ttc ctt gtg ttc ttc tgt gct gcc tgg tac atc      1296
His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile
            420                 425                 430

aag ggc agg ctg gtc cct ggt gcg gca tac gcc ttc tat ggc gtg tgg      1344
Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp
        435                 440                 445

ccg ctg ctc ctg ctt ctg ctg gcc tta cca cca cga gct tat gcc tagtaa    1395
Pro Leu Leu Leu Leu Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala
    450                 455                 460
<210> 46

<211> 463
<212> PRT
<213> 丙型肝炎病毒


<400> 46

Met Val Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr
  1               5                  10                  15

Ser Met Val Gly Asn Trp Ala Lys Val Leu Val Val Met Leu Leu Phe
             20                  25                  30
Ala Gly Val Asp Gly His Thr Arg Val Ser Gly Gly Ala Ala Ala Ser
         35                  40                  45

Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys
     50                  55                  60

Ile Gln Leu Val Asn Thr Asn Gly Ser Trp His Ile Asn Arg Thr Ala
 65                  70                  75                  80

Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe
                 85                  90                  95

Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser
            100                 105                 110

Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr
```

115 120 125

Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala
130 135 140

Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln Val Cys Gly Pro Val
145 150 155 160

Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Phe
165 170 175

Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp Ser Asp Val Leu Ile
180 185 190

Leu Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp
195 200 205

Met Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn
210 215 220

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

Arg Lys His Pro Glu Ala Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp
245 250 255

Leu Thr Pro Arg Cys Met Val His Tyr Pro Tyr Arg Leu Trp His Tyr
260 265 270

Pro Cys Thr Val Asn Phe Thr Ile Phe Lys Val Arg Met Tyr Val Gly
275 280 285

Gly Val Glu His Arg Phe Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu
290 295 300

Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu
305 310 315 320

Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu
325 330 335

Pro Ala Leu Ser Thr Gly Leu Ile His Leu His Gln Asn Ile Val Asp
340 345 350

Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val Val Ser Leu Val Ile
355 360 365

Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg
370 375 380

Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile Ala Gln Ala Glu Ala
385 390 395 400

Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala Ala Val Ala Gly Ala
405 410 415

His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile
420 425 430

Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp
435 440 445

Pro Leu Leu Leu Leu Leu Leu Ala Leu Pro Pro Arg Ala Tyr Ala
450 455 460

<210> 47

<211> 2082
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..2079

<220>
<221> mat_肽
<222> 1..2076
<400> 47

aat ttg ggt aag gtc atc gat acc ctt aca tgc ggc ttc gcc gac ctc 48
Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

gtg ggg tac att ccg ctc gtc ggc gcc ccc cta ggg ggc gct gcc agg 96
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
20 25 30

gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac ggc gtg aac tat gca 144

```
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
         35                  40                  45

aca ggg aat ttg ccc ggt tgc tct ttc tct atc ttc ctc ttg gct ttg      192
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
     50                  55                  60

ctg tcc tgt ctg acc gtt cca gct tcc gct tat gaa gtg cgc aac gtg      240
Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
 65                  70                  75                  80

tcc ggg atg tac cat gtc acg aac gac tgc tcc aac tca agc att gtg      288
Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                 85                  90                  95

tat gag gca gcg gac atg atc atg cac acc ccc ggg tgc gtg ccc tgc      336
Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

gtt cgg gag aac aac tct tcc cgc tgc tgg gta gcg ctc acc ccc acg      384
Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

ctc gca gct agg aac gcc agc gtc ccc acc acg aca ata cga cgc cac      432
Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt tcc gct atg tac gtg      480
Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145                 150                 155                 160

ggg gac ctc tgc gga tct gtc ttc ctc gtc tcc cag ctg ttc acc atc      528
Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
                165                 170                 175

tcg cct cgc cgg cat gag acg gtg cag gac tgc aat tgc tca atc tat      576
Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
            180                 185                 190

ccc ggc cac ata acg ggt cac cgt atg gct tgg gat atg atg atg aac      624
Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
        195                 200                 205

tgg tcg cct aca acg gcc ctg gtg gta tcg cag ctg ctc cgg atc cca      672
```

```
Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro
    210                 215                 220

caa gct gtc gtg gac atg gtg gcg ggg gcc cat tgg gga gtc ctg gcg      720
Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala
225                 230                 235                 240

ggc ctc gcc tac tat tcc atg gtg ggg aac tgg gct aag gtt ttg gtt      768
Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Val
                245                 250                 255

gtg atg cta ctc ttt gcc ggc gtc gac ggg cat acc cgc gtg tca gga      816
Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val Ser Gly
            260                 265                 270

ggg gca gca gcc tcc gat acc agg ggc ctt gtg tcc ctc ttt agc ccc      864
Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro
        275                 280                 285

ggg tcg gct cag aaa atc cag ctc gta aac acc aac ggc agt tgg cac      912
Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr Asn Gly Ser Trp His
    290                 295                 300

atc aac agg act gcc ctg aac tgc aac gac tcc ctc caa aca ggg ttc      960
Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe
305                 310                 315                 320

ttt gcc gca cta ttc tac aaa cac aaa ttc aac tcg tct gga tgc cca     1008
Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro
                325                 330                 335

gag cgc ttg gcc agc tgt cgc tcc atc gac aag ttc gct cag ggg tgg     1056
Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp
            340                 345                 350

ggt ccc ctc act tac act gag cct aac agc tcg gac cag agg ccc tac     1104
Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr
        355                 360                 365

tgc tgg cac tac gcg cct cga ccg tgt ggt att gta ccc gcg tct cag     1152
Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln
    370                 375                 380

gtg tgc ggt cca gtg tat tgc ttc acc ccg agc cct gtt gtg gtg ggg     1200
```

```
Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly
385                 390                 395                 400
acg acc gat cgg ttt ggt gtc ccc acg tat aac tgg ggg gcg aac gac      1248
Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp
                405                 410                 415

tcg gat gtg ctg att ctc aac aac acg cgg ccg ccg cga ggc aac tgg      1296
Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp
            420                 425                 430

ttc ggc tgt aca tgg atg aat ggc act ggg ttc acc aag acg tgt ggg      1344
Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly
        435                 440                 445

ggc ccc ccg tgc aac atc ggg ggg gcc ggc aac aac acc ttg acc tgc      1392
Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys
    450                 455                 460

ccc act gac tgt ttt cgg aag cac ccc gag gcc acc tac gcc aga tgc      1440
Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala Arg Cys
465                 470                 475                 480

ggt tct ggg ccc tgg ctg aca cct agg tgt atg gtt cat tac cca tat      1488
Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr Pro Tyr
                485                 490                 495

agg ctc tgg cac tac ccc tgc act gtc aac ttc acc atc ttc aag gtt      1536
Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe Lys Val
            500                 505                 510

agg atg tac gtg ggg ggc gtg gag cac agg ttc gaa gcc gca tgc aat      1584
Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala Cys Asn
        515                 520                 525

tgg act cga gga gag cgt tgt gac ttg gag gac agg gat aga tca gag      1632
Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu
    530                 535                 540

ctt agc ccg ctg ctg ctg tct aca aca gag tgg cag ata ctg ccc tgt      1680
Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys
545                 550                 555                 560

tcc ttc acc acc ctg ccg gcc cta tcc acc ggc ctg atc cac ctc cat      1728
Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly Leu Ile His Leu His
```

565 570 575

cag aac atc gtg gac gtg caa tac ctg tac ggt gta ggg tcg gcg gtt 1776
Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val
580 585 590

gtc tcc ctt gtc atc aaa tgg gag tat gtc ctg ttg ctc ttc ctt ctc 1824
Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu
595 600 605

ctg gca gac gcg cgc atc tgc gcc tgc tta tgg atg atg ctg ctg ata 1872
Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile
610 615 620

gct caa gct gag gcc gcc tta gag aac ctg gtg gtc ctc aat gcg gcg 1920
Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala
625 630 635 640

gcc gtg gcc ggg gcg cat ggc act ctt tcc ttc ctt gtg ttc ttc tgt 1968
Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys
645 650 655

gct gcc tgg tac atc aag ggc agg ctg gtc cct ggt gcg gca tac gcc 2016
Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala
660 665 670

ttc tat ggc gtg tgg ccg ctg ctc ctg ctt ctg ctg gcc tta cca cca 2064
Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu Leu Ala Leu Pro Pro
675 680 685

cga gct tat gcc tagtaa 2082
Arg Ala Tyr Ala
690

<210> 48

<211> 692
<212> PRT
<213> 丙型肝炎病毒

<400> 48

Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
1 5 10 15

```
Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg
            20                  25                  30

Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
        35                  40                  45

Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
    50                  55                  60

Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Glu Val Arg Asn Val
65                  70                  75                  80

Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val
                85                  90                  95

Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys
            100                 105                 110

Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr
        115                 120                 125

Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr Thr Ile Arg Arg His
    130                 135                 140

Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val
145                 150                 155                 160

Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu Phe Thr Ile
                165                 170                 175

Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr
            180                 185                 190

Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
        195                 200                 205

Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro
    210                 215                 220

Gln Ala Val Val Asp Met Val Ala Gly Ala His Trp Gly Val Leu Ala
225                 230                 235                 240

Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Leu Val
```

245 250 255

Val Met Leu Leu Phe Ala Gly Val Asp Gly His Thr Arg Val Ser Gly
260 265 270
Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val Ser Leu Phe Ser Pro
275 280 285

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

Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Gln Thr Gly Phe
305 310 315 320

Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn Ser Ser Gly Cys Pro
325 330 335

Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys Phe Ala Gln Gly Trp
340 345 350

Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser Asp Gln Arg Pro Tyr
355 360 365

Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile Val Pro Ala Ser Gln
370 375 380

Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val Val Gly
385 390 395 400

Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn Trp Gly Ala Asn Asp
405 410 415

Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro Pro Arg Gly Asn Trp
420 425 430

Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe Thr Lys Thr Cys Gly
435 440 445

Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn Asn Thr Leu Thr Cys
450 455 460

Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr Tyr Ala Arg Cys
465 470 475 480

Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met Val His Tyr Pro Tyr

485 490 495

Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr Ile Phe Lys Val
500 505 510

Arg Met Tyr Val Gly Gly Val Glu His Arg Phe Glu Ala Ala Cys Asn
515 520 525

Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp Arg Ser Glu
530 535 540

Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp Gln Ile Leu Pro Cys
545 550 555 560

Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly Leu Ile His Leu His
565 570 575

Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly Val Gly Ser Ala Val
580 585 590

Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu Leu Leu Phe Leu Leu
595 600 605

Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp Met Met Leu Leu Ile
610 615 620

Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val Val Leu Asn Ala Ala
625 630 635 640

Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe Leu Val Phe Phe Cys
645 650 655

Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro Gly Ala Ala Tyr Ala
660 665 670

Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu Leu Ala Leu Pro Pro
675 680 685

Arg Ala Tyr Ala
690

<210> 49

<211> 2433
<212> DNA
<213> 丙型肝炎病毒

<220>
<221> CDS
<222> 1..2430

<220>
<221> mat_肽
<222> 1..2427
<400> 49

```
atg agc acg aat cct aaa cct caa aga aaa acc aaa cgt aac acc aac        48
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
  1               5                  10                  15

cgc cgc cca cag gac gtc aag ttc ccg ggc ggt ggt cag atc gtt ggt        96
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
             20                  25                  30

gga gtt tac ctg ttg ccg cgc agg ggc ccc agg ttg ggt gtg cgc gcg       144
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
         35                  40                  45

act agg aag act tcc gag cgg tcg caa cct cgt ggg agg cga caa cct       192
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
     50                  55                  60

atc ccc aag gct cgc cga ccc gag ggt agg gcc tgg gct cag ccc ggg       240
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
 65                  70                  75                  80

tac cct tgg ccc ctc tat ggc aat gag ggc atg ggg tgg gca gga tgg       288
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp
                 85                  90                  95

ctc ctg tca ccc cgc ggc tct cgg cct agt tgg ggc cct aca gac ccc       336
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
            100                 105                 110

cgg cgt agg tcg cgt aat ttg ggt aag gtc atc gat acc ctt aca tgc       384
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
        115                 120                 125
```

```
ggc ttc gcc gac ctc gtg ggg tac att ccg ctc gtc ggc gcc ccc cta      432
Gly Phe Ala Asp Leu Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
    130                 135                 140

ggg ggc gct gcc agg gcc ctg gcg cat ggc gtc cgg gtt ctg gag gac      480
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145                 150                 155                 160

ggc gtg aac tat gca aca ggg aat ttg ccc ggt tgc tct ttc tct atc      528
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
                165                 170                 175

ttc ctc ttg gct ttg ctg tcc tgt ctg acc gtt cca gct tcc gct tat      576
Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr
            180                 185                 190

gaa gtg cgc aac gtg tcc ggg atg tac cat gtc acg aac gac tgc tcc      624
Glu Val Arg Asn Val Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser
        195                 200                 205

aac tca agc att gtg tat gag gca gcg gac atg atc atg cac acc ccc      672
Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro
    210                 215                 220

ggg tgc gtg ccc tgc gtt cgg gag aac aac tct tcc cgc tgc tgg gta      720
Gly Cys Val Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val
225                 230                 235                 240

gcg ctc acc ccc acg ctc gca gct agg aac gcc agc gtc ccc acc acg      768
Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr
                245                 250                 255

aca ata cga cgc cac gtc gat ttg ctc gtt ggg gcg gct gct ttc tgt      816
Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys
            260                 265                 270

tcc gct atg tac gtg ggg gac ctc tgc gga tct gtc ttc ctc gtc tcc      864
Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser
        275                 280                 285

cag ctg ttc acc atc tcg cct cgc cgg cat gag acg gtg cag gac tgc      912
Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys
    290                 295                 300
```

```
aat tgc tca atc tat ccc ggc cac ata acg ggt cac cgt atg gct tgg      960
Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp
305                 310                 315                 320

gat atg atg atg aac tgg tcg cct aca acg gcc ctg gtg gta tcg cag     1008
Asp Met Met Met Asn Trp Ser Pro Thr Thr Ala Leu Val Val Ser Gln
                325                 330                 335

ctg ctc cgg atc cca caa gct gtc gtg gac atg gtg gcg ggg gcc cat     1056
Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His
            340                 345                 350

tgg gga gtc ctg gcg ggc ctc gcc tac tat tcc atg gtg ggg aac tgg     1104
Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp
        355                 360                 365

gct aag gtt ttg gtt gtg atg cta ctc ttt gcc ggc gtc gac ggg cat     1152
Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His
    370                 375                 380

acc cgc gtg tca gga ggg gca gca gcc tcc gat acc agg ggc ctt gtg     1200
Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val
385                 390                 395                 400

tcc ctc ttt agc ccc ggg tcg gct cag aaa atc cag ctc gta aac acc     1248
Ser Leu Phe Ser Pro Gly Ser Ala Gln Lys Ile Gln Leu Val Asn Thr
                405                 410                 415

aac ggc agt tgg cac atc aac agg act gcc ctg aac tgc aac gac tcc     1296
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
            420                 425                 430

ctc caa aca ggg ttc ttt gcc gca cta ttc tac aaa cac aaa ttc aac     1344
Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn
        435                 440                 445

tcg tct gga tgc cca gag cgc ttg gcc agc tgt cgc tcc atc gac aag     1392
Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys
    450                 455                 460

ttc gct cag ggg tgg ggt ccc ctc act tac act gag cct aac agc tcg     1440
Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser
465                 470                 475                 480
```

```
gac cag agg ccc tac tgc tgg cac tac gcg cct cga ccg tgt ggt att      1488
Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile
                485                 490                 495

gta ccc gcg tct cag gtg tgc ggt cca gtg tat tgc ttc acc ccg agc      1536
Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser
            500                 505                 510

cct gtt gtg gtg ggg acg acc gat cgg ttt ggt gtc ccc acg tat aac      1584
Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn
        515                 520                 525

tgg ggg gcg aac gac tcg gat gtg ctg att ctc aac aac acg cgg ccg      1632
Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro
    530                 535                 540

ccg cga ggc aac tgg ttc ggc tgt aca tgg atg aat ggc act ggg ttc      1680
Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe
545                 550                 555                 560

acc aag acg tgt ggg ggc ccc ccg tgc aac atc ggg ggg gcc ggc aac      1728
Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn
                565                 570                 575

aac acc ttg acc tgc ccc act gac tgt ttt cgg aag cac ccc gag gcc      1776
Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala
            580                 585                 590

acc tac gcc aga tgc ggt tct ggg ccc tgg ctg aca cct agg tgt atg      1824
Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met
        595                 600                 605

gtt cat tac cca tat agg ctc tgg cac tac ccc tgc act gtc aac ttc      1872
Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe
    610                 615                 620

acc atc ttc aag gtt agg atg tac gtg ggg ggc gtg gag cac agg ttc      1920
Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe
625                 630                 635                 640

gaa gcc gca tgc aat tgg act cga gga gag cgt tgt gac ttg gag gac      1968
Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp
                645                 650                 655
```

agg gat aga tca gag ctt agc ccg ctg ctg tct aca aca gag tgg 2016
Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Ser Thr Thr Glu Trp
660 665 670

cag ata ctg ccc tgt tcc ttc acc acc ctg ccg gcc cta tcc acc ggc 2064
Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly
675 680 685

ctg atc cac ctc cat cag aac atc gtg gac gtg caa tac ctg tac ggt 2112
Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly
690 695 700

gta ggg tcg gcg gtt gtc tcc ctt gtc atc aaa tgg gag tat gtc ctg 2160
Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu
705 710 715 720

ttg ctc ttc ctt ctc ctg gca gac gcg cgc atc tgc gcc tgc tta tgg 2208
Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp
725 730 735

atg atg ctg ctg ata gct caa gct gag gcc gcc tta gag aac ctg gtg 2256
Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val
740 745 750

gtc ctc aat gcg gcg gcc gtg gcc ggg gcg cat ggc act ctt tcc ttc 2304
Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe
755 760 765

ctt gtg ttc ttc tgt gct gcc tgg tac atc aag ggc agg ctg gtc cct 2352
Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro
770 775 780

ggt gcg gca tac gcc ttc tat ggc gtg tgg ccg ctg ctc ctg ctt ctg 2400
Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu
785 790 795 800

ctg gcc tta cca cca cga gct tat gcc tagtaa 2433
Leu Ala Leu Pro Pro Arg Ala Tyr Ala
805 810

<210> 50

<211> 809
<212> PRT
<213> 丙型肝炎病毒

<400> 50

```
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
 1               5                  10                  15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25                  30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
        35                  40                  45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
    50                  55                  60

Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ala Trp Ala Gln Pro Gly
65                  70                  75                  80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Met Gly Trp Ala Gly Trp
                85                  90                  95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
            100                 105                 110

Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys
        115                 120                 125
Gly Phe Ala Asp Leu Val Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu
    130                 135                 140

Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145                 150                 155                 160

Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
                165                 170                 175

Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr
            180                 185                 190

Glu Val Arg Asn Val Ser Gly Met Tyr His Val Thr Asn Asp Cys Ser
        195                 200                 205

Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr Pro
    210                 215                 220
```

```
Gly Cys Val Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val
225                 230                 235                 240

Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro Thr Thr
                245                 250                 255

Thr Ile Arg Arg His Val Asp Leu Leu Val Gly Ala Ala Ala Phe Cys
            260                 265                 270

Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser
        275                 280                 285

Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val Gln Asp Cys
    290                 295                 300

Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp
305                 310                 315                 320

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

Leu Leu Arg Ile Pro Gln Ala Val Val Asp Met Val Ala Gly Ala His
            340                 345                 350

Trp Gly Val Leu Ala Gly Leu Ala Tyr Tyr Ser Met Val Gly Asn Trp
        355                 360                 365

Ala Lys Val Leu Val Val Met Leu Leu Phe Ala Gly Val Asp Gly His
    370                 375                 380

Thr Arg Val Ser Gly Gly Ala Ala Ala Ser Asp Thr Arg Gly Leu Val
385                 390                 395                 400

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

Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
            420                 425                 430

Leu Gln Thr Gly Phe Phe Ala Ala Leu Phe Tyr Lys His Lys Phe Asn
        435                 440                 445

Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Ser Ile Asp Lys
```

```
     450                 455                 460

Phe Ala Gln Gly Trp Gly Pro Leu Thr Tyr Thr Glu Pro Asn Ser Ser
465                 470                 475                 480

Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Pro Arg Pro Cys Gly Ile
                485                 490                 495

Val Pro Ala Ser Gln Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser
            500                 505                 510

Pro Val Val Val Gly Thr Thr Asp Arg Phe Gly Val Pro Thr Tyr Asn
        515                 520                 525

Trp Gly Ala Asn Asp Ser Asp Val Leu Ile Leu Asn Asn Thr Arg Pro
    530                 535                 540

Pro Arg Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Gly Thr Gly Phe
545                 550                 555                 560

Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn Ile Gly Gly Ala Gly Asn
                565                 570                 575

Asn Thr Leu Thr Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala
            580                 585                 590

Thr Tyr Ala Arg Cys Gly Ser Gly Pro Trp Leu Thr Pro Arg Cys Met
        595                 600                 605

Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe
    610                 615                 620

Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Phe
625                 630                 635                 640

Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp
                645                 650                 655

Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Glu Trp
            660                 665                 670

Gln Ile Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly
        675                 680                 685
```

```
Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly
    690                 695                 700

Val Gly Ser Ala Val Val Ser Leu Val Ile Lys Trp Glu Tyr Val Leu
705                 710                 715                 720

Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu Trp
                725                 730                 735

Met Met Leu Leu Ile Ala Gln Ala Glu Ala Ala Leu Glu Asn Leu Val
            740                 745                 750

Val Leu Asn Ala Ala Ala Val Ala Gly Ala His Gly Thr Leu Ser Phe
        755                 760                 765

Leu Val Phe Phe Cys Ala Ala Trp Tyr Ile Lys Gly Arg Leu Val Pro
    770                 775                 780

Gly Ala Ala Tyr Ala Phe Tyr Gly Val Trp Pro Leu Leu Leu Leu Leu
785                 790                 795                 800

Leu Ala Leu Pro Pro Arg Ala Tyr Ala
                805
```

<210> 51

<211> 17
<212> PRT
<213> 丙型肝炎病毒
<220>
<221> 修饰位点
<222> 1..17
<400> 51

```
Ser Asn Ser Ser Glu Ala Ala Asp Met Ile Met His Thr Pro Gly Cys
1               5                   10                  15

Val
```

<210> 52

<211> 22

<212> PRT
<213> 丙型肝炎病毒
<220>
<221> 修饰位点
<222> 1..22
<400> 52

Gly Gly Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
1 5 10 15

Ser Pro Thr Thr Ala Leu
20

<210> 53

<211> 37
<212> PRT
<213> 丙型肝炎病毒
<220>
<221> 修饰位点
<222> 1..37
<400> 53

Tyr Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys
1 5 10 15

Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr
20 25 30

Pro Gly Cys Gly Lys
35

<210> 54

<211> 25
<212> PRT
<213> 丙型肝炎病毒
<220>
<221> 修饰位点
<222> 1..25
<400> 54

Gly Gly Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Ala Thr
1 5 10 15

Gln Leu Arg Arg His Ile Asp Leu Leu
20 25

<210> 55

<211> 25
<212> PRT
<213> 丙型肝炎病毒
<220>
<221> 修饰位点
<222> 1..25
<400> 55

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

Thr Ile Arg Arg His Val Asp Leu Leu
20 25

<210> 56

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 56

Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr Gln Val Arg Asn
1 5 10 15

Ser Thr Gly Leu
20

<210> 57

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 57

Gln Val Arg Asn Ser Thr Gly Leu Tyr His Val Thr Asn Asp Cys Pro
1 5 10 15

Asn Ser Ser Ile
20

<210> 58

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 58

Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr Glu Ala His Asp Ala Ile
1 5 10 15

Leu His Thr Pro
20

<210> 59

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 59

Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Met Ile Met His Thr
1 5 10 15

Pro Gly Cys Val
20

<210> 60

<211> 19

<212> PRT
<213> 丙型肝炎病毒

<400> 60

His Asp Ala Ile Leu His Thr Pro Gly Val Pro Cys Val Arg Glu Gly
1 5 10 15

Asn Val Ser

<210> 61

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 61

Cys Val Arg Glu Gly Asn Val Ser Arg Cys Trp Val Ala Met Thr Pro
1 5 10 15

Thr Val Ala Thr
20

<210> 62

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 62

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

Gln Leu Arg Arg
20

<210> 63

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 63

Leu Pro Ala Thr Gln Leu Arg Arg His Ile Asp Leu Leu Val Gly Ser
1 5 10 15

Ala Thr Leu Cys
20

<210> 64
<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 64

Leu Val Gly Ser Ala Thr Leu Cys Ser Ala Leu Tyr Val Gly Asp Leu
1 5 10 15

Cys Gly Ser Val
20

<210> 65

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 65

Gln Leu Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Gly Cys
1 5 10 15

Asn Cys Ser Ile

20

<210> 66

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 66

Thr Gln Gly Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His
1 5 10 15

Arg Met Ala Trp
20

<210> 67

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 67

Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro
1 5 10 15

Thr Ala Ala Leu
20

<210> 68

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 68

Asn Trp Ser Pro Thr Ala Ala Leu Val Met Ala Gln Leu Leu Arg Ile

1 5 10 15

Pro Gln Ala Ile
20

<210> 69

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 69

Leu Leu Arg Ile Pro Gln Ala Ile Leu Asp Met Ile Ala Gly Ala His
1 5 10 15

Trp Gly Val Leu
20

<210> 70

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 70

Ala Gly Ala His Trp Gly Val Leu Ala Gly Ile Ala Tyr Phe Ser Met
1 5 10 15

Val Gly Asn Met
20

<210> 71

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 71

Val Val Leu Leu Leu Phe Ala Gly Val Asp Ala Glu Thr Ile Val Ser
1 5 10 15

Gly Gly Gln Ala
20

<210> 72

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 72

Ser Gly Leu Val Ser Leu Phe Thr Pro Gly Ala Lys Gln Asn Ile Gln
1 5 10 15

Leu Ile Asn Thr
20

<210> 73

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 73

Gln Asn Ile Gln Leu Ile Asn Thr Asn Gly Gln Trp His Ile Asn Ser
1 5 10 15

Thr Ala Leu Asn
20

<210> 74

<211> 21
<212> PRT
<213> 丙型肝炎病毒

<400> 74

Leu Asn Cys Asn Glu Ser Leu Asn Thr Gly Trp Trp Leu Ala Gly Leu
1 5 10 15

Ile Tyr Gln His Lys
20

<210> 75

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 75

Ala Gly Leu Ile Tyr Gln His Lys Phe Asn Ser Ser Gly Cys Pro Glu
1 5 10 15

Arg Leu Ala Ser
20

<210> 76

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 76

Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Pro Leu Thr Asp Phe Asp
1 5 10 15

Gln Gly Trp Gly
20

<210> 77

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 77

Thr Asp Phe Asp Gln Gly Trp Gly Pro Ile Ser Tyr Ala Asn Gly Ser
1 5 10 15

Gly Pro Asp Gln
20

<210> 78

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 78

Ala Asn Gly Ser Gly Pro Asp Gln Arg Pro Tyr Cys Trp His Tyr Pro
1 5 10 15

Pro Lys Pro Cys
20

<210> 79

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 79

Trp His Tyr Pro Pro Lys Pro Cys Gly Ile Val Pro Ala Lys Ser Val
1 5 10 15

Cys Gly Pro Val
20

<210> 80

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 80

Ala Lys Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val
1 5 10 15

Val Val Gly Thr
20

<210> 81

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 81

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

Tyr Ser Trp Gly
20

<210> 82

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 82

Gly Ala Pro Thr Tyr Ser Trp Gly Glu Asn Asp Thr Asp Val Phe Val
1 5 10 15

Leu Asn Asn Thr
20

<210> 83

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 83

Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys
1 5 10 15

Val Cys Gly Ala
20

<210> 84

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 84

Gly Phe Thr Lys Val Cys Gly Ala Pro Pro Val Cys Ile Gly Gly Ala
1 5 10 15

Gly Asn Asn Thr
20

<210> 85

<211> 19
<212> PRT
<213> 丙型肝炎病毒

<400> 85

Ile Gly Gly Ala Gly Asn Asn Thr Leu His Cys Pro Thr Asp Cys Arg
1 5 10 15

Lys His Pro

<210> 86

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 86

Thr Asp Cys Phe Arg Lys His Pro Asp Ala Thr Tyr Ser Arg Cys Gly
1 5 10 15

Ser Gly Pro Trp
20

<210> 87

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 87

Ser Arg Cys Gly Ser Gly Pro Trp Ile Thr Pro Arg Cys Leu Val Asp
1 5 10 15

Tyr Pro Tyr Arg
20

<210> 88

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 88

Cys Leu Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Ile
1 5 10 15

Asn Tyr Thr Ile
20

<210> 89

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 89

Pro Cys Thr Ile Asn Tyr Thr Ile Phe Lys Ile Arg Met Tyr Val Gly
1 5 10 15

Gly Val Glu His
20

<210> 90

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 90

Met Tyr Val Gly Gly Val Glu His Arg Leu Glu Ala Ala Cys Asn Trp
1 5 10 15

Thr Pro Gly Glu
20

<210> 91

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 91

Ala Cys Asn Trp Thr Pro Gly Glu Arg Cys Asp Leu Glu Asp Arg Asp
1 5 10 15

Arg Ser Glu Leu
20

<210> 92

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 92

Glu Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Thr Thr Thr
1 5 10 15

Gln Trp Gln Val
20

<210> 93

<211> 9
<212> PRT
<213> 丙型肝炎病毒

<400> 93

Tyr Gln Val Arg Asn Ser Thr Gly Leu

1 5

<210> 94

<211> 29
<212> DNA
<213> 丙型肝炎病毒

<400> 94

agctaattaa ttaagcttgc atgcctgca 29

<210> 95

<211> 60
<212> DNA
<213> 丙型肝炎病毒

<400> 95

gttaattaac tattagtgat ggtggtgatg gtgtctgccc tcgatcacgt gcaggcctcc 60

<210> 96

<211> 19
<212> DNA
<213> 丙型肝炎病毒

<400> 96

gtttaaccac tgcatgatg 19

<210> 97

<211> 20
<212> DNA
<213> 丙型肝炎病毒

<400> 97

gtcccatcga gtgcggctac 20

<210> 98

<211> 45
<212> DNA
<213> 丙型肝炎病毒

<400> 98

cgtgacatgg tacattccgg acacttggcg cacttcataa gcgga 45

<210> 99

<211> 42
<212> DNA
<213> 丙型肝炎病毒

<400> 99

tgcctcatac acaatggagc tctgggacga gtcgttcgtg ac 42

<210> 100

<211> 42
<212> DNA
<213> 丙型肝炎病毒

<400> 100

tacccagcag cgggagctct gttgctcccg aacgcagggc ac 42

<210> 101

<211> 42
<212> DNA
<213> 丙型肝炎病毒

<400> 101

tgtcgtggtg gggacggagg cctgcctagc tgcgagcgtg gg 42

<210> 102

<211> 48
<212>

<400> 102

cgttatgtgg cccgggtaga ttgagcactg gcagtcctgc accgtctc 48

<210> 103

<211> 42
<212> DNA
<213> 丙型肝炎病毒

<400> 103

cagggccgtt ctaggcctcc actgcatcat catatcccaa gc 42

<210> 104

<211> 26
<212> DNA
<213> 丙型肝炎病毒

<400> 104

ccggaatgta ccatgtcacg aacgac 26

<210> 105

<211> 24
<212> DNA
<213> 丙型肝炎病毒

<400> 105

gctccattgt gtatgaggca gcgg 24

<210> 106

<211> 23
<212> DNA
<213> 丙型肝炎病毒

<400> 106

gagctcccgc tgctgggtag cgc 23

<210> 107

<211> 25
<212> DNA
<213> 丙型肝炎病毒

<400> 107

cctccgtccc caccacgaca atacg 25

<210> 108

<211> 27
<212> DNA
<213> 丙型肝炎病毒

<400> 108

ctacccgggc cacataacgg gtcaccg 27

<210> 109

<211> 24
<212> DNA
<213> 丙型肝炎病毒

<400> 109

ggaggcctac aacggccctg gtgg 24

<210> 110

<211> 22
<212> DNA
<213> 丙型肝炎病毒

<400> 110

ttctatcgat taaatagaat tc 22

<210> 111

<211> 23
<212> DNA
<213> 丙型肝炎病毒

<400> 111

gccatacgct cacagccgat ccc 23

<210> 112

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 112

Tyr Glu Val Arg Asn Val Ser Gly Ile Tyr His Val Thr Asn Asp Cys
1 5 10 15

Ser Asn Ser Ser
20

<210> 113

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 113

Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp
1 5 10 15

Met Ile Met His Thr
20

<210> 114

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 114

Ala Ala Asp Met Ile Met His Thr Pro Gly Cys Val Pro Cys Val
1 5 10 15

Arg Glu Asn Asn Ser
20

<210> 115

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 115

Pro Cys Val Arg Glu Asn Asn Ser Ser Arg Cys Trp Val Ala Leu
1 5 10 15

Thr Pro Thr Leu Ala
20

<210> 116

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 116

Val Ala Leu Thr Pro Thr Leu Ala Ala Arg Asn Ala Ser Val Pro
1 5 10 15

Thr Thr Thr Ile Arg
20

<210> 117

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 117

Ser Val Pro Thr Thr Thr Ile Arg Arg His Val Asp Leu Leu Val
1 5 10 15

Gly Ala Ala Ala Phe
20

<210> 118

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 118

Leu Leu Val Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly
1 5 10 15

Asp Leu Cys Gly Ser
20

<210> 119

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 119

Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Ser Gln Leu
1 5 10 15

Phe Thr Ile Ser Pro
20

<210> 120

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 120

Ser Gln Leu Phe Thr Ile Ser Pro Arg Arg His Glu Thr Val Gln
1 5 10 15

Asp Cys Asn Cys Ser
20

<210> 121

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 121

Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr
1 5 10 15

Gly His Arg Met Ala
20

<210> 122

<211> 20
<212> PRT
<213> 丙型肝炎病毒

<400> 122

```
His Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp
  1               5                  10                  15
Ser Pro Thr Thr Ala
                 20
```

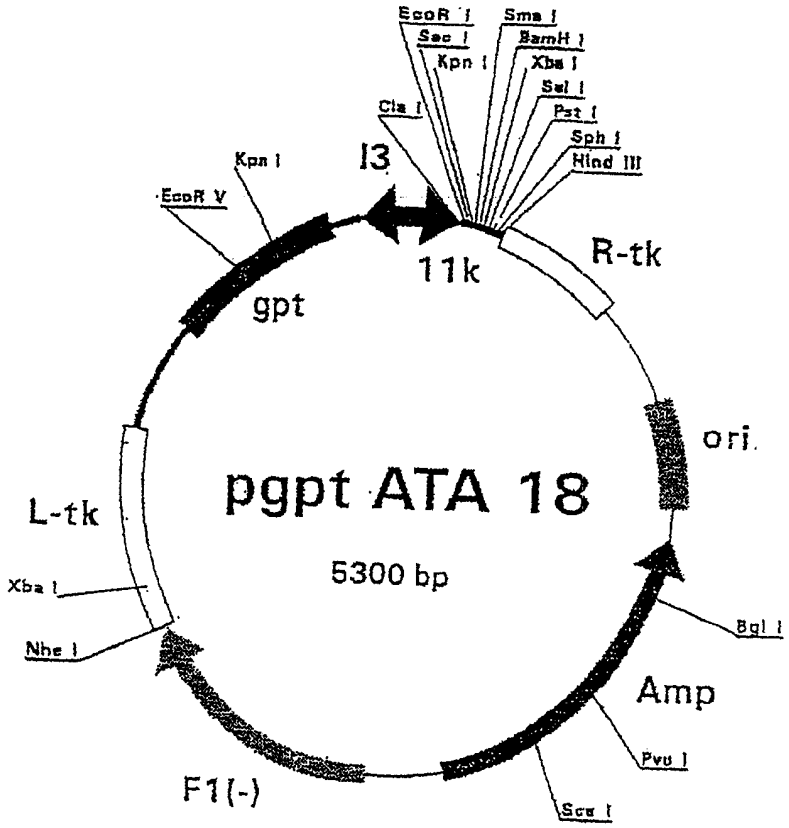

EcoR I
Sma I
Sac I
BamH I
Kpn I
Xba I
Sal I
Cla I
Pst I
Sph I
Hind III
I3
Kpn I
EcoR V
11k
R-tk
gpt
ori
pgpt ATA 18
5300 bp
L-tk
Xba I
Bgl I
Nhe I
Amp
Pvu I
F1(-)
Sca I

图 1

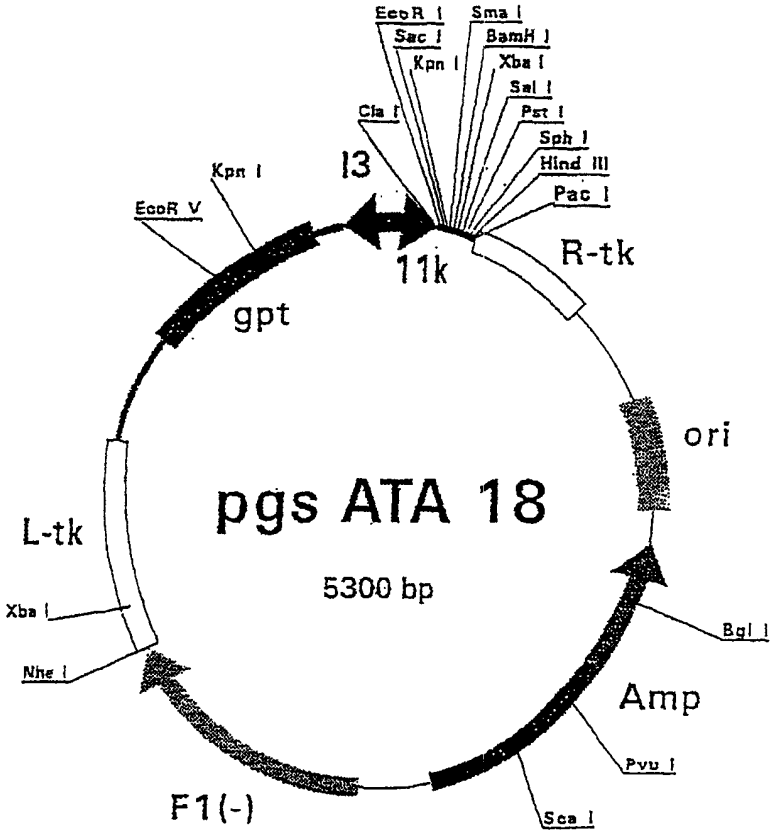

EcoR I
Sma I
Sac I
BamH I
Kpn I
Xba I
Sal I
Cla I
Pst I
Sph I
Hind III
Pac I
I3
Kpn I
EcoR V
11k
R-tk
gpt
ori
pgs ATA 18
5300 bp
L-tk
Xba I
Bgl I
Nhe I
Amp
Pvu I
F1(-)
Sca I

图 2

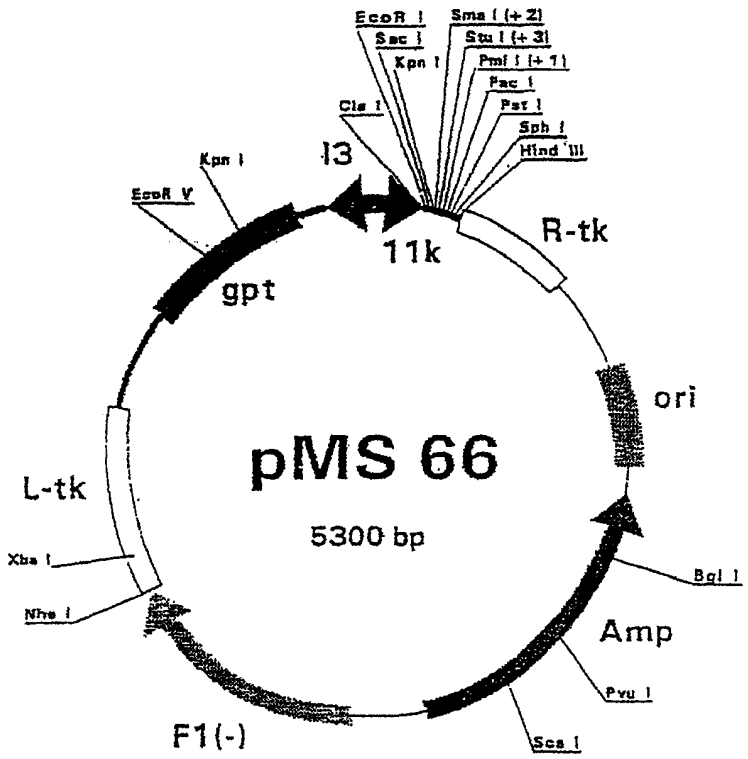

EcoR I
Sma I (+ 2)
Sac I
Stu I (+ 3)
Kpn I
Pmi I (+ 1)
Pac I
Cla I
Pst I
Sph I
Hind III
I3
Kpn I
EcoR V
11k
R-tk
gpt
ori
pMS 66
5300 bp
L-tk
Xba I
Bgl I
Nhe I
Amp
Pvu I
F1(-)
Sca I

图 3

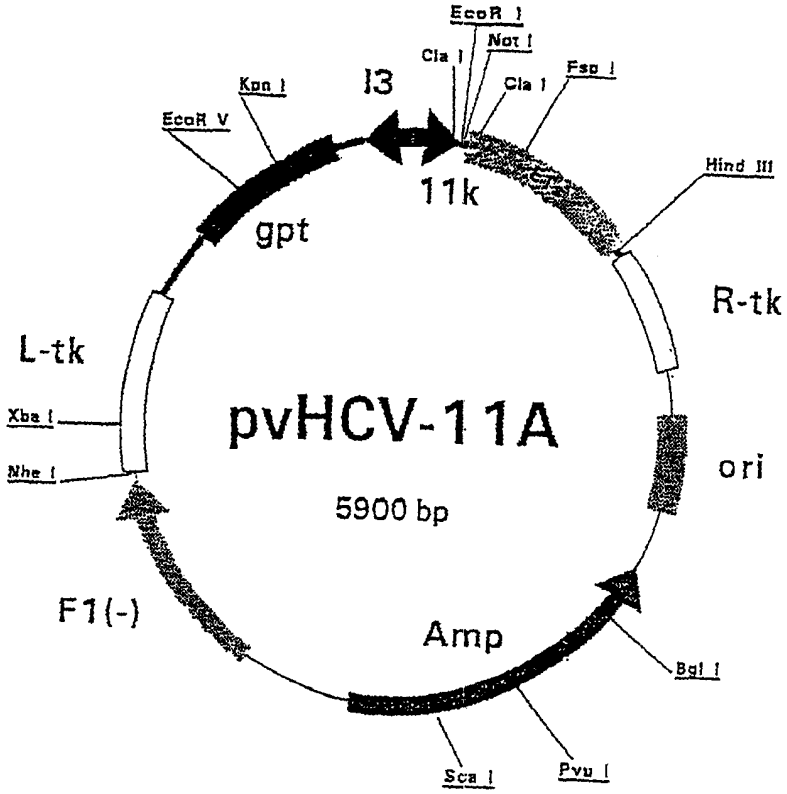

EcoR I
Not I
Cla I
I3
Cla I
Fsp I
Kpn I
EcoR V
Hind III
11k
gpt
R-tk
L-tk
pvHCV-11A
Xba I
5900 bp
Nhe I
ori
F1(-)
Amp
Bgl I
Sca I
Pvu I

图 4

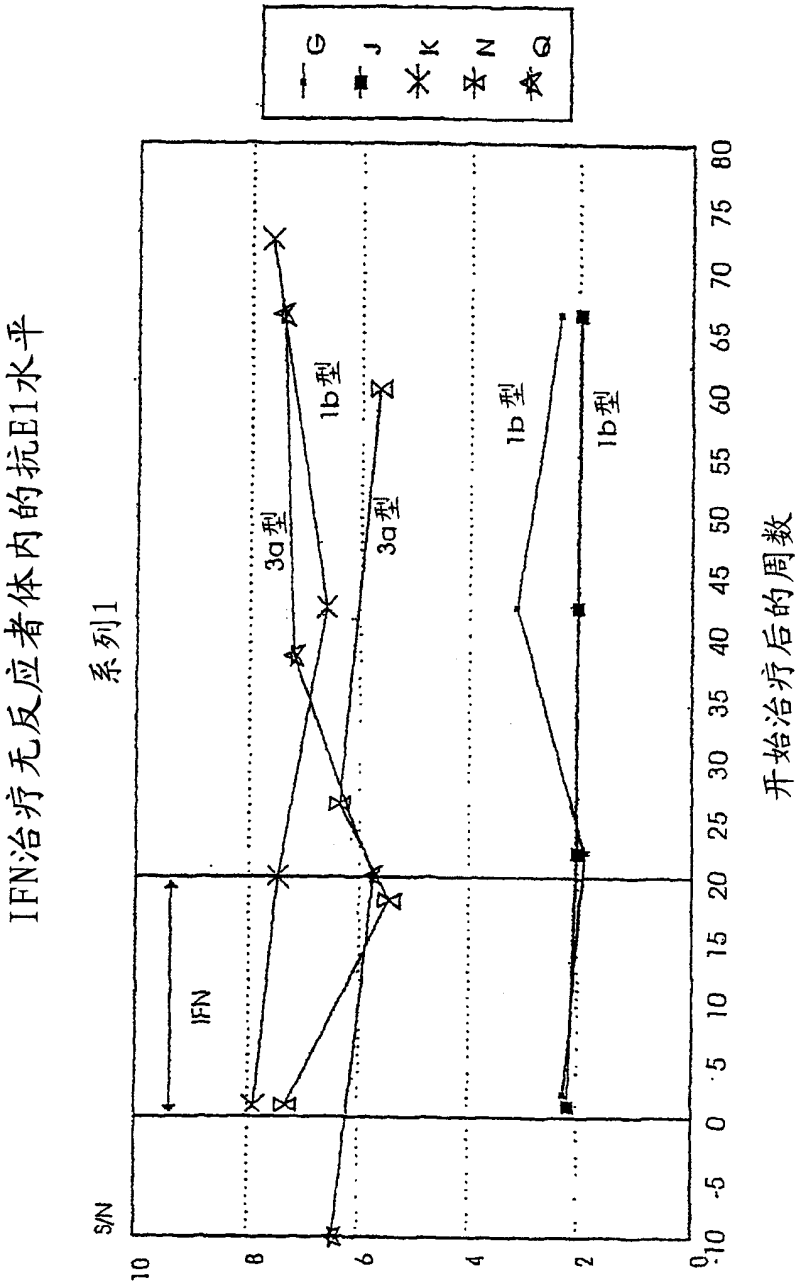

IFN治疗无反应者体内的抗E1水平
S/N
系列1
IFN
3a型
1b型
3a型
1b型
1b型
G
J
K
N
Q
开始治疗后的周数

图 5

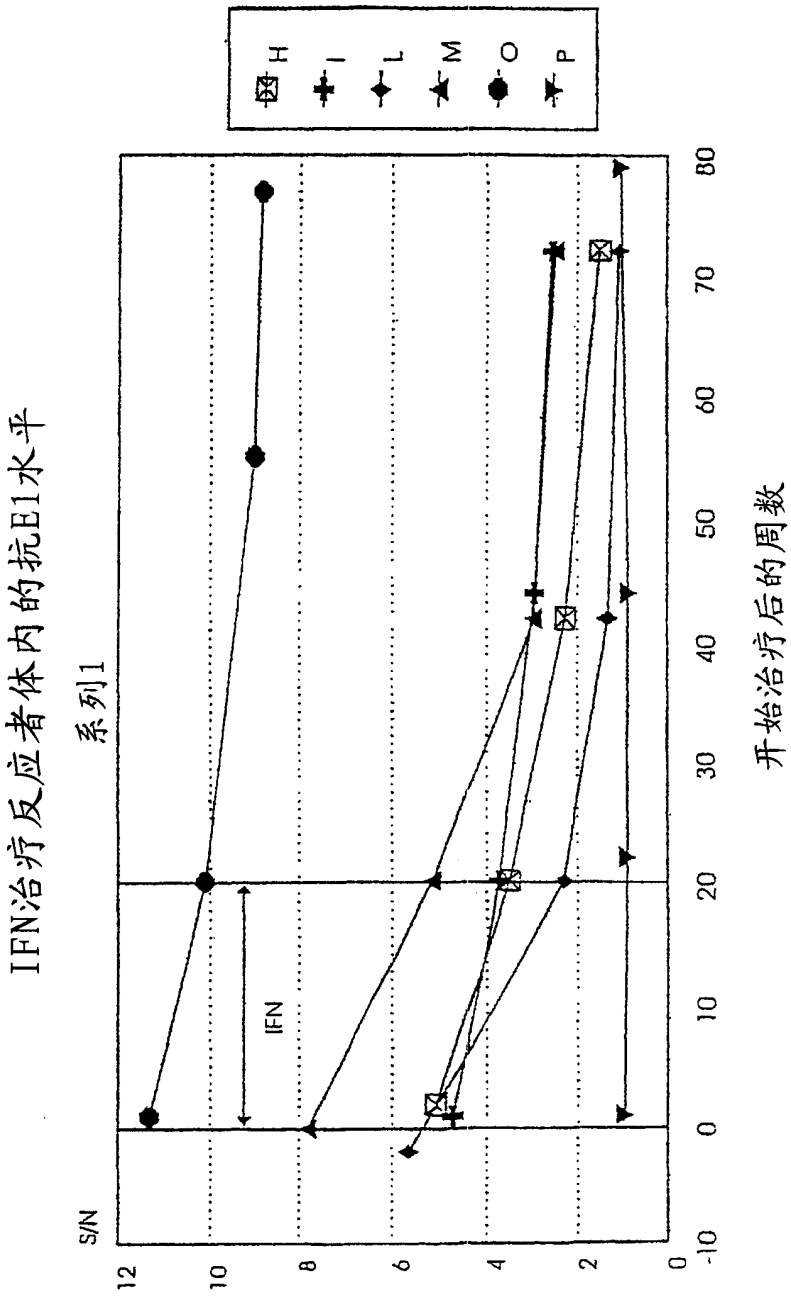

IFN治疗反应者体内的抗E1水平
系列1
S/N
12
10
8
6
4
2
0
-10
0
10
20
30
40
50
60
70
80
开始治疗后的周数
IFN
H
I
L
M
O
P

图 6

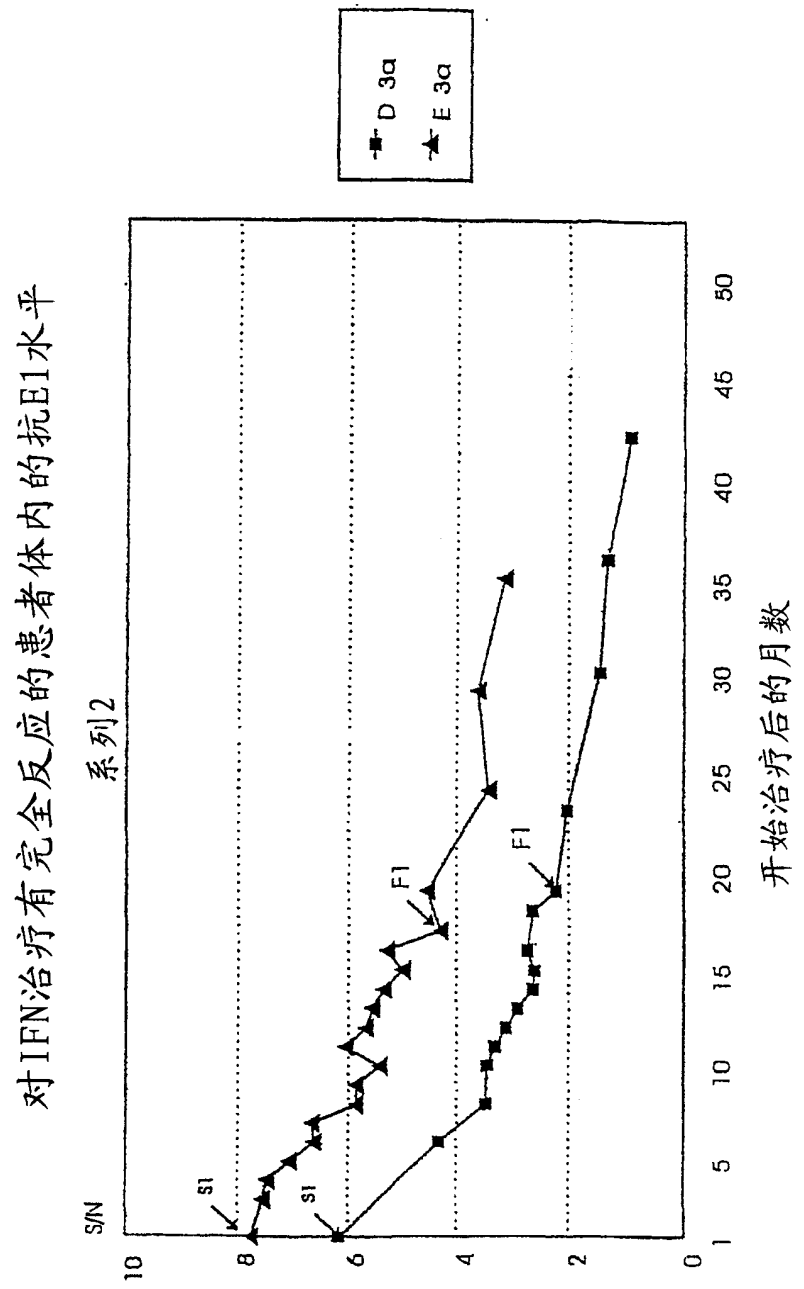

对IFN治疗有完全反应的患者体内的抗E1水平
系列2
D 3a
E 3a
S/N
10
8
6
4
2
0
1
5
10
15
20
25
30
35
40
45
50
S1
S1
F1
F1
开始治疗后的月数

图 7

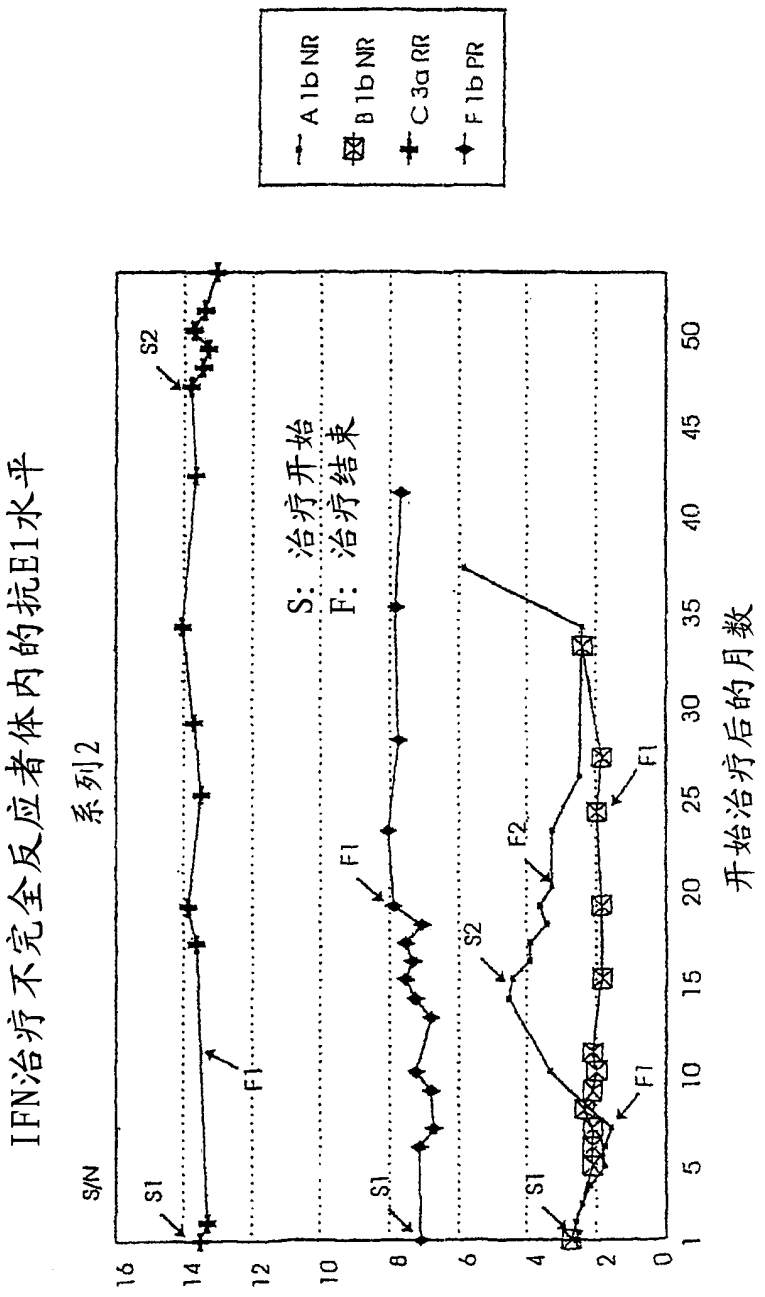

IFN治疗不完全反应者体内的抗E1水平
S/N
系列2
A 1b NR
B 1b NR
C 3a RR
F 1b PR
S: 治疗开始
F: 治疗结束
S1
S2
F1
F2
开始治疗后的月数
0
2
4
6
8
10
12
14
16
1
5
10
15
20
25
30
35
40
45
50

图 8

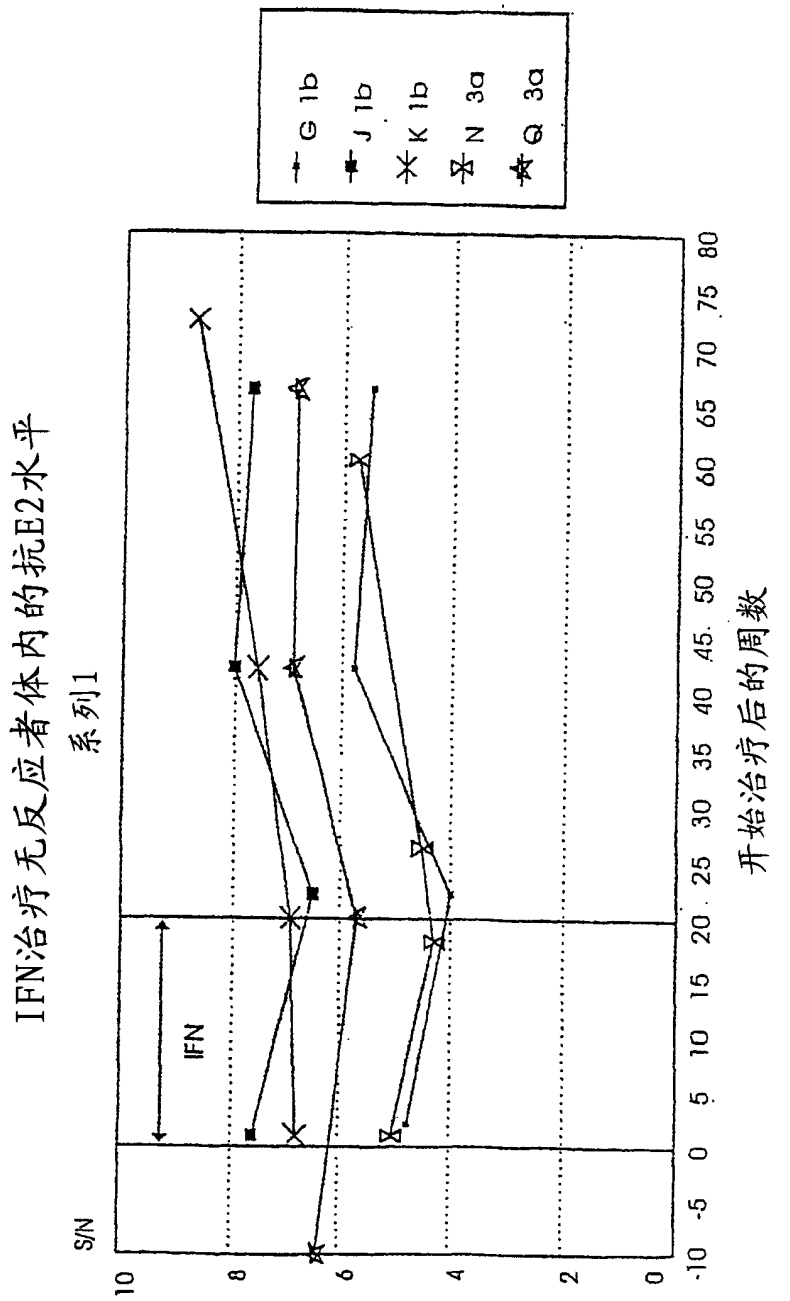

IFN治疗无反应者体内的抗E2水平
系列1
S/N
IFN
开始治疗后的周数
G 1b
J 1b
K 1b
N 3a
Q 3a
10
8
6
4
2
0
-10
-5
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

图 9

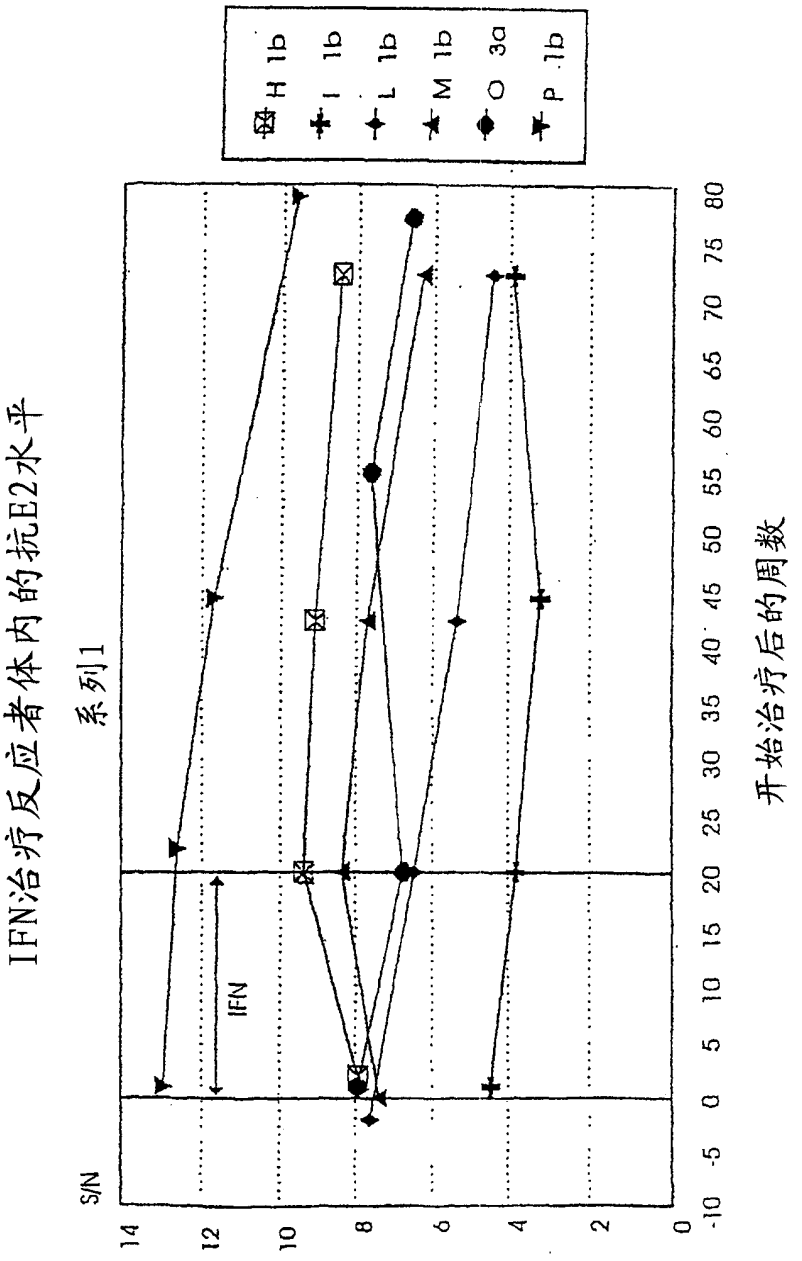

IFN治疗反应者体内的抗E2水平
系列1
S/N
IFN
开始治疗后的周数
H 1b
I 1b
L 1b
M 1b
O 3a
P 1b

图 10

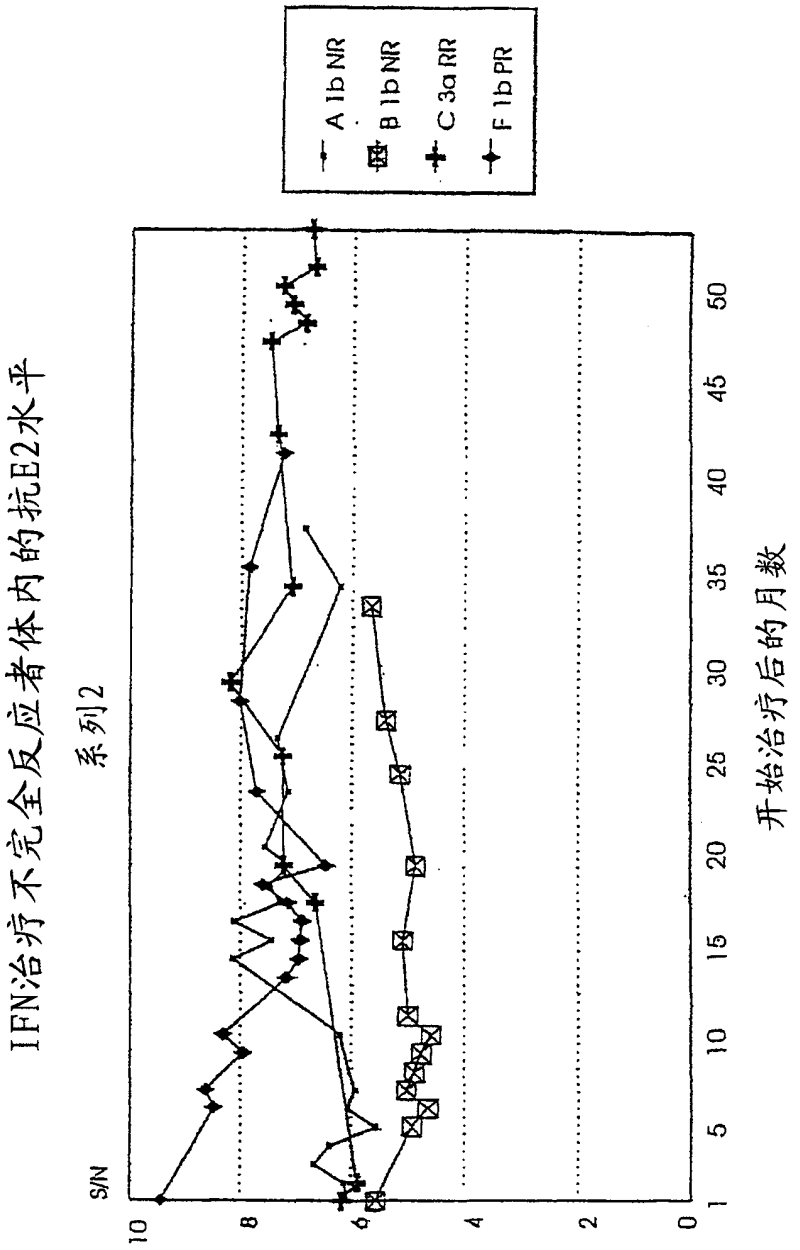

IFN治疗不完全反应者体内的抗E2水平
系列2
S/N
10
8
6
4
2
0
1
5
10
15
20
25
30
35
40
45
50
开始治疗后的月数
A 1b NR
B 1b NR
C 3a RR
F 1b PR

图 11

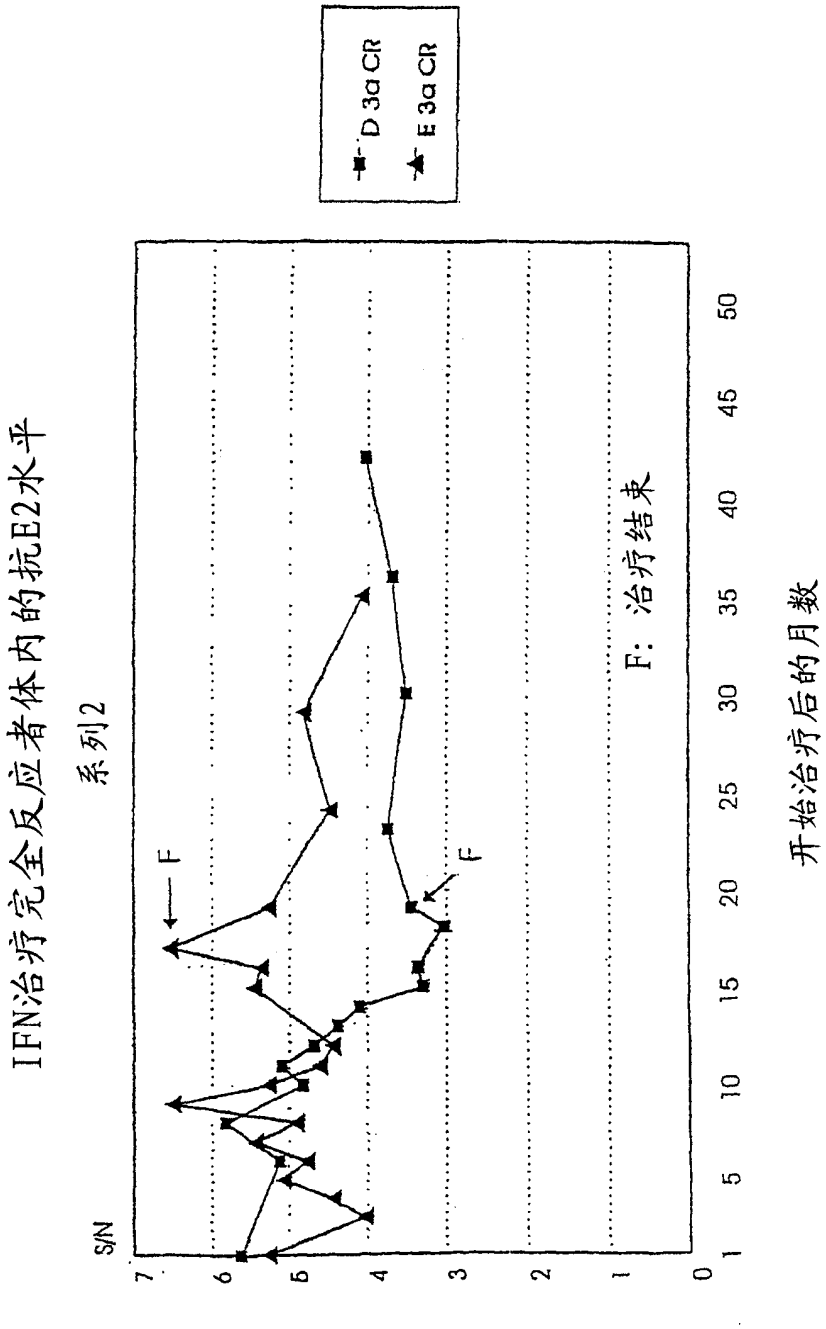

IFN治疗完全反应者体内的抗E2水平
系列2
D 3a CR
E 3a CR
S/N
F
F
F: 治疗结束
开始治疗后的月数
0
1
2
3
4
5
6
7
1
5
10
15
20
25
30
35
40
45
50

图 12

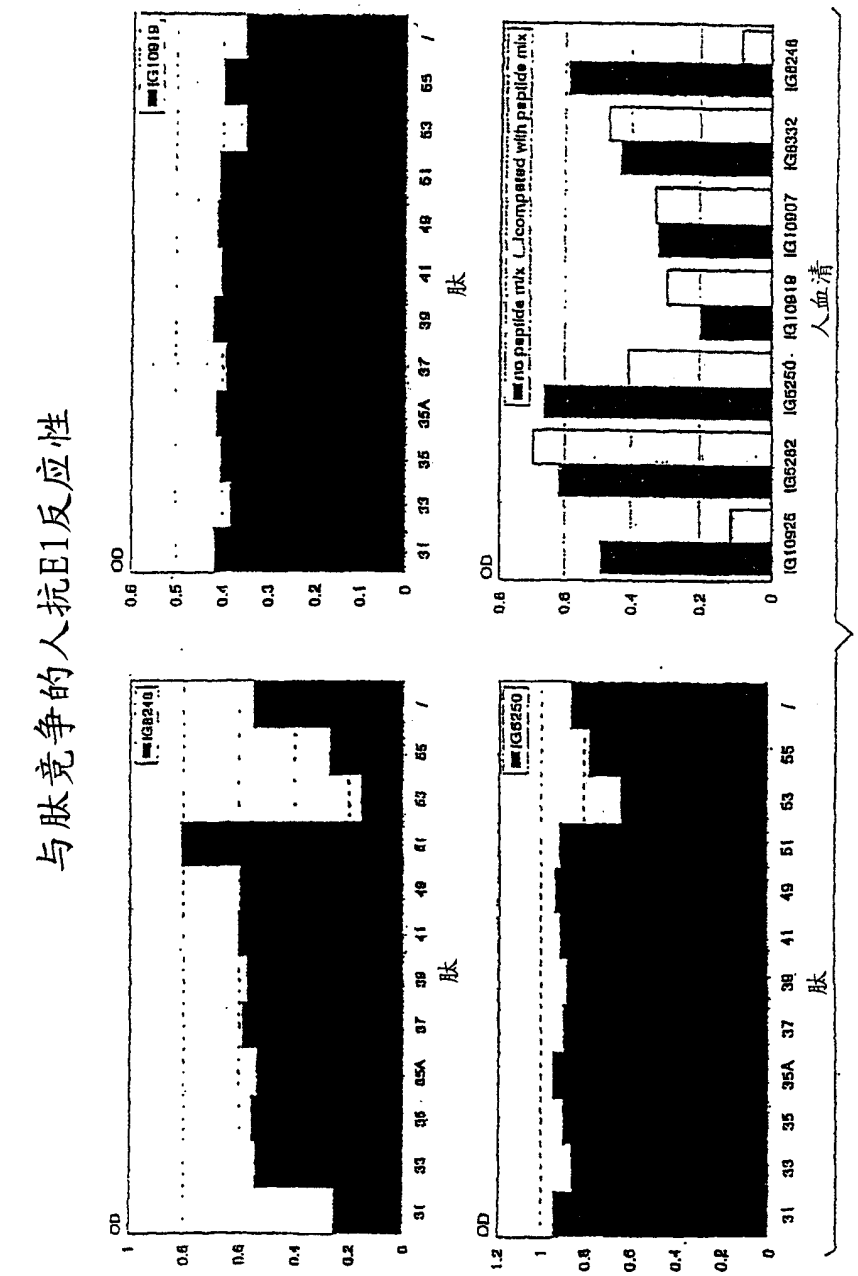

与肽竞争的人抗E1反应性
OD
肽
IG10918
IG8240
IG8250
no peptide mix
compated with peptide mix
IG10926
IG5282
IG5250
IG10918
IG10907
IG8332
IG8248
人血清

图 13

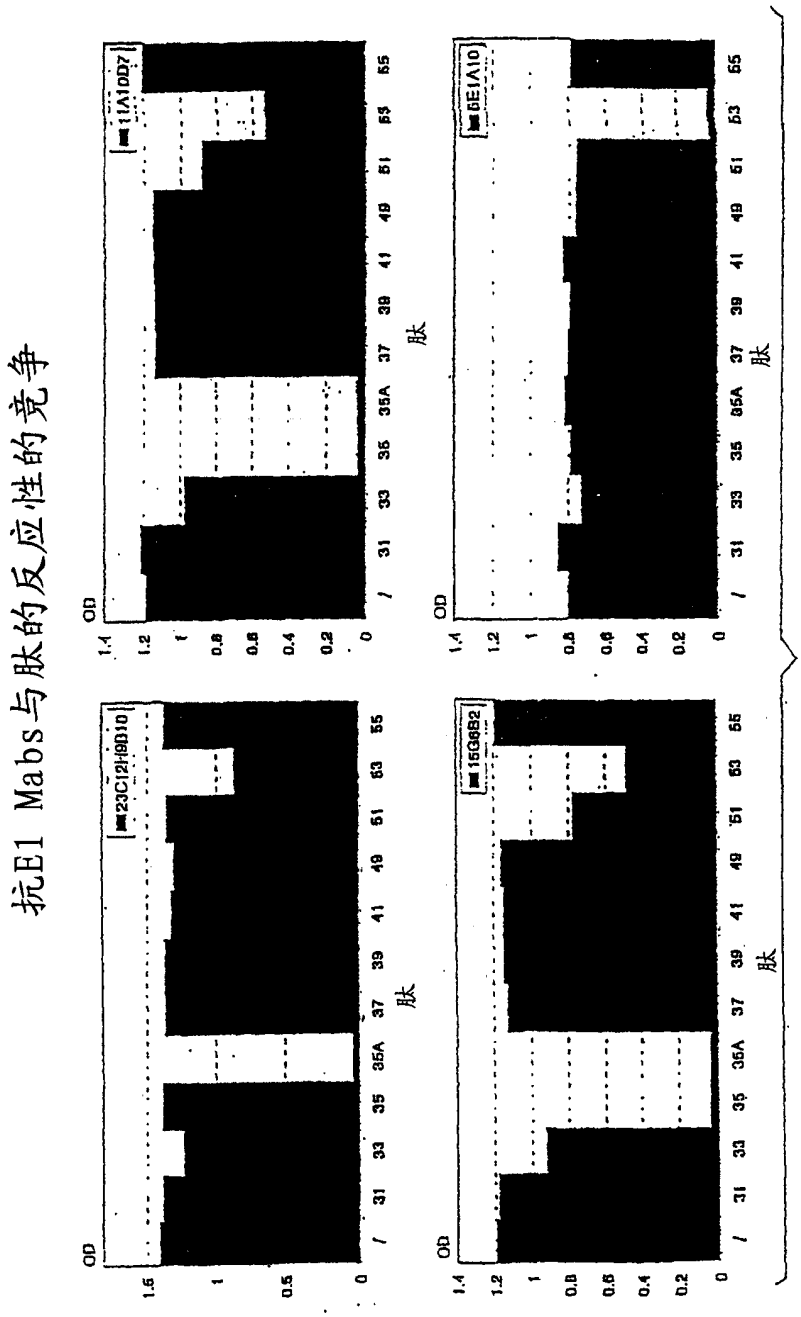


图 14

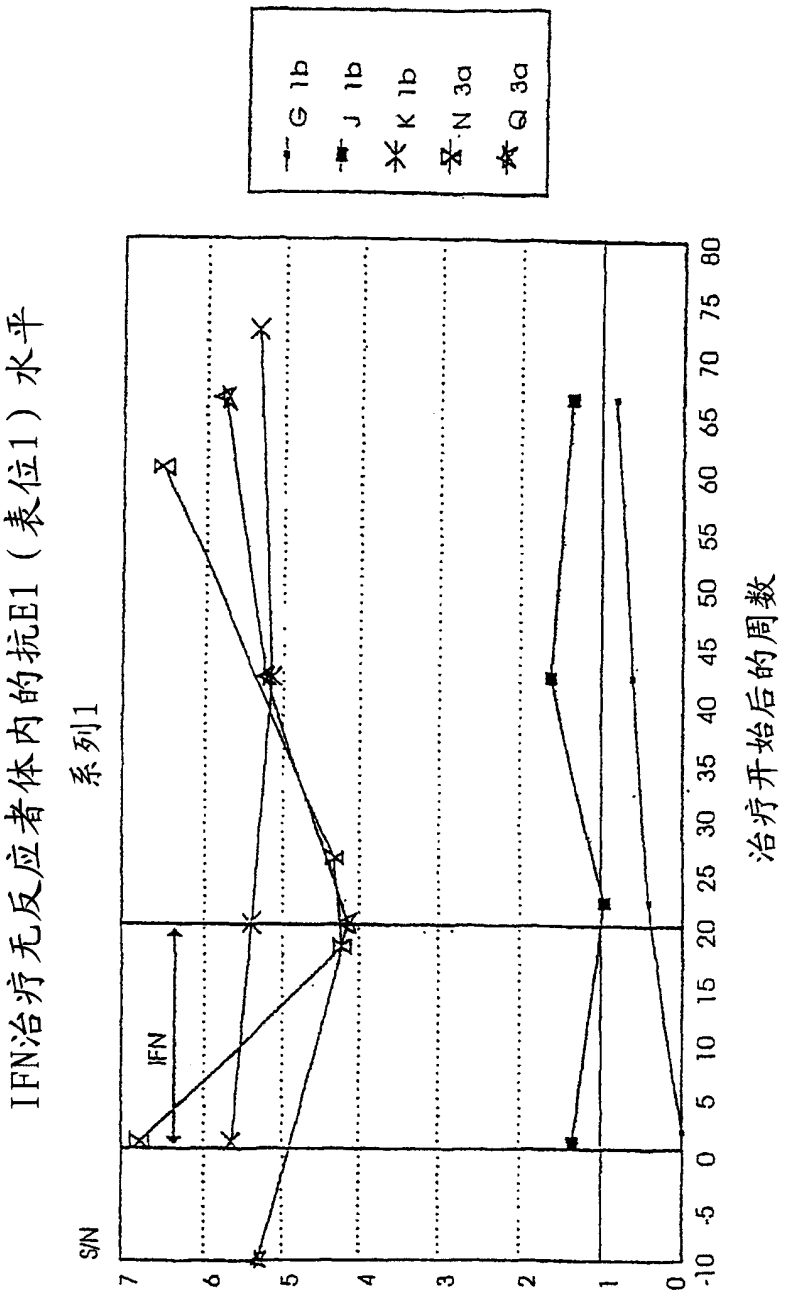

IFN治疗无反应者体内的抗E1（表位1）水平
系列1
S/N
IFN
治疗开始后的周数
G 1b
J 1b
K 1b
N 3a
Q 3a

图 15

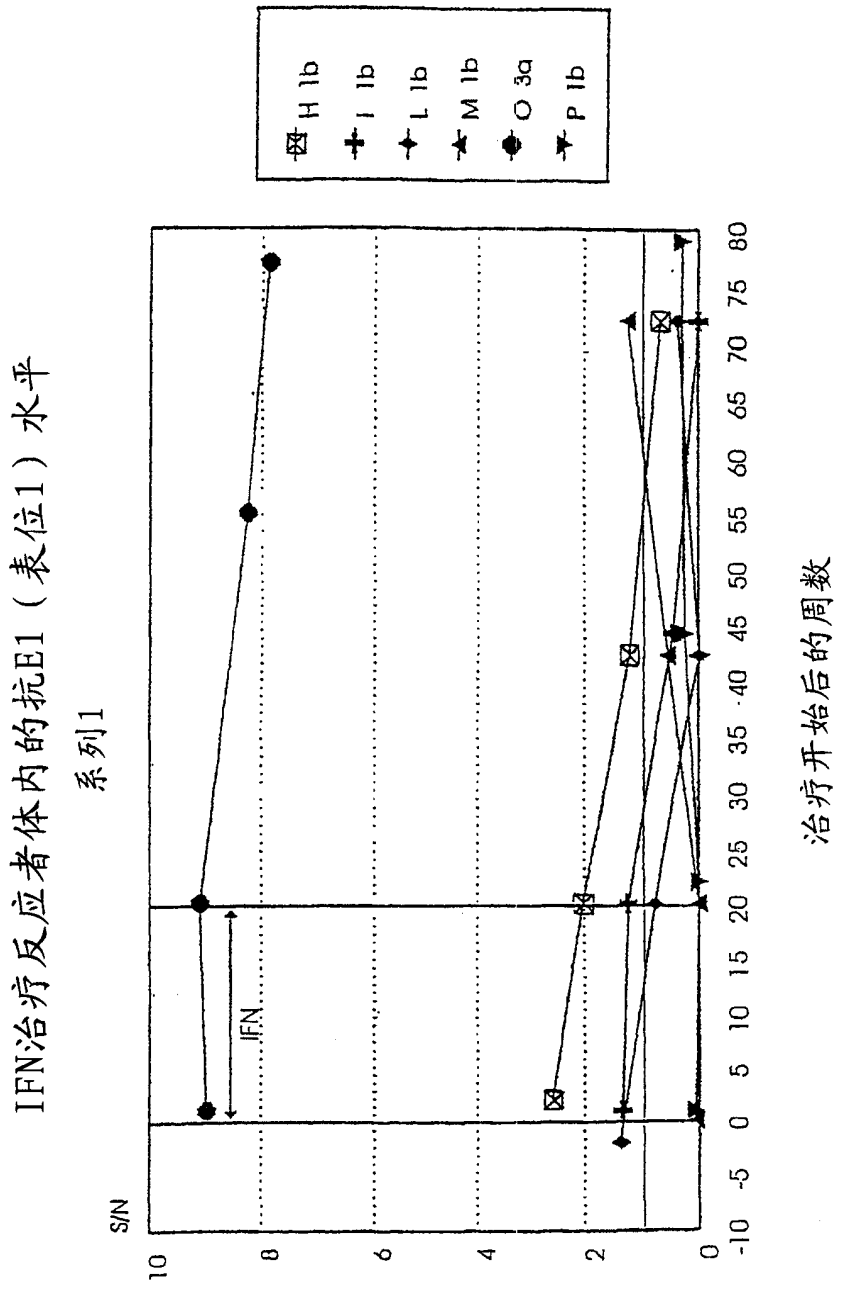

IFN治疗反应者体内的抗E1（表位1）水平
系列1
S/N
10
8
6
4
2
0
-10 -5 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
治疗开始后的周数
IFN
H 1b
I 1b
L 1b
M 1b
O 3a
P 1b

图 16

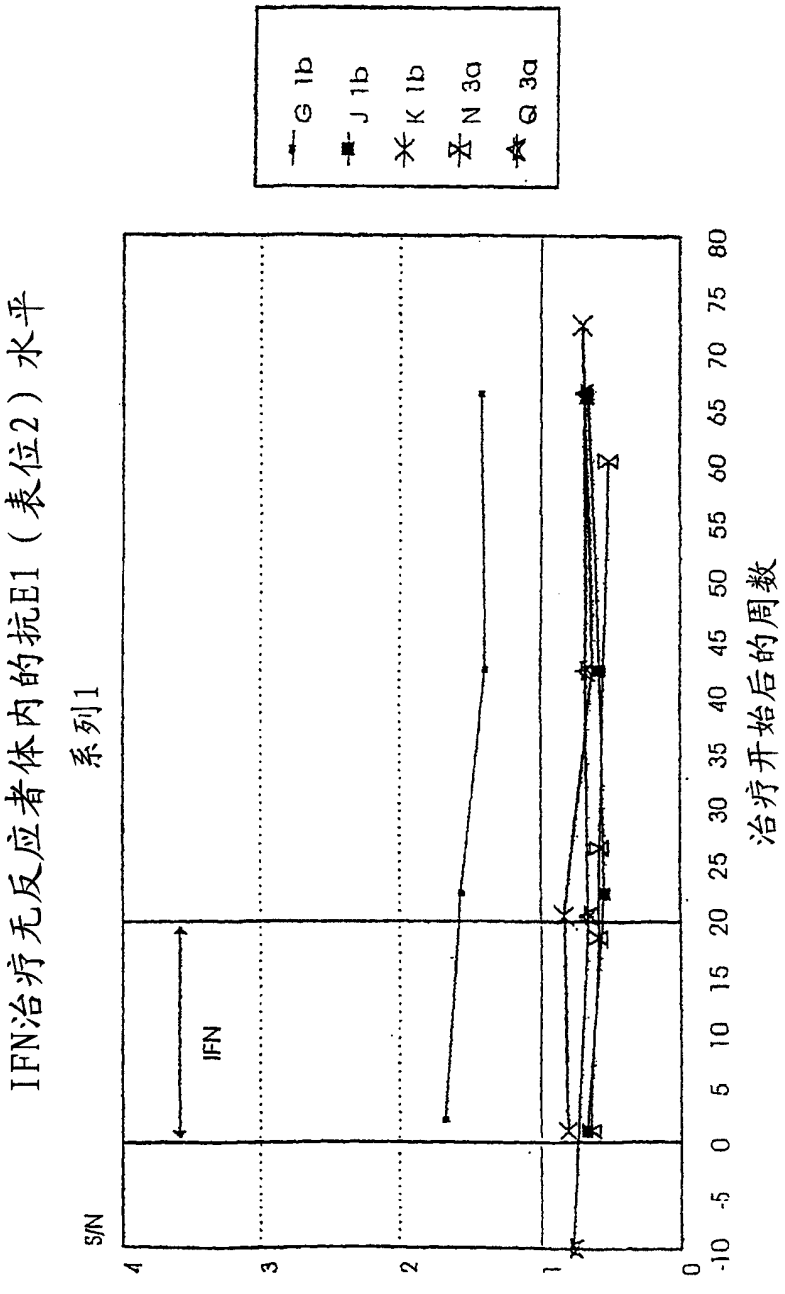

IFN治疗无反应者体内的抗E1（表位2）水平
系列1
S/N
IFN
4
3
2
1
0
-10 -5 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
治疗开始后的周数
G 1b
J 1b
K 1b
N 3a
Q 3a

图 17

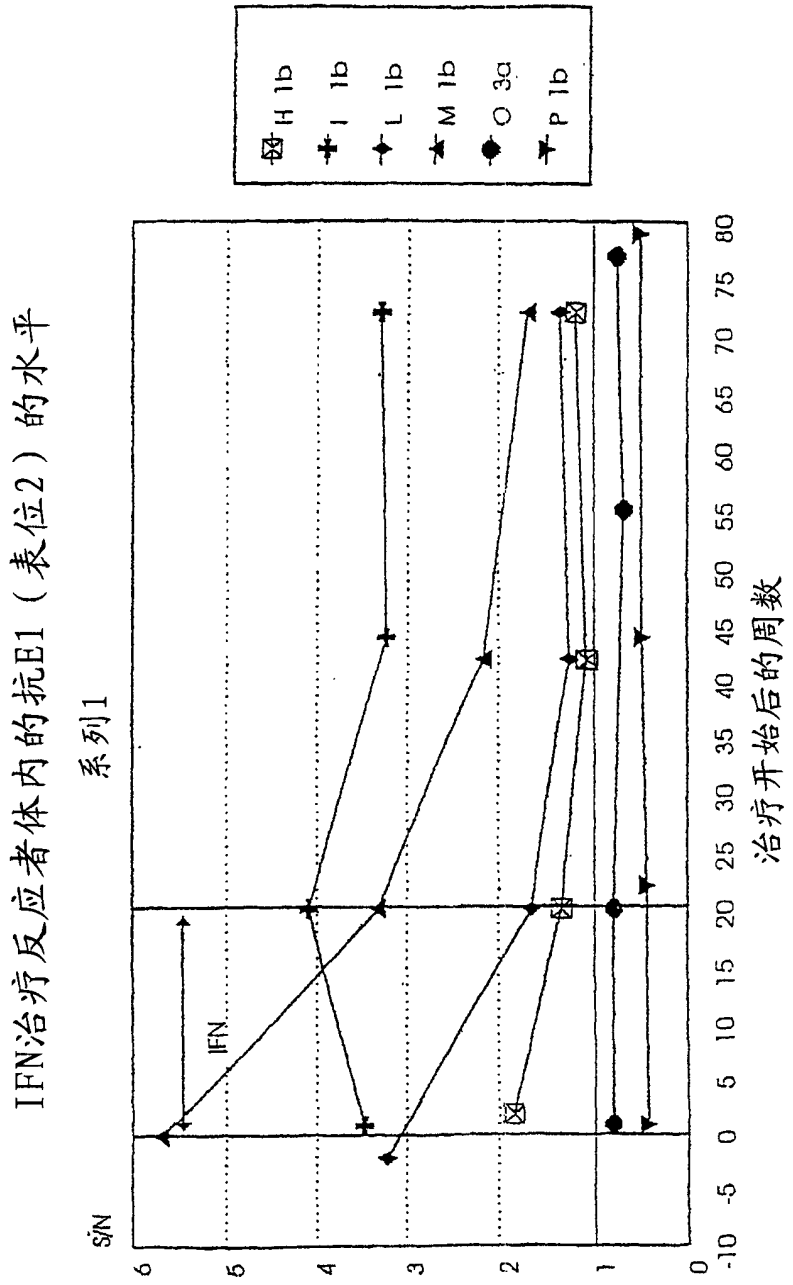

IFN治疗反应者体内的抗E1（表位2）的水平
系列1
S/N
IFN
治疗开始后的周数
H 1b
I 1b
L 1b
M 1b
O 3a
P 1b

图 18

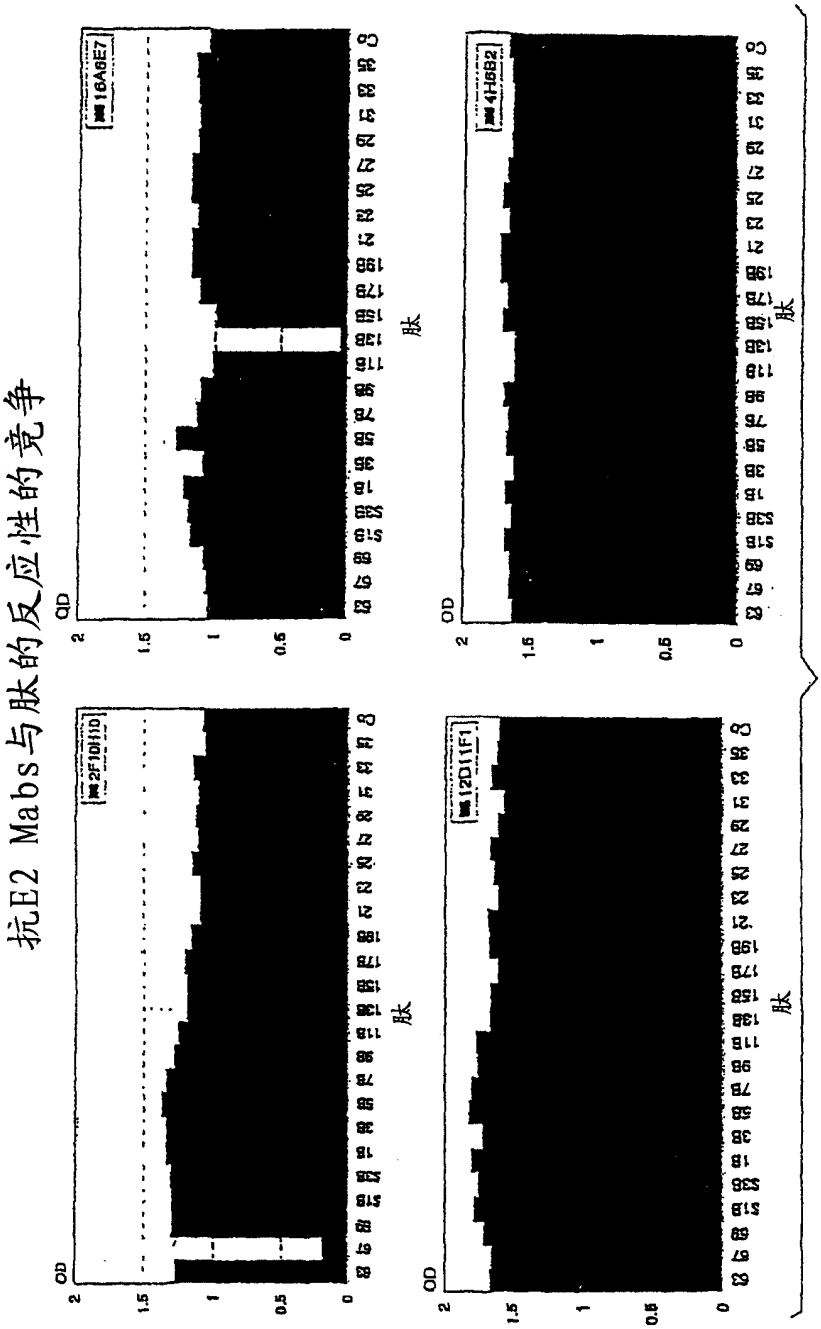


图 19

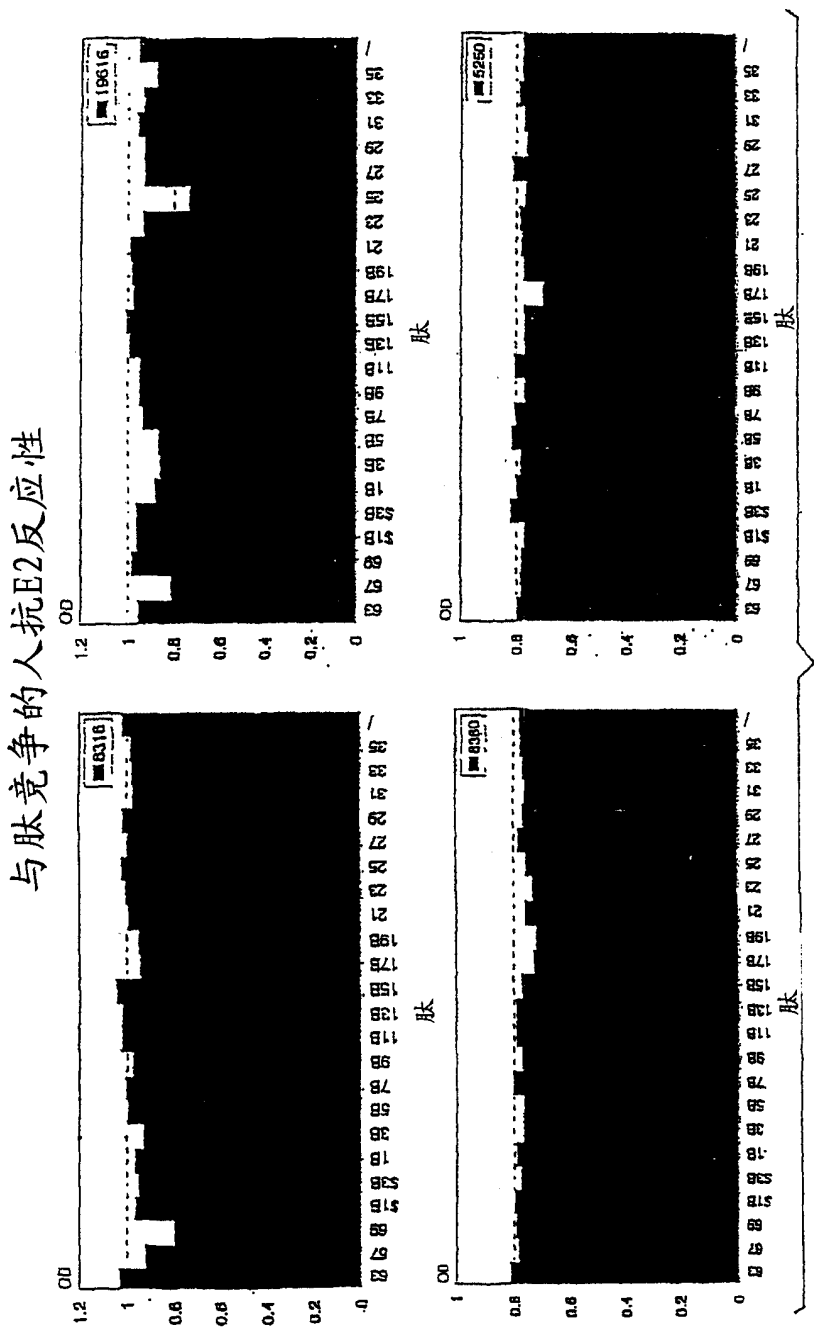


图 20

5' GGCATGCAAGCTTAATTAATT3' (SEQ ID NO 1)
3'ACGTCCGTACGTTCGAATTAATTAATCGA5' (SEQ ID NO 94)

5'CCGGGGAGGCCTGCACGTGATCGAGGGCAGACACCATCACCACCATCACTAATAGT
TAATTAACTGCA 3' (SEQ ID NO 2)
3'CCTCCGGACGTGCACTAGCTCCCGTCTGTGGTAGTGGTGGTAGTGATTATCAATTAATTG
5' (SEQ ID NO 95)

SEQ ID NO 3 (HCCI9A)
ATGCCCGGTTGCTCTTTCTCTATCTTCCTCTTGGCTTTACTGTCCTGTCTGACCATTCCA
GCTTCCGCTTATGAGGTGCGCAACGTGTCCGGGATGTACCATGTCACGAACGACTGCT
CCAACTCAAGCATTGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGT
GCCCTGCGTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTC
GCAGCTAGGAACGCCAGCGTCCCCACCACGACAATACGACGCCACGTCGATTTGCTCG
TTGGGGCGGCTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTCTTC
CTCGTCTCCCAGCTGTTCACCATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCA
ATTGCTCAATCTATCCCGGCCACATAACAGGTCACCGTATGGCTTGGGATATGATGAT
GAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCCACAAGCT
GTCGTGGACATGGTGGCGGGGGCCCATTGGGGAGTCCTGGCGGGCCTCGCCTACTATT
CCATGGTGGGGAACTGGGCTAAGGTTTTGATTGTGATGCTACTCTTTGCTCTCTAATAG

SEQ ID NO 5 (HCCI10A1)
ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCT
ATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCG
CAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTAT
GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAAC
AACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCG
TCCCCACCACGACAATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTG

图 21A

TTCCGCTATGTACGTGGGGGACCTCTGCGGATCTGTCTTCCTCGTCTCCCAGCTGTTCA
CCATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGG
CCACATAACGGGTCACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACG
GCCCTGGTGGTATCGCAGCTGCTCCGGATCCCACAAGCTGTCGTGGACATGGTGGCGG
GGGCCCATTGGGGAGTCCTGGCGGGTCTCGCCTACTATTCCATGGTGGGGAACTGGGC
TAAGGTTTTGATTGTGATGCTACTCTTTGCTCCCTAATAG

SEQ ID NO 7 (HCCI11A)

ATGTTGGGTAAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGTGCTGCCAGAGCCCTGGCGCATGGCGTCCG
GGTTCTGGAAGACGGCGTGAACTATGCAACAGGGAATTTGCCTGGTTGCTCTTTCTCTA
TCTTCCTCTTGGCTTTACTGTCCTGTCTGACCATTCCAGCTTCCGCTTATGAGGTGCGC
AACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTATG
AGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAACA
ACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCGT
CCCCACTACGACAATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTT
CCGCTATGTACGTGGGGGATCTCTGCGGATCTGTCTTCCTCGTCTCCCAGCTGTTCACC
ATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCC
ACATAACAGGTCACCGTATGGCTTGGGATATGATGATGAACTGGTAATAG

SEQ ID NO 9 (HCCI12A)

ATGCCCGGTTGCTCTTTCTCTATCTTCCTCTTGGCCCTGCTGTCCTGTCTGACCATACCA
GCTTCCGCTTATGAAGTGCGCAACGTGTCCGGGGTGTACCATGTCACGAACGACTGCT
CCAACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGT
GCCCTGCGTTCGGGAGGGCAACTCCTCCCGTTGCTGGGTGGCGCTCACTCCCACGCTC
GCGGCCAGGAACGCCAGCGTCCCCACAACGACAATACGACGCCACGTCGATTTGCTC
GTTGGGGCTGCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTT
CCTTGTTTCCCAGCTGTTCACCTTCTCACCTCGCCGGCATCAAACAGTACAGGACTGCA
ACTGCTCAATCTATCCCGGCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGAT
GAACTGGTCCTAATAG

SEQ ID NO 11 (HCCI13A)

ATGTCCGGTTGCTCTTTCTCTATCTTCCTCTTGGCCCTGCTGTCCTGTCTGACCATACCA
GCTTCCGCTTATGAAGTGCGCAACGTGTCCGGGGTGTACCATGTCACGAACGACTGCT
CCAACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGT

图 21B

GCCCTGCGTTCGGGAGGGCAACTCCTCCCGTTGCTGGGTGGCGCTCACTCCCACGCTC
GCGGCCAGGAACGCCAGCGTCCCCACAACGACAATACGACGCCACGTCGATTTGCTC
GTTGGGGCTGCTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTT
CCTTGTTTCCCAGCTGTTCACCTTCTCACCTCGCCGGCATCAAACAGTACAGGACTGCA
ACTGCTCAATCTATCCCGGCCATGTATCAGGTCACCGCATGGCTTGGGATATGATGAT
GAACTGGTAATAG

SEQ ID NO 13 (HCCl17A)
ATGCTGGGTAAGGCCATCGATACCCTTACGTGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAAGACGGCGTGAACTATGCAACAGGGAATTTGCCTGGTTGCTCTTTCTCTA
TCTTCCTCTTGGCTTTACTGTCCTGTCTAACCATTCCAGCTTCCGCTTACGAGGTGCGC
AACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTATG
AGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAACA
ACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCGGCTAGGAACGCCAGCAT
CCCCACTACAACAATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTGTT
CCGCTATGTACGTGGGGGATCTCTGCGGATCTGTCTTCCTCGTCTCCCAGCTGTTCACC
ATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCC
ACATAACGGGTCACCGTATGGCTTGGGATATGATGATGAACTGGTACTAATAG

SEQ ID NO 15 (HCPr51)
ATGCCCGGTTGCTCTTTCTCTATCTT

SEQ ID NO 16 (HCPr52)
ATGTTGGGTAAGGTCATCGATACCCT

SEQ ID NO 17 (HCPr53)
CTATTAGGACCAGTTCATCATCATATCCCA

SEQ ID NO 18 (HCPr54)
CTATTACCAGTTCATCATCATATCCCA

SEQ ID NO 19 (HCPr107)
ATACGACGCCACGTCGATTCCCAGCTGTTCACCATC

图 21C

SEQ ID NO 20 (HCPr108)

GATGGTGAACAGCTGGGAATCGACGTGGCGTCGTAT

SEQ ID NO 21 (HCCI37)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCT
ATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCG
CAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTAT
GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAAC
AACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCG
TCCCCACCACGACAATACGACGCCACGTCGATTCCCAGCTGTTCACCATCTCGCCTCG
CCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGT
CACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTAT
CGCAGCTGCTCCGGATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGCCCATTGGGG
AGTCCTGGCGGGTCTCGCCTACTATTCCATGGTGGGGAACTGGGCTAAGGTTTTGATTG
TGATGCTACTCTTTGCTCCCTAATAG

SEQ ID NO 23 (HCCI38)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCT
ATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCG
CAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTAT
GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAAC
AACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCG
TCCCCACCACGACAATACGACGCCACGTCGATTCCCAGCTGTTCACCATCTCGCCTCG
CCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGT
CACCGTATGGCTTGGGATATGATGATGAACTGGTAA
TAG

SEQ ID NO 25 (HCCI39)

ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCT

图 21D

ATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCG
CAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTAT
GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAAC
AACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCG
TCCCCACCACGACAATACGACGCCACGTCGATTCCCAGCTGTTCACCATCTCGCCTCG
CCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGT
CACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTAT
CGCAGCTGCTCCGGATCCTCTAATAG

SEQ ID NO 27 (HCCI40)
ATGTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCT
ATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCG
CAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTAT
GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAAC
AACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCG
TCCCCACCACGACAATACGACGCCACGTCGATTCCCAGCTGTTCACCATCTCGCCTCG
CCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGGCCACATAACGGGT
CACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTAT
CGCAGCTGCTCCGGATCGTGATCGAGGGCAGACACCATCACCACCATCACTAATAG

SEQ ID NO 29 (HCCI62)
ATGGGTAAGGTCATCGATACCCTTACGTGCGGATTCGCCGATCTCATGGGGTACATCC
CGCTCGTCGGCGCTCCCGTAGGAGGCGTCGCAAGAGCCCTTGCGCATGGCGTGAGGGC
CCTTGAAGACGGGATAAATTTCGCAACAGGGAATTTGCCCGGTTGCTCCTTTTCTATTT
TCCTTCTCGCTCTGTTCTCTTGCTTAATTCATCCAGCAGCTAGTCTAGAGTGGCGGAAT
ACGTCTGGCCTCTATGTCCTTACCAACGACTGTTCCAATAGCAGTATTGTGTACGAGGC
CGATGACGTTATTCTGCACACACCCGGCTGCATACCTTGTGTCCAGGACGGCAATACA
TCCACGTGCTGGACCCCAGTGACACCTACAGTGGCAGTCAAGTACGTCGGAGCAACCA
CCGCTTCGATACGCAGTCATGTGGACCTATTAGTGGGCGCGGCCACGATGTGCTCTGC
GCTCTACGTGGGTGACATGTGTGGGGCTGTCTTCCTCGTGGGACAAGCCTTCACGTTCA
GACCTCGTCGCCATCAAACGGTCCAGACCTGTAACTGCTCGCTGTACCCAGGCCATCT
TTCAGGACATCGAATGGCTTGGGATATGATGATGAACTGGTAATAG

图 21E

SEQ ID NO 31 (HCCI63)

ATGGGTAAGGTCATCGATACCCTAACGTGCGGATTCGCCGATCTCATGGGGTATATCC
CGCTCGTAGGCGGCCCCATTGGGGGCGTCGCAAGGGCTCTCGCACACGGTGTGAGGGT
CCTTGAGGACGGGGTAAACTATGCAACAGGGAATTTACCCGGTTGCTCTTTCTCTATCT
TTATTCTTGCTCTTCTCTCGTGTCTGACCGTTCCGGCCTCTGCAGTTCCCTACCGAAATG
CCTCTGGGATTTATCATGTTACCAATGATTGCCCAAACTCTTCCATAGTCTATGAGGCA
GATAACCTGATCCTACACGCACCTGGTTGCGTGCCTTGTGTCATGACAGGTAATGTGA
GTAGATGCTGGGTCCAAATTACCCCTACACTGTCAGCCCCGAGCCTCGGAGCAGTCAC
GGCTCCTCTTCGGAGAGCCGTTGACTACCTAGCGGGAGGGGCTGCCCTCTGCTCCGCG
TTATACGTAGGAGACGCGTGTGGGGCACTATTCTTGGTAGGCCAAATGTTCACCTATA
GGCCTCGCCAGCACGCTACGGTGCAGAACTGCAACTGTTCCATTTACAGTGGCCATGT
TACCGGCCACCGGATGGCATGGGATATGATGATGAACTGGTAATAG

SEQ ID NO 33 (HCPr109)

TGGGATATGATGATGAACTGGTC

SEQ ID NO 34 (HCPr72)

CTATTATGGTGGTAAKGCCARCARGAGCAGGAG

SEQ ID NO 35 (HCCL22A)

TGGGATATGATGATGAACTGGTCGCCTACAACGGCCCTGGTGGTATCGCAGCTGCTCC
GGATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGCCCATTGGGGAGTCCTGGCGG
GCCTCGCCTACTATTCCATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTC
TTTGCCGGCGTCGACGGGCATACCCGCGTGTCAGGAGGGGCAGCAGCCTCCGATACCA
GGGGCCTTGTGTCCCTCTTTAGCCCCGGGTCGGCTCAGAAAATCCAGCTCGTAAACAC
CAACGGCAGTTGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAAC
AGGGTTCTTTGCCGCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCCAGAG
CGCTTGGCCAGCTGTCGCTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTT
ACACTGAGCCTAACAGCTCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACC
GTGTGGTATTGTACCCGCGTCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCC
CTGTTGTGGTGGGGACGACCGATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAA
CGACTCGGATGTGCTGATTCTCAACAACACGCGGCCGCCGCGAGGCAACTGGTTCGGC
TGTACATGGATGAATGGCACTGGGTTCACCAAGACGTGTGGGGGCCCCCCGTGCAACA
TCGGGGGGGCCGGCAACAACACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCC
CGAGGCCACCTACGCCAGATGCGGTTCTGGGCCCTGGCTGACACCTAGGTGTATGGTT

图 21F

CATTACCCATATAGGCTCTGGCACTACCCCTGCACTGTCAACTTCACCATCTTCAAGGT
TAGGATGTACGTGGGGGGCGTGGAGCACAGGTTCGAAGCCGCATGCAATTGGACTCG
AGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTGCTG
TCTACAACAGAGTGGCAGATACTGCCCTGTTCCTTCACCACCCTGCCGGCCCTATCCA
CCGGCCTGATCCACCTCCATCAGAACATCGTGGACGTGCAATACCTGTACGGTGTAGG
GTCGGCGGTTGTCTCCCTTGTCATCAAATGGGAGTATGTCCTGTTGCTCTTCCTTCTCCT
GGCAGACGCGCGCATCTGCGCCTGCTTATGGATGATGCTGCTGATAGCTCAAGCTGAG
GCCGCCTTAGAGAACCTGGTGGTCCTCAATGCGGCGGCCGTGGCCGGGGCGCATGGC
ACTCTTTCCTTCCTTGTGTTCTTCTGTGCTGCCTGGTACATCAAGGGCAGGCTGGTCCC
TGGTGCGGCATACGCCTTCTATGGCGTGTGGCCGCTGCTCCTGCTTCTGCTGGCCTTAC
CACCACGAGCTTATGCCTAGTAA

SEQ ID NO 37 (HCCI41)

GATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGCCCATTGGGGAGTCCTGGCGGG
CCTCGCCTACTATTCCATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCT
TTGCCGGCGTCGACGGGCATACCCGCGTGTCAGGAGGGGCAGCAGCCTCCGATACCA
GGGGCCTTGTGTCCGTCTTTAGCCCCGGGTCGGCTCAGAAAATCCAGCTCGTAAACAC
CAACGGCAGTTGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAAC
AGGGTTCTTTGCCGCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCCAGAG
CGCTTGGCCAGCTGTCGCTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTT
ACACTGAGCCTAACAGCTCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACC
GTGTGGTATTGTACCCGCGTCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCC
CTGTTGTGGTGGGGACGACCGATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAA
CGACTCGGATGTGCTGATTCTCAACAACACGCGGCCGCCGCGAGGCAACTGGTTCGGC
TGTACATGGATGAATGGCACTGGGTTCACCAAGACGTGTGGGGGCCCCCCGTGCAACA
TCGGGGGGGCCGGCAACAACACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCC
CGAGGCCACCTACGCCAGATGCGGTTCTGGGCCCTGGCTGACACCTAGGTGTATGGTT
CATTACCCATATAGGCTCTGGCACTACCCCTGCACTGTCAACTTCACCATCTTCAAGGT
TAGGATGTACGTGGGGGGCGTGGAGCACAGGTTCGAAGCCGCATGCAATTGGACTCG
AGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTGCTG
TCTACAACAGAGTGGCAGAGTGGCAGAGCTTAATTAATTAG

SEQ ID NO 39 (HCCI42)

GATCCCACAAGCTGTCGTGGACATGGTGGCGGGGGCCCATTGGGGAGTCCTGGCGGG
CCTCGCCTACTATTCCATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCT

图 21G

TTGCCGGCGTCGACGGGCATACCCGCGTGTCAGGAGGGGCAGCAGCCTCCGATACCA
GGGGCCTTGTGTCCCTCTTTAGCCCCGGGTCGGCTCAGAAAATCCAGCTCGTAAACAC
CAACGGCAGTTGGCACATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAAC
AGGGTTCTTTGCCGCACTATTCTACAAACACAAATTCAACTCGTCTGGATGCCCAGAG
CGCTTGGCCAGCTGTCGCTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTT
ACACTGAGCCTAACAGCTCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACC
GTGTGGTATTGTACCCGCGTCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCC
CTGTTGTGGTGGGGACGACCGATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAA
CGACTCGGATGTGCTGATTCTCAACAACACGCGGCCGCCGCGAGGCAACTGGTTCGGC
TGTACATGGATGAATGGCACTGGGTTCACCAAGACGTGTGGGGGCCCCCCGTGCAACA
TCGGGGGGGCCGGCAACAACACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCC
CGAGGCCACCTACGCCAGATGCGGTTCTGGGCCCTGGCTGACACCTAGGTGTATGGTT
CATTACCCATATAGGCTCTGGCACTACCCCTGCACTGTCAACTTCACCATCTTCAAGGT
TAGGATGTACGTGGGGGGCGTGGAGCACAGGTTCGAAGCCGCATGCAATTGGACTCG
AGGAGAGCGTTGTGACTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTGCTG
TCTACAACAGGTGATCGAGGGCAGACACCATCACCACCATCACTAATAG

SEQ ID NO 41 (HCCI43)
ATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACG
GGCATACCCGCGTGTCAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCT
CTTTAGCCCCGGGTCGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCAC
ATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCAC
TATTCTACAAACACAAATTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTCG
CTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGC
TCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCG
CGTCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGAC
GACCGATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTG
ATTCTCAACAACACGCGGCCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATG
GCACTGGGTTCACCAAGACGTGTGGGGGCCCCCCGTGCAACATCGGGGGGGCCGGCA
ACAACACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGC
CAGATGCGGTTCTGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGG
CTCTGGCACTACCCCTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGG
GGGCGTGGAGCACAGGTTCGAAGCCGCATGCAATTGGACTCGAGGAGAGCGTTGTGA
CTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGG
CAGAGCTTAATTAATTAG

图 21H

SEQ ID NO 43 (HCC144)

ATGGTGGGGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACG
GGCATACCCGCGTGTCAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCT
CTTTAGCCCCGGGTCGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCAC
ATCAACAGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCAC
TATTCTACAAACACAAATTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTCG
CTCCATCGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGC
TCGGACCAGAGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCG
CGTCTCAGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGAC
GACCGATCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTG
ATTCTCAACAACACGCGGCCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATG
GCACTGGGTTCACCAAGACGTGTGGGGGCCCCCCGTGCAACATCGGGGGGGCCGGCA
ACAACACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGC
CAGATGCGGTTCTGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGG
CTCTGGCACTACCCCTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGG
GGGCGTGGAGCACAGGTTCGAAGCCGCATGCAATTGGACTCGAGGAGAGCGTTGTGA
CTTGGAGGACAGGGATAGATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGGTGAT
CGAGGGCAGACACCATCACCACCATCACTAATAG

SEQ ID NO 45 (HCCL64)

ATGGTGGCGGGGGCCCATTGGGGAGTCCTGGCGGGCCTCGCCTACTATTCCATGGTGG
GGAACTGGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACGGGCATAC
CCGCGTGTCAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCTCTTTAGC
CCCGGGTCGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCACATCAAC
AGGACTGCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCACTATTCT
ACAAACACAAATTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTCGCTCCAT
CGACAAGTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGCTCGGAC
CAGAGGCCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCTC
AGGTGTGCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACCGA
TCGGTTTGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTC
AACAACACGCGGCCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACT
GGGTTCACCAAGACGTGTGGGGGCCCCCCGTGCAACATCGGGGGGGCCGGCAACAAC
ACCTTGACCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGAT
GCGGTTCTGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGG
CACTACCCCTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCG

图 21I

TGGAGCACAGGTTCGAAGCCGCATGCAATTGGACTCGAGGAGAGCGTTGTGACTTGGA
GGACAGGGATAGATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGATA
CTGCCCTGTTCCTTCACCACCCTGCCGGCCCTATCCACCGGCCTGATCCACCTCCATCA
GAACATCGTGGACGTGCAATACCTGTACGGTGTAGGGTCGGCGGTTGTCTCCCTTGTC
ATCAAATGGGAGTATGTCCTGTTGCTCTTCCTTCTCCTGGCAGACGCGCGCATCTGCGC
CTGCTTATGGATGATGCTGCTGATAGCTCAAGCTGAGGCCGCCTTAGAGAACCTGGTG
GTCCTCAATGCGGCGGCCGTGGCCGGGGCGCATGGCACTCTTTCCTTCCTTGTGTTCTT
CTGTGCTGCCTGGTACATCAAGGGCAGGCTGGTCCCTGGTGCGGCATACGCCTTCTAT
GGCGTGTGGCCGCTGCTCCTGCTTCTGCTGGCCTTACCACCACGAGCTTATGCCTAGTAA

SEQ ID NO 47 (HCCI65)

AATTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACA
TTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCG
GGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCTTTCTCT
ATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAAGTGCG
CAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTGTGTAT
GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGAGAAC
AACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCCAGCG
TCCCCACCACGACAATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTTCTG
TTCCGCTATGTACGTGGGGGACCTCTGCGGATCTGTCTTCCTCGTCTCCCAGCTGTTCA
CCATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATCCCGG
CCACATAACGGGTCACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCTACAACG
GCCCTGGTGGTATCGCAGCTGCTCCGGATCCCACAAGCTGTCGTGGACATGGTGGCGG
GGGCCCATTGGGGAGTCCTGGCGGGCCTCGCCTACTATTCCATGGTGGGGAACTGGGC
TAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACGGGCATACCCGCGTGTCAG
GAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCTCTTTAGCCCCGGGTCGGC
TCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCACATCAACAGGACTGCCCT
GAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCACTATTCTACAAACACAAA
TTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTCGCTCCATCGACAAGTTCG
CTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGCTCGGACCAGAGGCCCTA
CTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCTCAGGTGTGCGGT
CCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACCGATCGGTTTGGTGT
CCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTCAACAACACGCGG
CCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGGGTTCACCAAGA
CGTGTGGGGGCCCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTTGACCTGCC

图 21J

CCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTCTGGGCC
CTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTACCCCTGCA
CTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGCACAGGTT
CGAAGCCGCATGCAATTGGACTCGAGGAGAGCGTTGTGACTTGGAGGACAGGGATAG
ATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGATACTGCCCTGTTCC
TTCACCACCCTGCCGGCCCTATCCACCGGCCTGATCCACCTCCATCAGAACATCGTGG
ACGTGCAATACCTGTACGGTGTAGGGTCGGCGGTTGTCTCCCTTGTCATCAAATGGGA
GTATGTCCTGTTGCTCTTCCTTCTCCTGGCAGACGCGCGCATCTGCGCCTGCTTATGGA
TGATGCTGCTGATAGCTCAAGCTGAGGCCGCCTTAGAGAACCTGGTGGTCCTCAATGC
GGCGGCCGTGGCCGGGGCGCATGGCACTCTTTCCTTCCTTGTGTTCTTCTGTGCTGCCT
GGTACATCAAGGGCAGGCTGGTCCCTGGTGCGGCATACGCCTTCTATGGCGTGTGGCC
GCTGCTCCTGCTTCTGCTGGCCTTACCACCACGAGCTTATGCCTAGTAAGCTT

SEQ ID NO 49 (HCCI66)

ATGAGCACGAATCCTAAACCTCAAAGAAAAACCAAACGTAACACCAACCGCCGCCCA
CAGGACGTCAAGTTCCCGGGCGGTGGTCAGATCGTTGGTGGAGTTTACCTGTTGCCGC
GCAGGGGCCCCAGGTTGGGTGTGCGCGCGACTAGGAAGACTTCCGAGCGGTCGCAAC
CTCGTGGGAGGCGACAACCTATCCCCAAGGCTCGCCGACCCGAGGGTAGGGCCTGGG
CTCAGCCCGGGTACCCTTGGCCCCTCTATGGCAATGAGGGCATGGGGTGGGCAGGATG
GCTCCTGTCACCCCGCGGCTCTCGGCCTAGTTGGGGCCCTACAGACCCCCGGCGTAGG
TCGCGTAATTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTCGCCGACCTCGTGG
GGTACATTCCGCTCGTCGGCGCCCCCCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGG
CGTCCGGGTTCTGGAGGACGGCGTGAACTATGCAACAGGGAATTTGCCCGGTTGCTCT
TTCTCTATCTTCCTCTTGGCTTTGCTGTCCTGTCTGACCGTTCCAGCTTCCGCTTATGAA
GTGCGCAACGTGTCCGGGATGTACCATGTCACGAACGACTGCTCCAACTCAAGCATTG
TGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTTCGGGA
GAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCCACGCTCGCAGCTAGGAACGCC
AGCGTCCCCACCACGACAATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTT
TCTGTTCCGCTATGTACGTGGGGGACCTCTGCGGATCTGTCTTCCTCGTCTCCCAGCTG
TTCACCATCTCGCCTCGCCGGCATGAGACGGTGCAGGACTGCAATTGCTCAATCTATC
CCGGCCACATAACGGGTCACCGTATGGCTTGGGATATGATGATGAACTGGTCGCCTAC
AACGGCCCTGGTGGTATCGCAGCTGCTCCGGATCCCACAAGCTGTCGTGGACATGGTG
GCGGGGGCCCATTGGGGAGTCCTGGCGGGCCTCGCCTACTATTCCATGGTGGGGAACT
GGGCTAAGGTTTTGGTTGTGATGCTACTCTTTGCCGGCGTCGACGGGCATACCCGCGT
GTCAGGAGGGGCAGCAGCCTCCGATACCAGGGGCCTTGTGTCCCTCTTTAGCCCCGGG

图 21K

TCGGCTCAGAAAATCCAGCTCGTAAACACCAACGGCAGTTGGCACATCAACAGGACT
GCCCTGAACTGCAACGACTCCCTCCAAACAGGGTTCTTTGCCGCACTATTCTACAAAC
ACAAATTCAACTCGTCTGGATGCCCAGAGCGCTTGGCCAGCTGTCGCTCCATCGACAA
GTTCGCTCAGGGGTGGGGTCCCCTCACTTACACTGAGCCTAACAGCTCGGACCAGAGG
CCCTACTGCTGGCACTACGCGCCTCGACCGTGTGGTATTGTACCCGCGTCTCAGGTGT
GCGGTCCAGTGTATTGCTTCACCCCGAGCCCTGTTGTGGTGGGGACGACCGATCGGTT
TGGTGTCCCCACGTATAACTGGGGGGCGAACGACTCGGATGTGCTGATTCTCAACAAC
ACGCGGCCGCCGCGAGGCAACTGGTTCGGCTGTACATGGATGAATGGCACTGGGTTCA
CCAAGACGTGTGGGGGCCCCCCGTGCAACATCGGGGGGGCCGGCAACAACACCTTGA
CCTGCCCCACTGACTGTTTTCGGAAGCACCCCGAGGCCACCTACGCCAGATGCGGTTC
TGGGCCCTGGCTGACACCTAGGTGTATGGTTCATTACCCATATAGGCTCTGGCACTAC
CCCTGCACTGTCAACTTCACCATCTTCAAGGTTAGGATGTACGTGGGGGGCGTGGAGC
ACAGGTTCGAAGCCGCATGCAATTGGACTCGAGGAGAGCGTTGTGACTTGGAGGACA
GGGATAGATCAGAGCTTAGCCCGCTGCTGCTGTCTACAACAGAGTGGCAGATACTGCC
CTGTTCCTTCACCACCCTGCCGGCCCTATCCACCGGCCTGATCCACCTCCATCAGAAC
ATCGTGGACGTGCAATACCTGTACGGTGTAGGGTCGGCGGTTGTCTCCCTTGTCATCA
AATGGGAGTATGTCCTGTTGCTCTTCCTTCTCCTGGCAGACGCGCGCATCTGCGCCTGC
TTATGGATGATGCTGCTGATAGCTCAAGCTGAGGCCGCCTTAGAGAACCTGGTGGTCC
TCAATGCGGCGGCCGTGGCCGGGGCGCATGGCACTCTTTCCTTCCTTGTGTTCTTCTGT
GCTGCCTGGTACATCAAGGGCAGGCTGGTCCCTGGTGCGGCATACGCCTTCTATGGCG
TGTGGCCGCTGCTCCTGCTTCTGCTGGCCTTACCACCACGAGCTTATGCCTAGTAA

图 21L

在450nm下测定的OD
构建体

级份	体积 稀释	39 型 1b	40 型 1b	62 型 3a	63 型 5a
开始	23 ml 1/20	2.517	1.954	1.426	1.142
流通	23 ml 1/20	0.087	0.085	0.176	0.120
1	0.4 ml 1/200	0.102	0.051	0.048	0.050
2		0.396	0.550	0.090	0.067
3		2.627	2.603	2.481	2.372
4		3	2.967	3	2.694
5		3	2.810	2.640	2.154
6		2.694	2.499	1.359	1.561
7		2.408	2.481	0.347	1.390
8		2.176	1.970	1.624	0.865
9		1.461	1.422	0.887	0.604
10		1.286	0.926	0.543	0.519
11		0.981	0.781	0.294	0.294
12		0.812	0.650	0.249	0.199
13		0.373	0.432	0.239	0.209
14		0.653	0.371	0.145	0.184
15		0.441	0.348	0.151	0.151
16		0.321	0.374	0.098	0.106
17		0.525	0.186	0.099	0.108
18		0.351	0.171	0.083	0.090
19		0.192	0.164	0.084	0.087

图 22

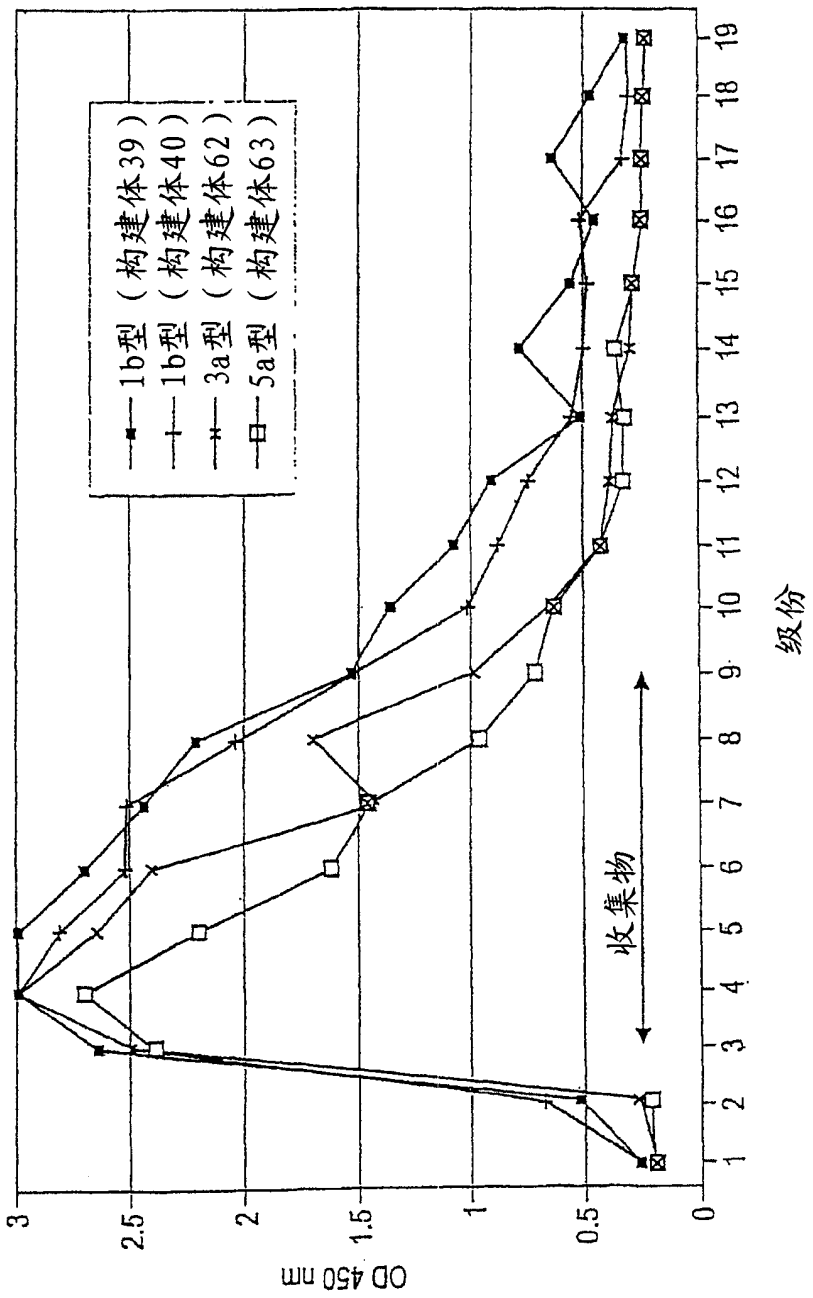

1b型（构建体39）
1b型（构建体40）
3a型（构建体62）
5a型（构建体63）
OD 450 nm
3
2.5
2
1.5
1
0.5
0
收集物
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
级份

图 23

			在450nm下测定的OD			
			构建体			
级份	体积	稀释	39 型 1b	40 型 1b	62 型 3a	63 型 5a
20	250 μl	1/200	0.072	0.130	0.096	0.051
21			0.109	0.293	0.084	0.052
22			0.279	0.249	0.172	0.052
23			0.093	0.151	0.297	0.054
24			0.080	0.266	0.438	0.056
25			0.251	0.100	0.457	0.048
26			3	1.649	0.722	0.066
27			3	3	2.528	0.889
28			3	3	3	2.345
29			3	3	2.849	2.580
30			2.227	1.921	1.424	1.333
31			0.263	0.415	0.356	0.162
32			0.071	0.172	0.154	0.064
33			0.103	0.054	0.096	0.057
34			0.045	0.045	0.044	0.051
35			0.043	0.047	0.045	0.046
36			0.045	0.045	0.049	0.040
37			0.045	0.047	0.046	0.048
38			0.046	0.048	0.047	0.057
39			0.045	0.048	0.050	0.057
40			0.046	0.049	0.048	0.049

图 24

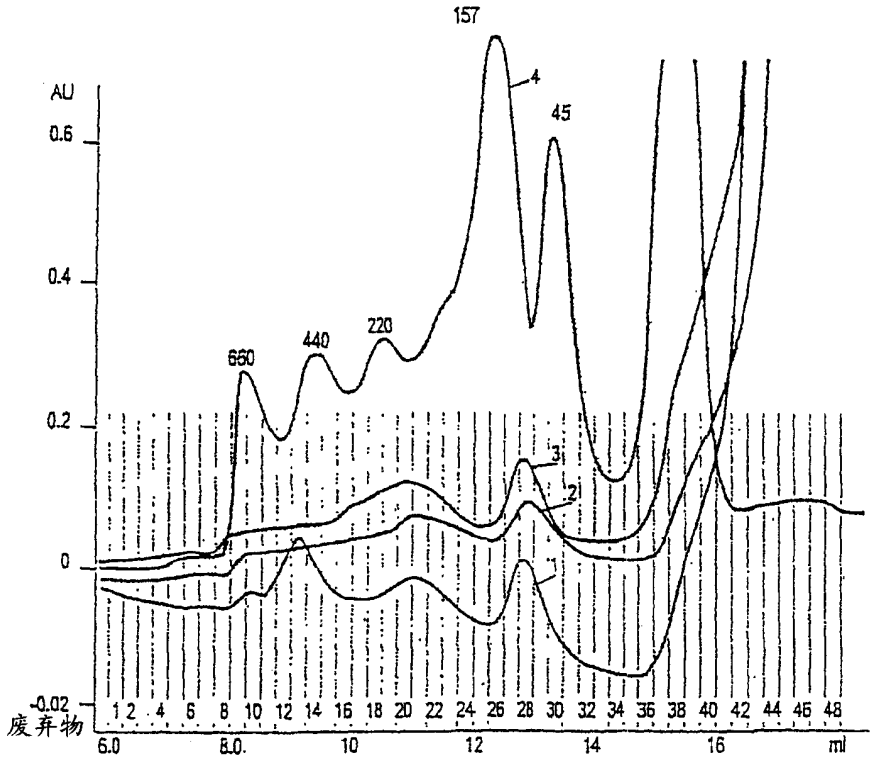

157
4
45
AU
0.6
0.4
0.2
0
-0.02
660
440
220
3
2
1
废弃物
1 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48
6.0
8.0
10
12
14
16
ml

图 25

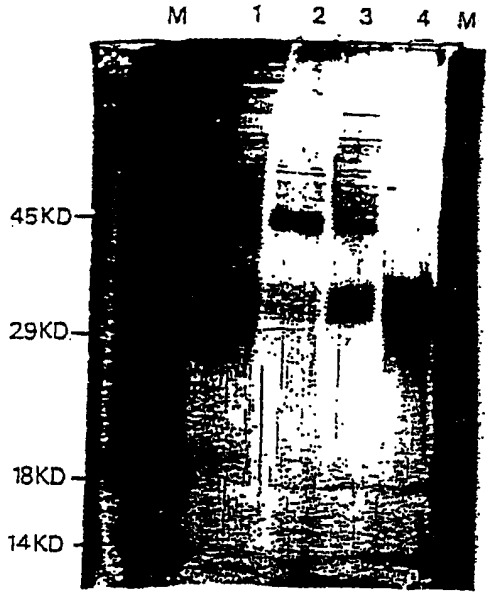

M
1
2
3
4
M
45KD
29KD
18KD
14KD

图 26

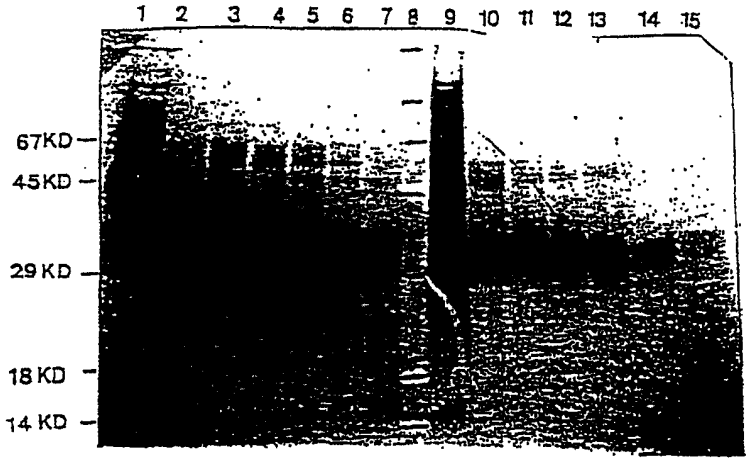

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
67KD
45KD
29 KD
18 KD
14 KD

图 27

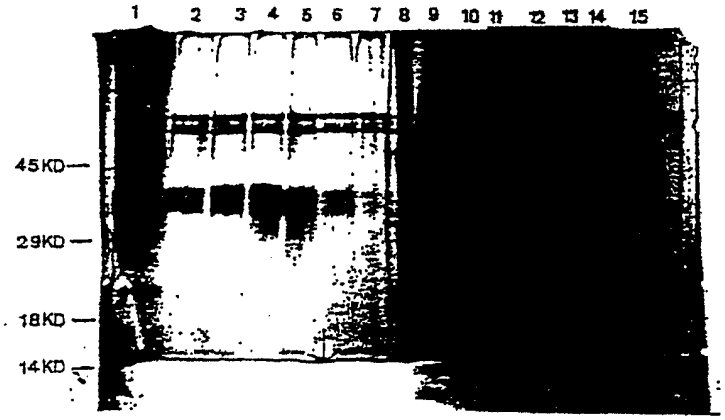
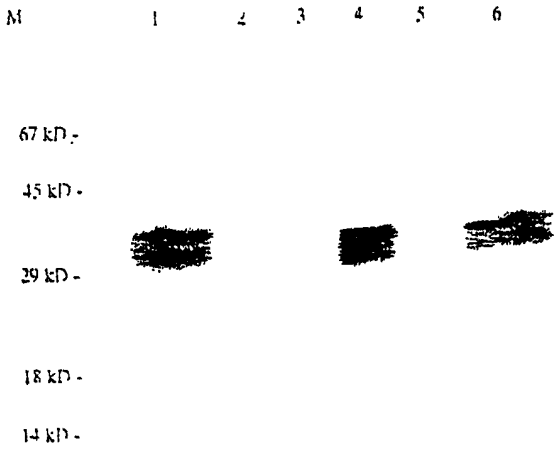


图 28



泳道1: 粗制裂解物
泳道2: 流通物Lentil层析
泳道3: 用EMPIGEN Lentil 层析洗涤
泳道4: 洗脱物Lentil层析
泳道5: 在浓缩Lentil洗脱物时的流通物
泳道6: 大小排阻层析后的E1收集物

图 29

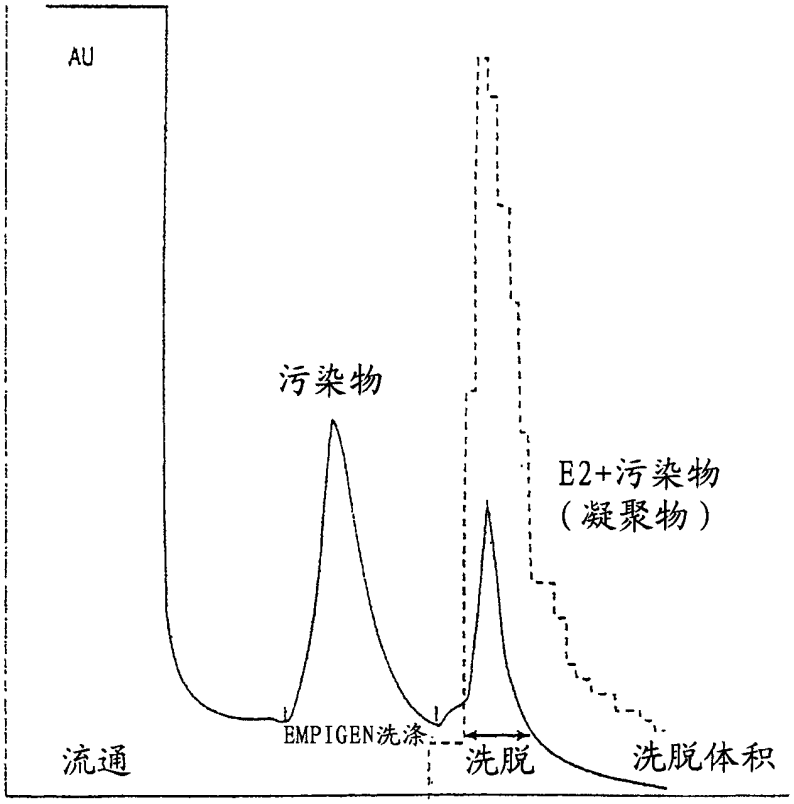


图 30

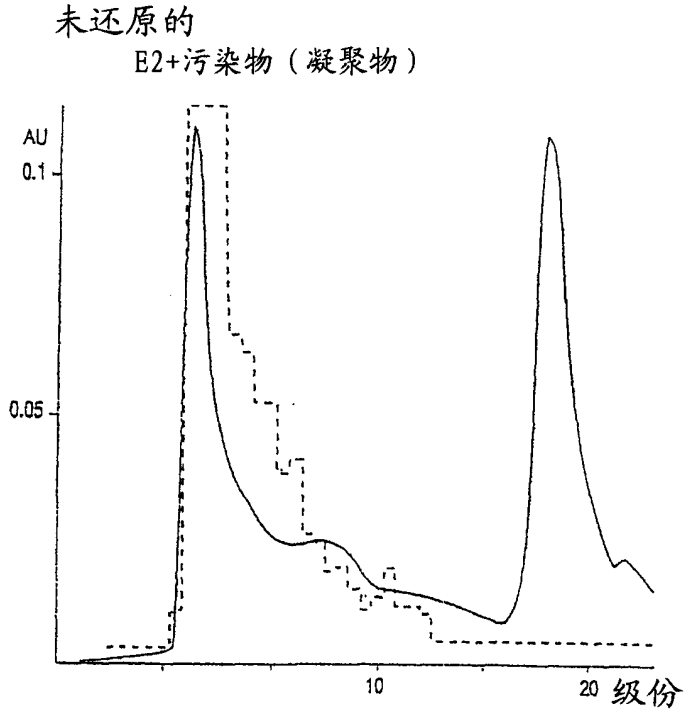

未还原的
E2+污染物（凝聚物）
AU
0.1
0.05
10
20
级份

图 31A

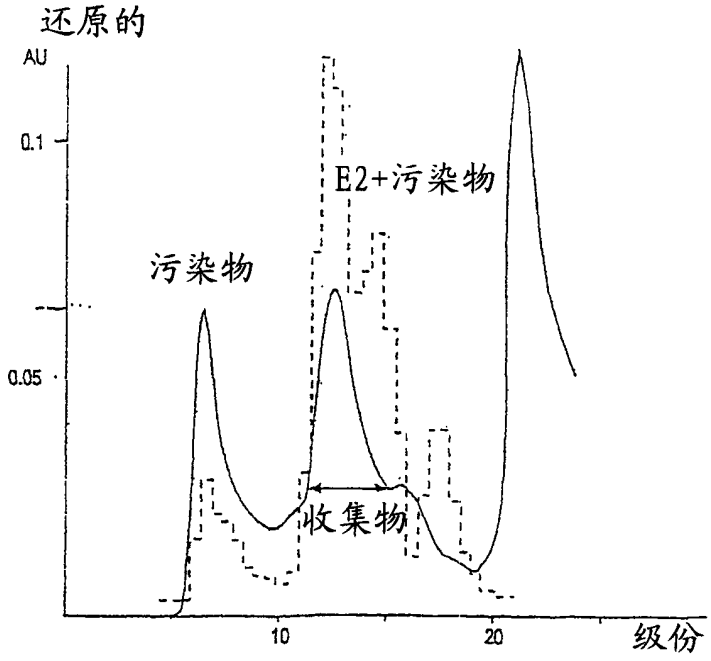

还原的
AU
0.1
0.05
E2+污染物
污染物
收集物
10
20
级份

图 31B

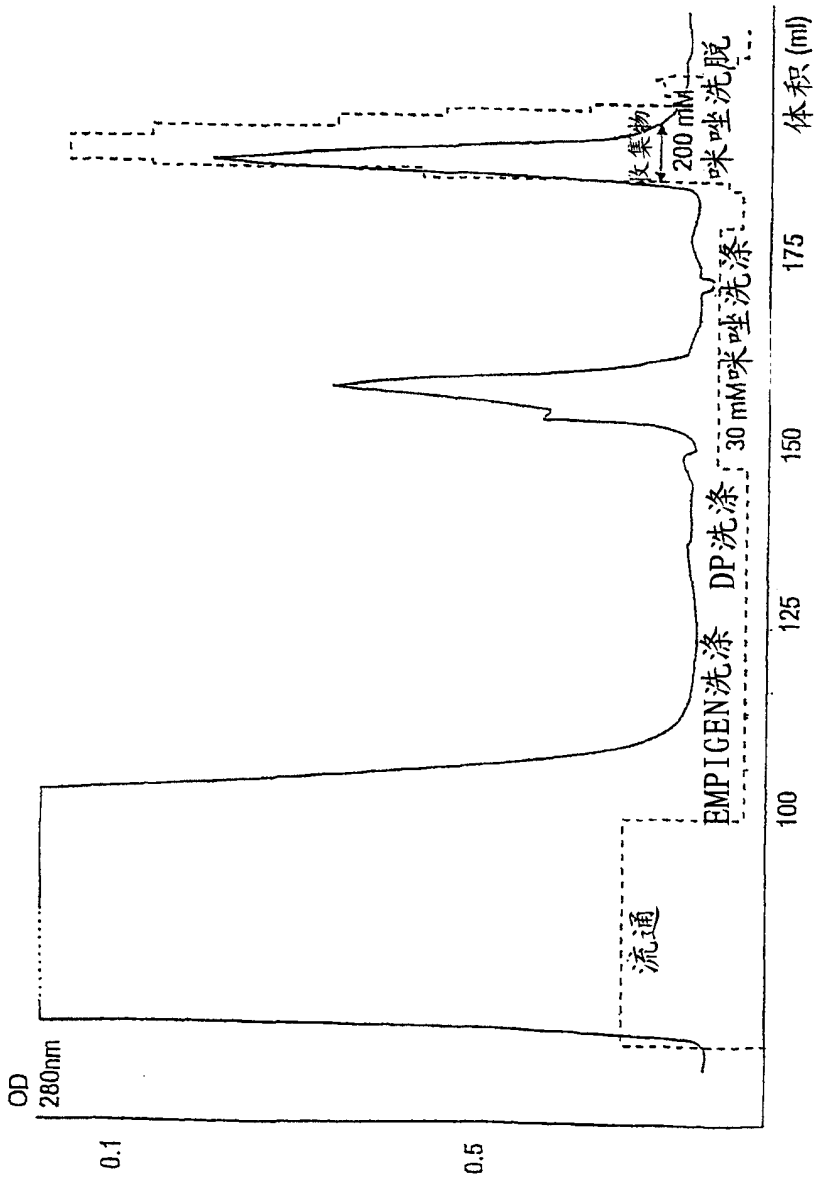
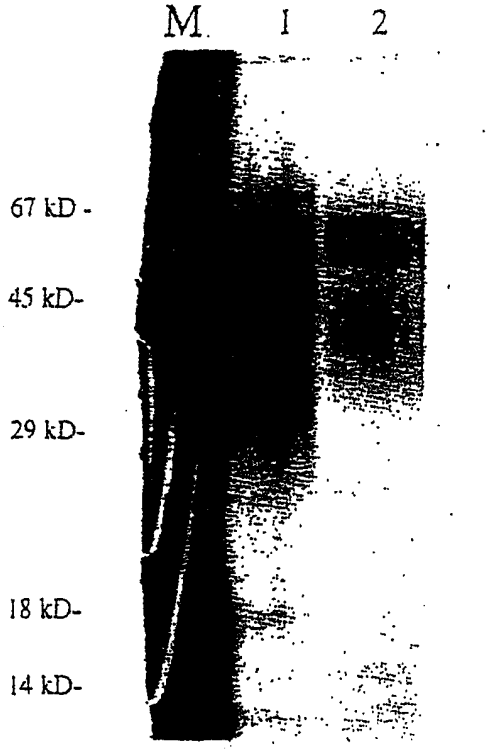

OD
280nm
0.1
0.5
流通
EMPIGEN洗涤
DP洗涤
30 mM咪唑洗涤
收集物
200 mM
咪唑洗脱
100
125
150
175
体积 (ml)

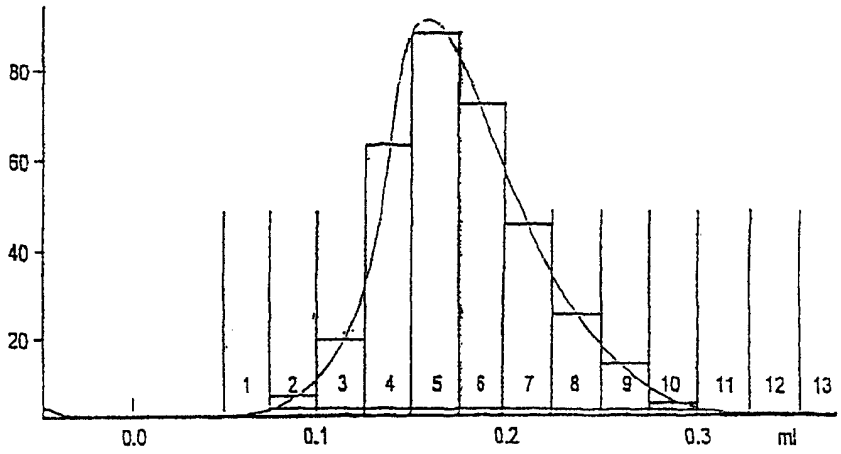
图 32

纯化E2的银染色



1. 30 mM咪唑洗涤Ni-IMAC
2. 0.5 ug E2

图 33



No.	Ret. (ml)	峰开始 (ml)	峰结束 (ml)	持续 (ml)	面积 (ml*mAU)	高度 (mAU)
1	-0.45	-0.46	-0.43	0.04	0.0976	4.579
2	1.55	0.75	3.26	2.51	796.4167	889.377
3	3.27	3.26	3.31	0.05	0.0067	0.224
4	3.35	3.32	3.33	0.02	0.0002	0.018

检测峰总数=4
基线上方的总面积=0.796522ml*AU
评估峰内的总面积=0.796521ml*AU
峰面积/总面积之比=0.999999
总的峰持续=2.613583ml

图 34

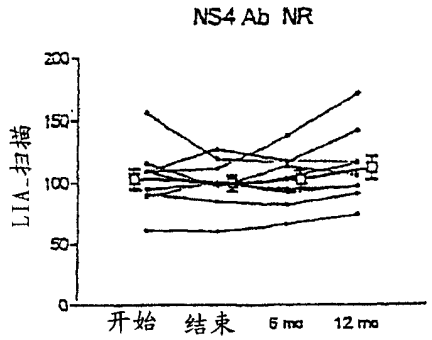


图 35A－1

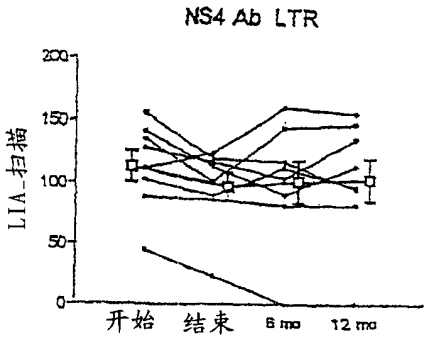


图 35A－2

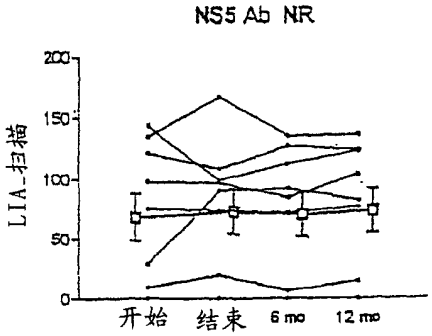


图 35A－3

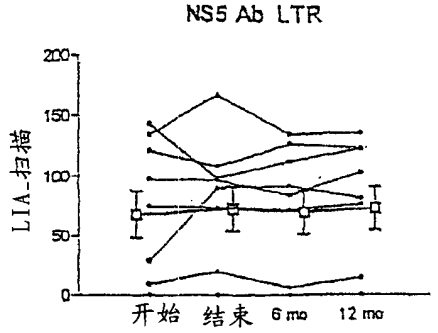


图 35A－4

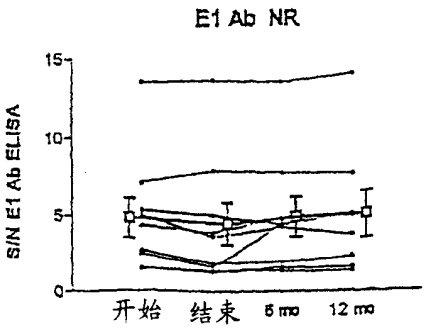


图 35A－5

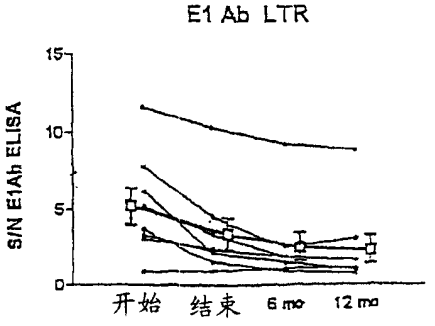


图 35A－6

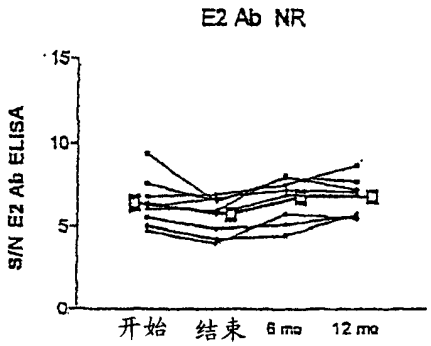

E2 Ab NR
S/N E2 Ab ELISA
15
10
5
0
开始
结束
6 mo
12 mo

图 35A－7

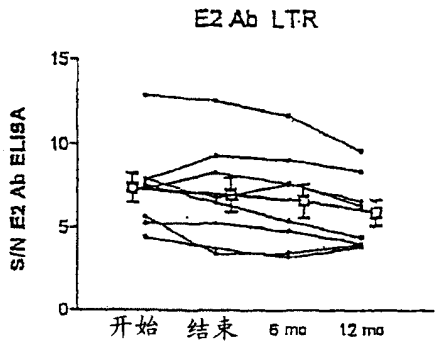

E2 Ab LTR
S/N E2 Ab ELISA
15
10
5
0
开始
结束
6 mo
12 mo

图 35A－8

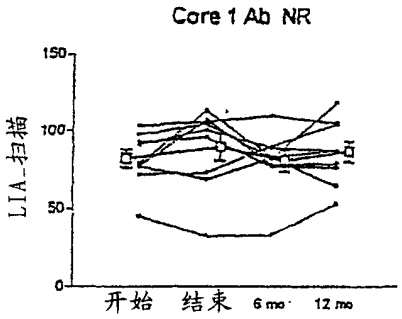


图 35B—1

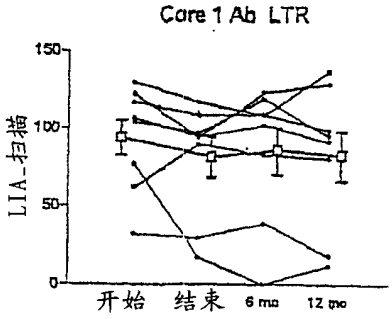


图 35B—2

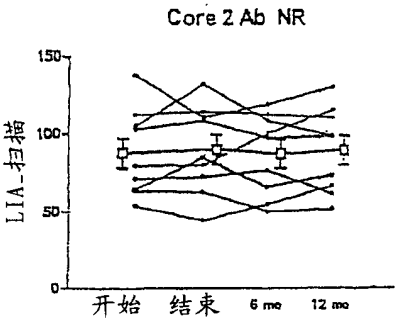


图 35B—3

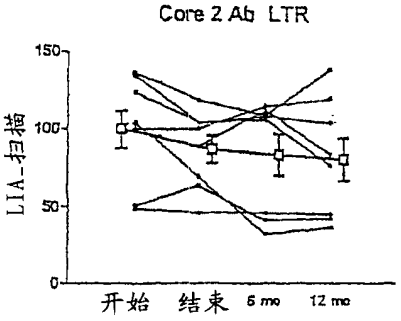


图 35B—4

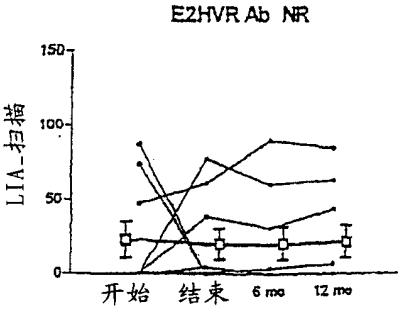


图 35B—5

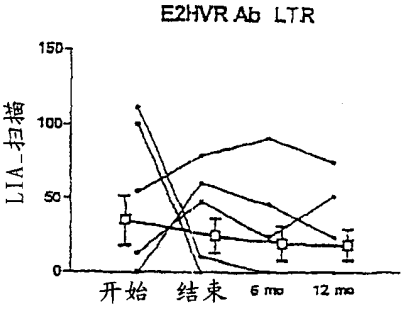


图 35B—6

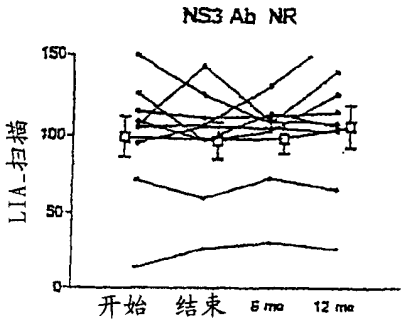

NS3 Ab NR
LIA-扫描
150
100
50
0
开始
结束
6 mo
12 mo

图 35B－7

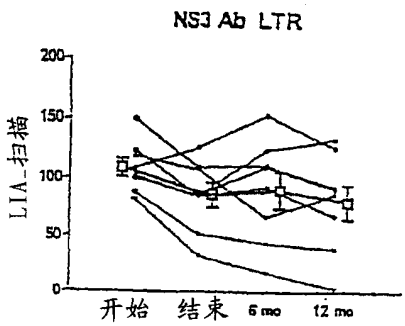

NS3 Ab LTR
LIA-扫描
200
150
100
50
0
开始
结束
6 mo
12 mo

图 35B－8

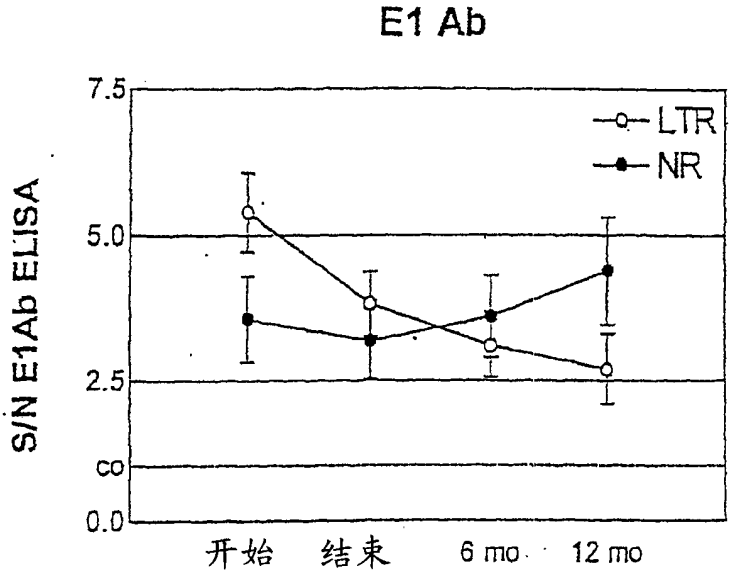

E1 Ab
S/N E1Ab ELISA
7.5
5.0
2.5
CO
0.0
LTR
NR
开始
结束
6 mo
12 mo

图 36A

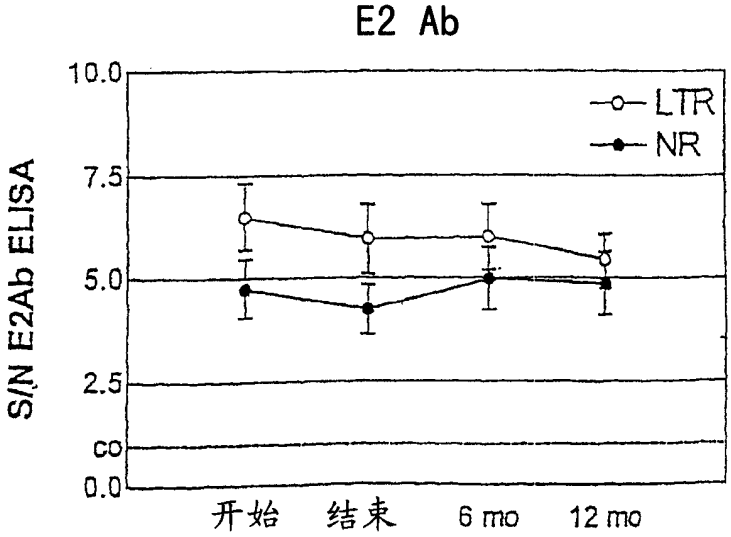

E2 Ab
S/N E2Ab ELISA
10.0
7.5
5.0
2.5
CO
0.0
LTR
NR
开始
结束
6 mo
12 mo

图 36B

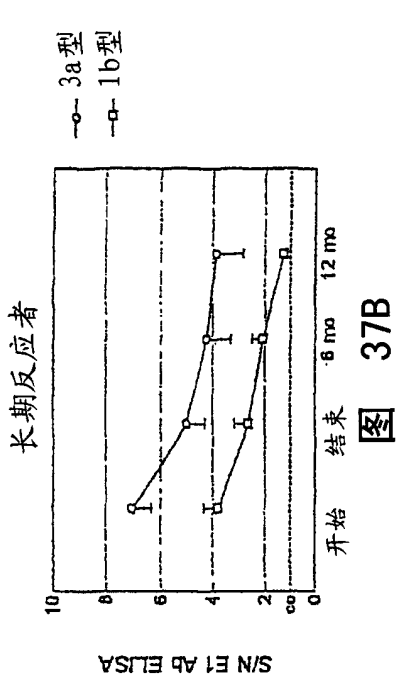


图 37A

图 37B

图 37C

图 37D

抗E2单克隆抗体的相对作图位置

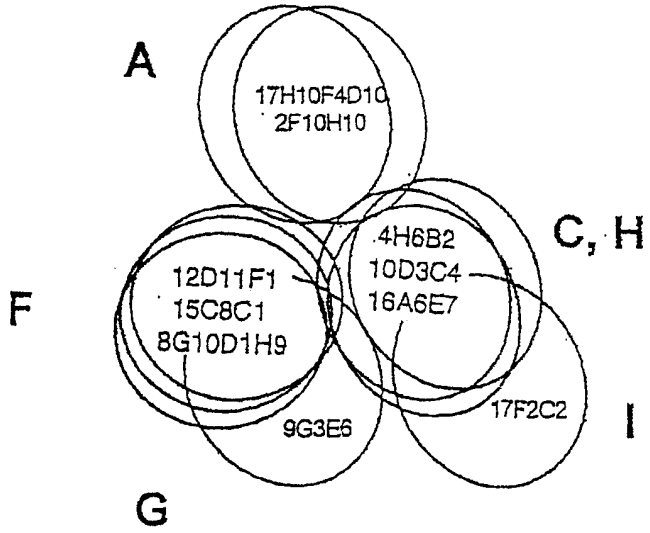


图 38

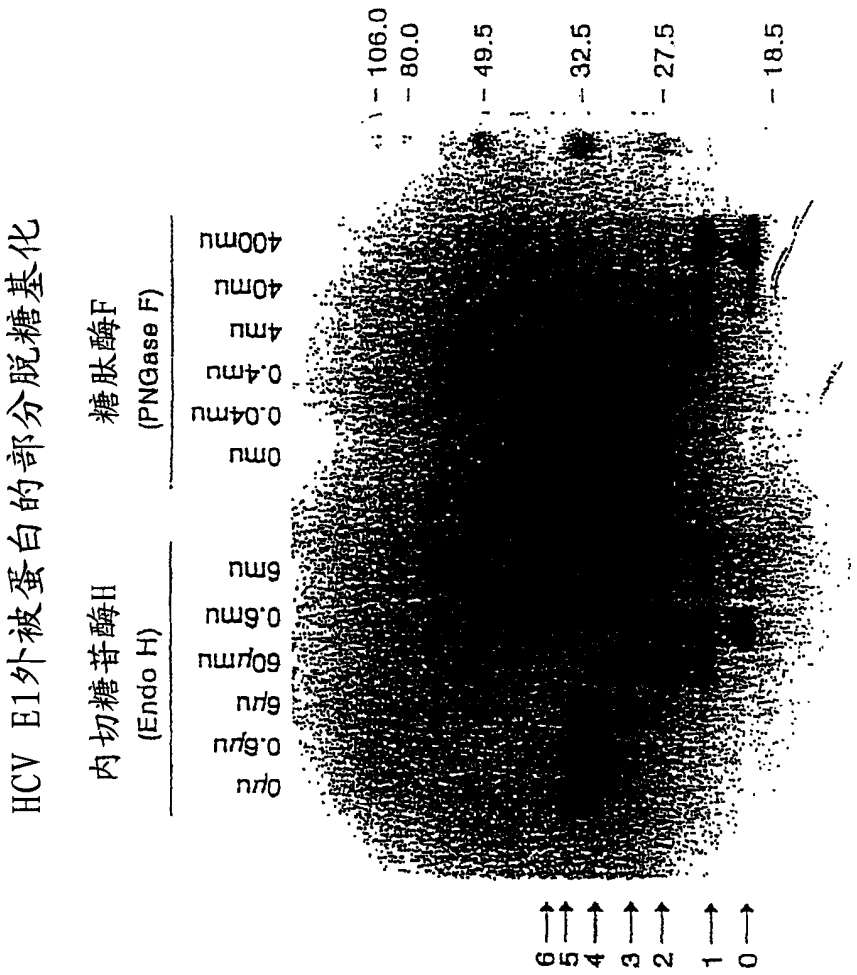

HCV E1外被蛋白的部分脱糖基化
内切糖苷酶H (Endo H)
糖肽酶F (PNGase F)
0μu
0.6μu
6μu
60μmu
0.6mu
6mu
0mu
0.04mu
0.4mu
4mu
40mu
400mu
106.0
80.0
49.5
32.5
27.5
18.5
6
5
4
3
2
1
0

图 39

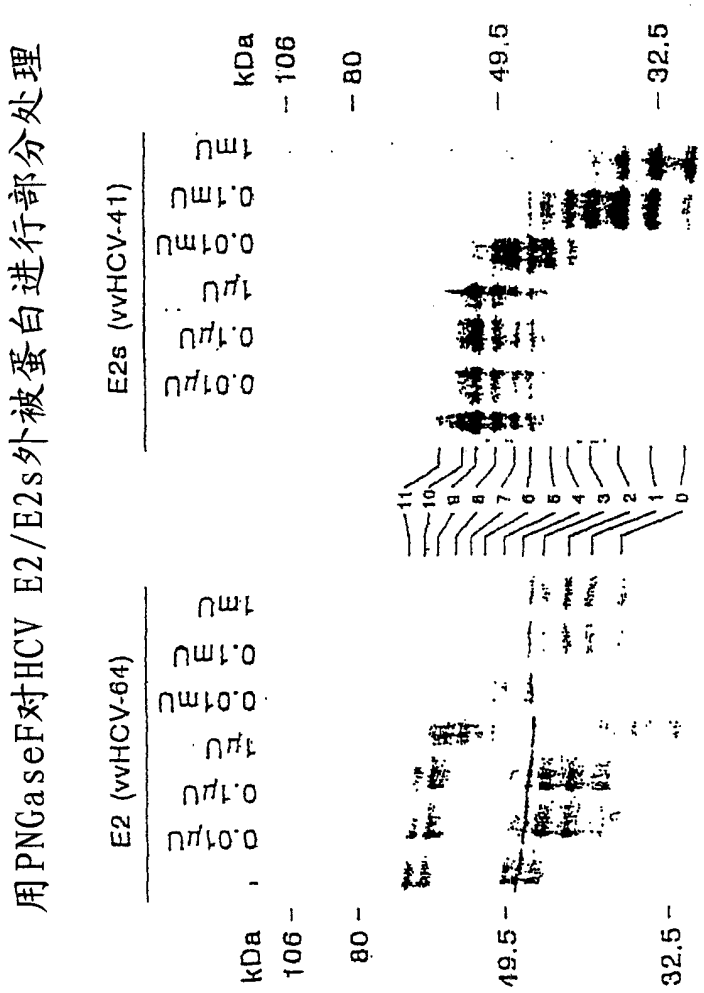

用PNGaseF对HCV E2/E2s外被蛋白进行部分处理
E2 (vvHCV-64)
E2s (vvHCV-41)
kDa
106
80
49.5
32.5
-
0.01μU
0.1μU
1μU
0.01mU
0.1mU
1mU
11
10
9
8
7
6
5
4
3
2
1
0

图 40

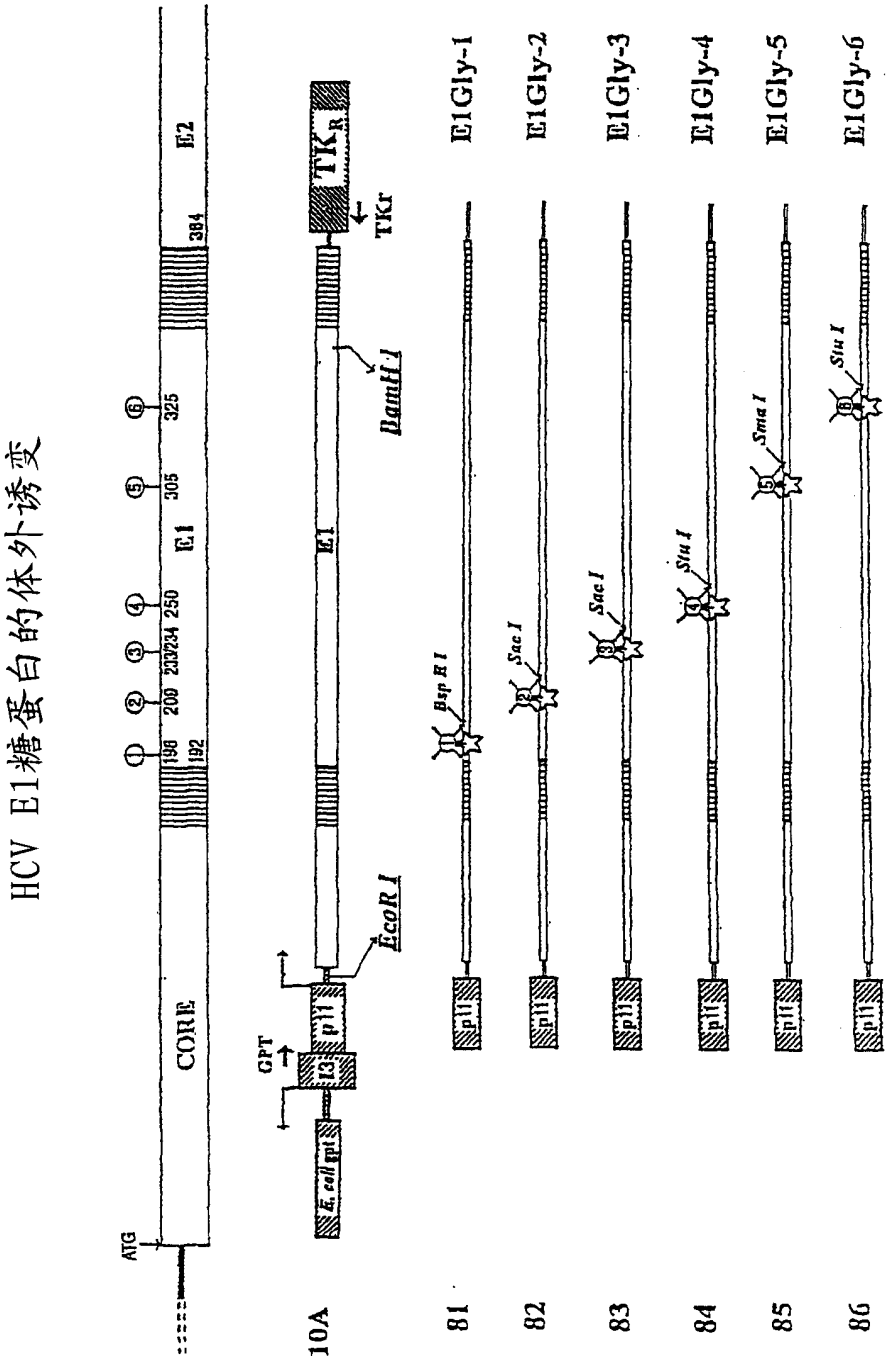

HCV E1糖蛋白的体外诱变
ATG
CORE
E1
E2
192
196
200
233234
250
305
325
384
①
②
③
④
⑤
⑥
10A
GPT
E. coli gpt
I3
p11
EcoR I
E1
BamH I
TK_R
TKr
81
82
83
84
85
86
E1Gly-1
E1Gly-2
E1Gly-3
E1Gly-4
E1Gly-5
E1Gly-6
Bsp R I
Sac I
Sac I
Stu I
Sma I
Stu I

图 41

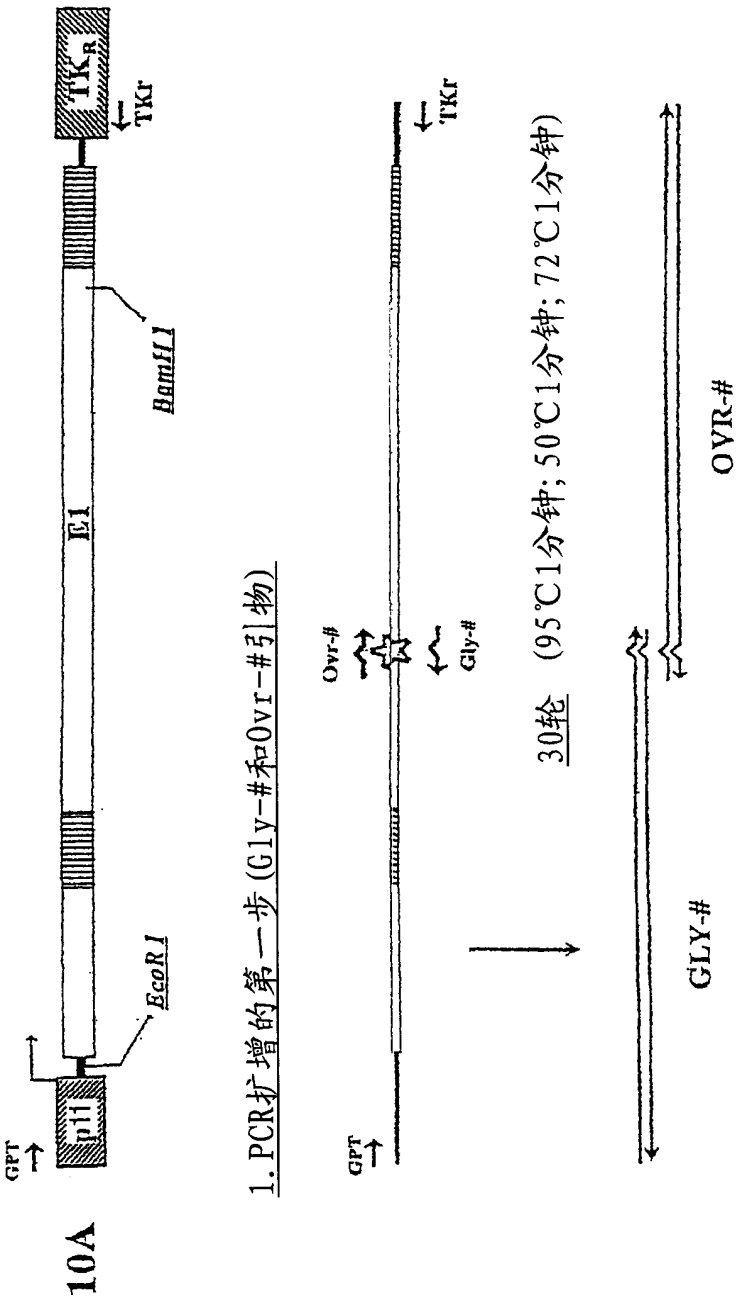

HCV E1糖蛋白的体外诱变
10A
GPT
p11
EcoR I
E1
BamH I
TKr
1. PCR扩增的第一步 (Gly-#和Ovr-#引物)
GPT
Ovr-#
Gly-#
TKr
30轮 (95℃1分钟; 50℃1分钟; 72℃1分钟)
GLY-#
OVR-#

图 42A

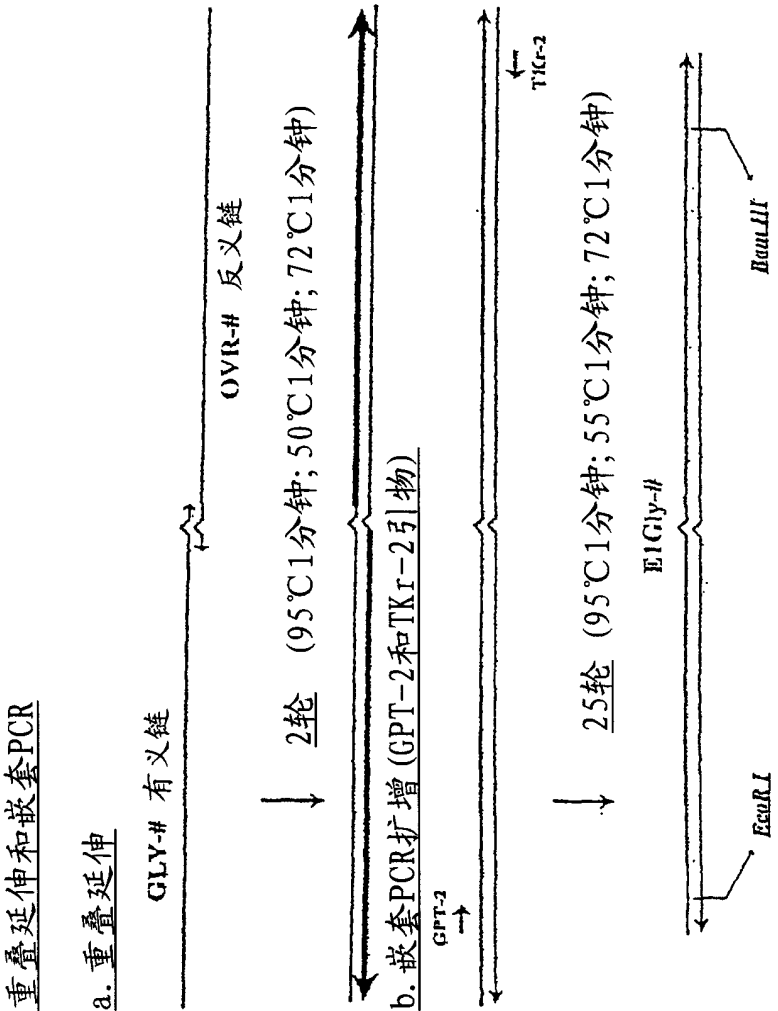

2. 重叠延伸和嵌套PCR
a. 重叠延伸
GLY-# 有义链
OVR-# 反义链
2轮 (95℃1分钟; 50℃1分钟; 72℃1分钟)
b. 嵌套PCR扩增(GPT-2和TKr-2引物)
GPT-2
TKr-2
25轮 (95℃1分钟; 55℃1分钟; 72℃1分钟)
E1Gly-#
BamHI
EcoRI

图 42B

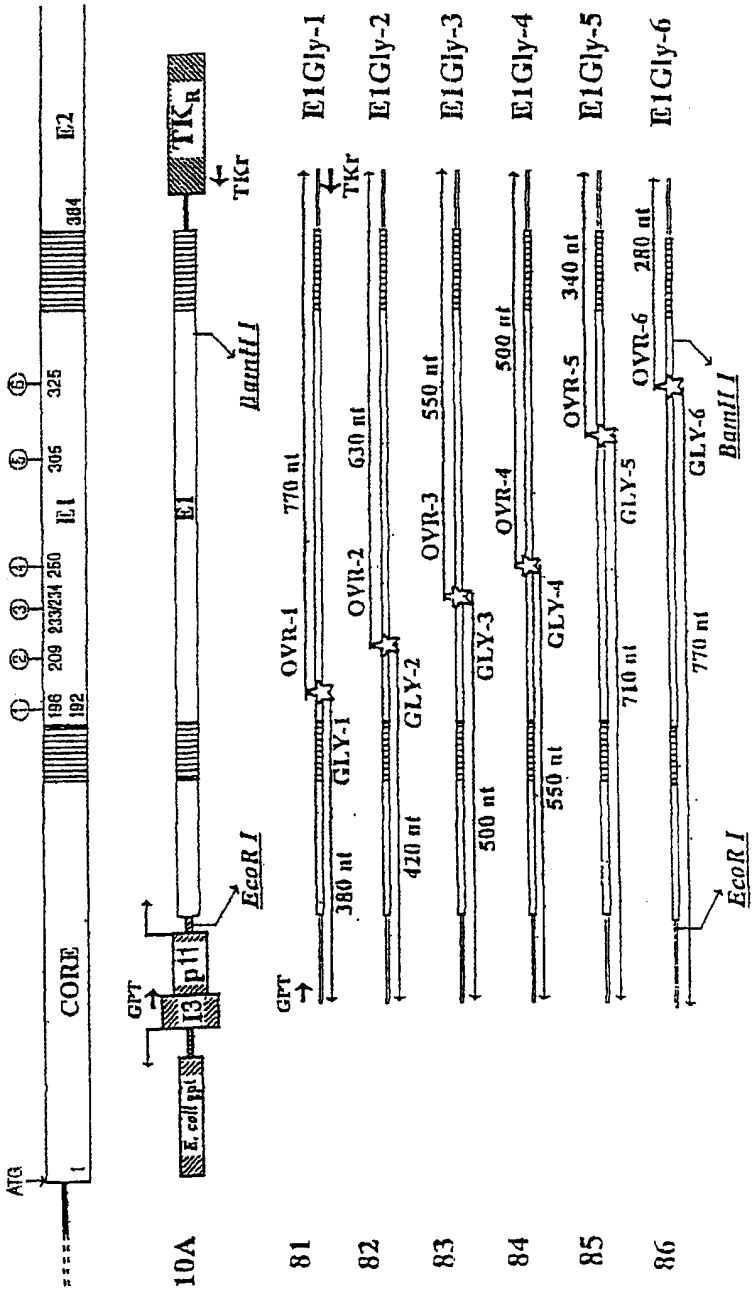

HCV E1糖蛋白的体外诱变
ATG
CORE
E1
E2
1
192
196
209
233
234
250
305
325
384
①
②
③
④
⑤
⑥
E. coli tpt
I3
p11
GPT
EcoR I
BamH I
TKr
TKR
10A
81
82
83
84
85
86
E1Gly-1
E1Gly-2
E1Gly-3
E1Gly-4
E1Gly-5
E1Gly-6
OVR-1
GLY-1
380 nt
770 nt
OVR-2
GLY-2
420 nt
630 nt
OVR-3
GLY-3
500 nt
550 nt
OVR-4
GLY-4
550 nt
500 nt
OVR-5
GLY-5
710 nt
340 nt
OVR-6
GLY-6
770 nt
280 nt

图 43

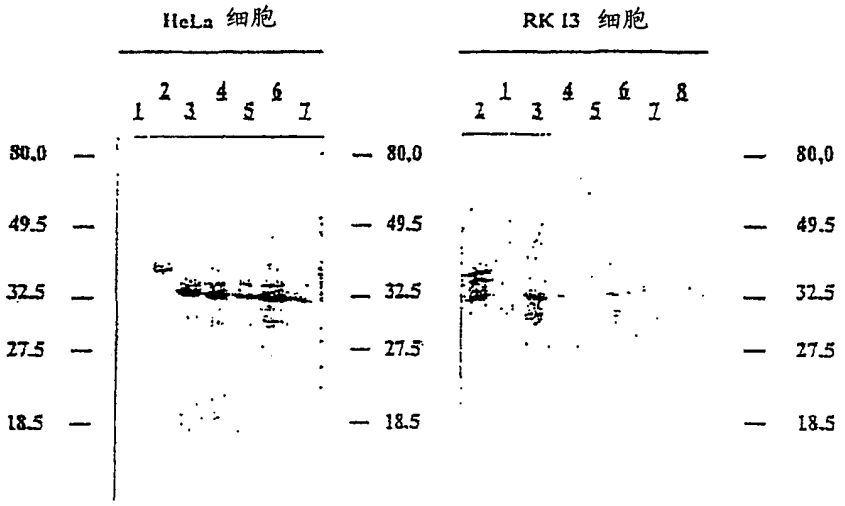

HeLa 细胞
RK 13 细胞
1 2 3 4 5 6 7
1 2 3 4 5 6 7 8
80.0
49.5
32.5
27.5
18.5

图 44A

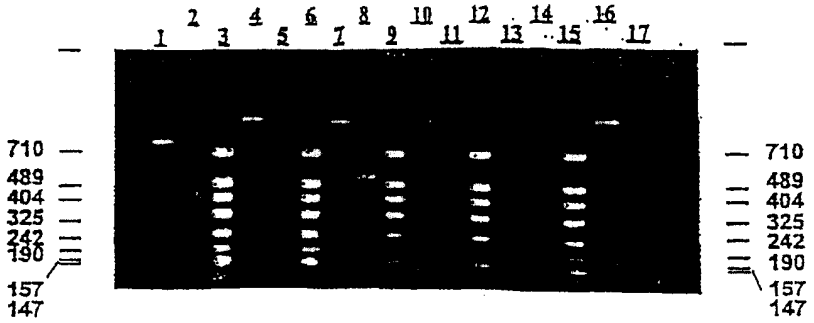

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
710
489
404
325
242
190
157
147

图 44B

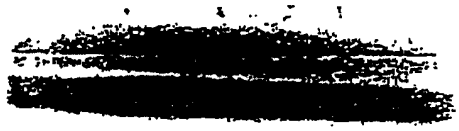


图 45

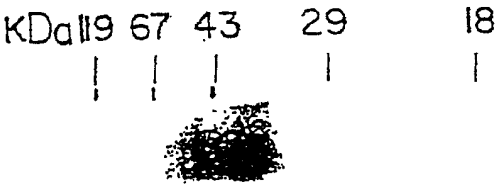


图 46

	年龄 (岁)	HCV感染 (年数)	基因型
Marcel	17	9	1a
Peggy	21	16.5	1b
Femma	15	9	1a
Yoran	12	无	
Marti	12	无	

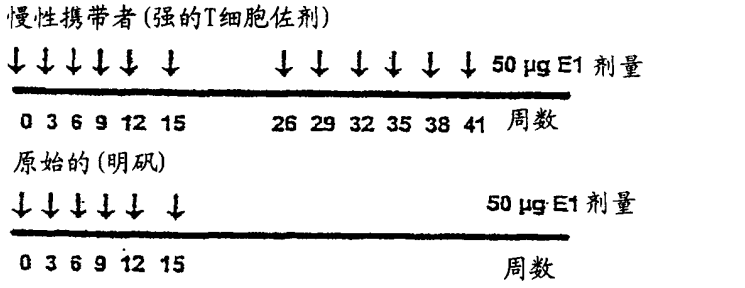


图 47

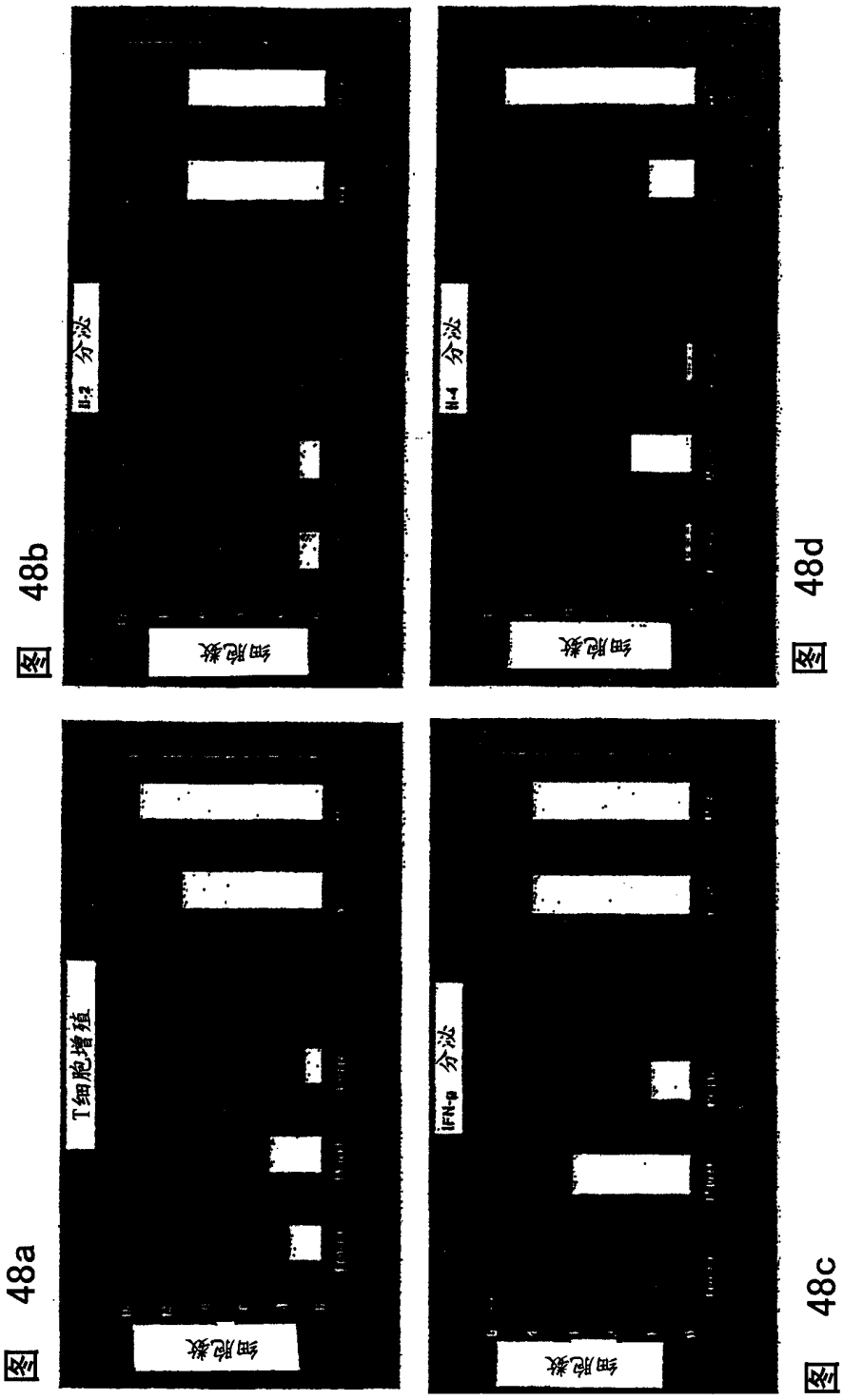


图 48a 图 48b 图 48c 图 48d

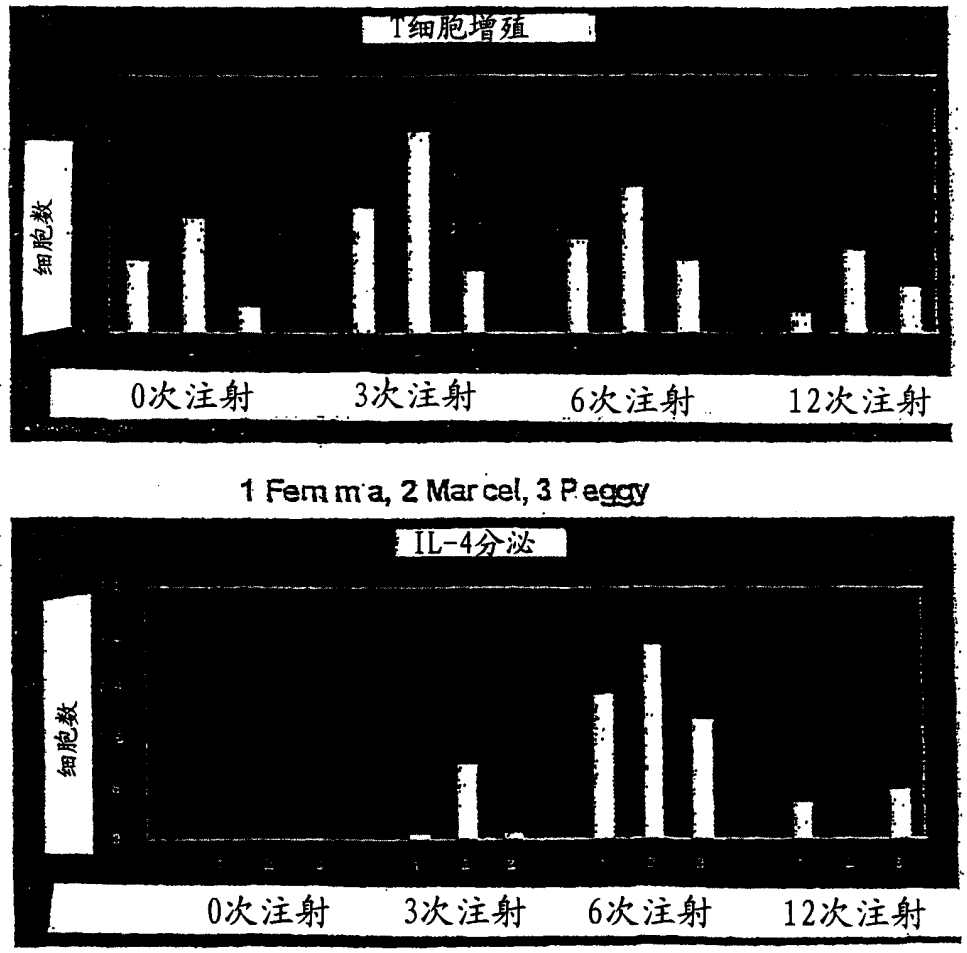

T细胞增殖
细胞数
0次注射
3次注射
6次注射
12次注射
1 Femm'a, 2 Marcel, 3 Peggy
IL-4分泌
细胞数
0次注射
3次注射
6次注射
12次注射

图 49

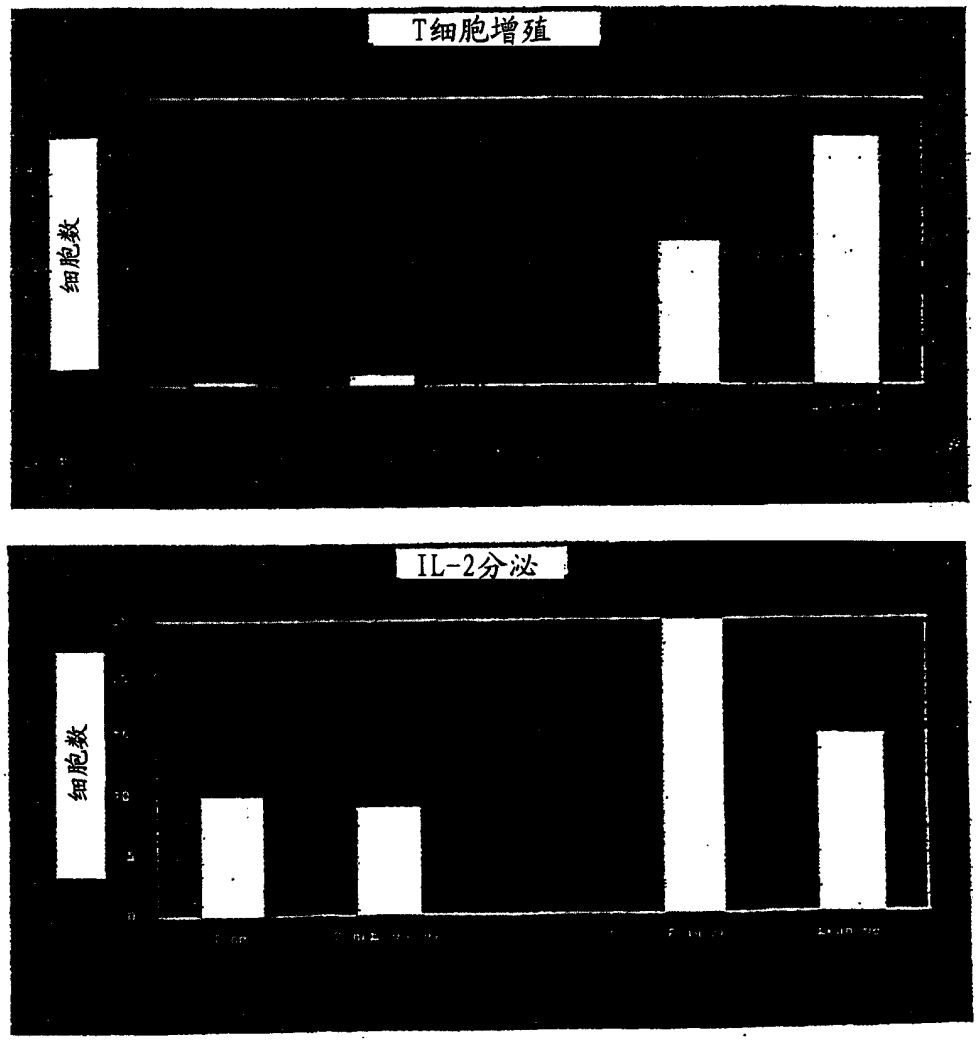


图 50

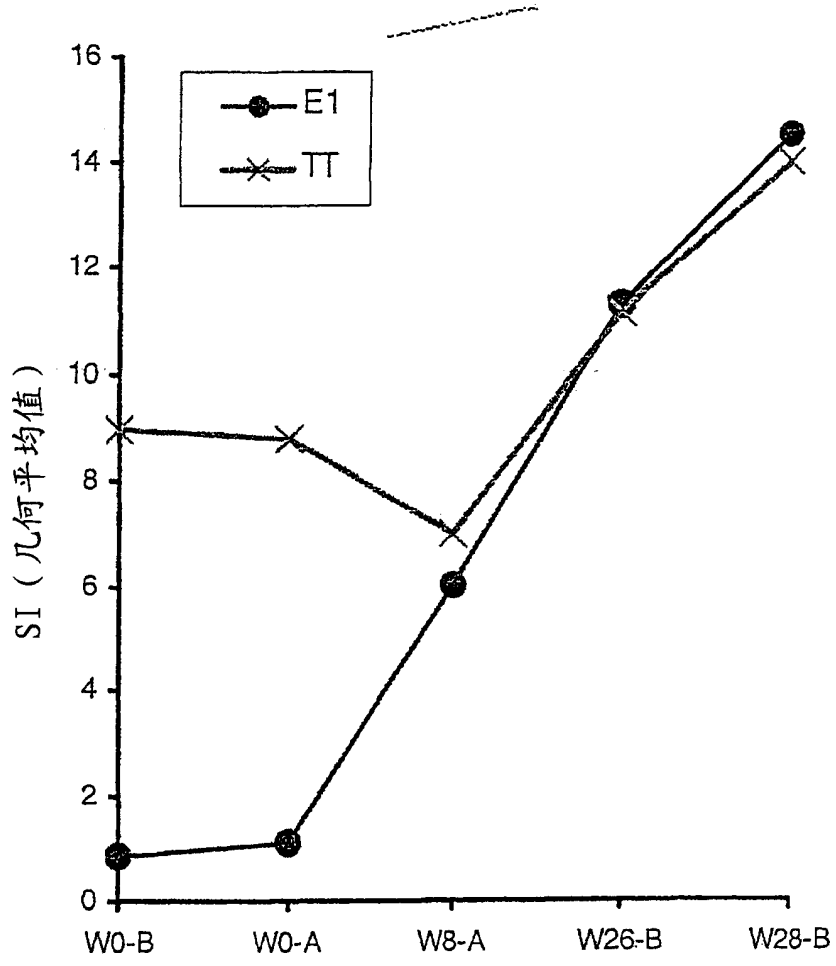


图 51

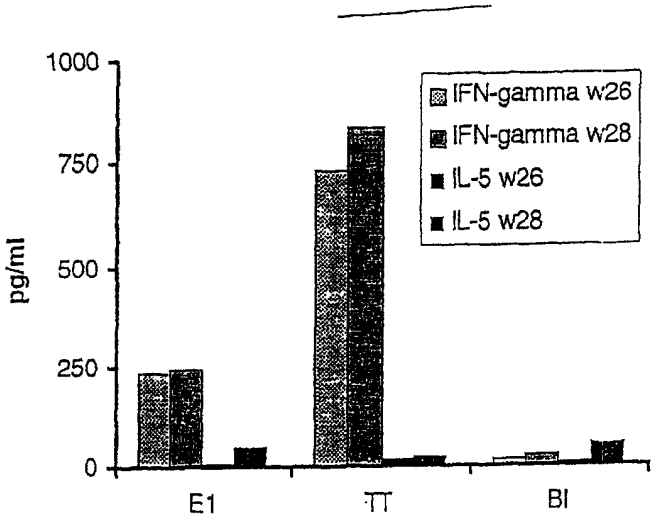
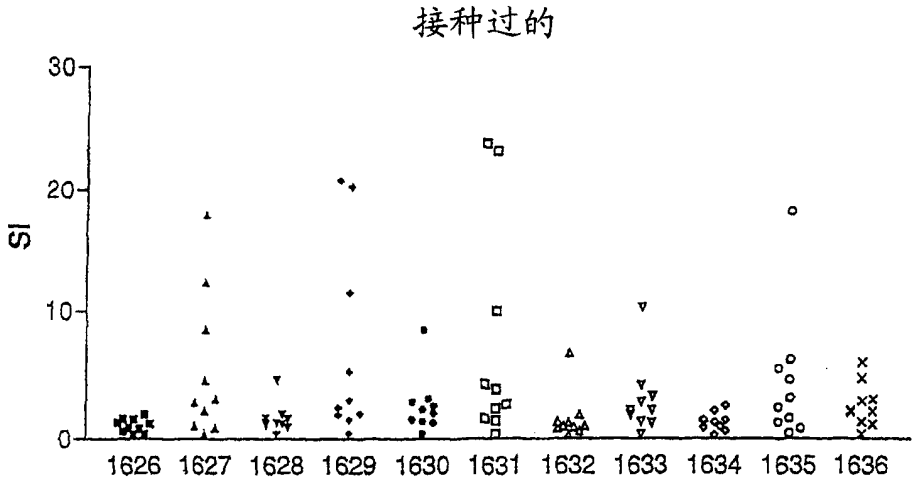

1000
750
500
250
0
pg/ml
IFN-gamma w26
IFN-gamma w28
IL-5 w26
IL-5 w28
E1
TT
BI

图 52


接种过的
SI
30
20
10
0
1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636

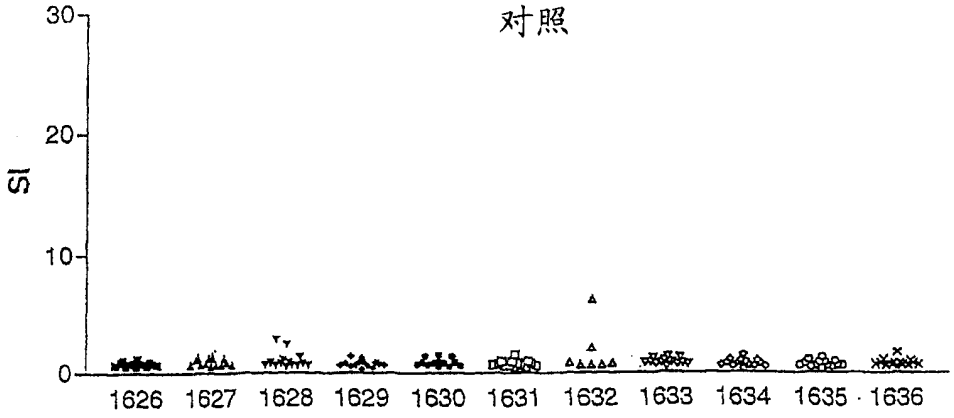

对照
SI
30
20
10
0
1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636

图 53