

[19] 中华人民共和国国家知识产权局



[12] 发 明 专 利 说 明 书

专利号 ZL 02807758. X

[51] Int. Cl.

C12N 15/33 (2006.01)

A61K 45/00 (2006.01)

A61P 31/12 (2006.01)

[45] 授权公告日 2007 年 3 月 14 日

[11] 授权公告号 CN 1304575C

[22] 申请日 2002.4.4 [21] 申请号 02807758. X

[30] 优先权

[32] 2001. 4. 4 [33] US [31] 09/826,115

[86] 国际申请 PCT/US2002/010764 2002.4.4

[87] 国际公布 WO2002/081754 英 2002.10.17

[85] 进入国家阶段日期 2003.9.30

[73] 专利权人 美国政府健康及人类服务部;疾病控制
和预防中心

地址 美国佐治亚州

[72] 发明人 张光正

[56] 参考文献

US5494671A 1996.2.27

US6165477A 2000.12.26

CN1253506A 2000.5.17

审查员 吴文英

[74] 专利代理机构 中原信达知识产权代理有限责
任公司

代理人 丁业平 王维玉

权利要求书 3 页 说明书 171 页 附图 8 页

[54] 发明名称

预防黄病毒感染的核酸疫苗

[57] 摘要

本发明包括含有转录单位的分离核酸, 该转录单位编码第一种黄病毒的信号序列和第二种黄病毒的免疫原性黄病毒抗原或者含有来自多种黄病毒序列的嵌合免疫原性黄病毒抗原。 本发明进一步包括核酸和蛋白疫苗以及该疫苗免疫受试者抗黄病毒感染的应用。 本发明还提供了由本发明的核酸编码的抗原, 响应于抗原引发产生的抗体以及抗原和/或抗体在检测黄病毒或诊断黄病毒感染中的应用。

1. 一种含有转录单位的分离的核酸，所述转录单位编码第一种黄病毒结构蛋白质的信号序列和第二种黄病毒的免疫原性黄病毒抗原，其中转录单位指导抗原的合成，以及其中信号序列是日本脑炎病毒信号序列。

2. 如权利要求 1 所述的核酸，其中信号序列通过缩短 n-区，改变 c-区或上述两种而修饰。

3. 如权利要求 1 所述的核酸，其中信号序列包括 SEQ ID NO:14。

4. 如权利要求 1 所述的核酸，其中信号序列包括 SEQ ID NO:27。

5. 如权利要求 1 所述的核酸，其中免疫原性黄病毒抗原属于一种选自黄热病毒、登革血清型 1 病毒、登革血清型 2 病毒、登革血清型 3 病毒、登革血清型 4 病毒、玻瓦桑病毒和西尼罗病毒的黄病毒。

6. 如权利要求 1 所述的核酸，其中转录单位编码西尼罗病毒的 E 蛋白和 M 蛋白质。

7. 如权利要求 1 所述的核酸，其中转录单位编码黄热病毒的 E 蛋白和 M 蛋白质。

8. 如权利要求 1 所述的核酸，其中转录单位编码圣路易斯脑炎病毒的 E 蛋白和 M 蛋白质。

9. 如权利要求 1 所述的核酸，其中转录单位编码玻瓦桑病毒的 E 蛋白和 M 蛋白质。

10. 如权利要求 1 所述的核酸，其中抗原选自黄病毒的 M 蛋白，黄病毒的 E 蛋白质，黄病毒的 M 蛋白和 E 蛋白质，黄病毒的部分 M 蛋白，黄病毒的部分 E 蛋白质以及黄病毒的部分 M 蛋白和黄病毒的部分 E 蛋白质或其任意组合。

11. 如权利要求 10 所述的核酸，其中抗原是黄病毒的 M 蛋白和 E 蛋白质。

12. 如权利要求 1 所述的核酸，其中抗原为圣路易斯脑炎病毒抗原。

13. 如权利要求 1 所述的核酸，其中抗原为黄热病毒抗原。

14. 如权利要求 1 所述的核酸，其中抗原为登革病毒抗原。

15. 如权利要求 1 所述的核酸，其中抗原为西尼罗病毒抗原。

16. 如权利要求 1 所述的核酸，其中核酸是 DNA。

17. 如权利要求 16 所述的核酸，其包含一个选自 SEQ ID NO:15, SEQ ID NO:19, SEQ ID NO:21 和 SEQ ID NO:23 的核苷酸序列。

18. 如权利要求 1 所述的核酸，其中转录单位包含一个适当排列的控制序列，从而其可操作地控制抗原的合成。

19. 如权利要求 18 所述的核酸，其中控制序列是巨细胞病毒立即早期启动子。

20. 如权利要求 1 所述的核酸，其包含一个位于多肽的翻译起始位点的 Kozak 共有序列，所述多肽含有由 TU 编码的抗原。

21. 如权利要求 1 所述的核酸，其中转录单位包含一个 poly-A 终止子。

22. 一种细胞，其含有如权利要求 1-21 任一项所述的核酸。

23. 一种组合物，其含有如权利要求 1-21 任一项所述的核酸和一种可药用的载体。

24. 如权利要求 1-21 任一项所述的核酸在制备用于免疫受试者抗黄病毒感染的药物中的应用。

25. 如权利要求 24 所述的应用，其中药物以单剂量施用。

26. 如权利要求 24 所述的应用，其中药物通过肠胃外途径给药。

27. 权利要求 1-21 任一项所述的核酸转录单位用于体外生成黄病毒抗原。

预防黄病毒感染的核酸疫苗

本申请是 2001 年 4 月 4 日提交的在审美国申请 No.09/826,115 的部分继续并且要求其权利，09/826,115 是 2000 年 11 月 29 日提交的在审美国申请 09/701,536 的部分继续并要求其权利，09/701,536 是 1999 年 6 月 3 日提交的国际申请 No.PCT/US99/12298 的国家审查阶段，并且要求 1998 年 6 月 4 日提交的美国临时申请 No.60/087,908 的权利，这些申请此处全部引作参考。

发明领域

本发明涉及新的疫苗、诊断试剂及其在治疗和预防由黄病毒引起的疾病中应用两者的方法。尤其是，所述疫苗为含有诸如日本脑炎病毒（JEV），西尼罗脑炎病毒（WNV）或相关黄病毒的黄病毒结构蛋白基因的重组核酸。这些疫苗在体内施用时用作病毒蛋白抗原生物合成的转录单位。诊断试剂为含有抗原的组合物，所述抗原从可用来检测黄病毒感染的重组核酸制备而来。

发明背景

黄病毒为黄病毒属的成员，其被分类在黄病毒科中。黄病毒主要对人类和其它哺乳动物致病。给人类和动物造成疾病的黄病毒包括 Alfuy, Apoi, Aroa, Bagaza, Banzi, Batu Cave, Bouboui, Bukalasa 蝙蝠, Bussuquara, Cacipacore, Carey Island, Cowbone Ridge, 达喀尔蝙蝠, 登革(血清型 1, 2, 3 和 4), Edge Hill, Entebbe 蝙蝠, Gadgets Gully, Iguape, Ilheus, 以色列火鸡脑膜脑炎 (Israel turkey meningoencephalitis), 日本脑炎 (Japanese encephalitis), Jugra, Jutiapa, Kadam, Karshi, Kedougou, Kokobera, Koutango, Kunjin, Kyasanur 森林出血热 (Kyasanur Forest disease), Langat, Meaban, Modoc, Montana myotis 白质脑炎 (leukoencephalitis), 墨累谷 (Murray Valley) 脑炎,

Naranjal, Negishi, Ntaya, Omsk 出血热, 金边蝙蝠 (Phnom Penh bat), Potiskum, 玻瓦桑 (Powassan), Rio Bravo, 罗西奥 (Rocio), Royal Farm, 俄国春夏脑炎 (Russian spring summer encephalitis), Saboya, Sal Vieja, San Perlita, Saumarez Reef, Sepik, Sokuluk, Spondweni, 圣路易脑炎, Stratford, 壁虱性脑炎-中欧亚型 (Tick-borne encephalitis-central European subtype), 壁虱性脑炎-远东亚型 (Tick-borne encephalitis-far eastern subtype), Tembusu, THCAr, Tyuleniy, Uganda S, Usutu, 西尼罗脑炎 (West Nile), Yaounde, 黄热病 (Yellow fever), Yokose, Ziki, 细胞融合试剂和其它的相关黄病毒列在 Kuno 等(*J. Virol.* 72:73-83 (1998))中。

黄病毒含有下列三种结构蛋白: prM/M, 前膜蛋白和膜蛋白; E, 包膜蛋白; 和 C, 衣壳蛋白。(Monath, *Virology* (Fields 编), Raven Press, New York, 1990, pp. 763-814; Heinz 和 Roehrig, 病毒免疫化学 II: 血清学诊断和疫苗基础 (Immunochemistry of Viruses II: The Basis for Serodiagnosis and Vaccines) (van Regenmortel 和 Neurath 编), Elsevier, Amsterdam, 1990, pp. 289-305)。M 的分子量(MW)大约为 7-8 千道而顿 (kDa), E 的分子量(MW)大约为 55-60kDa。M 的较大合成前体为 prM。在成熟病毒粒子中, prM 经加工形成 M 蛋白, prM 的 pr 部分被去除。M 和 E 位于病毒颗粒的膜上, 并且长期以来被认为构成了病毒的重要免疫原性部分。

黄病毒为 RNA 病毒, 含有单链 RNA, 在不同种中, 长度大约为 10 千碱基(kb)。MW 为 12-14kDa 的 C 蛋白和 RNA 复合形成核壳复合体。很多非结构蛋白也由 RNA 基因组编码, 其命名为 NS1, NS2A, NS2B, NS3, NS4A, NS4B 和 NS5。基因组在宿主细胞中被翻译成多蛋白, 然后通过病毒或者宿主特异性蛋白酶进行共翻译或者翻译后加工形成单独的基因产物 (图 1)

几种黄病毒的基因组核苷酸序列已知概述在美国专利 No.

5,494,671 中。JEV 序列由 Sumiyoshi 等(*Virology* 161:497-510 (1987)) 和 Hashimoto 等(*Virus Genes* 1:305-317 (1988))提供。JEV 的毒株 SA-14 和在中国作为疫苗使用的减毒株 SA-14-14-2 的核苷酸序列在 Nitayaphan 等人(*Virology* 177:541-552 (1990))的工作中进行了比较。

编码其它黄病毒种的结构蛋白的核苷酸序列也已经公知。在很多例子中, 已经报道了全基因组序列。报道的序列包括登革病毒血清型 1, 登革病毒血清型 2(Deubel 等, *Virology* 155:365-377 (1986); Gruenberg 等, *J. Gen. Virol.* 69:1391-1398 (1988); Hahn 等, *Virology* 162:167-180 (1988)), 登革病毒血清型 3(Osatomi 等, *Virus Genes* 2:99-108 (1988)), 登革病毒血清型 4(Mackow 等, *Virology* 159:217-228 (1987), Zhao 等, *Virology* 155:77-88 (1986)), 西尼罗病毒(Lanciotti 等, *Science* 286:2331-2333 (1999)), 玻瓦桑病毒(Mandl 等 *Virology* 194:173-184 (1993))和黄热病毒(YFV) (Rice 等, *Science* 229:726-733 (1985))。

包括圣路易脑炎病毒 (SLEV), WNV 和 JEV 的很多黄病毒通过蚊子传播给人类和其它宿主动物。因此这些病毒在很广泛的地区传播流行而且很难阻断或者预防其传播。

西尼罗热是一种蚊媒的黄病毒感染, 主要通过多种 *Culex mosquitoes* (库蚊) 传播给脊椎动物。类似于其它黄病毒在日本脑炎 (JE) 抗原复合物, 包括 JE, SLE 和墨累谷脑炎 (MVE) 病毒,, WNV 在节肢动物传播介体和鸟类之间保持了一种自然循环。该病毒最初在 1937 年分离自乌干达西尼罗地区的一个发热病人(Smithburn 等, *Am. J. Trop. Med. Hyg.* 20:471-492 (1940))。很快该病毒就被公认为一种分布最广泛的黄病毒, 其地理分布区包括非洲, 中东, 西亚, 欧洲和澳洲 (Hubalek 等, *Emerg. Infect. Dis.* 5:643-50 (1999))。临床上, 人类的西尼罗热是一种自限急性发热疾病, 伴随着头痛, 肌痛, 多关节病, 皮疹和淋巴结肿大 (Monath 和 Tsai, Clinical Virology, (Richman, Whitley 和 Hayden 编), Churchill-Livingstone, New York, 1997, pp.

1133-1186)。偶尔也报道急性肝炎或胰腺炎,在 WNV 感染的老年患者的例子中有时还并发脑炎或脑膜炎(Asnis 等, *Clin. Infect. Dis.* 30:413-418 (2000))。因此 WNV 感染在世界上的很多地区都是受到广泛关注的健康问题。

该疾病的地理传播,尤其是 1999 年 WNV 进入美国,大大提高了对于患该病的人类和动物的健康关注意识。在 1999 年 8 月底和 9 月初,纽约和周边地区经历了病毒性脑炎的爆发,62 例确诊病例中有 7 例死亡。与这次爆发同时,地方卫生官员观察到鸟类(特别是乌鸦)和马死亡率的提高。后来根据人类,鸟和蚊子物种的单克隆抗体(Mab)图谱和基因组序列检测发现这次爆发是由 WNV 引起的(Anderson 等, *Science* 286:2331-2333 (1999); Jia 等, *Lancet* 354:1971-1972 (1999); Lanciotti 等, *Science* 286:2333-2337 (1999))。随后在冬天的几个月内检测到的病毒活性显示该病毒已经在北美定居 (*Morb. Mortal. Wkly. Rep.* 49:178-179 (2000); Asni 等, *Clin. Infect. Dis.* 30:413-418 (2000); Garmendia 等, *J. Clin. Micro.* 38:3110-3111 (2000))。在 2000 年间来自东北和大西洋中部州报道的监视数据证实了动物流行/流行传播的增强以及记录了鸟类,蚊子和马,以及人类的大量病毒感染个例地理分布的延伸(*Morb. Mortal. Wkly. Rep.* 49:820-822 (2000))。

目前,还没有现成的人用或者兽用疫苗预防 WNV 的感染,蚊子的控制是唯一抗击疾病蔓延的实用策略。

日本脑炎病毒(JEV)感染成人和儿童,并在热带和亚热带的亚洲地区的婴儿,儿童和老人中有很高的死亡率(Tsai 等, *Vaccines* (Plotkin 编) W. B. Saunders, Philadelphia, Pa, 1999, pp. 672-710。在幸存者中,有涉及脑炎症状感染后遗留的严重神经症状。在这个地区的发达国家,诸如日本,台湾和韩国,JEV 通过使用灭活 JEV 疫苗已经大部分得到控制。但是在这个地区的其它国家仍在流行。

合适用于抗 JEV 感染的疫苗包括通过诸如福尔马林方法处理的灭活病毒，以及减毒病毒(Tsai 等, Vaccines (Plotkin 编) W. B. Saunders, Philadelphia, Pa, 1994, pp. 671-713)。全病毒疫苗，虽然有效，但确实存在一定的问题和/或缺点。病毒培养在小鼠的脑部或采用哺乳动物细胞作为宿主的细胞培养物中。这种培养方法很烦琐并很昂贵。此外，也伴随着将宿主细胞，即脑或者其它宿主的抗原掺入最终疫苗产物的危险，可能导致疫苗受体不想要和不需要的变态反应。也存在参与疫苗生产的工人被不经意感染的危险。最后，还存在病毒没有完全或者全部灭活或者减毒，因而疫苗可能实际上引起疾病的危险。

登革热和登革出血热（DF/DHF）由登革病毒引起，也是蚊媒黄病毒。有四种抗原性相关，但是不同的登革病毒血清型(DEN-1, DEN-2, DEN-3 和 DEN-4)，所有都能引起 DF/DHF。DF 属于登革相关疾病的温和型，其症状包括发热，皮疹，严重头痛和关节痛。这些患有 DF 的个体死亡率很低，但是患有 DHF 的个体的死亡率可能高达 5%。可靠的证据显示，在过去的 40 年间，多于 3 百万的 DHF 个例和 58,000 的死亡个例是由于 DHF 引起，使得 DHF 成为一种主要的新出现疾病（Halstead, 登革和登革出血热（Dengue and Dengue Hemorrhagic Fever）(Gubler 和 Kuno 编) CAB International, New York, NY, (1997) pp 23-44)。然而，尽管进行了几十年的努力，还没有出现适合的防止登革病毒感染的有效的疫苗。

黄热病在南美洲热带地区和非洲撒哈拉沙漠以南地区流行，并由蚊子传播。感染导致发热，寒战，严重头痛和其它疼痛，厌食，恶心和呕吐，并伴随黄疸的出现。培养在感染鸡胚中的活病毒疫苗 17D，被认为是安全有效的。但是，仍然需要在诸如在非洲热带地区和最需要疫苗的美洲地区通常遇到的恶劣条件下稳定存在的疫苗。

两种黄病毒嵌合体的重组黄病毒公开于 PCT 申请 WO 93/06214 中。该嵌合体是融合了一种“类型”，或血清型的登革病毒或黄病毒

的结构蛋白和另一种不同“类型”，或血清型的登革病毒或其它黄病毒的结构蛋白的构建体。

近年来，一些重组亚单位和病毒疫苗已被设计。美国专利 No.4,810,492 描述了生产 JEV 的 E 糖蛋白作为疫苗的抗原。将相应的 DNA 克隆到表达系统，目的是在合适的宿主细胞，诸如，大肠杆菌，酵母，或者高等生物细胞培养物中表达抗原蛋白。美国专利 No.5,229,293 公开了包含 JEV E 蛋白基因的重组杆状病毒。该病毒用于感染培养物中的昆虫细胞，以产生 E 蛋白并回收用作疫苗。

美国专利 No. 5,021,347 公开了一种整合了 JEV E 蛋白基因的重组牛痘病毒基因组。活重组牛痘病毒作为疫苗免疫抗 JEV。整合了登革血清型 2，登革血清型 4 和 JEV 的 E 蛋白平截 C-末端基因的重组牛痘病毒和杆状病毒公开于美国专利 5,494,671。美国专利 5,514,375 公开了多种重组牛痘病毒，其表达从 prM 延伸到 NS2B 的部分 JEV 可读框。这些痘病毒诱导形成含有加工后 M 蛋白和 E 蛋白的细胞外颗粒。编码这些 JEV 蛋白的两种重组病毒在小鼠体内产生了高滴度的中和抗体和血凝素—抑制抗体，以及保护性免疫。两种免疫处理后的影响程度高于只进行一种免疫处理的情况。当含有 JEV 的 prM/M 和 E 蛋白的基因的重组痘病毒给小鼠施用后赋予小鼠保护性免疫（Konishi 等，*Virology* 180:401-410 (1991)）。用携带 JEV prM 和 E 基因的重组痘病毒感染 HeLa 细胞后显示产生亚病毒颗粒（Konishi 等，*Virology* 188:714-720 (1992)）。Dmitriev 等人报道了用编码壁虱性脑炎病毒（tick-borne encephalitis virus）结构蛋白和特定非结构蛋白的重组痘病毒免疫小鼠（*J. Biotechnology* 44:97-103 (1996)）。

重组病毒载体也已被制备用作登革热的病毒疫苗。Zhao 等人（*J. Virol.* 61:4019-4022 (1987)）制备了携带登革血清型 4 结构蛋白和 NS1 的重组痘病毒并且用重组病毒感染哺乳动物细胞后获得了表达。用重组杆状病毒感染靶昆虫细胞后也获得了类似表达（Zhang 等，*J. Virol.* 62:3027-3031 (1988)）。Bray 等（*J. Virol.* 63:2853-2856 (1989)）也报道了

一种基于 E 蛋白基因的重组痘登革疫苗，其能够对小鼠实施保护性免疫抵抗登革脑炎的攻击。Falgout 等(*J. Virol* 63:1852-1860 (1989))和 Falgout 等(*J. Virol.* 64:4356-4363 (1990))报道了类似的结果。Zhang 等(*J. Virol* 62:3027-3031 (1988))的研究显示编码登革 E 和 NS1 蛋白的重组杆状病毒同样保护小鼠免遭登革脑炎的攻击。其它结构基因和非结构基因整合到重组病毒疫苗的组合不能产生显著免疫 (Bray 等, *J. Virol.* 63:2853-2856 (1989))。同样，用表达 E 蛋白的重组杆状病毒免疫猴子时，对登革病毒攻击不能产生完全的保护性免疫(Lai 等(1990) pp. 119-124, F. Brown, R. M. Chancock, H. S. Ginsberg 和 R. Lerner (编) 疫苗 90: 生产包括预防 AIDS 的新疫苗的现代方法 (Vaccines 90:Modern approaches to new vaccines including prevention of AIDS), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY。

使用重组 DNA 制备物的免疫已被报道于使用断乳小鼠作为模型的 SLEV 和登革病毒-2 (Phillpotts 等, *Arch. Virol.* 141:743-749 (1996); Kochel 等, *Vaccine* 15:547-552 (1997))。通过使用单剂量或双剂量的 DNA 免疫，编码 SLEV 的 prM 和 E 基因的质粒 DNA 提供了抗 SLEV 攻击的部分保护。在这些试验中，对照小鼠显示大约 25% 的存活率，DNA 免疫的小鼠中没有检测到保护性抗体 (Phillpotts 等, *Arch. Virol.* 141:743-749 (1996))。在接受 3 次皮内注射含 prM 的重组登革-2 质粒 DNA 的小鼠中，100% 产生了抗登革-2 中和抗体，并且 92% 接受相应 E 基因的小鼠同样产生了中和抗体(Kochel 等, *Vaccine* 15:547-552 (1997))。但是采用双剂量方案的攻击试验没能保护小鼠免于致命登革-2 病毒的攻击。

用来定期免疫抗 JEV, SLEV, 登革病毒和其它黄病毒感染的疫苗在应用中有很多缺点和问题。灭活疫苗的制备费用昂贵也很困难。此外，任一这种疫苗都必须承担来自制备病毒中所使用宿主细胞蛋白变态反应的危险。另外，这些疫苗对于从事生产的工人也有很大的危险性。候选减毒 JEV 疫苗正在进行临床试验，但是自 1996 年以来还

未在中国以外的地方没有获得广泛的接受(Hennessy 等, *Lancet* 347:1583-1586 (1996))。

基于仅仅使用诸如 JEV 黄病毒特定蛋白的重组疫苗, 其生产方法是在细胞培养物中生物合成表达后纯化或处理抗原, 并不诱导高滴度的抗体。同样, 类似于全病毒制备物, 这些疫苗具有对宿主或者载体抗原不利的变态反应的危险。抗登革病毒和 WNV 疫苗的开发还很落后, 且这些以病毒为基础的或者以重组蛋白为基础的疫苗面临着如上所述疫苗的相同问题。

因此需要开发直接抗诸如黄热病毒, 登革病毒, JEV, SLEV 和 WNV 的黄热病毒的疫苗或者改良疫苗, 其制备要便宜, 而且对于参与生产的工人呈现很小的危险性, 由于杂质或者不定的免疫原性部分而导致不利变态反应的危险性最低, 并且高效诱导中和抗体和保护性免疫。此外还需要抗 JEV, WNV 和相关黄病毒的疫苗使需要的免疫剂量数最小化。

如详述的生产疫苗的现有技术的很多缺点也适于免疫诊断试剂的抗原和抗体的生产中。尤其是, 在生产病毒抗原时并发的危险性和费用, 以及大多数目前可用的重组表达抗原诱导有效免疫应答的失败, 在免疫诊断试剂领域与相同的危险性, 高昂的费用和相应灵敏度的缺乏并行。因此, 由于用活病毒意外感染的高昂费用和危险性, 而且灵敏度低于从前试验所需要的水平, 所以目前需要开发快速, 简单, 高敏感度的诊断试剂用于黄病毒感染和/或污染的检测。

本发明通过提供高效免疫原性重组抗原用于检测针对待选黄病毒的抗体的诊断方法满足了这些需要。本发明还提供了衍生自黄病毒, 黄病毒基因或者其类似物的重组抗原在检测针对黄病毒蛋白的抗体的免疫诊断方法中的应用。

发明概述

本发明提供了含有免疫原性黄病毒抗原转录单位(TU)的核酸分子。TU 整合到细胞后指导细胞合成抗原。本发明的一个重要方面,黄病毒可以为黄热病毒(YFV),登革病毒血清型 1(DEN-1),登革病毒血清型 2(DEN-2),登革病毒血清型 3(DEN-3),登革病毒血清型 4(DEN-4),圣路易脑炎病毒(SLEV),日本脑炎病毒(JEV),西尼罗病毒(WNV),玻瓦桑病毒或者任一其它黄病毒。在本发明的重要实施方案中,抗原可以为黄病毒 prM/M 蛋白, E 蛋白, 或两者。在本发明的重要实施方案中,抗原可以为嵌合黄病毒蛋白。尤其是,当 TU 包括 prM/M 和 E 蛋白时,宿主细胞分泌含有 prM/M 和 E 抗原的亚病毒颗粒。本发明的另一个重要方面,核酸为 DNA 分子。在另一个重要实施方案中,核酸 TU 包括适当排列的控制序列,其可操作地控制 prM/M 和 E 抗原的表达,这个控制序列可以为巨细胞病毒立即早期启动子。在另一个实施方案中,TU 的核苷酸序列经改造后优化在真核细胞中的翻译,操作方法是使由 TU 产生的 mRNA 的 5'-端非翻译区的大发夹结构最小化和/或在由 TU 产生的 mRNA 的翻译起始位点包含 Kozak 共有序列。在另外的实施方案中,转录单位也包括聚腺苷酸(poly-A)终止子。

本发明也提供了含有核酸分子的宿主细胞,该核酸分子包含免疫原性黄病毒抗原的转录单位指导宿主细胞合成免疫原性抗原。黄病毒可以为 YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, 玻瓦桑病毒或其它黄病毒。在重要的实施方案中。抗原可以为 prM/M 蛋白, E 蛋白, 或 prM/M 和 E 蛋白两者。在后者情况下,细胞分泌含有 prM/M 和 E 抗原的亚病毒颗粒。

此外,本发明提供了免疫接种受试者抗黄病毒感染的组合物,其包括含有免疫原性黄病毒抗原转录单位的核酸分子。该转录单位在受试者体内细胞中整合之后,就指导该细胞合成免疫原性抗原。该组合物还含有药学上可接受的载体。在重要的实施方案中,黄病毒可以为

YFV, DEN-1, DEN-2, DEN3, DEN-4, SLEV, JEV, WNV, 玻瓦桑病毒或其它黄病毒。此外,抗原可以为 prM/M 蛋白,E 蛋白,或 prM/M 和 E 蛋白两者。在后者的情况下,细胞分泌含有黄病毒 prM/M 和 E 抗原的亚病毒颗粒。这些亚病毒颗粒也称为非传染性的重组抗原 (NRA)。在重要的实施方案中,核酸分子为 DNA 分子。在其它的重要实施方案中,转录单位还含有适当排列的控制序列,当核酸导入个体细胞后,所述控制序列可操作地控制 prM/M 和 E 抗原的合成。这个控制序列可以为巨细胞病毒立即早期启动子。在另外的实施方案中,转录单位也包括聚腺苷酸终止子。

本发明提供的用于免疫接种个体抗黄病毒的组合物可以包括一种或多种核酸分子,其包括一种以上免疫原性黄病毒抗原的转录单位。一种以上免疫原性黄病毒抗原可以来自不同的黄病毒种,株或者任一组合的分离物。在重要的实施方案中,黄病毒可以包括两种或更多种,三种或更多种,四种或更多种,五种或更多种,或者七种或更多种黄病毒。这些黄病毒的例子包括,但不限于 YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, 玻瓦桑病毒或其它黄病毒。可以设计组合疫苗以对特定地理地区常见的黄病毒疾病提供免疫。直接针对热带和亚热带的亚洲的特定实施方案中,可以选择 DEN-1, DEN-2, DEN-3, DEN-4, WN, 和 JE 病毒。在直接针对非洲的特定实施方案中,可以选择 DEN-1, DEN-2, DEN-3, DEN-4, WN 和 YF。在直接针对拉丁美洲的特定实施方案中,可以选择 DEN-1, DEN-2, DEN-3, DEN-4, Rocio 和 YF 病毒。

本发明进一步提供了免疫受试者抗黄病毒感染的方法。该方法包括对受试者施用有效剂量的免疫接种组合物,该组合物含有包括免疫原性黄病毒抗原转录单位的核酸分子。转录单位被细胞吸收后指导受试者体内的细胞合成免疫原性抗原。组合物还包括药学上可接受的载体。在方法的重要实施方案中,黄病毒可以为 YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, 玻瓦桑病毒或其它黄病毒。在

方法的另外一个重要方面，抗原可以为 prM/M 蛋白，E 蛋白，或 prM/M 和 E 蛋白两者。当抗原为 prM/M 和 E 蛋白两者时，受试者体内的细胞，在其中整合了核酸后，分泌含有黄病毒 prM/M 和 E 抗原的亚病毒颗粒。此外，在方法的重要实施方案中，免疫接种的组合物以单剂量通过肠胃外途径施用给受试者。该方法的另一方面，核酸为 DNA 分子。在该方法的其它实施方案中，转录单位还包括合适排列的控制序列，其可操作地控制 prM/M 和 E 抗原的合成，以及在该实施方案的重要方面，这个控制序列为巨细胞病毒立即早期启动子。此外，转录单位也包括聚腺苷酸终止子。

本发明的这些方面和实施方案是其明显特征和优点的基础。作为只含有部分黄病毒基因组而非包括全基因组序列的核酸构建体，含有核酸 TU 的疫苗是完全非存活的。因此对于参与其生产或者接受疫苗的受试者不存在被黄病毒感染的危险。核酸疫苗制备简单，施用简单，且在使用前储存稳定。令人意外地已发现，只给哺乳动物施用单剂量的本发明的核酸疫苗后，基本上 100% 成功地提供了保护性免疫。另外一个令人意外的结果是核酸 TU 能够在雌性哺乳动物体内产生对于黄病毒的免疫可通过乳液传递给后代。不希望被理论所限制，本发明人认为核酸成功提供保护性免疫的可能机制是因为含有核酸的宿主细胞，诸如施用疫苗的受试者的细胞，产生了含有黄病毒 prM/M 和 E 抗原的亚病毒颗粒。这些颗粒具有类似于天然的黄病毒粒子的免疫原性特征。

本发明也提供了非感染性抗原多肽，抗原多肽片段和含有黄病毒 prM/M 和/或 E 蛋白的 NRA，其中跨膜信号序列衍生自第一种黄病毒，而 M 和/或 E 蛋白来自第二种黄病毒。此外，prM/M 蛋白可以含有来自第一种和第二种黄病毒的氨基酸序列。另外，E 蛋白可以包含来自第一种和第二种黄病毒的氨基酸序列。此处应用的“嵌合体”是指含有来自一种以上黄病毒序列的任一蛋白或核酸。此处所用的“无毒”是指本发明的抗原或者疫苗不能引起疾病。更具体而言，重组蛋白抗

原不含有来自黄病毒的对于黄病毒感染，复制和致病所必须的污染的基因组材料。

本发明的多肽可以含有此处定义，或者本领域技术人员公知的所选择黄病毒 prM，M 和/或 E 蛋白的氨基酸序列。本发明的核酸可以含有编码所选择黄病毒 prM，M 和/或 E 蛋白的核苷酸序列。

本发明的抗原可以为非连接的，或者可以与载体分子结合以有助于将抗原放置在固相上。载体分子可以为一种与抗原共轭结合的分 子，并且载体分子不与人血清中的抗体反应。这种载体的一个例子是小牛血清白蛋白（BSA）。

本发明的抗原也可以为重组蛋白，其通过在能够产生抗原的表达系统中表达编码抗原的核酸而获得。

本发明抗原的氨基酸序列含有 prM，M 和/或 E 抗原的免疫反应部分。这些抗原可以另外结合到其它序列，这些序列就是设计用来提供一些其它特征的，诸如去除/添加能够进行二硫键连接的氨基酸提供更坚固的二级结构以提高表位的反应活性，从而提高其生物存活期或者改变其毒性或者预防感染。在任一情况下，抗原必须具有免疫反应性和/或免疫原性。

附图简述

图 1 为黄病毒多蛋白加工的示意图。中间水平区提供了病毒基因组的示意图。划线表示 5'和 3'非-翻译区，盒区代表结构蛋白(左侧和上部)和非结构蛋白（右侧和下部）的可读框。宿主细胞信号酶在 E 蛋白的 C-末端的剪切和翻译同时发生，将结构区和非结构区分开。弗林蛋白酶（*furin*）为类枯草杆菌酶（*subtilase-like*）细胞酶，可能负责 prM 剪切。病毒多蛋白的可能跨膜区由阴影区表示。

图 2 是 JEV 基因组(上面)的图谱，以及用于逆转录酶-聚合酶链

式反应 (RT-PCR) 的寡核苷酸 DNA 序列 (中间) 以构建表达 prM-E 蛋白编码区 (下面) 转录单位。病毒多蛋白的可能跨膜区由阴影区表示。

图 3 显示了质粒载体 pCDNA3, pCBamp 和 pCIBamp 以及它们之间关系的示意图。这些质粒 CMV(巨细胞病毒)启动子/增强子元件, BGHp(A)(牛生长激素多腺苷酸化信号和转录终止序列), 包括用于大肠杆菌中选择和维持的氨苄青霉素抗性基因和 ColEI 复制起点。将大肠杆菌中的单链拯救复制起点 fl, SV40 复制起点(SV40 ORI), 新霉素抗性编码区和 SV40p(A)序列从 pCDNA3 中删除产生 pCBamp。内含子序列插入到 pCBamp 的 NcoI-KpnI 位点产生质粒 pCIBamp。

图 4 显示了来自 JE-4B COS-1 培养液的蔗糖梯度纯化的亚病毒颗粒 SDS-PAGE-免疫印迹分析 (4B, 每对泳道的右侧泳道)。来自 JEV 感染的 C6/36 细胞培养物的密度梯度纯化的 JE 病毒粒子作为阳性对照 (JEV, 每对泳道的左侧泳道)。JE HIAF (超免疫腹水); 4G2, 抗-E 单克隆抗体; JM01, 抗-M 单克隆抗体; NMAF (正常小鼠腹水)。

图 5 显示了差速区带蔗糖梯度分析的 E 抗原谱, E 抗原从用或不用 Triton X-100 处理的 JE-4B 细胞培养基的 PEG 沉降制备而来。

图 6 显示了由 SignalP-HMM 软件程序 (A) 预测的 pCBJE1-14 (pCBJE)信号肽的概率。通过在-4 和-2 位置(C-4G 和 G-2S) (板 B, JELSS-M) 改变 c-区的序列, 通过缩短 n-区(板 C, JE-SS-ORI), 或者通过这两种修饰的组合(板 D, JE-SS-M), 信号肽概率得到提高。

图 7 显示了质粒载体 pCBD2-14-16(100% DEN-2 E), pCBD2-1J-4-3 (90% DEN-2 E:10% JEV E), 和 pCB8D2-2J-2-9-1 (80% DEN-2 E:20% JEV E)的示意图。这些质粒包括人巨细胞病毒(CMV)早期基因启动子, JE 病毒信号序列; DEN-2 病毒 prM 和 E 基因区(氨基末端分别为 100%, 90%, 或 80%); JE 病毒 E 基因区(分别为无, 10%或 20%); 以及牛生长激素 poly A 信号(BGH)。

图 8 显示了通过 western 印迹比较分泌型和膜结合的重组蛋白。(A)PEG-沉降和用乙醇从培养液中提取 DEN-2 质粒 pCB8D2-2J-29-1, pCB9D2-1J-4-3, pCBD2-14-16, 和对照质粒 pEGFP 后, 分析分泌的重

组抗原。泳道 1(V)，金印迹 (Gold Blot) (Owl Separation Systems, Portsmouth, NH)染色的纯化 DEN-2 病毒。来自每种质粒分泌的重组抗原与下列物质的反应性：a，抗-包膜(E)特异性 Mab 1A6A-8；b，MAB 1A6A-8a, 抗-衣壳(C)特异性 Mab 1A2A-1 和特异于 DEN-2 病毒前膜(prM)蛋白的抗-血清的混合物；和 c，正常小鼠腹水。(B)重组质粒-转化的细胞疏水膜蛋白的分析。泳道 1(V)，通过金印迹染色的纯化的 DEN-2 病毒；泳道 2(V)，纯化 DEN-2 病毒和 Mab 1A6A-8，Mab 1A2A-1，特异于 DEN-2 病毒 M 蛋白的抗-血清，和特异于 DEN-2 病毒 prM 蛋白的抗-血清的混合物的反应性。自每种质粒-转化的细胞系分离的疏水性膜蛋白与下列物质的反应性：a，Mab 1A6A-8；b，Mab 1A6A-8，Mab 1A2A-1，特异于 DEN-2 病毒 M 蛋白的抗-血清，和特异于 DEN-2 病毒 prM 蛋白的抗-血清的混合物；和 c，正常小鼠腹水。

本发明的详细描述

本发明包括编码诸如 prM/M 和 E 蛋白抗原的黄病毒抗原蛋白的核酸转录单位。当核酸被合适的细胞吸收时，尤其是当细胞为受试者的细胞时，核酸的功能是表达 prM/M 和 E 蛋白抗原。本发明也包括一种疫苗，其活性剂为核酸转录单位 (TU)。本发明进一步包括含有 TU 的细胞。此外，本发明包括通过给受试者施用有效剂量的含核酸 TU 分子的疫苗免疫受试者抗黄病毒感染的方法。

本发明提供了包含转录单位的分离核酸，该转录单位编码第一种黄病毒结构蛋白的信号序列和第二种黄病毒的免疫原性黄病毒抗原，其中的转录单位指导抗原的合成。本发明进一步包括利用核酸转录单位 (TU) 产生黄病毒抗原以及由核酸 TU 产生的黄病毒抗原。本发明包括的黄病毒抗原包括整合了第一种黄病毒和至少另一种黄病毒的氨基酸序列的嵌合黄病毒抗原。本发明还包括利用由本发明的 TU 编码的黄病毒抗原产生黄病毒-特异性抗体以及检测黄病毒-特异性抗体的存在。

在一个实施方案中，本发明分离的核酸可以含有编码日本脑炎病毒信号序列的转录单位。

在另一个实施方案中，本发明的转录单位可以编码一种免疫原性黄病毒抗原，其可来自一种或多种下列黄病毒：黄热病毒；登革病毒血清型 1，登革病毒血清型 2，登革病毒血清型 3，登革病毒血清型 4，日本脑炎病毒，玻瓦桑病毒和西尼罗病毒。

在另一个实施方案中，本发明的转录单位可以编码一种免疫原性嵌合黄病毒抗原，其可包括来自一种以上下列病毒的序列：黄热病毒；登革病毒血清型 1，登革病毒血清型 2，登革病毒血清型 3，登革病毒血清型 4，日本脑炎病毒，玻瓦桑病毒和西尼罗病毒。

在一个具体的实施方案中，本发明的核酸可编码日本脑炎病毒的信号序列和西尼罗病毒，SLEV，YFV 和/或玻瓦桑病毒的 M 蛋白和 E 蛋白。核酸还可编码免疫原性抗原，其可以是一种黄病毒的 M 蛋白，一种黄病毒的 E 蛋白，一种黄病毒的 M 蛋白和 E 蛋白，一种黄病毒的部分 M 蛋白，一种黄病毒的部分 E 蛋白和/或一种黄病毒的部分 M 蛋白和一种黄病毒的部分 E 蛋白。在一个优选的实施方案中，分离的核酸编码一种黄病毒的 M 蛋白和 E 蛋白。另外，本发明的核酸可以为 DNA，并可包括核苷酸序列 SEQ ID NO:15，SEQ ID NO:19，SEQ ID NO:21，SEQ ID NO:23 或 SEQ ID NO:42。

在另一个具体的实施方案中，本发明的核酸可以编码日本脑炎病毒的信号序列，第二种病毒的 M 蛋白和嵌合 E 蛋白，所述嵌合 E 蛋白通过用编码相应部分的 JEV E 蛋白的核酸取代编码第二种病毒 E 蛋白的核酸部分取代形成。或者，相应于第二种病毒 E 蛋白的缺失部分的部分序列可以被选自第三种病毒的其它序列取代，或者为非病毒序列。第二种蛋白可以为西尼罗病毒，SLEV，YFV，玻瓦桑和/或登革病毒的一种血清型的蛋白。嵌合 E 蛋白可以包括那些其 C 末端部分来

自一种黄病毒，而嵌合 E 蛋白的剩余部分来自另一种黄病毒。C 末端部分可以为，例如嵌合 E 蛋白的 5，10，15，20，25，30，35，40，45，50，60 或 75%。本发明的核酸可以为 DNA，并可以包括来自核苷酸序列 SEQ ID NO:44 或 SEQ ID NO:46 的蛋白编码序列。本发明的核酸可以包括核苷酸序列 SEQ ID NO:44 或 SEQ ID NO:46。

本发明的转录单位还可以包括适当排列的控制序列，其可操作地控制抗原的合成。控制序列可以为，例如，巨细胞病毒立即早期启动子。本发明的核酸还可以包括 Kozak 共有序列，所述共有序列位于含有由转录单位编码的抗原多肽的翻译起始位点。本发明的转录单位还可以包括 polyA 终止子。

本发明进一步提供了含有本发明核酸的细胞。

本发明还提供了含有药学上可接受的载体和本发明的核酸或者细胞或抗原的组合物。本发明还提供了免疫受试者抗黄病毒感染的方法，包括给受试者施用有效剂量的本发明的组合物。在具体的实施方案中，用来免疫受试者的组合物指导黄病毒的 M 蛋白和 E 蛋白两者的合成，以及受试者体内的细胞在其中整合核酸后，分泌含有 M 蛋白和 E 蛋白的亚病毒颗粒。或者，组合物可以含有一种黄病毒的 M 蛋白和/或 E 蛋白或者含有 M 蛋白和 E 蛋白的亚病毒颗粒。在本发明的方法中，免疫部分可以以单剂量通过肠胃外给药的途径给受试者施用。

本发明进一步提供了产自本发明分离核酸的抗原。作为例子，由 TU 核苷酸序列编码的来自第二种黄病毒的抗原可以为 M 蛋白，其可来自，例如，西尼罗病毒。抗原也可以为来自登革病毒，圣路易脑炎病毒，日本脑炎病毒，玻瓦桑病毒和/或黄热病毒的蛋白。在进一步的实施方案中，抗原包括一种 prM/M 蛋白，所述蛋白含有来自第一种黄病毒的跨膜信号序列和来自第二种黄病毒 prM/M 蛋白剩余部分的其它

氨基酸序列，第二种黄病毒可选自 SLEV, JEV, YFV, WNV, 和/或玻瓦桑病毒。来自第一种黄病毒的跨膜信号序列可以为改良的或者修饰的信号序列，其中信号序列被赋予了需要的特征，诸如高度信号序列可能性。通过设计或者选择完成这些目的可以通过使用机器学习计算机程序，包括但不限于使用隐马尔科夫模型（hidden Markov Model）的程序。

由 TU 的核苷酸序列编码的抗原可以为西尼罗病毒抗原，登革病毒抗原，圣路易脑炎病毒抗原，日本脑炎病毒抗原，玻瓦桑病毒和/或黄热病毒抗原。

由 TU 的核苷酸序列编码的抗原也可以为 E 蛋白，其可以为来自西尼罗病毒，登革病毒，圣路易脑炎病毒，日本脑炎病毒，玻瓦桑病毒和/或黄热病毒的 E 蛋白。这些编码的抗原也可以为包含选自一种以上黄病毒的氨基酸序列的嵌合 E 蛋白。

此外，由 TU 的核苷酸序列编码的抗原可以为 M 蛋白和 E 蛋白，可以选自西尼罗病毒，登革病毒，圣路易脑炎病毒，日本脑炎病毒，玻瓦桑病毒和/或黄热病毒。

此处所用的“M 蛋白”或者“pr/M 蛋白”或者“prM/M 蛋白”是指黄病毒 M 蛋白或黄病毒 prM 蛋白。例子包括但不限于，含有一种或者多种黄病毒 prM 蛋白的氨基酸序列的 prM 蛋白，不含有其它氨基酸序列的 M 蛋白以及含有其它氨基酸序列的蛋白，它们经体内或者体外加工产生成熟的 M 蛋白。

此处所用的“核酸转录单位”或“核酸转录单位分子”是指编码一种或者多种特异性蛋白的核酸。TU 具有被导入合适的细胞后，核酸就能诱导合成一种或者多种由该核酸编码的特异性基因产物的生物活性。一种或多种基因产物为其它的生物大分子，诸如蛋白，与 TU

化学上不相关。核酸 TU 诱导细胞利用其细胞组分产生由 TU 的核酸编码的特异性基因产物。虽然，任一核酸都可以用作 TU，但是在优选的实施方案中，TU 为质粒或者类似载体的 DNA，其中质粒或者载体包含标志基因或者其它序列构建体的编码序列，所述编码序列有助于使用 TU 进行试验和生物合成。

此处所用的“控制序列”是整合到 TU 中的调控核苷酸序列，其与细胞的合适细胞组分相互作用并增强或者激活由 TU 编码的基因产物的生物合成。因此，合适的控制序列为细胞组分可以与其相互作用后合成基因产物的序列。当根据特定编码序列可操作地排列在核酸中时，控制序列可以有效地控制特定核酸的表达以产生基因产物。

此处所用的“启动子”为 TU 中的一段核苷酸序列作为控制序列。

此处所用的“Kozak 序列”或者“Kozak 共有序列”是在翻译起始位点的核苷酸序列，用来优化真核 mRNA 的翻译(Kozak, *Mol. Cell. Biology* 9:5134-5142 (1989))。

此处所用的“终止子”是一段延伸的核苷酸序列，用来诱导成熟 mRNA 的 3'端聚腺苷酸化。终止子序列发现在特定编码序列的后面或者下游。

此处所用的“细胞”是含有编码一种或者多种基因产物的原核或者真核细胞，或者 TU 被导入的细胞。因此，细胞含有外来的或者异源物质 TU，其不是自然或者内源性细胞组分。合适的细胞是在导入 TU 后能够生物合成基因产物的细胞。尤其是，合适的细胞为能够响应控制序列和终止子序列的细胞，如果有的话，所述序列包含在 TU 之中。在本发明重要的实施方案中，细胞为哺乳动物细胞。在本发明尤其重要的实施方案中，细胞为在施用了 TU 作为疫苗组分的人类或者非人类受试者体内自然存在的细胞。或者，在分析或者诊断应用中，

包括制备抗原用作疫苗或者用于免疫诊断分析，或者为了说明的目的，细胞可以为在体外培养的人类或者非人类的细胞。

此处所用的特异于特定病原体的“疫苗”或者“用于免疫接种受试者的组合物”是指一种制备物，当施用给受试者后引起受试者的免疫原性反应。此处所用的“免疫原性反应”是指赋予受试者抗病原体的保护性免疫。不希望受限于理论，相信免疫原性反应可能来自中和抗体的产生（即，体液免疫反应）或者来自免疫系统的胞毒细胞（即，细胞免疫反应）。此处所用的“免疫原性抗原”为当将其导入受试者后，或者当其在宿主或者受试者体内合成后诱导免疫原性反应的抗原。此处所用的“有效剂量的”疫苗或者免疫接种组合物是指当施用给受试者后使用量足以给受试者赋予保护性免疫。从历来的观点看，疫苗被理解为含有作为有效成分的一种或者多种含有病原体，尤其是其表面的特异性分子组分或者结构。这些结构可以包括通常在病原微生物中发现的诸如蛋白，复合碳水化合物，和/或复合脂类的表面组分。

此处所用的，但是必须强调的术语“疫苗”或者“免疫接种受试者的组合物”延伸了前段概述的传统含义。此处所用的这些术语也涉及本发明的 TU 或者含有 TU 的组合物。TU 诱导了一种或者多种由受试者的细胞中 TU 所编码的特异性基因产物的生物合成，其中基因产物为病原体的特异性抗原。生物合成的抗原作为免疫原。和已经说明的一样，TU，也就是疫苗可以为编码特异性免疫原性抗原的任一核酸。在本发明优选的实施方案中，疫苗的 TU 为 DNA。TU 可以包括整合了为了分子生物学领域，细胞生物学领域，和病毒免疫学领域的技术人员操作方便的其它基因或者特定序列的质粒或载体(参见 Molecular Cloning: A Laboratory Manual, 第 2 版, Sambrook, Fritsch 和 Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; 和 Current Protocols in Molecular Biology, Ausubel 等, John Wiley 和 Sons, New York 1987 (最新的季刊), 此处引作参考)。

本发明的 TU 分子包括核酸，或者核酸衍生物，具有编码与黄病毒的抗原相关的特定基因产物的核苷酸序列，这些黄病毒包括但不限于，WNV，JEV，登革病毒，黄热病毒和 SLEV。虽然任一核酸都可以作为 TU，但是在重要实施方案中，TU 为 DNA。或者，核酸可以为 RNA 分子。核酸还可以为任一种具有磷酸二酯键主链的 DNA 或者 RNA 的若干种衍生物，该主链进行了化学修饰提高了 TU 作为药剂的稳定性。预想的修饰包括但不限于，硫代磷酸酯衍生物或者磷酸酯衍生物。衍生物的这些和其它例子是核酸化学领域技术人员公知的。

JEV 的基因组已经被鉴定和测序（图 1 和 2）。M 结构蛋白表达成包括 pre-M 序列(pr)多蛋白的一部分。该 pr 序列，紧临 M 蛋白序列的氨基末端，预防了多蛋白加工中构象的问题。尤其是，pr 序列的存在对于阻止 E 蛋白的错误折叠非常重要。因此，prM 的存在允许 JEV 颗粒的装配。一旦形成了病毒粒子或者颗粒，pr 序列就可以从 prM 蛋白剪切掉形成含有 M 蛋白的成熟病毒颗粒，虽然 prM 蛋白的剪切产生 M 蛋白并不是产生感染性颗粒所必须的。来自多种不同相关黄病毒的 prM 序列剪切到几乎很短的长度，但是黄病毒自身却有传染性。这些具有相似基因组结构和功能的相关黄病毒的例子包括但不限于 WNV，YFV，登革病毒和 SLEV。

在一个实施方案中，本发明中编码黄病毒 M 和 E 蛋白的 TU 为 DNA。根据上段中的讨论，该 DNA 含有编码 M 蛋白的核苷酸序列，包括 pre-M 序列，和编码 E 蛋白的核苷酸序列。以这种方式，目的基因产物能够在细胞内形成亚病毒颗粒。然后 pre-M 序列以类似于完整病毒粒子发生的方式被剪切。

为了在体内作为疫苗有效发挥功能，TU 中优选包括控制序列，可以增强或者促进编码抗原的核苷酸序列的转录。这些启动子的应用是分子生物学领域，细胞生物学领域和病毒免疫学领域的技术人员公知的（参见 *Molecular Cloning: A Laboratory Manual*, 第 2 版, Sambrook,

Fritsch 和 Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; 和 *Current Protocols in Molecular Biology*, Ausube 等, John Wiley 和 Sons, New York 1987 (最新季刊)。当 TU 作为疫苗用于哺乳动物宿主时, 采用的启动子优选为能够在哺乳动物细胞中有效操作的启动子。这种启动子根据转录受到其促进的编码序列进行排列, 排列位置可操作性地促进这种转录。在本发明的一个重要实施方案中, 启动子为巨细胞病毒早期启动子。此外, 在本发明进一步优选的实施方案中, 在 TU 核酸中, 编码序列后连接了终止子序列 (Sambrook 等)。本发明的特定实施方案涉及原核生物和真核生物细胞。很多启动子序列已知可用于原核或者真核细胞。(参见 Sambrook 等)。

本发明的核酸还包括本领域技术人员公知的 DNA 序列作为免疫刺激元件。这些元件的例子包括但不限于, 细菌 DNA 中的某些 CpG 基元(Sato 等, *Science* 273:352-354 (1996); Klinman 等, *Vaccine* 17:19-25 (1998))。

本发明的 TU 的制备可以通过分子生物学领域的技术人员公知的方法很容易地实现。涉及的步骤阐述在, 例如 Molecular Cloning :A Laboratory Manual, 第 2 版, Sambrook, Fritsch 和 Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989 和 Current Protocols in Molecular Biology, Ausubel 等, John Wiley 和 Sons, New York 1987 (最新季刊)。黄病毒 RNA 分子可以通过熟悉例如黄病毒, 和其它病毒组的病毒学家公知的方法分离自活病毒样品。JEV 使用的方法概述在 Kuno 等 (J. Virol. 72:73-83 (1998)) 中。RNA 作为模板利用逆转录酶合成 cDNA。从 cDNA, 含有从 pre-M 到 E 编码区(图 2)的片段通过已知用来合适地剪切 cDNA 提供这种片段的限制性核酸酶的消化而获得。JEV 限制性酶切的例子提供在 Nitayaphan 等(1990)和 Konishi 等(1991)中。整合启动子, 例如巨细胞病毒启动子, 促进有效翻译的序列, 例如 Kozak 序列, 和聚腺苷酸化信号, 也同样是分子生物学领域和重组 DNA 工程领域技术人员公知的(Kozak, Mol. Cell. Biology

9:5134-5142 (1989); Azevedo 等, *Braz. J. Med. Biol. Res.* 32:147-153 (1999)). 当含有包括目的编码序列和控制序列的核酸被制备后, 可以通过扩增核酸的方法大量获得产物。这些方法是分子生物学领域和重组 DNA 工程领域技术人员公知的。这些方法的例子包括将核酸整合到复制质粒中, 其步骤是将质粒在诸如原核细胞的细胞中培养, 培养完成后收集质粒, 并通过本领域公知的诸如 PCR 和其它扩增方法扩增核酸。这些例子并不是想要限制获得含有 TU 的核酸的可能方式。

本发明的含有 TU 的核酸分子可以通过在分子生物学领域和病毒免疫学领域公知的多种方式导入合适的细胞。导入方式的例子包括但不限于, 导入被细胞吸收的质粒或者类似的核酸载体, 或者包裹在诸如脂质体, 特别是含有阳离子脂类的脂质体的泡囊状脂类结构中, 或者吸附通过胞吞作用导入细胞中的颗粒。

通常, 本发明的细胞为含有 TU 或者其中已导入 TU 的原核或者真核细胞。本发明的 TU 诱导胞内生物合成编码的 prM/M 和 E 抗原。合适的细胞为导入核酸后能够生物合成基因产物的细胞。在本发明的一个具体实施方案中, 合适的细胞为能够响应控制序列和终止序列的细胞, 若有的话, 所述序列包含在 TU 内。为了以这种方式响应, 这种细胞包含与控制序列和终止序列相互作用并且用来执行各自的启动和终止功能的组分。当细胞体外培养时, 可以为原核生物, 单细胞真核生物细胞或者多细胞真核生物细胞。在本发明的一个具体实施方案中, 细胞为哺乳动物细胞。在这些例子中, 合成的 prM/M 和 E 蛋白基因产物适用于分析或诊断, 包括制备用作疫苗的抗原或者, 用于免疫诊断分析, 或者用于说明的目的。

在一些情况下, 诸如, 当细胞为培养的哺乳动物细胞时, prM/M 和 E 抗原以亚病毒颗粒的形式分泌。它们是 prM/M 和 E 蛋白的聚集物, 类似于活病毒的表面超微结构形态和免疫原性特征。因为本发明的 TU 不包括黄病毒基因组的其它部分, 所以没有整合的衣壳, 最重

要的是没有传染性的病毒 RNA。

在本发明的另一个重要实施方案中，细胞为已经施用了 TU 作为疫苗的受试者自然细胞组分。当 TU 施用给受试者后，被受试者的细胞吸收。如果存在任一启动子序列和终止子，受试者的细胞具有响应所述启动子序列和终止子的能力。在任一情况下，TU 诱导受试者的细胞合成黄病毒 prM/M 和 E 基因产物。不希望被理论所限制，相信受试者的细胞体内产生由 prM/M 和 E 抗原组成的亚病毒颗粒，如同体外培养的哺乳动物细胞中发生的一样。这些亚病毒颗粒，相信作为体内免疫原，刺激受试者的免疫系统产生免疫反应，对受试者赋予保护性免疫。依然不希望被理论所限制，产生的保护性免疫可以通过体液或者细胞免疫产生，即分别通过 MHC II 型或 I 型限制的机制，或者同时利用两个机制产生。

根据本发明，受试者通过施用有效剂量的含编码 prM 和/或 E 抗原的核酸的 TU 免疫抗黄病毒，诸如 JEV，YFV，登革病毒，SLEV，WNV 或者其它黄病毒的感染。此核酸被导入受试者的细胞后引起黄病毒 prM 和/或 E 抗原的合成。

为了将 TU 施用给受试者，可以使 TU 掺入到含有可药用载体的组合物中。术语“可药用载体”不是指生物学的材料或者是其它不合适的材料，而是可将其连同免疫原性材料一起施用给受试者的材料（即，重组黄病毒蛋白抗原或其部分），同时所述材料不引起不需要的生物效果或者以有害的方式和疫苗中所含的其它任一组分相互作用。可药用载体或者其组分的例子包括水，生理盐水和普通生理缓冲液（其它例子，参见 Arnon, R. (编) *Synthetic Vaccines I*: pp. 83-92, CRC Press, Inc, Boca Raton, Florida, 1987）。

本领域的技术人员应该明白，所述疫苗接种剂量的临界值为诱导宿主保护性反应所需的免疫原的总量，所述宿主易发生由毒性或者野

生型黄病毒感染引起的传染性疾病。剂量的数目和体积可以变化，由操作者根据诸如年龄，体重，性别，种族，施用疫苗的类型，施用的方式，受试者的全面条件等，以及本领域技术人员意识到的其他重要因素的参数来确定。

TU 可以通过口服，肠胃外（例如，静脉内），肌内注射，腹膜内注射，经皮，体外，鼻内，局部等给受试者施用。通过插管可以直接到达呼吸系统（例如，肺部）的任一区域。所需 TU 的准确量在受试者之间各不相同，依赖于种族，年龄，体重和受试者的整体条件，使用疫苗的免疫原性，受试者免疫所抗的黄病毒株或者种，施用的方式等等。因此，不可能明确说明本发明的每个实施方案中的准确剂量。但是，通过本领域的普通技术人员仅仅利用这里教导给出的常规试验程序和本领域现成的程序就可确定合适的剂量。

如果肠胃外施用本发明的疫苗，通常其特征在于注射给药。注射剂可以制备成常规的形式，作为液体溶液或者悬浮液，在注射前适合溶于或者悬浮于液体中的固体，或者作为乳化剂。最近肠胃外给药的改进方法包括使用慢慢释放或者持续释放系统，这样就保持了恒定的剂量。参见例如，美国专利 No.3,610,795，此处引作参考。

对于固体组合物而言，传统非毒性固相载体包括，例如，药用级的甘露醇，乳糖，淀粉，硬脂酸镁，糖精钠，滑石，纤维素，葡萄糖，蔗糖，碳酸镁等等。药学上可施用的液体组合物的制备可以通过，例如将此处描述的活性化合物和药用佐剂溶解，分散等到赋形剂，例如水，盐水，含水葡萄糖，甘油，乙醇等中形成溶液或者悬浮液。如果需要，待给药的药用组合物还可以含有少量无毒的辅助物质，例如增湿剂或者乳化剂，pH 缓冲剂等，例如乙酸钠，失水山梨糖醇单月桂酸酯，三乙醇胺乙酸钠，油酸合三乙醇胺等等。制备这些剂型的实际方法是公知的，或者对于本领域技术人员是显而易见的，例如参见 Remington's Pharmaceutical Sciences (Martin, E. W. (编)，最新版 Mack

Publishing Co, Easton, PA)。

在一个实施方案中,本发明的 TU 可以通过利用电转移介导的体内基因递送的方法施用给受试者,其中在 TU 施用给受试者后立即将经皮的电脉冲施加给受试者,从而更加高效可重复的将核酸体内转移给受试者的组织(Mir 等, *Proc. Nat. Acad. Sci USA* 96:4262-4267 (1999))。

在本发明的方法中,描述了通过将本发明的疫苗施用给受试者从而免疫受试者的过程,免疫的功效可以根据本领域技术人员公知的临床方法进行监控,从而监测受试者的免疫状况。

本领域的技术人员很容易确定接种组合物的有效剂量,就是当施用给受试者后给受试者赋予保护性免疫的剂量。为了进行这种确定,技术人员可以估计在施用了疫苗的受试者的血液中诱导出黄病毒 prM/M-和 E-特异性抗体和/或黄病毒 prM/M-和 E-特异性胞毒 T 淋巴细胞的能力。通过用相应于用来免疫试验受试者的抗原性组合物的活黄病毒攻击,还可以确定赋予给试验受试者的保护性免疫的水平。这种攻击试验是本领域技术人员公知的。

通常,根据本发明,为了免疫受试者抗 WNV, JEV, YFV, 登革病毒, SLEV, 或者其它的黄病毒,并且意识到这种方法中所用的 TU 有不同的大小,可以使用的剂量范围从大约 0.1 μ g/kg 体重到大约 50 μ g/kg 体重。

令人意外的发现,本发明为 DNA 的 TU 在通过肌内注射或者通过电转移对受试者仅仅施用单次有效剂量的 TU 后,就赋予受试者保护性免疫的有效性水平为大约 100%。这个结果和利用传统疫苗实施的许多免疫方法(如上描述)形成对照,后者需要一次或者多次加强接种并且不能赋予有效性接近 100%的保护性免疫。

进一步令人意外的发现，保护性免疫可以从免疫接种的母体受试者传递给受试者的后代。利用本发明的 TU DNA 免疫接种母体后，大部分初生小鼠表现出抗病毒攻击的保护。不希望被理论所限制，已知可以对初生哺乳动物赋予被动免疫，因为母亲的乳液中存在特异于多种病原体的中和抗体。可能这些初生儿中发现的抗 JEV 保护性免疫是以这种方式传递给他们的。

在本发明的另一个实施方案中，TU 编码第一种黄病毒结构蛋白的信号序列和第二种黄病毒的免疫原性黄病毒抗原。因此，在一个实施方案中，例如第一种黄病毒结构蛋白的信号序列被第二种黄病毒结构蛋白的信号序列替换，其产生新多肽的正确折叠，在宿主中的正确加工，和/或加工蛋白的正确折叠。

在本发明的另一个实施方案中，TU 可以编码免疫原性黄病毒抗原，其中抗原含有一种或者多种黄病毒的序列。信号序列可以为改良的信号肽。信号序列的改良，或多种最适信号序列的选择可以通过应用实施例 18 和其引用的参考文献教导的原理和技术来完成，每个参考文献在此处引作参考，专门用于在分别涉及选择，鉴定和设计具有需要特征和功能的信号序列中的教导。通常，这些需要的特征和功能将包括较高的信号序列概率。

在本发明的另一个实施方案中，一个单独的组合物中包括了多种 TU 或者包括编码来自多种黄病毒的免疫原性黄病毒抗原的一种 TU。因此，在一个实施方案中，例如，TU 可以编码新生多肽或被加工成来自多种黄病毒的蛋白的多肽。优选地，加工的蛋白形成引发抗蛋白的免疫反应的亚病毒颗粒。亚病毒颗粒可以形成自源于同一黄病毒，多种黄病毒组合的序列的加工蛋白，或者嵌合黄病毒蛋白。含有多种 TU 或者一种编码来自多种黄病毒的免疫原性黄病毒抗原的 TU 的联合疫苗可被定制用于特定的地理区域，其方法是包含本地区流行或者可

能碰到的黄病毒的蛋白。例如，地处热带或者亚热带的亚洲的疫苗可以包括编码来自 DEN 的四种血清型，WN 和 JE 病毒疫苗蛋白的 TU。类似地，非洲和拉丁美洲的有用疫苗可以包括分别编码来自 DEN 的四种血清型，WN 和 YF 病毒和来自 DEN 的四种血清型，罗西奥（Rocio）和 YF 病毒的蛋白的 TU。

在另一个实施方案中，TU 编码第一种黄病毒的信号序列和包括来自多种黄病毒的氨基酸序列的免疫原性嵌合黄病毒抗原。信号序列可以为日本脑炎病毒的信号序列。嵌合黄病毒抗原可以包括日本脑炎病毒抗原的序列。在特定实施方案中，嵌合抗原为 E 蛋白。E 蛋白的 C 末端部分可以为来自日本脑炎病毒的 E 蛋白序列。C 末端部分可以为，例如，嵌合 E 蛋白的 5，10，15，20，25，30，40，50 或者 75%。在优选的实施方案中，TU 编码日本脑炎病毒结构蛋白的信号序列，登革病毒的 prM 蛋白以及含有日本脑炎病毒和登革病毒序列的嵌合 E 蛋白。嵌合蛋白可以为 E 蛋白，其中 C 末端部分含有日本脑炎病毒的序列。TU 的例子包括 SEQ ID NO:44 和 SEQ ID NO:46 显示的核酸序列，其指导诸如 SEQ ID NO:45 和 SEQ ID NO:47 所示的黄病毒抗原的合成。

本发明进一步提供了在可药用的载体中含有本发明多肽的免疫原性组合物用作蛋白疫苗。产自本发明的转录单位的抗原可以用来引发受试者的有效免疫反应。用于该目的的抗原可以含有黄病毒 prM 蛋白，黄病毒 M 蛋白，黄病毒 E 蛋白或者它们的任一组合，包括蛋白免疫原性片段。特别优选的实施方案采用此处描述的 NRA。更加优选的实施方案是含有一种黄病毒的信号序列和一种或多种不同黄病毒的结构蛋白的嵌合蛋白。在尤其优选的实施方案中，抗原的信号序列为日本脑炎病毒的信号序列。在其它优选的实施方案中，信号序列为改良的信号肽。信号序列的改良或者选择多种最适的信号序列，可以通过实施实施例 18 和其引用的参考文献所教导的原理和技术完成，每个参考文献在此处引作参考专门用于在选择，鉴定和设计具有需要特

征和功能的信号序列中的教导。通常，这些需要的特征和功能将包括较高的信号序列概率。

在其它的实施方案中，本发明的蛋白疫苗还包括合适的佐剂。此处所用的“佐剂”为免疫反应的增效剂或者增强剂。术语“合适”是指包括任一能够用来与疫苗免疫原（即，黄病毒 prM 蛋白，黄病毒 M 蛋白，黄病毒 E 蛋白或者它们的任一组合）组合以增强免疫反应，同时对于免疫接种的受试者不产生副反应的物质。特定佐剂的有效剂量可以通过优化佐剂对于接种受试者免疫反应的潜在效果很容易确定。在优选的实施方案中，本发明疫苗的佐剂的辅助作用是一个 2-阶段过程，首先利用 2% 的氢氧化铝溶液然后使用矿物油。在具体的实施方案中，合适的佐剂可以选自：矿物质，带有水乳液的植物油或者鱼油，不完全弗氏佐剂，E.coli J5，葡聚糖硫酸酯，氧化铁，藻酸钠，Bacto-佐剂，特定的合成聚合物，诸如 Carbopol (BF Goodrich Company, Cleveland, Ohio)，聚合氨基酸和氨基酸的共聚物，皂苷，角叉藻聚糖，REGRESSIN (Vetrepharm, Athens, GA)，AVRIDINE(N,N-双十八烷基-N',N'-双(2-羟乙基)丙二胺)，长链多分散的 β (1,4)连接的点缀有 O-乙酰化基团的甘露聚糖聚合物(例如，ACEMANNAN)，源自分支杆菌种非病原菌株去蛋白的高度纯化的细胞壁提取物(例如，EQUIMUNE，Vetrepharm Research Inc, Athens GA)，单油酸甘露醇酯，石蜡油和胞壁酰二肽。

另一方面，本发明提供了用免疫原性量的本发明蛋白疫苗免疫受试者引发受试者有效免疫反应的方法。可以通过口服，肠胃外，鼻内，气管内，肌内，乳房内，皮下，静脉内，和/或真皮内施用方法实施免疫。含有黄病毒 prM 蛋白，黄病毒 M 蛋白和/或黄病毒 E 蛋白的疫苗可以通过注射，吸入，摄取，或者灌输进行给药。可以施用单剂量和/或重复剂量的疫苗制备物，即“加强剂”可以通过定期间隔给药加强最初的免疫反应或者从上次给药后间隔较长的时间进行给药。免疫接种的时间间隔可以根据受试者的年龄和条件改变。

术语“免疫原性量”是指免疫原或者其部分的使用量，其足以诱导接种受试者的免疫反应并保护受试者免于由通过接触野生型或者强毒黄病毒感染所导致的疾病，或者具有降低黄病毒感染对于接种受试者影响的治疗或者商用的有益效果。

本发明进一步提供了由本发明的抗原免疫反应产生的抗体。本发明的抗体可包括多克隆和单克隆抗体，其可以为完整的免疫球蛋白分子，嵌合免疫球蛋白分子，“人源化抗体”或者 Fab 或 F(ab')₂ 片段。这种抗体和抗体片段可以通过本领域技术人员公知的技术生产，这些技术包括描述于 Harlow 和 Lane (Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989) 和 Kohler 等人 (Nature 256:495-97, 1975) 和美国专利 5,545,806, 5,569,825 和 5,625,126 的技术，此处引作参考。抗体可以为任一同种型的 IgG, IgA, IgD, IgE 和 IgM。

本发明也可包括含有相连的 V_H 和 V_L 结构域的单链抗体 (ScFv)，并且其保持了原来独特型抗体的构象和特异性结合活性。这种单链抗体是本领域公知的，并可以通过标准方法制备。(参见，例如 Alvarez 等, Hum. Gene Ther. 8:229-242 (1997))。

可以制备抗本发明的抗原的抗体，所述抗原通过编码一种或多种黄病毒的 prM, M 和/或 E 抗原的免疫原性氨基酸序列和另一种不同黄病毒（例如，JEV）的信号序列的核酸序列合成。使用这些嵌合构建体合成的免疫原性肽，采用本领域公知的鉴定氨基酸序列中的免疫原性区域的方法可以很容易地得以鉴定，并用于制备本发明的抗体。

抗原/抗体复合物能够形成的条件，以及检测抗原/抗体复合物的形成和被检蛋白定量的方法是本领域公知的。这些方法包括但不限于，Western 印迹，免疫沉淀，免疫荧光，免疫细胞化学，免疫组化，

荧光激活细胞分类术(FACS), 荧光原位杂交(FISH), 免疫磁性测定法, ELISA, ELISPOT(Coligan 等编, 1995. *Current Protocols in Immunology*. Wiley, New York), 凝集试验, 絮凝试验, 细胞筛选等, 正如本领域技术人员所公知的。

此处所用的术语“结合”是指很好鉴定的抗体和抗原的结合, 以及和抗原的其它非随机结合。此处所用的“特异性结合”描述了一种抗体或者其他配体除一种特异性抗原以外基本上不与其它任一抗原进行交叉反应, 所述一种特异性抗原在此情形下为本发明抗原。

本发明的抗体或者配体可以结合到基质上(例如, 珠子, 管子, 载玻片, 培养皿, 硝酸纤维板等)或者与可检测部分共轭结合, 或者即结合到基质上又共轭结合。预期用于本发明的可检测部分包括, 但不限于, 免疫荧光部分(例如, 荧光素, 罗丹明), 放射性部分(例如 ^{32}P , ^{125}I , ^{35}S), 酶部分(例如, 辣根过氧化物酶, 碱性磷酸酶), 胶体金部分和生物素部分。这些结合技术是本领域标准技术(例如, Harlow 和 Lane, *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989); Yang 等, *Nature* 382:319-324 (1996))。

本发明进一步提供了检测样品中黄病毒抗体的方法, 包括将样品和本发明的黄病毒抗原在抗原/抗体复合物可以形成的条件下接触; 并检测复合物的形成, 从而检测样品中的黄病毒抗体。

本发明进一步提供了检测样品中黄病毒抗原的方法, 包含将样品和本发明的抗体在抗原/抗体复合物可以形成的条件下接触; 并检测复合物的形成, 从而检测样品中的黄病毒抗原。

检测样品中黄病毒抗原的方法可以例如通过将来自受试者的流体样品或者组织样品和本发明的抗体结合并检测抗体和抗原的结合来进行。可以设想抗原会在完整的黄病毒粒子上, 可以是显示在表达抗原

的黄病毒感染细胞表面上的黄病毒编码的蛋白，或者可以为抗原片段。该方法的流体样品可以包括含有抗原或者含有抗原的细胞的任一生物流体，诸如，脑脊液，血液，胆汁，血浆，血清，唾液和尿液。其它可能的体液例子包括痰，粘液等等。

检测样品中黄病毒抗体的方法可以通过将来自受试者的流体样品或者组织样品和本发明的抗原结合并检测抗原和抗体的结合来进行。该方法的流体样品可以包括含有抗体的任一生物流体，诸如，脑脊液，血液，胆汁，血浆，血清，唾液和尿液。其它可能的体液例子包括痰，粘液等等。

酶免疫测定，诸如免疫荧光测定(IFA)，酶联免疫吸附测定(ELISA)和免疫印迹可以很容易适应根据本发明的方法来检测黄病毒抗体。有效用于检测抗体的 ELISA 方法可以为，例如如下方法：(1)将抗原结合到基质上；(2)将结合的抗原和含有抗体的流体或者组织样品接触；(3)将上述抗原抗体复合物和结合到可检测部分的二级抗体接触，可检测部分可以和结合的抗体反应（例如，辣根过氧化物酶或碱性磷酸酶）；(4)将上述反应物和酶的底物反应；(5)将上述反应物和染色剂接触；和(6)观察/测定颜色变化或者显影。

用于检测黄病毒抗体的其它免疫技术采用单克隆（MAbs）检测在竞争性抑制测验中与黄病毒抗原特异性反应的抗体。简而言之，将样品和结合到基质(例如 ELISA 96-孔板)的本发明抗原接触。多余的样品被完全清洗掉。接着，标记的（例如，酶联，荧光，放射性等）单克隆抗体与任一前面形成的抗原/抗体复合物接触并测定单克隆抗体结合的量。抑制单克隆抗体结合的量相对于对照（没有抗体）进行测定，同时检测并测定样品中的抗体。当基于黄病毒特定类型或病毒株的单克隆抗体特异性结合时，单克隆抗体抑制程度可以对检测特定黄病毒类型和毒株而言是非常特异的分析方法。MAbs 也可以通过常规方法，例如免疫荧光测定（IFA）直接用于检测细胞中的黄病毒抗原。

作为其它的例子，微量凝集试验也可用于检测样品中黄病毒抗体的存在。简而言之，乳胶珠，红细胞或者其它可凝集的颗粒用本发明的抗原包被，并与样品混合，这样样品中与抗原特异性反应的抗体与抗原交联，引起凝集。凝集的抗原/抗体复合物形成用肉眼可以观察或者通过分光光度计可以测定的沉降。在上述测验的修饰中，本发明的抗体能被结合到样品中可凝集的颗粒和抗原上，然后进行检测。

本发明进一步提供了诊断受试者黄病毒感染的方法，包括将来自受试者的样品和本发明的抗原在抗原/抗体复合物可以形成的条件下接触；并检测抗原/抗体复合物的形成，由此诊断受试者黄病毒的感染。

本发明进一步提供了诊断受试者黄病毒感染的方法，包括将来自受试者的样品和本发明的抗原在抗原/抗体复合物可以形成的条件下接触；并检测抗原/抗体复合物的形成，由此诊断受试者黄病毒的感染。

在此处教导的诊断方法中，本发明的抗原可以与基质结合并与诸如血液，血清，尿液或者唾液的流体样品接触。这些样品可以直接来自病人或者以部分纯化的形式获得。以这种方式，抗原的特异性抗体（一级抗体）将与结合的抗原特异性反应。因此，结合到或者标记有可检测部分的二级抗体可被加入以增强一级抗体的检测。通常，与抗原的不同表位特异性反应或者与配体或反应抗体非特异性反应的二级抗体或者其它配体，可根据其与一级抗体的多个位点反应的能力加以选择。因此，例如，二级抗体的若干分子可以与每种一级抗体反应，使一级抗体更易检测到。

可检测部分能够可视性检测到沉降或者颜色的变化，还可通过显微术可视性检测，或者通过光谱测定，放射测定等方法自动检测。可检测部分的例子包括荧光素和罗丹明（用于荧光显微术），辣根过氧化物酶（用于光或电显微术和生化检测），生物素-链霉亲和素（用于

光或者电显微术)和碱性磷酸酶(通过颜色变化用于生化检测)。

本发明的具体实施方案在下列的实施例中详细描述。这些实施例并不旨在限制说明书中公开的本发明的范围。

实施例

采用涉及制备和表达本发明的核酸 TU 分子的分子生物学和重组 DNA 技术的通用方法描述在, 例如, Current Protocols in Molecular Biology, Ausubel 等, John Wiley 和 Sons, New York 1987 (最新季刊), 和 Molecular Cloning: A Laboratory Manual 第 2 版, Sambrook, Fritsch 和 Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989。

实施例 1. 制备含有编码 JEV prM 和 E 抗原转录单位的重组质粒。
采用 QIAamp™ Viral RNA 试剂盒(Qiagen, Santa Clarita, CA)从接种生长在小鼠大脑的 150μL 的 JEV 病毒株 SA14 病毒中提取基因组 RNA。将吸附到硅胶膜上的 RNA 稀释在 80μL 不含核酸酶的水中, 作为扩增 JEV prM 和 E 基因编码序列的模板。根据 Nitayaphan 等人(*Virology* 177:541-552 (1990))的工作获得引物序列。含有基因组核苷酸序列 389-2478 区的单链 cDNA 片段通过逆转录酶-聚合酶链式反应(RT-PCR)进行扩增。限制性位点 KpnI 和 XbaI, 共有 Kozak 核糖体结合序列, 以及翻译起始位点通过扩增引物 14DV389 经改造连接到 cDNA 的 5' 末端(核苷酸序列, SEQ ID NO:1; 氨基酸序列, SEQ ID NO:2)。在 NotI 限制性位点之后的符合读框的翻译终止密码子通过扩增引物 c14DV2453(SEQ ID NO:3)插入到 cDNA 的 3'末端(图 2)。通过 Titan RT-PCR 试剂盒(Boehringer Mannheim, Indianapolis, IN)进行一管 RT-PCR。10μL 的病毒 RNA 和各 1μL 的 14DV389(50μM) 和 c14DV2453(50μM)和 18μL 无核酸酶的水混合, 并将混合物在 85℃ 加热 5 分钟, 然后冷却到 4℃。加入 75μl 的反应混合物[20μL 的 5×缓冲液, 2μL 的 dNTP 混合物(各 10mM), 5μL 的二硫苏糖醇(0.1mM),

0.5 μ L 的 RNasinTM (40U/ μ L,Boehringer Mannheim), 2 μ L 的聚合酶混合物, 和 45.5 μ L 的无核酸酶的水], 并进行如下的 RT-PCR: 1 个循环(50 $^{\circ}$ C 30 分钟, 94 $^{\circ}$ C 3 分钟, 50 $^{\circ}$ C 30 秒钟, 68 $^{\circ}$ C 2.5 分钟), 9 个循环(94 $^{\circ}$ C 30 秒钟, 50 $^{\circ}$ C 30 秒钟, 68 $^{\circ}$ C 2.5 分钟), 20 个循环(94 $^{\circ}$ C 30 秒钟, 50 $^{\circ}$ C 30 秒钟, 在第一个循环中 68 $^{\circ}$ C 2.5 分钟, 随后的每个循环增加 5 秒钟), 最后在 68 $^{\circ}$ C 延伸 15 分钟。RT-PCR 产物用 QIAquickTM PCR 纯化试剂盒(Qiagen)进行纯化, 并用 50 μ L 的 1mM pH7.5 Tris-HCl 稀释。

所有载体的构建和分析通过标准技术进行(Sambrook 等, 1989)。RT-PCR 扩增的 cDNA 用 KpnI 和 NotI 核酸酶进行酶切, 然后插入真核表达质粒载体(pCDNA3, Invitrogen, Carlsbad, CA)的 KpnI-NotI 位点。电穿孔-感受态大肠杆菌 XL1-Blue 细胞(Stratagene, La Jolla, CA)通过电穿孔进行转化(Gene Pulser, Bio-Rad, Hercules, CA), 然后涂布到含 100 μ g/mL 羧苄青霉素(Sigma Chemical Co, St. Louis, MO)的 LB 琼脂培养皿中。挑选克隆并接种到含有 100 μ g/mL 羧苄青霉素的 3 mL LB 肉汤中。用 QIAprepTM 旋转微量制备试剂盒(Qiagen)从培养了 14 个小时的培养物中提取质粒 DNA。按推荐的方法进行 DNA 自动测序(Applied Biosystems/Perkin Elmer, Foster City, CA)。cDNA 的两条链均被测序并显示与原始 SA14 毒株的序列(Nitayaphan 等, 1990)相同。

含有 fl 起始点 (ori), SV40 起始点, 新霉素抗性基因, 和 SV40 的 poly(A)元件的质粒 pCDNA3(Invitrogen, Carlsbad, CA), 将其核苷酸(nt)1289 到 nt3455 的片段用 PvuII 酶切缺失, 然后连接产生 pCBamp 质粒。包含嵌合内含子的载体 pCIBamp, 该嵌合内含子插入到 pCBamp 的 NcoI/KpnI 位点, 通过用 NcoI 和 KpnI 酶切将内含子序列从 pCI (Promega, Madison, WI)剪切下来构建而成。产生的 566-bp 的片段通过用 NcoI-KpnI 酶切克隆到 pCBamp 中, 替换其 289-bp 的片段。图 3 说明了质粒 pCDA3, pCBamp 和 pCIBamp 之间的关系。

含有编码 JEV prM 和 E 蛋白转录单位的质粒从这些质粒进行制备。在重组质粒 pCDJE2-7(核苷酸序列 SEQ ID NO:10; 氨基酸序列, SEQ ID NO:11)中含有 JEV prM 和 E 编码区的 cDNA 片段衍生自 pCDNA3 载体, 其通过 NotI 和 KpnI 或 XbaI 进行酶切并克隆到 pCBamp, pCIBamp, pCEP4(Invitrogen, Carlsbad, CA) 或者 pREP4(Invitrogen, Carlsbad, CA)的 KpnI-NotI 位点, 或者克隆到 pRc/RSV (Invitrogen, Carlsbad, CA)表达载体的 SpeI-NotI 位点以分别构建 pCBJE1-14(核苷酸序列, SEQ ID NO:17; 氨基酸序列 SEQ ID NO:18), pCIBJES14, pCEJE, pREFE, 和 pRCJE。来自每种质粒克隆的双链 cDNA 被测序, 并鉴定含有正确核苷酸序列的重组克隆。在体外转化哺乳动物细胞或者在小鼠免疫试验中使用的质粒 DNA 用 EndoFree™ Plasmid Maxi 试剂盒(Qiagen)通过阴离子交换层析法进行纯化。

实施例 2.采用间接免疫荧光抗体测定法评价由多种重组质粒表达的 JEV prM 和 E 蛋白。在瞬时转化细胞系 COS-1, COS-7 和 SVT2(ATCC, Rockville MD; 分别为 1650-CRL, 1651-CRL, 和 163.1-CCL)中由多种重组表达质粒表达的 JEV 特异性基因产物采用间接免疫荧光抗体测定法 (IFA) 进行评价。SV-T2 细胞系在后面的试验中被排除, 因为预实验结果显示只有 1-2%的转化 SV-T2 细胞为 JEV 抗原阳性。为了转化, 细胞在 150cm² 的培养瓶中生长到 75%汇合时, 用胰蛋白酶处理, 在 4℃重悬浮在磷酸盐缓冲液(PBS)中, 最终细胞数为 5×10⁶/mL。10μg 的质粒 DNA 通过电穿孔利用设定在 150V, 960μF 和 100 Ω 电阻的 BioRad Gene Pulse™ (Bio-Rad)导入 300μL 的细胞悬浮液中。电穿孔 5 分钟后, 用 25 mL 的新鲜培养基稀释细胞并接种到 75cm² 的培养瓶中。转化 48 小时后, 将培养基从细胞移走, 细胞用胰蛋白酶处理, 并重悬浮在含 3%标准山羊血清的 5mL PBS 中。10μL 等分试样点到玻片上, 空气干燥, 并用丙酮在-20℃固定 20 分钟。使用荧光素异硫氰酸盐共轭的山羊抗小鼠免疫球蛋白 G(Sigma Chemical Co.)和 JEV HIAF 对丙酮固定的质粒转化细胞进行 IFA 测定。

为了确定各种启动子和 poly(A)元件对于 JEV prM 和 E 蛋白表达的影响,用等量的 pCDJE2-7 (SEQ ID NO:10), pCEJE, pREJE, 或 pRCJE 质粒 DNA 瞬间转化 COS-1 和 COS-7 细胞系。JEV 抗原在由所有四种重组质粒转化的两种细胞系中均有表达,从而证明了 CMV 或者 RSV(劳氏肉瘤病毒)启动子和 BGH 或者 SV40 poly (A)元件具有功能活性。但是,分别由 IFA 阳性细胞数目和 IFA 强度确定的转化细胞的百分比以及 JEV 抗原的表达水平在不同质粒之间差别很大(表 1)。由 pCDJE2-7(SEQ ID NO:10), pCBJEI-14(SEQ ID NO:17)和 pCIBJES14 转化的 COS-1 细胞中有百分比显著高的细胞表达 JEV 抗原,而且蛋白的表达水平和 JEV 感染的细胞的表达水平一致。另一方面,在用 pCEJE, pREJE, 或 pRCJE 载体转染的细胞中抗原表达细胞的百分比低,而且荧光的强度也低,显示抗原表达弱。

为了确定 pCDJE2-7 (SEQ ID NO:10)表达 JEV 蛋白的增强是否受到 SV40-编码的真核复制起点的影响,构建质粒 pCBJEI-14 (SEQ ID NO:17)使得含有 fl 起始点, SV40 起始点, 新霉素抗性基因, 和 SV40poly(A)的元件的 2166bp 的片段从 pCDJE2-7 上被去除。然后将嵌合内含子插入 pCBJEI-14 产生 pCIBJES14。pCIBJES14 质粒被用来确定 JEV 蛋白的表达是否可以被内含子的序列增强。转化后,含有 pCBJEI-14 和 pCIBJES14 两载体的细胞表达的 JEV 抗原的水平与 pCDJE2-7 中观察到的相类似(表 1)。这个结果显示 JEV prM 和 E 抗原由重组载体的表达仅仅受到转录调控元件的影响。真核复制起点和内含子序列都不增强 JEV 抗原在所用细胞中的表达。含有 CMV 启动子和 BGH poly(A)的载体(图 3)被选择用于进一步分析。

实施例 3.选择体外转化的组成型表达 JEV 特异性基因产物的稳定细胞系。用 10 μ g 的 pCDJE2-7 DNA 通过前面例子中描述的电穿孔转化 COS-1 细胞。在非选择性培养基中培养 24 小时后,用新霉素(0.5 mg/mL, Sigma Chemical Co.)处理细胞。在 2 到 3 星期后可见的新霉

素抗性克隆通过限度稀释在含新霉素的培养基中进行克隆。利用 JEV HIAF 通过 IFA, 载体编码的 JEV 基因产物的表达被首先筛选出来。选择一个 JEV-IFA 阳性克隆(JE-4B)和一个阴性克隆(JE-5A)用于进一步的分析, 并保存在含 200 μ g/mL 新霉素的培养基中。

由 JE-4B 克隆表达的 JEV E 蛋白的真实性通过采用一组 JEV E-特异性鼠源单克隆抗体(Mab)的 IFA 的表位图谱加以说明(Kimura-Kuroda 等, *J. Virol.* 45:124-132 (1983); Kimura-Kuroda 等, *J. Gen. Virol.* 67:2663-2672 (1986); Zhang 等, *J. Med. Virol.* 29:133-138 (1989); 和 Roehrig 等, *Virol.* 128:118-126 (1983))。JEV HIAF 和标准小鼠血清分别作为阳性或者阴性对照。四种 JEV-特异性 Mabs, 六种黄病毒-亚型特异性 Mabs, 和两种黄病毒-组反应的 Mabs 同样和 4B 克隆或者 JEV-感染的 COS-1 细胞反应(表 2)。

实施例 4.由 JE-4B COS-1 细胞系分泌的亚病毒颗粒的抗原特征和免疫检测。

a. 制备亚病毒颗粒。 JE-4B COS-1 细胞生长并维持在含 200 μ g/mL 新霉素的培养基中。培养基按常规收集并保存在 4 $^{\circ}$ C, 每星期两次补充培养基, 每 7-10 天细胞分裂成 1:5。用 Sorvall F16/250 转子在 10,000 rpm 离心 30 分钟, 然后通过 5%的蔗糖垫(w/w,用 10mM 的 Tris HCl, pH 7.5, 100mM 的 NaCl(TN 缓冲液)进行制备)用 Sorvall TH641 转子在 4 $^{\circ}$ C, 39,000rpm 离心 4 小时对培养基进行澄清。含有亚病毒颗粒的沉淀重悬浮在 TN 缓冲液中并储存在 4 $^{\circ}$ C。或者, 加入 7% 或 10% 的 PEG-8000 (w/v)到澄清的培养基中。混合物在 4 $^{\circ}$ C 至少搅拌 2 小时, 沉降的颗粒在 10,000rpm 离心 30 分钟进行收集。沉降物重悬浮在 TN 缓冲液中并保存在 4 $^{\circ}$ C。通过 TN 中 5-25%的连续蔗糖梯度以 38,000 rpm, 4 $^{\circ}$ C 下差速区带离心 90 分钟, 从沉淀部分和 PEG-沉降的制备物中纯化亚病毒颗粒。从梯度上层收集 1-mL 级分, 通过抗原捕捉 ELISA(参见下面)进行试验, 阳性部分添加到 TN 中的 25-50%蔗糖梯度上。在 35,000 rpm, 4 $^{\circ}$ C 平衡密度梯度离心过夜。从底部收集 0.9-

mL 平衡梯度的级分。通过抗原捕捉 ELISA 进行试验，并在 pH6.6 对红血球凝集(HA)活性进行评价。每级分称取 100 μ L 等分试样以确定其密度。合并 ELISA-阳性级分，并在 39,000 rpm，4 $^{\circ}$ C 离心 3-4 小时进行沉淀，将沉淀重悬浮在 TN 缓冲液中。在沉淀样品上测定抗原-捕捉 ELISA 和 HA 滴度。JEV-感染的 COS-1 细胞上清液也进行如上文详细描述的程序并作为梯度分析的阳性对照。也通过甘油/酒石酸盐平衡梯度中沉降从感染后 5-6 天的感染 C6/36 细胞中纯化 JE 病毒粒子。

b. 亚病毒颗粒的 Western 印迹。亚病毒颗粒梯度纯化的样品和电泳样品缓冲液混合，然后如 Laemmli (*Nature* 277:680-685 (1970))描述的方法在含 10%或 12.5%十二烷基硫酸钠的聚丙烯酰胺凝胶(SDS-PAGE)上电泳。将蛋白转移到硝酸纤维素膜上，然后用多克隆 JEV HIAF，黄病毒交叉反应抗-E Mab 4G2(Henchal 等, *Amer. J. Trop. Med. Hyg.* 31:830-836 (1982))，或者小鼠抗-prM 肽超免疫血清(JM01)进行免疫化学检测。图 4 显示由 JEV 感染的 C6/36 和 JE-4B COS-1 细胞产生的 M 和 E 蛋白的比较。在标准小鼠腹水和 Jmol 抗肽血清中观察到了一些针对 E 蛋白的非特异性反应。在大小上与 M 和 E 相同的蛋白从亚病毒颗粒中分泌出来，并分别通过 E-特异性 Mab 4G2 和 prM-特异性 JM01 抗血清可以进行检测。

c. 在培养基中 JEV 亚病毒颗粒的密度梯度检测。为了进行 ELISA，抗原捕捉抗体(4G2)稀释在 0.1M，pH9.6 的碳酸钠缓冲液中，并通过在 4 $^{\circ}$ C 的过夜温育用来包被 96-孔微量滴定板(Immulon II, Dynatech, Chantilly, VA)。用 3%标准山羊血清的 PBS 封闭反应后，将两倍系列稀释的样品加入 4G2-包被的微滴板，然后在 37 $^{\circ}$ C 温育 1.5 小时。通过辣根过氧化物酶共轭的 6B6C-1 Mag，在 37 $^{\circ}$ C 温育 1 小时检测捕捉抗原。通过 TMB (3,3',5,5'-四甲基联苯胺)-ELISA (Life Technologies, Grand Island, NY)检测固相上的酶活。

细胞接种四天后从 15 $\times$ 150cm² 的培养瓶中收集大约 500mL JE-4B 细胞的细胞培养基。PEG-沉降的亚病毒颗粒重悬浮在 2mL 的 pH 7.5 的 TN 缓冲液中；将 0.7mL 等分试样的重悬浮沉淀加入 5-25%蔗糖梯

度上。将能够破裂亚病毒颗粒的 Triton X-100 加入另外 0.7mL 等分试样, 终浓度达到 0.1%, 然后加到由含有 0.1% Triton X-100 的 TN 缓冲液制备的 5-25%蔗糖梯度上。从含 Triton X-100 的梯度上部观察到一条大约 2.5cm 的明确的混浊带, 但在不含表面活性剂的梯度中未观察到。从上层到下层对每层梯度收集级分(1mL)(图 5)。通过抗原捕捉 ELISA 分析各个收集的级分。在级分 4-6 检测抗原, 显示亚病毒颗粒相对快速的沉降特征。用 Triton X-100 处理 JE-4B 培养基的 PEG 沉淀物将 ELISA-反应材料转移到了梯度的上层。因此, 用 Triton X-100 处理仅仅产生了慢速沉降分子。相似的发现由 Konishi 等(*Virol.* 188:714-720 (1992))报道。这些结果显示含有 prM/M 和 E 的亚病毒颗粒的快速沉淀可以被表面活性剂的处理破坏。

在 pH 6.1 到 7.0 的范围内通过 Clarke 和 Casals(*Amer. J. Trop. Med. Hyg.* 7:561-573 (1958))的方法确定了红血球凝集(HA)活性。由 JE-4B 细胞分泌的亚病毒颗粒和由 JEV 感染的 COS-1 细胞产生的病毒粒子颗粒在确定的最适 pH6.6 时, 有相似的 HA 图谱。

实施例 5.用本发明的 pCDJE2-7 核酸疫苗接种的小鼠的免疫反应和用市售 JEV 疫苗接种的小鼠的免疫反应之间的比较。每组 5 只 3 周大的雌性 ICR 远交小鼠在左侧和右侧四头肌肌肉内注射 100 μ g pCDJE2-7 质粒的 100 μ L 蒸馏水, 或者皮下施用 JE-VAX (由 the Research Foundation for Microbial Disease of Osaka University 生产, 并由 Connaught Laboratories, Swiftwater, PA.销售), 剂量为人剂量的 1/5。编码和表达不相关蛋白的质粒 pCDNA3/CAT (Invitrogen) 作为阴性接种对照。除了一组 pCDJE2-7 接种的小鼠, 所有的动物在 3 周后又施用额外剂量的质粒或者 JE-VAX 而加强接种。小鼠在接种 3, 6, 9, 23, 40 和 60 周后自后眼窝处取血。通过酶联免疫吸附试验(ELISA), 或者通过噬斑减少中和实验(PRNT)(Roehrig 等, *Virol.* 171:49-60 (1989); 以及 Hunt 和 Calisher, *Amer. J. Trop. Med. Hyg.* 28:740-749 (1979))测定抗纯化 JEV 的 JEV 抗体滴度。

pCDJE2-7 核酸疫苗和 JE-VAX，第一次接种 3 星期内，所有三组小鼠都 100%地提供了血清转化(表 3)。JEV ELISA 和 PRNT 抗体滴度分别在免疫 6 周和 8 周后达到了最高水平。接受单剂量 DNA 疫苗的小鼠和接受了双剂量疫苗的小鼠有相似的抗体反应。可比较的 ELISA 抗体滴度在 DNA-接种的小组中，实验结束后保持长达 60 周。但是，JE-VAX 小组中仅仅四分之一在接种 60 周后为 JEV 抗体阳性。pCDNA3/CAT 对照组不含有任一可检测的 JEV 抗体。这些结果显示，单剂量的 JEV-特异性核酸疫苗在小鼠体内比市售 FDA-认证的 JE-VAX 疫苗更有效地保持 JEV 抗体。

实施例 6.在不同年龄接种本发明的多种核酸疫苗构建体和市售 JEV 疫苗有效性的比较。通过 pCDJE2-7, pCBE1-14, 或 pCIBJES14 转化的 COS-1 细胞表达相似水平的 JEV 蛋白。由这些核酸构建体诱导的 JEV 抗体在两个不同接种年龄和 JE-VAX 市售疫苗进行比较。3-天(混合性别)或者 3-周-龄(雌性)ICR 远交小鼠，每组 10 只，用 50 或 100 μ g 的质粒 DNA 肌肉接种，或者用十分之一或者五分之一剂量的 JE-VAX 皮下接种。免疫后 3 和 7 周收集血清样品，1:1600 稀释后通过 ELISA 采用纯化 JEV 作为抗原进行测定。结果显示在表 4。

质粒 pCBE1-14 在两个接种年龄段均表现出最大程度的血清转化，即，抗体滴度大于 1:1600，达到 80-100%。3 天大接种 pCDJE2-7 或 pCIBJES14 的小鼠至 7 周时表现适中的血清转化(每种 60%)，但是接种 3 周后测定时，表现较弱的血清转化(分别为 40%和 10%)。但是，3 周龄接种这些质粒时，接种后 3 周和 7 周血清转化达到 90%或 100%。相比而言，市售疫苗，JE-VAX，3 天大给药后没有血清转化，3 周龄给药达到 100%。因此，JEV prM 和 E 的核酸 TU 比非常高剂量的市售疫苗提供了更好的血清转化的程度，并在幼年动物和成熟动物体内都提供了出乎意料的高水平血清转化。

实施例 7. 由本发明的核酸疫苗赋予的保护性免疫。实施例 6 中的三天大的接种组接种 7 周后用腹膜内注射 50,000pfu/100 μ L 的小鼠适应性 JEV 毒株 SA14 进行攻击, 观察 3 周。在接受了多种含核酸 TU 的疫苗构建体的小组中获得了长达 21 天的 100%保护 (表 5)。相比之下, 60%的 JE-VAX-接种小鼠, 和 70%的 pCDNA3/CAT-接种的阴性对照, 至 21 天时无一幸免于病毒攻击。这些结果显示, 本发明的核酸 TU 对于接种小鼠提供了出人意料的有效保护。这暗示可以采用本发明的核酸疫苗作为人类儿童疫苗的可能性。相比之下, 目前使用的人用 JE-VAX 灭活疫苗对于幼年动物无效。

实施例 8.与母体抗体滴度相关的初生小鼠的被动保护。3 周龄的雌性 ICR 小鼠用单剂量或者剂量间之间间隔两天的双剂量的 100 μ g/100 μ L pCDJE2-7 质粒 DNA, 或者用五分之一人剂量的双剂量 JE-VAX 接种。阴性对照组接受双剂量的 100 μ g/100 μ L pCDNA-3/CAT 质粒。来自母体抗体的被动保护通过幼崽来评价, 幼畜产自试验用雌性小鼠和非免疫的雄性小鼠之间的交配, 交配发生在第一次接种后 9 周, 或者第二次接种后 6 周。幼崽出生后 3-15 天通过腹膜内接种 5,000 pfu/100 μ L 的小鼠适应的 SA14 病毒进行攻击试验, 每天观察共 3 周(表 6)。存活率相关于母体中和抗体的滴度。带有 1:80 PRNT 的母亲哺乳的幼崽 100%幸免于病毒感染, 但是对照母亲哺乳的幼崽没有存活的 (表 6)。带有 1:20 和 1:40 PRNT 滴度的母亲哺乳的较大幼崽中分别观察到 45%和 75%的部分保护。存活率也相关于由免疫母亲哺育的时间长度。如上所述, 13-15 天大的幼崽有更高的存活率。母亲带有 1:20 和 1:40 PRNT 滴度时 3-4 天大的幼崽无一幸免于病毒攻击。因此, 母体抗体给后代提供了部分到完全的保护性免疫。此外, 在 97% (29/30) 的攻击后幼崽的血清中, 通过 ELISA 检测到了 JEV 抗体。

用 1 或 2、100 μ g 剂量的质粒 DNA 肌内接种小鼠, 或者用两次 1/5 人剂量的 JE-VAX 疫苗接种小鼠。接种后 9 周收集血清用于与非免疫的雄性小鼠交配前的 PRNT 试验。

实施例 9.制备含编码 WNV prM 和 E 抗原转录单位的重组质粒。

从用 NY99-6480 病毒株感染的 150 μ L Vero 细胞培养基中, 利用 QIAamp™ Viral RNA 试剂盒(Qiagen, Santa Clarita, CA)提取基因组 RNA, 所述的 NY99-6480 病毒株分离自 1999 年纽约爆发的毒株。提取的 RNA 经过洗脱悬浮在 80 μ l 不含核酸酶的水中, 并作为 WNV prM 和 E 基因编码序列的扩增模板。引物序列获自 Lanciotti 等人(*Science* 286:2333-2337 (1999))的工作。含有基因组核苷酸区的 cDNA 片段通过逆转录酶-聚合酶链式反应(RT-PCR)进行扩增。利用扩增引物 WN466(核苷酸序列 SEQ ID NO:12)将限制性位点 BsmBI 和 KasI 改建到 cDNA 的 5'端。NotI 限制性位点后面的符合读框翻译终止密码子利用扩增引物 cWN2444(SEQ ID NO:13)导入在 cDNA 的 3'末端。利用 QIAquick™ PCR 纯化试剂盒(Qiagen)纯化 RT-PCR 产物。

上述两个扩增引物(SEQ ID NO:12 和 SEQ ID NO:13)产生的双链扩增子用 KasI 和 NotI 酶切产生 998 bp (nt-1470 到 2468)的 DNA 片段, 并插入 pCBJESS 载体的 KasI 和 NotI 位点形成中间质粒, pCBINT。PCBJESS 衍生自 pCBamp 质粒, 其含有巨细胞病毒早期基因启动子和翻译控制序列元件以及改造的 JE 信号序列元件(Chang 等, *J. Virol.* 74:4244-4252 (2000))。JE 信号序列元件含有 JE 信号序列(SEQ ID NO:14)。

随后 cDNA 扩增子用 BsmBI 和 KasI 酶消化, 剩下的 1003bp 片段(nt-466 到 1470)插入 pCBINT 的 KasI 位点形成 pCBWN (核酸序列, SEQ ID NO:15; 氨基酸序列 SEQ ID NO:16)。利用 ABI prism377 测序仪(Applied Biosystems/Perkin Elmer, Foster City, CA)自动测序 DNA 序列, 以证实重组质粒具有如 Lanciotti 等(*Science* 286:2333-2337 (1999))限定的正确的 prM 和 E 序列。

用于哺乳动物细胞体外转化和小鼠免疫试验的质粒 DNA 通过实

施例 1 描述的阴离子交换层析法进行纯化。

实施例 10.免疫化学鉴定和评价由 pCBWN 表达的 WNV prM 和 E 蛋白。由 pCBWN 质粒编码的 WNV 特异性基因产物表达在 COS-1 细胞中。根据 Chang 等人(*J. Virol.* 74:4244-4252 (2000))的方法对细胞进行电穿孔,并用 pCBWN 转化。电穿孔的细胞接种在 75 cm² 的培养瓶或者含有一个无菌盖玻片/孔的 12-孔组织培养皿中。所有培养瓶和 12-孔板培养置于 37℃, 5% CO₂ 的培养箱内。电穿孔 40 小时后,含有吸附细胞的盖玻片从小孔中移走,用 PBS 简单清洗,用丙酮在室温下固定 2 分钟,然后空气干燥。

采用实施例 2 所述的间接免疫荧光抗体测定法 (IFA) 检测蛋白表达。利用黄病毒 E-蛋白特异性单克隆抗体(Mab)4G2, WNV 小鼠超免疫腹水(HIAF)和以 1:200 稀释于 PBS 的标准小鼠血清(NMS)作为一级抗体检测蛋白表达(Henchal 等, *Am. J. Trop. Med. Hyg.* 31:830-836 (1982))。

电穿孔后 40 和 80 小时收集组织培养基。利用抗原捕捉(Ag-捕捉)ELISA 检测瞬时转化的 COS-1 细胞培养基中分泌的 WN 病毒抗原。Mab 4G2 和辣根过氧化物酶-共轭的 Mab 6B6C-1 分别用来捕捉 WN 病毒抗原并检测捕捉的抗原(Chang 等, *J. Virol.* 74:4244-4452 (2000); Henchal 等, *Am. J. Trop. Med. Hyg.* 31:830-836 (1983); Roehrig 等, *Virology* 128:118-126 (1983))。

培养基中的 WN 病毒抗原用 10%的聚乙二醇(PEG)-8000 沉淀进行浓缩。沉淀物重悬浮在 TNE 缓冲液 (50 mM Tris, 100 mM NaCl, 10 mM EDTA, pH 7.5)中,通过离心澄清,并保存在 4℃。或者,将沉淀物重悬浮在冷冻的缓冲液(0.1 M TRIZMA 和 0.4%牛血清白蛋白的硼酸盐缓冲液中, pH9.0)中,冻干并保存在 4℃。冻干的制备物用作评价 MAC-和间接 IgG ELISA 的抗原。

通过 IFA 检测瞬时转化的 COS-1 细胞中的 WN 病毒-特异性蛋白。在这些细胞中表达的 E, prM 和 M 蛋白分泌到培养基中。通过 PEG 沉淀浓缩的 WN 病毒抗原用 70% 的乙醇提取, 去除残余的 PEG(Aizawa 等, *Appl. Environ. Micro.* 39:54-57 (1980))。乙醇提取的抗原和梯度纯化的 WN 病毒粒子用 Excel Plus 电泳装置(Invitrogen Corp, Carlsbad, CA) 在 NuPAGE, 4-12% 梯度 Bis-Tris 凝胶上进行分析, 然后利用 Excel Plus Blot Unit (Invitrogen Corp.) 电印迹到硝酸纤维膜上。由瞬时转化的 COS-1 细胞产生的 WN 病毒-特异性蛋白通过 WN 病毒特异性小鼠 HIAF 或者黄病毒 E 蛋白反应性 Mab 4G2 在 Western 印迹分析中进行检测, 利用 NMS 作为阴性血清对照。蛋白表现出和来自 WN 病毒感染的乳鼠大脑 (SMB) 对应的梯度纯化的病毒粒子 E, prM 和 M 蛋白相似的反应性和相同的分子量。

在 NRA 作为抗原用于诊断 ELISA 的分析中, 一小瓶冻干 NRA 代表从 40 ml 的组织培养液中收集的抗原, 重新溶于 1.0ml 的蒸馏水中, 并且和重新溶解的 WN 病毒感染的乳鼠大脑 (SMB) 抗原比较, WN 病毒感染的乳鼠大脑 (SMB) 抗原为冻干的 β -丙醇酸内酯灭活的蔗糖-丙酮提取物(Clarke 等, *Am. J. Trop. Med. Hyg.* 7:561-573 (1958))。所有重组蛋白, prM, M 和 E 具有和梯度纯化的病毒粒子 E, prM 和 M 蛋白相似的反应性。

在试验阶段, 相同的检测中同时用抗原测试编码的人类的样品。所用的 MAC-和 IgG ELISA 方法和公开的方法(Johnson 等, *J. Clin. Microbiol.* 38:1827-1831 (2000); Martin 等, *J. Clin. Microbiol.* 38:1823-1826. (2000))相同。人血清样品获自我们实验室的血清库, 其由 1999 年爆发时送到 DVVID 用于 WN 病毒确认的样品组成。在这些试验中, 筛选 MAC-和 IgG ELISA 以 1:400 的样品稀释比进行。产生阳性/阴性 (P/N) OD 比在 2 和 3 之间的样品被认为是可疑阳性。可疑血清样品通过 ELISA 终点滴定和噬斑减少中和实验 (PRNT) 证明

均为阳性。所有其 (P/N) OD 比大于 3.0 的样品均认为是阳性，不需要其它的确证试验。

采用黄病毒-组反应性，抗-E Mab, 4G2 和 6B6C-1 的 Ag-捕捉 ELISA 用于检测分泌到 pCBWN 转化的 COS-1 细胞培养液中的 NRA。转化一天后可以在培养基中检测到抗原；而在第 2 到第 4 天之间可以观察到无需进一步浓缩的培养液中最大的 ELISA 滴度(1:32-1:64)。通过 PEG 沉淀浓缩 NRA，重悬浮在冷干的缓冲液中，冻干保存。为了进行诊断实验，一瓶冻干的 NRA 重新溶解在 1.0 ml 的蒸馏水中，然后利用 WN 病毒阳性和阴性对照人血清在 MAC-或者间接 IgG ELISA 中进行滴定 (Johnson 等, *J. Clin. Microbiol.* 38:1827-1831 (2000); Martin 等, *J. Clin. Microbiol.* 38:1823-1826 (2000))。发现 NRA 的 1:320 和 1:160 稀释比分别是用于 MAC-和 IgG ELISA 的最佳浓度。这些稀释比对于 MAC-和 IgG 试验分别产生 4.19 和 4.54 的 P/N OD₄₅₀ 比。由 NY-6480 和 Eg101 毒株产生的 WN 病毒 SMB 抗原在 MAC-ELISA 试验和 IgG ELISA 试验中分别采用 1:320 和 1:640 以及 1:120 和 1:320 的稀释比。阴性对照抗原，标准 COS-1 细胞的培养基的 PEG 沉降物和标准 SMB 抗原使用和 NRA 和 SMB 各自抗原相同的稀释比。稀释比为 1:400 的人血清样品用病毒特异性和阴性对照抗原一式三份同时进行试验。为了阳性实验结果的有效性，和病毒抗原 (P) 反应的试验血清的 OD₄₅₀ 必须至少大于相应的和阴性对照抗原 (N) 反应的相同血清的光密度值两倍。

NRA 和 NY-06480 的反应性，Eg101 和 SLE 病毒 SMB 的反应性通过 MAC-和 IgG ELISAs，利用 21 个编码的人血清样品进行比较。21 个样品中，有 19 个在全部三种抗原上产生相同的结果 (8 个阴性，11 个可疑阳性或者阳性)。用 SLE SMB 抗原分别对 18 个样品进行试验。13 个 Eg-101-SMB 阳性样品中只有三个在 SLE MAC-ELISA 中呈阳性 (表 1)。通过 SLE MAC-ELISA 没有一个 WN 抗原阴性样品呈阳性。该结果证实了前面的观察，即，抗-WN 病毒 IgM 不与其它黄病毒有显著的交叉反应 (Tardei 等, *J. Clin. Microbiol.* 38:2232-2239 (1940))而

且特异性诊断急性 WN 病毒感染，不管 NRA 或者 SMB 抗原是否用于试验中。所有的样品也通过间接 IgG ELISA 同时试验。使用 3 种抗原中任一种，21 个样品中有 10 个为阳性。

来自相同病人的两个差异血清样品(7 和 9)，分别在疾病发作第 4 天和第 44 天收集，在最初试验中，用 NRA 和 SMB NY 抗原试验为 IgM-阴性，用 Eg-101 SMB 抗原试验为 IgM-阳性。为了进一步调查二者之间的差异，从该病人按序收集的 6 个样品通过终点 MAC-和 IgG ELISA 重新试验。在第 3 天和第 15 天之间，MAC-ELISA 的滴度显示大于 32 倍的系列增加可以用使用的所有抗原证明。疾病发作后第 9 天收集的脑脊液也证实了这个病人在采集样品前不久确实感染了 WN。脑脊液抗 Eg-101-和 SLE-SMB 抗原的 IgM P/N 读数分别为 13.71 和 2.04。第 31 天和第 44 天的样品用 NY-SMB 抗原时为阴性(<1:400)，但是使用 NRA 和 Eg101-SMB 时为阳性。在这个试验中使用的所有 3 种抗原均观察到了一致的 IgG 滴度。

实施例 11 评价接种有 pCBWN 的动物的免疫反应。研究中使用了每组 10 只 3 周龄的雌性 ICR 小鼠。小鼠肌肉(i.m.)注射单剂量的 pCBWN 或者绿色荧光蛋白表达质粒(pEGFP) DNA(Clontech, San Francisco, CA.)。通过 EndoFree Plasmid Giga Kits(Qiagen)从 XL-1 blue 细胞纯化 pCBWN 质粒 DNA，然后重悬浮于 pH7.5 的 PBS 中，浓度为 1.0 μ g/ μ l。注射 100 μ g 的 pEGFP 的小鼠作为未接种对照。小鼠用 pCBWN 质粒注射，剂量分别为 100 μ l 体积中 100, 10, 1.0, 或者 0.1 μ g。接受了 10, 1.0, 或者 0.1 μ g 的 pCBWN 的小组利用 EMC-830 方形波电穿孔仪(Genetronics Inc. San Diego, CA.)通过电转移介导的体内基因转运的方法进行接种。电转移方法基于 Mir 等人的方法(*Proc. Natl. Acad. Sci. USA* 96:4262-4267 (1999))。DNA 注射后，立即通过两个不锈钢板电极施加经皮的电脉冲，在每侧腿部间距 4.5-5.5mm 放置。用 PBS 完全润湿腿部保证腿部皮肤的电接触。实施两组四个 40 伏/mm 的 25 毫秒的脉冲，两个脉冲之间间隔 200 毫秒。在每组脉冲之间颠倒电极极

性以提高电转移的效率。

注射后每 3 周对小鼠取血。WN 病毒特异性抗体反应通过 Ag-捕捉 ELISA 和噬斑减少中和实验 (PRNT) 进行评价。单独血清用 IgG-ELISA 进行试验, 来自每组 10 只小鼠的合并血清通过 PRNT 进行试验。在接种 3 周后, 所有用 pCBWN 接种的小鼠的 IgG ELISA 滴度范围为 1:640 到 1:1280。在 3 和 6 周收集的合并血清 Nt 抗体滴度为 1:80。来自 pEGFP 对照小鼠的血清样品无一显示对 WN 病毒的任何 ELISA 或者 Nt 滴度。

为了确定单独 i.m.接种 pCBWN 是否能够保护小鼠免受 WN 病毒的感染, 小鼠通过腹膜内注射或者暴露于病毒感染的库蚊叮咬下用 NY-6480 病毒进行攻击试验。半数小鼠组在接种 6 周后用 1,000 LD₅₀(1,025 PFU/100 μ l)的 NY99-6480 病毒进行腹膜内 (ip) 攻击。剩下的小鼠分别暴露于在攻击试验前 7 天感染有 NY99-6480 病毒的三带喙库蚊 (*Culex tritaeniorhynchus* mosquitoes) 的叮咬下。让蚊子叮咬小鼠直至蚊子充分饱食。在攻击试验后每天观察小鼠 2 次, 共三周。

很显然, Nt 抗体的出现相关于保护性免疫, 因为所有用 WN 病毒 DNA 免疫的小鼠在病毒攻击后都保持健康, 而所有对照小鼠在病毒攻击 4-6 天后均表现出 CNS 感染的症状, 并且腹膜内或者感染性蚊子攻击后分别平均 6.9 和 7.4 天后死亡。在接种组中, 病毒攻击(免疫后 9 周)3 周后收集的合并血清的 Nt 抗体滴度为 1:640 或 1:320。在 Western 印迹分析中, 合并的接种小鼠的血清仅仅和 E 蛋白反应。

通过电转移用每只动物 10.0 到 0.1 μ g 的 pCBWN 免疫每组 10 只小鼠。所有接受 pCBWN 的组都完全免于病毒攻击。免疫后 6 周, 所有电转移小鼠的 Nt 滴度低于接受常规 i.m.注射, 而非电转移 100 μ g pCBWN 的动物的 4 倍。这些证明有效免疫的结果暗示电转移方法增提高了本发明 DNA 疫苗的免疫原性和保护功效 (当按(Mir 等, *Proc.*

Natl. Acad. Sci. USA. 96:42624267 (1999))等描述的方法进行时)。

本研究中所用的混合繁殖的各种年龄的母马和被阉割的公马通过 ELISA 和 PRNT 试验显示为 WN 病毒和 SLE 病毒抗体阴性。四匹马 i.m 注射单剂量的 pCBWN 质粒(1,000 μ g/1,000 μ l 的 PBS, pH 7.5)。在病毒攻击前每隔 1 天收集血清样品,共 38 天,用 MAC-或 IgG ELISA 和 PRNT 评价 WN 病毒特异性抗体反应。

病毒攻击前两天,将 12 匹马(4 匹接种,8 匹对照)重新放入位于科罗拉多州立大学的生物安全级(BSL)-3 动物房。8 匹未接种的对照马是设计用来调查马中 WN 病毒诱导的发病机理以及马作为扩增宿主的潜力研究的亚群。在对马进行攻击实验前 12 天用 NY99-6425 或者 BC787 病毒感染的 14 或者 15 只白纹伊蚊 (*Aedes albopictus* mosquitoes) 叮咬攻击每匹试验马。让蚊子叮咬马 10 分钟。每天两次检查马的发病信号。测量体温,从第 0 天(感染的日子)到第 10 天每天两次收集血清样品,从第 11 天到第 14 天每天一次。攻击后每天记录脉搏和呼吸作用。通过噬斑滴定测定收集的血清样品用于检测病毒血症,以及通过 MAC-或者 IgG ELISA 和 PRNT 用于抗体反应。

在任一接种的马中没有观察到全身性或者局部反应。通过 PRNT 测定单个马血清。接种的马在第 14 和第 31 天之间 Nt 抗体大于或者等于 1:5。在第 37 天时(蚊子攻击前 2 天)接种马#5, #6, #7, 和#8 的终点滴度分别为 1:40, 1:5, 1:20, 和 1:20。用 pCBWN 质粒接种的马在病毒攻击后依然保持健康。从第 1 到第 14 天无一表现出可检测的病毒血症或者发烧。所有未接种的对照马在暴露于感染蚊子的叮咬下都被 WN 病毒感染。八匹未接种的马中有 7 匹在病毒攻击后 6 天内表现出病毒血症。患病毒血症的马在病毒攻击后第 7 和第 9 天之间出现 Nt 抗体。所有研究中表现疾病临床症状的唯一一匹马为 #11 马,其在感染 8 天后开始发烧并表现出神经症状。这匹马在 48 小时内发展成严重的临床疾病,并在第 9 天对其实施了安乐死。在表现 0, 2,

4 或者 6 天的病毒血症的四匹代表性马 # 9, 10, 14 和 15 被选择作为本实施例的例子。病毒滴度范围从最低可检测水平的 $10^{1.0}$ PFU/ml 血清(# 10 马)到 $10^{2.4}$ /ml(# 9 马)。在实验期间 # 14 马没有表现可检测的病毒血症。但是, 第 12 天后的 Nt 抗体检测证实这匹马感染了病毒。

在实验期间, 在接种的马中没有观察到如 Nt 滴度逐渐增加所显示的 Nt 抗体回忆应答。在蚊子攻击前接种马中现存的 Nt 抗体能抑制最初的病毒感染和复制。没有病毒复制, 由感染蚊子提供的攻击病毒抗原可能不含有足够的抗原量以刺激接种马中的免疫回忆应答。所有的接种马在病毒攻击 14 天后实施安乐死。没有观察到指示 WN 病毒感染的宏观病理学和组织病理学损害。

实施例 12. 制备含有黄热病毒 (YFV) 或者圣路易脑炎病毒 (SLEV) prM 和 E 蛋白编码序列的重组质粒。用类似于构建 pCDJE2-7 重组质粒的路线制备 YFV 和 SLEV 重组质粒。采用 QIAampTM Viral RNA 试剂盒(Qiagen, Santa Clarita, CA)从 150 μ L 的 YFV 病毒株 TRI-788379 或者 SLE 毒株 78V-6507 的病毒种子中提取基因组 RNA。病毒 RNA 作为扩增 YFV 或 SLEV 的 prM 和 E 基因编码区的模板。引物 YFDV389 (核苷酸序列 SEQ ID NO:4; 氨基酸序列, SEQ ID NO:5), cYFDV2452 (SEQ ID NO:6), SLEDV410 (核苷酸序列 SEQ ID NO:7; 氨基酸序列, SEQ ID NO:8)和 cSLEDV2449 (SEQ ID NO:9)用来产生如上所述用于制备 JEV 和 WNV 重组质粒的相应重组核酸。用 KpnI 和 NotI 酶消化的 RT-PCR 扩增 cDNA 插入真核表达质粒载体 pCDNA3(Invitrogen)的 KpnI-NotI 位点。测序 cDNA 的两条链, 并验证和来自 YFV 毒株 TRI-788379 或者 SLEV 毒株 78V-6507 的序列的同一性。分别含有 YFV 或者 SLEV 的 prM 和 E 编码区的核苷酸序列的重组质粒 pCDYF2 和 pCDSLE4-3 利用 EndoFreeTM Plasmid Maxi 试剂盒(Qiagen)进行纯化, 并用于体外转化或者小鼠免疫。

YFV 或者 SLEV 特异性抗原分别表达在被 pCDYF2 或者

pCDSLE4-3 转化的 COS-1 细胞中。蛋白的表达水平和 YFV-或者 SLEV-感染的 COS-1 细胞对照相似。如同在 JEV 模型中, 被携带病毒抗原基因的载体转化获得的 COS-1 细胞系组成型表达 YFV 或者 SLEV 抗原蛋白。利用一组 YFV 或者 SLEV E-特异性 Mabs 通过 IFA 进行的表位图谱显示真实的 E 蛋白由 pCDYF2-或者 pCDSLE4-3-转化的 COS-1 细胞表达。初步研究显示 100% 的三周龄雌性 ICR 小鼠在用单剂量的 100 μ g/100 μ L pCDSLE4-3 质粒的去离子水肌内接种后形成抗血清。

实施例 13.制备含有圣路易脑炎病毒的 prM 和 E 蛋白编码序列和 JEV 信号序列的重组质粒。采用 QIAampTM Viral RNA 试剂盒(Qiagen, Santa Clarita, CA)从 150 μ L 的用圣路易脑炎病毒 MSI-7 毒株感染的 Vero 细胞培养基中提取基因组 RNA。提取的 RNA 洗脱后悬浮在 80 μ L 不含核酸酶的水中, 并作为扩增圣路易脑炎病毒 prM 和 E 基因编码序列的扩增模板。引物序列获自 Trent 等人(*Virology* 156:293-304 (1987))的工作。含有基因组核苷酸区的 cDNA 片段通过逆转录酶-聚合酶链式反应(RT-PCR)进行扩增。利用扩增引物 SLE463 (SEQ ID NO:30)将限制性位点 AfeI 改建到 cDNA 的 5'端。在 NotI 限制性位点后面的符合读框翻译终止密码子利用扩增引物 cSLE2447 (SEQ ID NO:31)导入 cDNA 的 3'末端。利用 QIAquickTM PCR 纯化试剂盒(Qiagen)纯化 RT-PCR 产物。

上述两个扩增引物(SEQ ID NO :30 和 SEQ ID NO:31)产生的双链扩增子用 AfeI 和 NotI 酶消化产生 2004bp(463 到 2466nt)的 DNA 片段, 并插入 pCBJESS-M 载体的 AfeI 和 NotI 位点形成 pCBSLE(核苷酸序列, SEQ ID NO:21; 氨基酸序列, SEQ ID NO:22)。pCBJESS-M 衍生自 pCBamp 质粒, 其含有巨细胞病毒早期基因启动子和翻译控制序列元件以及改造的修饰 JE 信号序列元件(SEQ ID NO:27)。JE 信号序列元件含有在原始 pCBJESS 质粒的-4(Cys 到 Gly)和-2(Gly 到 Ser)位置上修饰的 JE 信号序列。

利用 ABI prism377 测序仪(Applied Biosystems/Perkin Elmer, Foster City, CA)自动测序 DNA 序列,以证实重组质粒具有按 Trent 等人(Virology 156:293-304 (1987))所限定的正确 prM 和 E 序列。

实施例 14.制备含有黄热病毒(YFV)的 prM 和 E 蛋白编码序列和 JEV 信号序列的重组质粒。采用 QIAamp™ Viral RNA 试剂盒(Qiagen, Santa Clarita, CA)从 150μL 的用黄热病毒 17D-213 毒株感染的 Vero 细胞培养基中提取基因组 RNA。提取的 RNA 洗脱后悬浮在 80μl 不含核酸酶的水中,并作为扩增黄热病毒 prM 和 E 基因编码序列的扩增模板。引物序列获自 Santos 等人(*Virus Research* 35:35-41 (1995))的工作。含有基因组核苷酸区的 cDNA 片段通过逆转录酶-聚合酶链式反应(RT-PCR)进行扩增。利用扩增引物 YF482(SEQ ID NO:28)将限制性位点 AfeI 改建到 cDNA 的 5'端。在 NotI 限制性位点后面的符合读框翻译终止密码子利用扩增引物 cYF2433 (SEQ ID NO:29)导入 cDNA 的 3'末端。利用 QIAquick™ PCR 纯化试剂盒(Qiagen)纯化 RT-PCR 产物。

上述两个扩增引物(SEQ ID NO :28 和 SEQ ID NO:29)产生的双链扩增子用 AfeI 和 NotI 酶消化产生 1971bp(482 到 2452nt)的 DNA 片段,并插入 pCBJESS-M 载体的 AfeI 和 NotI 位点形成 pCBYF(核苷酸序列, SEQ ID NO:23;氨基酸序列, SEQ ID NO:24)。pCBJESS-M 衍生自 pCBamp 质粒,其含有巨细胞病毒早期基因启动子和翻译控制序列元件以及改造的 JE 信号序列元件(SEQ ID NO:27)。JE 信号序列元件含有在 pCBJESS 质粒的 JESS 的-4(Cys 到 Gly)和-2(Gly 到 Ser)位置上修饰的 JE 信号序列。

利用 ABI prism377 测序仪(Applied Biosystems/Perkin Elmer, Foster City, CA)自动测序 DNA 序列,以证实重组质粒具有如 Santos 等人(*Virus Research* 35:35-41 (1995))所限定的正确 prM 和 E 序列。

实施例 15.制备含有玻瓦桑病毒的 prM 和 E 蛋白编码序列和 JEV

信号序列的重组质粒。采用 QIAamp™ Viral RNA 试剂盒(Qiagen, Santa Clarita, CA)从 150μL 的用玻瓦桑 LB 毒株感染的 Vero 细胞培养基中提取基因组 RNA。提取的 RNA 洗脱后悬浮在 80μl 不含核酸酶的水中, 并作为扩增玻瓦桑病毒 prM 和 E 基因编码序列的扩增模板。引物序列获自 Mandl 等人(*Virology* 194:173-184 (1993))的工作。含有基因组核苷酸区的 cDNA 片段通过逆转录酶-聚合酶链式反应(RT-PCR)进行扩增。利用扩增引物 POW454 (SEQ ID NO:25)将限制性位点 AfeI 改建到 cDNA 的 5'端。在 NotI 限制性位点后面的符合读框翻译终止密码子利用扩增引物 cPOW2417 (SEQ ID NO:26)导入 cDNA 的 3'末端。利用 QIAquick™ PCR 纯化试剂盒(Qiagen)纯化 RT-PCR 产物。

上述两个扩增引物(SEQ ID NO :25 和 SEQ ID NO:26)产生的双链扩增子用 AfeI 和 NotI 酶消化产生 1983bp(454 到 2436nt)的 DNA 片段, 并插入 pCBJESS-M 载体的 AfeI 和 NotI 位点形成 pCBPOW (核苷酸序列, SEQ ID NO:19; 氨基酸序列, SEQ ID NO:20)。pCBJESS-M 衍生自 pCBamp 质粒, 其含有巨细胞病毒早期基因启动子和翻译控制序列元件以及改造的 JE 信号序列元件(SEQ ID NO:27)。JE 信号序列元件含有在 pCBJESS 质粒的 JESS 的-4(Cys 到 Gly)和-2(Gly 到 Ser)位置上修饰的 JE 信号序列。

利用 ABI prism377 测序仪(Applied Biosystems/Perkin Elmer, Foster City, CA)自动测序 DNA 序列, 以证实重组质粒具有如 Mandl 等人(*Virology* 194:173-184 (1993))所限定的正确 prM 和 E 序列。

实施例 16.含有登革血清型 2 结构蛋白质编码序列的质粒的制备。遵照如实施于其它黄病毒的那些方法(参见实施例 1, 9 和 12-15)来制备含有登革血清型 2 抗原的核酸 TU 的载体。依照实施例, 用于载体构建的扩增引物可被选择来改造正常的登革病毒信号序列或被选择以改造其它黄病毒的信号序列, 例如修饰的日本脑炎病毒信号序列。

一个含有从 prM 到 E 的登革血清型 2 基因区的质粒被构建。登革血清型 2 prM 和 E 基因(Deubel 等, *Virology* 155:365-377 (1986); Gruenberg 等, *J.Gen.Virol.*69:1301-1398 (1988); Hahn 等, *Virology* 162:167-180 (1988))被连接到质粒, 诸如 pCDNA3 中, 然后被切下并克隆到诸如 pCBamp, pCEP4, pREP4 或 pRc/ RSV 的载体(由 Invitrogen, Carlsbad,CA 提供)中, 以使其表达。如果必要, cDNA 序列中编码的登革病毒血清型 2- 特异性序列可利用诸如聚合酶链式反应(PCR)的方法来扩增。或者, 如果病毒 RNA 是基因区域来源的, DNA 序列可通过 RT-PCR 方法来扩增。一个在 5'端包含一个起始密码子以及在 3'端包含一个终止密码子的 DNA 片段将被克隆到一个表达载体的合适限制性核酶特异性位点中, 通过如此的方式使得巨细胞病毒(CMV)立即早期(IE)启动子, 一个起始密码子和一个终止子, 被可操作地连接于登革病毒血清型 2 的序列中。

实施例 17.使用登革血清型 2 DNA 疫苗接种小鼠。

实施例 16 中制备的编码从 prM 到 E 的基因区域的登革血清型 2 核酸 TU 疫苗悬浮在适当的可药用载体中, 诸如注射用水中或缓冲生理盐水中, 并且肌肉注射到刚断奶小鼠组的体内。对照组接受缺少登革血清型 2 特异性基因的可比较的质粒制备物。登革血清型 2 特异性抗体, 和/或登革血清型 2-特异性免疫系统胞毒细胞的产生,在其后以固定时间间隔评价, 例如以周间隔。在核酸 TU 疫苗给药后大约二到四个月, 小鼠用登革血清型 2 病毒攻击。其后以合适的间隔评价病毒血症的水平, 例如每隔一天进行评价。母体抗体赋予的被动保护的评价显示在实施例 8 中。

实施例 18.改良的信号肽的设计和构建。 信号肽能决定插入蛋白质的易位和方向, 因而确定 prM 和 E 蛋白的拓扑学。真核细胞信号肽的最基本特征由被称为 h-区域的一段 8 到 12 个疏水氨基酸组成(von Heijne, “信号序列, 变化的限制”(“Signal sequences.The limits of

variation") *J.Mol.Biol.*184:99-105 (1985))。在启动子 Met 和 h-区域之间的区域, 已知为 n-区域, 通常具有一到五个氨基酸, 且通常带有正电的氨基酸。在 h- 区域和切割位点之间是 c-区域, 其由三到七个极性的但是最通常不带电的氨基酸残基组成。在病毒多蛋白合成的过程中, 在 C 和 prM 蛋白的结合区信号酶切割位点从隐蔽到可切割构象的调节依赖于 C 蛋白通过病毒蛋白酶复合物 NS2B/NS3 的在先去除 ((Lobigs, "黄病毒膜蛋白的切割和尖峰异源二聚体的分泌需要病毒蛋白酶 NS3 的功能" ("Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3,") *Proc.Natl.Acad.Sci.U S A.*90:6218-6222 (1993))。因此, 当 prM 和 E 蛋白质通过表达质粒单独表达时, 考虑病毒信号序列的有效性是很关键的。

不同质粒构建体中的信号肽的区别可用来至少部分解释, 蛋白质易位, 切割位点的出现以及正确的拓扑学的差异, 因此说明 prM 和 E 分泌和 VLP 形成。这些特性的调整或优化可通过带有能够赋予所期望特征的信号序列的选择或应用而得以改进。其可通过使用例如, 一个在真核生物上进行的隐马尔可夫模型(HMM)的计算机程序学习机器来完成(参见 Henrik Nielsen 等, "通过隐马尔可夫模型预测信号肽和信号锚定" ("Prediction of signal peptides and signal anchors by a hidden Markov model"), 分子生物学智能系统第六届国际会议录 (In Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6)), AAAI Press, Menlo Park, California, pp.122--130 (1998); Nielsen 等, "预测信号肽和其它蛋白分选信号的机器学习方法" ("Machine learning approaches to the prediction of signal peptides and other protein sorting signals,") *Protein Engineering* 12:3-9 (1999); Nielsen 等, "神经网络方法用于鉴定原核和真核生物信号肽并预测其切割位点" ("A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites,") *Int. J. Neural Sys.*8:581-599 (1997); "从序列到分选: 信号肽的

预测”(“From sequence to sorting: Prediction of signal peptides,”) Henrik Nielsen, 博士论文, Defended at Department of Biochemistry, Stockholm University, Sweden (May 25 1999); 每个文献均在此处引作参考, 尤其是对使用计算机运算法则优化信号序列相关教导的参考文献)。

所用程序类型的一个例子可在 <http://www.cbs.dtu.dk/services/SignalP-2.0/> 上自 2002 年 4 月 3 日起找到。所描述的 HMM 和所引用的参考文献被用于计算 prM 信号肽序列在不同质粒构建体中的信号肽概率(表 7)。SignalP- HMM 搜索正确地预测在所有构建体中信号肽酶的切割位点。然而, 观察到的切割概率(范围在 0.164 和 1.000 之间)和信号肽概率(范围在 0.165 到 1.00 之间)的区别很大(表 7)。这不令人惊讶, 因为切割位点和信号肽概率已知也被 n-区域的正电荷氨基酸, 构建体中 h-区域的疏水氨基酸长度和 c-区域的氨基酸组成所影响((Chang 等, “黄病毒 DNA 疫苗: 现状及前景”(“Flavivirus DNA vaccines: current status and potential,”) *Annals of NY Acad. Sci.* 951:272-285 (2001); Sakaguchi 等, “通过疏水区段和 N-末端电荷之间的平衡确定信号和信号锚定序列的功能”(“Functions of Signal and Signal-Anchor Sequences are Determined by the Balance Between the Hydrophobic Segment and the N-Terminal Charge,”) *Proc. Natl. Acad. Sci. USA* 89:16-19 (1992))。

各起源于不同 JE 病毒毒株的 3 个 JE 病毒质粒构建体, 显示出不同的疫苗潜力(Lin 等, “用日本脑炎病毒非结构蛋白 NS1 的 DNA 免疫引发小鼠的保护性免疫”(“DNA immunization with Japanese encephalitis virus nonstructural protein NS1 elicits protective immunity in mice,”) *J. Virol.* 72:191-200 (1998); Konishi 等, “通过用编码日本脑炎病毒前膜蛋白和包膜蛋白的基因的质粒免疫诱导小鼠抗日本脑炎的保护性免疫”(“Induction of protective immunity against Japanese encephalitis in mice by immunization with a plasmid encoding Japanese encephalitis virus premembrane and envelope genes,”) *J. Virol.* 72:4925-

4930 (1998); Chang 等, “单次肌内注射重组质粒 DNA 诱导小鼠的保护性免疫和预防日本脑炎” (“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,”) *J. Virol.* 74:4244-4252 (2000)). 这些构建体中的信号肽序列区别在可能含有或不含有带电荷氨基酸的 n-区域的长度(表 7)。包含正电荷氨基酸的 n-区域在细胞质一侧形成一个短环, 其导致 h-区域 (跨膜螺旋) 从尾部方向被插入, 暴露出信号酶切割位点。在我们的研究中, 包含 prM/M 和 E 蛋白的分泌性 VLP 可由 pCDJE2-7 转化的细胞系, JE4B 或 pCBJE1-14 瞬间转化的 COS-1 细胞的培养基中纯化而来。梯度纯化的 VLP 和病毒粒子具有相同的免疫学和生物化学特性。VLP 和病毒粒子之间从 prM 到成熟 M 蛋白的加工效率, 这是黄病毒形态发生的标记, 也是相似的。因此, 通过 pCDJE2-7 和 pCBJE1-14 表达的 prM 和 E 蛋白可以类似于病毒粒子 prM 和 E 的方向被表达成 I 型跨膜蛋白(Chang 等, “单次肌内注射重组质粒 DNA 诱导小鼠的保护性免疫和预防日本脑炎” (“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,”) *J. Virol.* 74:4244-4252 (2000)). 相反, pcDNA3JEME 的 prM 蛋白质可被表达为 II 型跨膜蛋白, 且因为其 n-区域正电荷氨基酸缺失使其跨膜 h-区域以头的方向插入(Konishi 等, “通过用编码日本脑炎病毒前膜蛋白和包膜蛋白的基因的质粒免疫诱导小鼠抗日本脑炎的保护性免疫” (“Induction of protective immunity against Japanese encephalitis in mice by immunization with a plasmid encoding Japanese encephalitis virus premembrane and envelope genes,”) *J. Virol.* 72:4925-4930 (1998)). 有效的蛋白质合成与具有正确的拓扑学的表达蛋白组合, 尤其是与表达的 prM 和 E 组合, 能增强 VLP 的形成和分泌, 并因此提高了 DNA 疫苗的免疫原性(Chang 等, “单次肌内注射重组质粒 DNA 诱导小鼠的保护性免疫和预防日本脑炎” (“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,”) *J. Virol.* 74:4244-4252 (2000)).

基于计算机计算的使用，如上所述，已经被应用于优化表达质粒的设计。尤其是，SignalP- HMM 程序的预测能力被用于设计 WN 病毒表达质粒(表 2) (Davis 等, “西尼罗病毒重组 DNA 疫苗保护小鼠和马免受病毒的攻击，且体外表达非感染的重组抗原能够用于酶联免疫吸附试验”(“West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays,”) *J. Virol.* 75:4040-4047 (2001))。pCBWN 质粒的组成是短形式的 JE 病毒信号肽，其后跟着 WN 病毒 prM- E 基因序列。此构建体的疫苗潜能已被充分地证明，因为单次 i.m.注射 pCBWN DNA 不仅诱导保护免疫而且也预防了小鼠和马的 WN 病毒感染。

如更早所讨论及如实施例 13-15 所证明，来自相同病毒的病毒编码信号序列如同抗原编码区域一样不必是有效的最佳的信号肽。此外，非修饰的信号序列不是必须最佳的。例如，质粒 pCBJEI-14 中编码的信号肽可被改进，如由信号序列概率测定一样，通过缩短 n-区域，改变 c-区域序列，或将两个修饰组合(图 6)。通过例证的方式，JE 病毒信号肽的缩短的形式已被用于 WN 病毒 prM 和 E 基因的表达，如在此所述，或在此引作参考的文献教导所述(Davis 等, “西尼罗病毒重组 DNA 疫苗保护小鼠和马免受病毒的攻击，且体外表达非感染的重组抗原能够用于酶联免疫吸附试验”(“West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzymelinked immunosorbent assays,”) *J. Virol.* 75:4040-4047 (2001))。由单次 i.m.接种的剂量滴定研究表明 pCBWN 在鼠中比 pCBJEI-14 的免疫原性多至少 2-4 倍。

实施例 19.多价疫苗。

设计用来免疫抗多种黄病毒的多价和/或联合疫苗也可被制备。

在多价疫苗的制备中，单价疫苗成分的制备包括与目的病原体相关的元件，诸如 YF，不同血清型的 DEN，JE，WN，SLE 和 TBE(RSSE 和 CEE)病毒或任一其它黄病毒的组合。DNA 构建体的设计和制备如其它实施例和说明书中所描述的进行。可制备适当疫苗的组合，以提供多价的或组合疫苗防御多种病原体。来自我们小组的初步数据证明 pCBJEI-14 和 pCBWN DNA 联合疫苗 i.m.注射小鼠，诱导 JE 病毒和 WN 病毒特异性 Nt 抗体的产生(表 8)。每种单价组分，即使使用同一转录和翻译调控子进行构建的，也应当优选在类似模型系统中测试来保证其疫苗潜力。然后以此配制组合的鸡尾酒疫苗。这种鸡尾酒疫苗可被特异性定制以适用于特定的地理区域。例如，一种用于热带和亚热带亚洲的鸡尾酒疫苗应包括四种血清型的 DEN，WN 和 JE 病毒疫苗。同样地，对于非洲和拉丁美洲有用的鸡尾酒疫苗应分别包括四种血清型的 DEN，WN 和 YF 病毒疫苗以及包括四种血清型的 DEN，Rocio 和 YF 病毒疫苗。

实施例 20.重组登革病毒 2 型疫苗的制备和试验。

a. 实施例的概述。 构建一系列编码登革病毒 2 型 (DEN-2) 前膜(prM)和外膜(E)蛋白的质粒。这些质粒包括一个真正的编码由 SEQ ID NO:43 所描述的蛋白的 DEN-2 prM- E 构建体(pCBD2-14-6)(SEQ ID NO:42)，一个 90% DEN-2E-10%日本脑炎(JE)病毒 E 嵌合构建体(pCB9D2-1J-4-3)(SEQ ID NO:44)，其编码由 SEQ ID NO:45 所描述的蛋白，以及一个 80%DEN-2E-20%JE E 嵌合构建体(pCB8D2-2J-2-9-1)(SEQ ID NO:46)，其编码由 SEQ ID NO:47 所描述的蛋白。单克隆抗体(MAb)反应性表明所有三个质粒均表达与一组 1，2 和 3 个抗体的结构域反应的 E 蛋白表位。然而，只有 pCB8D2-2J-2-9-1 构建体(SEQ ID NO:46)分泌高水平的 prM，M(成熟 prM)和 E 到质粒转化的 COS-1 细胞的培养基中。用 pCBD2-14-6 质粒(SEQ ID NO:42)转化的 COS-1 细胞表达的和用 pCB9D2-4-3 质粒(SEQ ID NO:44)转化的 COS-1 细胞表达的 prM 和 E 蛋白质的主要部分保持了膜结合活性。20%的编码 DEN-2E 的 E 蛋白的序列用编码相应的 JE E 蛋白序列替换对 MAbs

反应性没有影响。

在试验中，每组小鼠在 0 和 3 周接受两次肌肉免疫所选的质粒，且通过测定特异性中和和 ELISA 抗体评价免疫应答。表达分泌的 prM 和 E 的质粒，其可以形成亚病毒颗粒（SVP），在刺激抗体免疫应答上优于其它的构建体。90% 的中和滴度范围从 1:40 到 >1:1000 在 pCB8D2-2J-2-9-1 免疫小鼠的 9 个血清标本中有 7 个被观察到。

b. DEN-2 病毒和疫苗的重要性。

登革(DEN)热是一个在亚热带和热带地区发生的急性感染。它是人类最重要的黄病毒疾病之一。如较早所述，有四种不同的登革病毒的 DEN 血清型 (DEN-1, DEN-2, DEN-3 和 DEN-4)。由任意这些病毒导致的感染通常无症状或仅仅引起自限的发热，已知为登革热(DF)。然而，在百分比小的病例中，登革病毒感染造成一种比较严重的，威胁生命的登革热出血发烧或登革热休克综合症(DHF/DSS)。因此，虽在全球每年大约有一亿例相对轻微的 DF，受到有限的关注，同时据估计每年还有 500,000 就医的 DHF/ DSS 病例被报道。为了防御这种疾病，需要将有效抗所有四种血清型的安全和有效的 DEN 疫苗施用给 DEN 地方病区和传染区的孩子和非免疫的成人。

安全的疫苗必须使有毒病毒的严重感染的潜在危险最小化。此种有毒力的病毒可通过衍生自减毒疫苗病毒的某些类型疫苗的基因回复突变或重组而出现。此种情况确实在消灭脊髓灰质炎病毒的活动中出现过(Guillot 等, “人类疫苗和野生型脊髓灰质炎毒株之间的自然基因交换”(“Natural Genetic Exchanges between Vaccine and Wild Poliovirus Strains in Humans,”) *J. Virol.* 74:8434-8443 (2000); Liu 等, “1 型野生型脊髓灰质炎重组疫苗在中国广泛流行的分子演化”(“Molecular Evolution of a Type 1 Wild-Vaccine Poliovirus Recombinant during Widespread Circulation in China,”) *J. Virol.* 74:11153-11161 (2000))。此外，一个黄热病毒的美国毒株 TRINID79A 的基因测序,表明在这一毒

株和减毒的黄热病疫苗病毒 FNV 之间有广泛的相似性(Chang 等,“包膜蛋白基因核苷酸序列的变异鉴定两种不同的黄热病毒基因型”

("Nucleotide sequence variation of the envelope protein gene identifies two distinct genotypes of yellow fever virus,") J. Virol. 69:5773-5780 (1995); Pisano 等,“美国黄热病毒株 TRINID79A 的全核苷酸序列和系统发育” ("Complete nucleotide sequence and phylogeny of an American strain of yellow fever virus, TRINID79A,") Arch. Virol.144:1837-1843 (1999)。虽然不考虑其本身,相似性强烈地表明 TRINID 79 A 起源于 FNV 疫苗病毒。

基于 DNA 的疫苗的应用是一种新的和有前景的用于开发黄病毒疫苗的免疫方法(如这里所描述的,Chang 等,“黄病毒 DNA 疫苗:现状及前景” ("Flavivirus DNA vaccines:current status and potential,") *Ann. NY Acad. Sci.* 951:272-285 (2001), 在此处引用作为参考)。在此实施例中,大量的 DEN -2 疫苗被生产且在 DEN-2 构建体 i.m 免疫小鼠后,其免疫应答与 prM/ M 和 E 分泌的效率相关联。一种导致 prM/ M 和 E 抗原显著分泌的构建体显示能够在质粒接种的小鼠体内刺激高滴度的中和抗体。

c. 材料和方法

i.细胞培养物和病毒毒株。 COS-1 细胞(ATCC, Manassas,VA; 1650-CRL)在 37°C、5%CO₂ 的条件下生长在 Dulbecco 氏改良 Eagle 基础培养基(DMEM, GIBCO, Grand Island, NY)中,该基础培养基补充 10% 热灭活的胎牛血清(FBS, Hyclone Laboratories, Inc, Logan, UT), 1 mM 丙酮酸钠盐, 1mM 非必需氨基酸, 30 ml/ L 7.5%的 NaHCO₃, 100 单位/毫升的青霉素, 和 100 μg/ 毫升的链霉素。Vero 和 C6/36 细胞在与 COS-1 细胞相同的条件下培养。 DEN -2 病毒, 毒株-16681 用于 cDNA 克隆, IgG ELISA 和噬斑还原中和试验(PRNT)。病毒在 C6/36 细胞培养物中繁殖。用于免疫学或生物化学研究的病毒通过 7% 聚乙二醇(PEG-8000; Fisher Scientific, Fair Lawn, NJ)沉淀,然后用 30%

甘油-45% 钾-酒石酸盐梯度上的超高速离心进行纯化(Obijeski 等, “La Crosse 病毒的区段化基因组和核衣壳” (“Segmented genome and nucleocapsid of La Crosse virus,”) *J. Virol.* 20:664-675 (1976))。

ii. 质粒构建。通过利用 QIAamp™ 病毒 RNA 试剂盒 (Qiagen, Santa Clarita, CA)从 150µl 的 DEN-2 16681 毒株感染的 C6/36 细胞培养基中提取基因组 RNA。提取的 RNA 重悬浮在 80 µl 的焦碳酸二乙酯处理的水(DEPC, Sigma, ST. Louis, MO)中, 然后用作 DEN-2 病毒 prM 和 E 基因的 RT-PCR 扩增中的模板。引物序列(表 9)的设计基于已公开的序列(Gadkari 等,“蝙蝠中发现的森林病毒中和抗体的临界评价(初步报告)” (“Critical evaluation of Kyasanur Forest disease virus neutralizing antibodies found in bats (a preliminary report),”) *Indian J. Med. Res.* 64:64-67 (1976); Kinney 等, “构建登革 2 型病毒的感染 cDNA 克隆: 毒株 16681 及其减毒疫苗衍生物, 毒株 PDK-53” (“Construction of infectious cDNA clones for dengue 2 virus:strain 16681 and its attenuated vaccine derivative, strain PDK-53,”) *Virology* 230:300-308 (1997))。限制性酶 KasI 的识别和切割位点被插入在 cDNA 扩增子的 5'末端。接着一个 NotI 限制性位点的一个框内终止密码子导入 cDNA 扩增子的 3'末端。DEN-2 病毒 cDNA 扩增子用 KasI 和 NotI 酶消化, 然后被插入到 pCBJESS 载体的 KasI 和 NotI 位点以形成 100% DEN -2 E 质粒, pCBD2-14-6(SEQ ID NO:42)。

为了构建 90%和 80%的 DEN-2E 质粒, 该 100% DEN-2 质粒, pCBD2-14-6(SEQ ID NO:42), 和 JE 质粒, pCBJE1-14(SEQ ID NO:17) 被用作 PCR 模板来分别扩增 DEN-2 和 JE DNA 序列。用于扩增反应来获得 DEN-2 和 JE 基因片段的一系列引物列在表 9 中。T7 和 SP6 引发位点发现在 pCBamp 质粒中, 其衍生自原始 pCDNA-3 质粒 (Invitrogen, Carlsbad,CA), 且可按期望或需要使用。90% DEN-2-10% JE E 蛋白基因的 PCR 扩增 DNA 片段用 BstXI 限制性核酸内切酶消化, 用 T4 DNA 连接酶连接, 用 KasI 和 NotI 酶消化, 并插入到 pCBJESS 载

体的 KasI 和 NotI 位点中获得质粒 pCB9D2-1J-4-3(SEQ ID NO:44)。80% DEN-2-20% JE E 基因的 PCR 扩增 DNA 片段用 BsmBI 消化, 用 T4 DNA 连接酶连接, 用 KasI 和 NotI 酶消化, 并插入到 pCBJESS 载体的 KasI 和 NotI 位点中以获得质粒 pCB8D2-2J-2-9-1(SEQ ID NO:46)。三个质粒构建体的示意图显示在图 7 中。90%DEN-2-10%JE E 和 80%DEN-2-20%JE E 蛋白的结合区域, 分别显示在表 9 中。

自动化的 DNA 测序在 ABI Prism 377 测序仪 (Applied Biosystems/Perkin Elmer, Foster City, CA) 上依照厂商推荐的方法进行。带有正确 prM 和 E 序列的重组质粒通过序列分析得以鉴定。

iii. DEN-2 重组抗原通过电穿孔在 COS-1 细胞中的瞬间表达。

COS-1 细胞分别用 DEN-2 质粒或绿色荧光蛋白表达质粒对照 (pEGFP, Clontech, San Francisco, CA) 通过在其实施例中和 Chang 等的方法 (“单次肌肉注射重组质粒 DNA 诱导小鼠保护性免疫和预防日本脑炎” (“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,”) J. Virol. 74:4244-4252 (2000)) 进行电穿孔。电穿孔细胞接种到 75cm² 培养瓶上, 并保持在 37°C 和 5% CO₂。在电穿孔之后的六个小时培养基用含有 2% 胎牛血清的维持培养基替换。电穿孔 48 小时后收集组织培养基和细胞用于抗原的鉴定。

iv. 利用 DEN-2 E 特异性单克隆抗体进行表位绘图。

电穿孔 48 小时后, 吸附细胞被胰蛋白酶化, 重悬浮在含有 5% 山羊血清的 PBS 中, 点在 12 孔玻片上并空气干燥。吸附在点玻片上的细胞在 -20°C 与丙酮一起混合 10 分钟并使其空气干燥。E 蛋白特异性单克隆抗体 (MAb) 用于通过间接免疫荧光抗体试验 (IFA) 检测蛋白表达, 如前描述 (表 10; Chang 等, 单次肌肉注射重组质粒 DNA 诱导小鼠保护性免疫和预防日本脑炎 (“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents

Japanese encephalitis in mice,") *J. Virol.* 74:4244-4252 (2000))。

v. 重组 DEN-2 病毒抗原的鉴定。 电穿孔 48 小时后收集组织培养基。抗原-捕获 (Ag-capture)ELISA 用于检测瞬间转化的 COS-1 细胞的培养基中分泌的 DEN-2 病毒抗原。Mab 4G2 和辣根过氧化物酶结合的 Mab 6B6C-1 分别用于捕获 DEN 病毒抗原和检测捕获的抗原 (Chang 等, "单次肌肉注射重组质粒诱导小鼠的保护性免疫并预防日本脑炎" ("A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,") *J. Virol.* 74:4244-4252 (2000); Hunt 等, "在稳定转化的细胞中产生的日本脑炎病毒的重组颗粒性抗原是有效的非感染性抗原和亚基免疫原" ("A recombinant particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen,") *J. Virol. Methods.* 97:133-149 (2001))。

电穿孔 48 小时后, 每一质粒转化的细胞被胰蛋白酶化并重悬浮在 PBS 中使等分试样含有 5×10^6 个细胞。这些细胞样品使用 Mem-PER 哺乳动物膜蛋白质提取试剂盒(Pierce, Rockford, IL), 依照生产商建议的方法, 经处理后用于膜蛋白质提取。疏水和亲水性蛋白质被分开。此方法被发展成富集疏水相中的完整的膜蛋白。疏水和亲水性部分均通过 Ag-capture ELISA 用于 DEN -2 重组抗原的分析。

培养基中的重组抗原通过 10%聚乙二醇(PEG)-8000 沉淀而浓缩。沉淀物重悬浮在 TNE 缓冲液(50 mM Tris, 100 mM NaCl, 10 mM EDTA, pH 7.5)中至原始体积的 1/100, 通过离心使其澄清, 并 4℃下保存。通过 PEG 沉淀浓缩并重悬浮在 TNE 缓冲液中的重组抗原用 4.0%的乙醇提取以除去残余的 PEG (Hun 等, "在稳定转化的细胞中产生的日本脑炎病毒重组颗粒抗原是有效的非感染性抗原和亚基免疫原" ("A recombinant particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit

immunogen," "J. Virol. Methods. 97:133-149 (2001)). 乙醇提取的抗原, 来自转化细胞的疏水膜蛋白和梯度纯化的 DEN-2 病毒粒子利用 Excel Plus Electrophoresis Apparatus™(Invitrogen Corp, Carlsbad, CA) 在 NuPAGE, 4-12% Bis-Tris 梯度凝胶上进行分析, 然后使用 Excel Plus Blot Unit (Invitrogen Corp.) 在硝化纤维膜上进行电印迹。DEN-2 病毒-特异性蛋白通过 western 印迹使用 DEN-2 病毒特异性 MAbs1A6A-8(E 特异性)和 1A2A-1(衣壳特异性), 以及对 DEN-2 prM 特异性的兔抗血清和对包含 DEN-2M 蛋白的氨基酸 1-34 的肽特异性的小鼠血清进行检测, 以正常小鼠腹水作为阴性对照(Murray 等, "登革 2 型病毒蛋白 prM 和 C-prM 的加工" ("Processing of the dengue virus type 2 proteins prM and C-prM,") *J. Gen. Virol.* 74 (Pt 2):175-182 (1993); Roehrig 等, "牙买加登革 2 型病毒 Jamaica 包膜糖蛋白的单克隆抗体图谱" ("Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus,") *Virology* 246:317-328(1998))。

vi. 小鼠接种。 每组 10 只 3 周大的雌性 ICR 远交小鼠用于研究中。小鼠在第 0 周和第 3 周用 pCBD2-14-6, pCB9D2-1J-4-3, pCB8D2-2 J-2-9-1 或 pEGFP 以每只小鼠 100 μ g/100 μ l 的剂量进行 i.m 注射。质粒 DNA 用 EndoFree Plasmid Giga kits™(Qiagen)纯化自 XL-1 蓝细胞并重悬浮在 pH7.5 的 PBS 中, 浓度为 1.0 μ g/ μ l。接受 100 μ g pEGFP 的小鼠被用作质粒接种的对照。小鼠在注射之后每 3 三周取血一次, 且 DEN-2 病毒特异性抗体应答通过利用间接 ELISA 和 PRNT 进行评价。

vii. 血清学试验。 接种疫苗之前和之后的血清标本通过 ELISA 测试其抗体结合纯化的 DEN-2 病毒粒子的能力, 通过 PRNT 测试中和(Nt)抗体, 通过 Western 印迹测试识别纯化的 DEN-2 病毒蛋白的抗体。如前面所述(Chang 等, "单次肌肉注射重组质粒诱导小鼠的保护性免疫并预防日本脑炎病毒" ("A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents

Japanese encephalitis in mice,") *J. Virol.* 74:4244-4252 (2000)), 使用 DEN-2(毒株-16681)和 JE(毒株-Nakayama)病毒对 Vero 细胞进行 PRNT。测定终点在 90%的噬斑减少水平(Hunt 等,“在稳定转化的细胞中产生的日本脑炎病毒重组颗粒抗原是有效的非感染性抗原和亚基免疫原”(“A recombinant particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen,") *J. Virol. Methods.* 97:133-149 (2001))。

d. 结果。

i. DEN-2 病毒重组抗原的瞬间表达。通过分别转化 3 种重组 DEN-2 DNA 质粒到 COS-1 细胞中来完成 DEN-2 病毒的 prM 和 E 基因以及来自 DEN -2 病毒和 JE 病毒序列组合(80% DEN-20%JE 或 90% DEN -10%JE)的嵌合 E 基因的表达。基本质粒设计基于在前的用 JE 病毒和 WN 病毒重组质粒进行的研究结果,其中质粒转化的细胞表达并分泌真正的病毒蛋白到细胞培养液中(Chang 等,“单次肌肉注射重组质粒诱导小鼠的保护性免疫并预防日本脑炎病毒”(“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,") *J. Virol.* 74:4244-4252 (2000); Davis 等,“西尼罗病毒重组 DNA 疫苗保护小鼠和马抗病毒的攻击,并且体外表达可用作酶联免疫吸附试验的非感染型重组抗原”(“West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays,") *J. Virol.* 75:4040-4047 (2001))。DEN-2 重组蛋白质的瞬间表达通过细胞培养物上清液的 Ag-捕捉 ELISA 和通过丙酮-固定的转化 COS-1 细胞的 IFA 被首先评价(Chang 等,“单次肌肉注射重组质粒诱导小鼠的保护性免疫并预防日本脑炎病毒”(“A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice,") *J. Virol.* 74:4244-4252 (2000))。最佳抗原表达时间确定在电穿孔后 48 小时。

ii. 通过瞬间转化 COS-1 细胞表达的 E 蛋白表位图谱。

由每种重组质粒表达的 DEN-2 蛋白通过 IFA 利用一组带有已知与 DEN-2 病毒反应的鼠源 MAbs 进行评价(表 10; Henchal 等, “利用单克隆抗体表位分析登革-2 病毒粒子表面的抗原决定簇”(“Epitopic analysis of antigenic determinants on the surface of dengue-2 virions using monoclonal antibodies,”) *Am.J.Trop.Med.Hyg.* 34:162-169 (1985); Roehrig 等, “牙买加登革 2 病毒包膜糖蛋白的单克隆抗体图谱”(“Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica,”) *Virology* 246:317-328 (1998))。MAb 试验组包括与黄病毒的 E 蛋白以及 prM 和 C 蛋白的 3 个抗原结构域的每种反应的抗体(Mandl 等, “使用壁虱性脑炎病毒作为模型, 黄病毒包膜蛋白 E 分子水平上的抗原结构”(“Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model,”) *J. Virol.* 63:564-571 (1989); Rey 等, “2Å 分辨率的壁虱性脑炎病毒的包膜糖蛋白”(“The envelope glycoprotein from tick-borne encephalitis virus at 2Å resolution,”) *Nature* 375:291-298 (1995))。对黄病毒抗原结构域 2 和 3 特异的 MAbs 显示与 DEN-2 病毒和 3 种质粒表达蛋白每一个有几乎相同的定性反应。结构域 1-特异性 MAbs 之一, 1B4C-2 也显示出与所有表达蛋白相似的反应模式。然而, 结构域 1-特异性 MAbs, 2B3A-1 和 9A4D-1 中的两个与由质粒 pCBD2-14-6 和 pCB9D2-1J-4-3 表达的 E 蛋白如终点滴定显示有较少的反应(值在括号中, 表 10)。终点滴定的比较显示了表位 C3 和 C4 在含有 100% DEN-2 E 和 90% DEN-2 E-10% JE E 的构建体中表达水平明显低。对 prM 特异的 MAb 2H2, 具有与所有 3 种质粒表达的抗原相同的反应性。抗 C MAb 1A2A-1 与 DEN-2 病毒反应较好并且与质粒表达的病毒蛋白有低水平的非特异性反应, 所述病毒蛋白包括 prM 和 E, 但不包括 C。

iii. 由三种 DEN-2 重组质粒分别产生的分泌蛋白和膜结合蛋白的比较。

从每种重组 DEN-2 质粒转化 48 小时后的 COS-1 细胞中收集相似量的细胞培养液。在培养液中发现的分泌的重组抗原通过 PEG 沉淀浓缩 100 倍，然后通过乙醇提取除去干扰随后聚丙烯酰胺凝胶电泳分析的 PEG。由每种质粒表达的分泌抗原的相对量通过对 PEG-沉淀和乙醇抽提的细胞培养液制备物进行 Ag-捕捉 ELISA 分析来测定(表 11)。分泌的抗原仅从用 pCB8D2-2J-2-9-1 (SEQ ID NO:34)转染的细胞中检测到，其包含 80% DEN-2 E 和 20% JE E 基因。尽管努力浓缩表达的蛋白，但含有 100%DEN-2E 或 90% DEN-2 E-10% JE E 基因的重组质粒在培养液中不产生 ELISA-可检测的抗原。

Western 印迹分析也被用于评价由每种 DEN-2 重组质粒产生的分泌抗原。为了比较的目的，将等体积的 PEG 沉淀和乙醇提取的细胞培养上清液在 NuPAGE 梯度凝胶上电泳，电印迹到硝化纤维膜上，并利用 MAbs 或能与所有 DEN-2 结构蛋白反应的多克隆抗血清进行分析(图 8A)。在检测重组抗原中，Western 印迹分析显示比 Ag-捕获 ELISA 有较高的灵敏度，这是因为 DEN-2 特异性蛋白在来自两种质粒 pCB8D2-2J-2-9-1 和 pCB9D2-1J-4-3 (分别为 SEQ ID NO:46 和 44)的培养液中被检测到。质粒 pCB8D2-2J-2-9-1(SEQ ID NO:46)表达了大量的分泌抗原，其包括 E, prM 和 M 蛋白。由 pCB9D2-1J-4-3 (SEQ ID NO:44)生产相对较少的分泌抗原，且对 pCBD2-14-6 (SEQ ID NO:42)制备物而言几乎未发现可检测的水平，其显示含有相对低的表达 E 蛋白，特别是如果考虑到 E 特异性 MAb, 1A6A-8 和对照 pEGFP 的非特异性反应(图 8A, 14-6 和 GFP 泳道 a, b)。

由于 E, prM 和 M 在其胞内合成中是与膜结合的蛋白，因此这些蛋白通过 3 种重组 DEN-2 质粒表达的评价应包括来自质粒转化细胞的细胞膜制备物的评估。Mem-PER 哺乳动物膜蛋白提取试剂盒(Pierce)用于分离来自相同数量被每种重组质粒转化的细胞的完整膜蛋白。疏水蛋白通过相分离与亲水蛋白分开。通过 Ag-捕捉 ELISA 的初步分析显示亲水性蛋白部分是非反应性的；然而，来自用每一重组

DEN-2 质粒转化的 COS-1 细胞的疏水蛋白部分与 ELISA 试验有相似的滴度(表 11)。这些结果表明由所有 3 种质粒编码的重组抗原在转化后被表达,但表达的重组抗原并非全部在相同的水平上分泌。

疏水蛋白部分的 Ag-捕捉 ELISA 结果的证实通过 western 印迹完成(图 8B)。相同体积的来自每一质粒转化细胞的疏水蛋白部分依据厂商对 SDS-聚丙烯酰胺凝胶电泳的建议进行稀释,以减少带和泳道的扭曲。与 E-, prM-, C-, 和 M 特异性 MAbs 或多克隆抗血清的免疫印迹证实所有 3 种重组 DEN-2 质粒诱导产生相似量的含有 E 和 prM 的重组抗原。没有 M 蛋白被检测到,因为其不是从自 prM 加工而来,或者因为其水平太低检测不到。尽管努力减少带扭曲,与不含有高浓度表面活性剂的样品比较,疏水蛋白样品中高水平的表面活性剂显然引起 E 和 prM 以轻微异常的方式(较慢的迁移)进行电泳(图 8A 和 8B 中比较 E 和 prM 迁移)。

iv.用 3 种不同 DEN-2 重组 DNA 质粒接种小鼠的免疫应答比较.

三周龄的 ICR 小鼠在第 0 和第 3 周时用 100 μ g 的 pCB8D2-2J-2-9-1 (SEQ ID NO:46), pCB9D2-1J-4-3(SEQ ID NO:44), pCBD2-14-6(SEQ ID NO:42)或 pEGFP 通过 i.m 注射免疫。小鼠在初次免疫后的 3, 6, 和 9 周取血。单独和合并血清在接种 3 和 6 周后使用 1:100 和 1:400 的筛选稀释比通过间接 ELISA 测定,且在接种 9 周后进行终点滴定。9-周血清也通过 PRNT 利用 DEN-2 和 JE 病毒进行测定。ELISA 结果显示在一次免疫后(3 周血清),所有接受 pCB8D2-2J-2-9-1 的小鼠形成血清转化,然而只有 50%的 pCB9D2-1J-4-3 和 20% 的 pCBD2-14-6 接种小鼠与 DEN-2 病毒反应(表 12)。至接种后 9 周,所有用 pCB8D2-2J-2-9-1 或 pCBD2-1J-4-3 接种的小鼠显示了抗 DEN-2 ELISA 反应;然而,滴度的几何平均数显著不同(分别为滴度 1:20,000 对 1:708)。只有 40%的 pCBD2-14-6 免疫小鼠的抗 DEN-2 ELISA 的滴度大于 1:100。来自 pCB8D2-2J-2-9-1 免疫小鼠的合并 9 周血清对纯化的 DEN-2 病毒的 western 印迹显示对 E 糖蛋白免疫显性应答。也检测

到对 prM 和 M 的轻微反应。

更显然，根据 3 种 DEN-2 质粒疫苗潜力的评价，基于 90% 噬斑减少终点，9 只用 pCB8D2-2J-2-9-1 (SEQ ID NO:46) 免疫的小鼠中有 7 只观察到诱导的病毒中和抗体(表 10)。然而，如果使用 50% 的中和终点，则所有 9 只小鼠血清均具有 $31:40$ 的 PRNT 滴度。7 份带有中和活性的血清，90% 的中和滴度范围是从 $1:40$ 到 $>1:1000$ 。用 pCB9D2-IJ-4-3 免疫的小鼠无一产生中和抗体，且 10 只 pCBD2-14-6 接种小鼠的血清中只有一份产生中和病毒，其滴度仅为 $1:8$ 。

因为两个重组质粒，即 pCB9D2-IJ-4-3(SEQ ID NO:44) 和 pCB8D2-2J-2-9-1(SEQ ID NO:46) 含有 JE 病毒 E 基因序列，所有的血清可用于评价 JE 病毒中和活性的存在。然而，对任意免疫小组中的小鼠而言，90% 中和终点没有检测到这种活性。并不令人惊讶，用对照质粒 pEGFP 免疫的小鼠对 DEN-2 或 JE 病毒没有表现出活性。

e.讨论。 较早用于 JE 和 WN 疫苗生产的相同步骤被首先用于构建由确定的 DEN-2 prM 和 E 基因区域组成的重组 DEN-2 质粒，pCBD2-14-6(SEQ ID NO:42)。此质粒转化的 COS-1 细胞所表达的 DEN-2 蛋白的抗原性图谱用一组 MAb 通过 IFA 显示 prM 和 E 蛋白具有相容的荧光强度和与病毒感染细胞类似的 MAb 反应性(表 10)。然而，这些由编码真正 DEN-2 prM 和 E 区域的质粒所转化的 COS-1 细胞不能分泌可检测的 DEN-2 抗原到培养液中(如通过抗原-捕获 ELISA 所测定)。此外，利用编码确定的 DEN-2 prM 和 E 区域的质粒接种不能刺激 i.m. 免疫小鼠产生抗 DEN-2 病毒中和抗体(表 13)。有趣的是，通过 pCBD2-14-6 转化细胞产生带标点的球状荧光染色，表明 DEN-2 的 E 蛋白质的 C-末端可能有助于蛋白的膜保留信号。这种 IFA 染色模式在 JE 或 WN 构建体转化的细胞中没有观察到(Chang 等, “单次肌肉注射重组质粒诱导小鼠的保护性免疫并预防日本脑炎病毒” (“A single intramuscular injection of recombinant plasmid DNA induces protective

immunity and prevents Japanese encephalitis in mice,") *J. Virol.* 74:4244-4252 (2000); Davis 等, “西尼罗病毒重组 DNA 疫苗保护小鼠和马免受病毒的攻击, 且体外表达非感染的重组抗原可用于酶联免疫吸附试验” (“West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzymelinked immunosorbent assays,") *J. Virol.* 75:4040-4047 (2001)). 因此, 根据本发明教导所进行的观察, 其它两种质粒 pCB9D2-1J-4-3(SEQ ID NO:44)和 pCB8D2-2J-2-9-1(SEQ ID NO:46), 其中对 DNA 序列进行适当的操作, 将 DEN-2 的 C 末端的 10% 或 20% 分别用 JE 病毒 E 蛋白的相应区域取代。不同构建体在刺激免疫小鼠的可检测抗 DEN -2 ELISA 抗体中的相对效果显示在表 13 中。

这些结果与 prM 和 E 之间的交互作用能够影响粒子组装和分泌过程的模型相符。对此模型的支持发现在对壁虱性脑炎病毒的研究中, 其假设 prM 和 E 的胞外域之间的交互作用参与 prM 介导的 prM-E 胞内运输, 因而参与病毒类颗粒的分泌 (Allison 等, “壁虱性脑炎病毒包膜蛋白 E 的茎锚区功能元件的图谱” (“Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E,") *J. Virol.* 73:5605-5612 (1999))。

在本实施例中, 用 JE E 蛋白质替换 DEN-2 E 蛋白的 C-末端部分, 相应于 TBE H1^{pred} 替换 TM2, 导致 DEN-2 prM 蛋白质和嵌合 E 蛋白质的分泌。然而, 相比之下, TM1 和 TM2 在 TBE 中的替换, 仅引起抗原分泌的较小改进。由 pCBD2-14-6 和 pCB9D24-3 质粒转化的 COS-1 表达的 prM 和 E 蛋白的主要部分保持膜结合(表 13)。这些结果表明未确定的膜保留序列位于 DEN-2E 蛋白的 C 末端茎区。此 C 末端茎区用来自 JE 病毒的序列替换去除了此保留序列或使此保留序列无效。

其他人主张 prM 蛋白对于在 prM-E 成熟加工过程中维持 E 蛋白质的适当构象和分泌是必要的(Aberle 等, “以不同物理形式用表达壁

虱性脑炎病毒包膜蛋白 E 的构建体进行 DNA 免疫模式研究”(“A DNA immunization model study with constructs expressing the tick-borne encephalitis virus envelope protein E in different physical forms,”) *J. Immunol.* 163:6756-6761 (1999), Allison 等, “以可溶性或者颗粒形式合成和分泌重组壁虱性脑炎病毒蛋白 E”(“Synthesis and secretion of recombinant tick-borne encephalitis virus protein E in soluble and particulate form,”) *J. Virol.* 69:5816-5820 (1995)). 此外, 也已经证明 E 蛋白的外区域与 prM 互相作用。此相互作用已经被估计涉及墨累谷脑炎病毒中 E 的氨基酸残基 200-327 中的氨基酸序列 (Guirakhoo 等, “墨累谷脑炎病毒 prM 蛋白对病毒颗粒赋予了抗酸特征并且改变了 E 糖蛋白的 R2 结构域中表位的表达”(“The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein,”) *Virology* 191:921-931 (1992))。

适当的 prM 和 E 相互作用以及 E 蛋白结构保持完整可能保持所有 3 种 Den-2 构建体表达的蛋白中, 至少在免疫反应所要求的范围内。此外, pCB8D2-2J-2-9-1 中的 C 末端 20% E 的取代产生保持真正 DEN-2E 的 395 个氨基酸的蛋白。任意的这种修饰预计具有对 E 和 prM-E 相互作用的最小影响以及它们对嵌合 E 蛋白抗原特性的影响。由于用 JE 茎锚定序列替换 DEN-2E 的 C-末端区域对 MAbs 反应性没有影响(表 10), 如此替换的 DEN-2 序列的保持仅仅在获得 DEN-2 特异性免疫应答中是任选的。

以前, 已经显示一个编码壁虱性脑炎病毒 prM 和 E 蛋白的分泌亚病毒颗粒的质粒构建体, 在抗体应答的程度和功能性以及在对病毒攻击的应答方面, 优于其它编码 C 末端平截的分泌的可溶性 E 二聚物, 不分泌的全长 E, 没有有效分泌的平截 E 的构建体(Aberle 等, “以不同物理形式用表达壁虱性脑炎病毒包膜蛋白 E 构建体进行 DNA 免疫模式研究”(“A DNA immunization model study with constructs

expressing the tick-borne encephalitis virus envelope protein E in different physical forms,") *J. Immunol.* 163:6756-6761 (1999)). 然而, 在这里我们已经说明了 DEN -2 DNA 的疫苗效力和 prM/ M 和 E 的分泌相关(表 13)。但是, 分泌的 prM 和 E 的形态和物理特征在本研究中没有说明。然而, 由 pCB8D2-2J-2-9-1 构建体分泌的 prM 和 E 确实可能形成类病毒颗粒。相信在粒子表面呈递的多价保护性抗原提高了该构建体的疫苗效力。

早先开发 DEN-2 病毒 DNA 疫苗的尝试已经取得了不同程度的成功(Kochel 等, “接种表达登革-2 包膜基因的质粒引发小鼠的中和抗体” (“Inoculation of plasmids expressing the dengue-2 envelope gene elicit neutralizing antibodies in mice,") *Vaccine* 15:547-552 (1997); Konishi 等, “表达登革 2 型病毒前膜和包膜基因的 DNA 疫苗诱导小鼠的中和抗体和记忆 B 细胞” (“A DNA vaccine expressing dengue type 2 virus premembrane and envelope genes induces neutralizing antibody and memory B cells in mice,") *Vaccine* 18:1133-1139 (2000))。为了提高有效性水平, 已经采用不同的策略。例如, 已采用含有 pUC19 质粒的免疫刺激 CpG 基元的共免疫, 该质粒在疫苗系统中表达鼠源 GM-CSF, 或用溶酶体相关的膜保留序列替换 C-末端的 43 个氨基酸提高了对 DEN-2 疫苗的抗体反应(Porter 等, “登革 2 DNA 疫苗对小鼠的保护性功效以及 CpG 免疫刺激基元对于抗体应答的影响” (“Protective efficacy of a dengue 2 DNA vaccine in mice and the effect of CpG immuno-stimulatory motifs on antibody responses,") *Arch. Virol.* 143:997-1003 (1998); Raviprakash 等, “通过引入溶菌酶相关膜蛋白序列和使用表达 GM-CSF 的质粒对于登革 2 型病毒 DNA 疫苗的协同中和抗体应答” (“Synergistic Neutralizing Antibody Response to a Dengue Virus Type 2 DNA Vaccine by Incorporation of Lysosome-Associated Membrane Protein Sequences and Use of Plasmid Expressing GM-CSF,") *Virology* 290:74-82 (2001))。未甲基化的 CpG 基元直接激活巨噬细胞, 自然杀伤细胞和淋巴细胞以分泌细胞因子和趋化因子,

并支持通过 Th1 细胞因子介导的免疫应答的发生(Manders 等, “DNA 疫苗的免疫: CpG 基元和抗原呈递” (“Immunology of DNA vaccines:CpG motifs and antigen presentation,”) *Inflamm. Res.* 49:199-205 (2000))。然而, CpG 的包含可能偏离宿主细胞因子图形, 从而有助于发生 Th-1 介导的器官特异性自身免疫失调和干扰免疫平衡 (Smith 等, “DNA 疫苗的调节” (“The regulation of DNA vaccines,”) *Curr. Opin. Biotech.* 12:299-303 (2001))。也有证据显示小鼠体内过高水平的细胞因子虽然提高了特定 T-辅助细胞的应答, 但是能够降低或者关闭其它因素在免疫应答中的应答, 从而导致全面的免疫抑制或慢性炎症 (Robertson 等, “确保 DNA 疫苗的质量, 安全性, 和功效” (“Assuring the quality, safety, and efficacy of DNA vaccines,”) *Mol. Biotechnol.* 17:143-149 (2001))。因此, 黄病毒 DNA 免疫的安全性和功效可以通过操作质粒表达以增强转录和翻译以及靶向 prM 和 E 蛋白的分泌以促进多蛋白的正确加工和装配而获益(Chang 等, “黄病毒 DNA 疫苗: 现状及前景” (“Flavivirus DNA vaccines:current status and potential,”) *Ann. NY Acad. Sci.* 951:272-285 (2001))。今后的发展可能集中于加强抗原呈递细胞和肌肉细胞对于 DNA 的吸收(Rodriguez 等, “加强 DNA 免疫” (“Enhancing DNA immunization,”) *Virology* 268:233-238 (2000))。

表 1 在两个转化的细胞系中多种重组质粒瞬时表达 JE prM 和 M 蛋白

载体				重组质粒	阳性抗原细胞*的 IFA 强度/百分比	
	启动子	内含子	Poly(A)		COS-1	COS-7
pCDNA3	CMV	否	BGH	SV40	3+/40	3+/35
pCBamp	CMV	否	BGH	否	3+/45	nd
pCIBamp	CMV	是	BGH	否	3+/39	nd
pCEP4	CMV	否	SV40	OriP	2+/4	2+/3
pREP4	RSV	否	SV40	OriP	1+/3	1+/2
pRe/RSV	RSV	否	BGH	SV40	1+/3	1+/3
pCDNA3	CMV	否	BGH	SV40	-	-

*用 pCDNA3/CAT(阴性对照), pCDJE2-7, pCBJE1-14, pCIBJES14, pCEJEm pREJE, 或 pRCJE 转化不同的细胞系, 然后细胞用胰蛋白酶作用 48 小时, 并用 JE 病毒特异性 HIAF 通过间接免疫荧光抗体测定法 (IFA) 检测。数据表示为 IFA 阳性细胞的强度 (1+到 4+的范围) 和百分比。采用 pCDNA3/CAT 转化细胞作为阴性对照。

表 2.由带有 JE 病毒-反应抗体的 COS-1 细胞的 pCDJE2-7 稳定转化克隆(JE-4B)表达的蛋白的鉴定

Mab 或 抗血清	Mab 的生物活性		细胞的免疫荧光强度	
	特异性	生物功能	JEV 感染	4B
Mab:				
MC3	JEV 特异性		2+	2+
2F2	JEV 特异性	HI,N	4+	4+
112	JEV 特异性		4+	4+
503	JEV 特异性	N	4+	3+
109	亚型	HI	2+	1+
N.04	亚型	HI,N	3+	4+
201	亚型		1+	1+
203	亚型		4+	3+
204	亚型		2+	2+
301	亚型	HI	2+	2+
504	黄病毒		4+	4+
6B6C-1	黄病毒		2+	2+
3B4C-4	VEE		-	-
H1AF:				
Anti-JEV			4+	3+
Anti-WEE			-	-
PBS			-	-

表 3.用 pCDJE2-7 或 JE-VEX 疫苗免疫的小鼠免疫反应的持久性

	ELISA 滴度(log ₁₀)						PRNT _{90%} 滴度		
	3 周	6 周	9 周	23 周	40 周	60 周*	3 周	6 周	9 周
1xpCDJE2-7	2.6-3.2	3.8-5.0	3.8-4.4	>3.2	>3.2	2.4,2.4,3.8,4.4	<20	20	40-160
2xpCDJE2-7	2.6-3.8	4.4	3.8-4.4	>3.2	>3.2	2.6,3.8,3.8	<20	20-40	40-160
2xJE-VAX	2.6-3.8	4.4-5.0	3.8-5.6	>3.2	>3.2	<2,<2,<2,4.4	<20	20-40	20-160
2x pCDNA3/CAT	<2	<2	<2	ND	ND	<2	<20	<20	<20

小鼠接种 1 或 2, 100 μ g/剂量的质粒 DNA 或 1/5 人剂量的 JE-VAX 疫苗。在第二次免疫之前收集血清用于测试。

*单个血清滴度。

表 4. 用各种不同的 JEV 疫苗接种后小鼠年龄依赖的血清阳性率百分数

	3-天大		3-周龄	
	3 周 PV	7 周 PV	3 周 PV	7 周 PV
JE-VAX	0	0	100	100
pCDNA3/CAT	0	0	0	0
pCDJE2-7	40	60	90	90
C1BJES14	10	60	80	100
pCBE1-14	80	100	100	100

表 5. 在 3 天大时用各种不同 JEV 疫苗接种后 8 周龄的小鼠免于对 JEV 攻击

疫苗	攻击前 JEV 血清转化	各天攻击后的存活率(%)				
		6	7	8	9	21
JE-VAX	0	100	100	60	40	40
pCDNA3/CAT	0	100	80	30	30	30
pCDJE2-7	60	100	100	100	100	100
pC1BJES14	60	100	100	100	100	100
pCBE1-14	100	100	100	100	100	100

表 6.评价用 JEV-核酸接种的雌性小鼠产生的母亲抗体保护它们的幼崽免受致死 JEV 脑炎感染的能力

免疫接种的母亲		JEV 攻击的幼崽		
疫苗	PRNT _{90%}	攻击年龄 (天)	幸存者的数目 ¹	ELISA ²
1xpCDJE2-7	40	4	0/11	
2xpCDJE2-7	80	4	12/12	12/12
2xJE-VAX	20	3	0/16	
2xpCDNA-3/CAT	<10	5	0/14	
1xpCDJE2-7	20	15	5/11	5/5
2xpCDJE2-7	40	14	8/12	7/8
2xJE-VAX	80	13	5/5	5/5
2xpCDNA-3/CAT	<10	14	0/14	

小鼠用 1 或 2,100 μ g 剂量的质粒 DNA 肌内接种，或用 2, 1/5 人剂量的 JE-VAX 疫苗皮下接种。接种 9 周后收集血清在与非免疫的雄性交配前用于 PRNT 试验。

¹: 每一滴度的幸存者的数目/总数。

²: JEV ELISA-抗体-阳性动物的数目(滴度 $\geq$ 1:400)/幸存者的数目；在受攻击 12 周后收集血清用于试验。

表 7. 信号肽的特性及其在黄病毒 DNA 疫苗构建体中的疫苗潜力

质粒	prM 蛋白前的信号肽序列	信号肽概率 ^a			免疫方式/保护
		SP	AP	C 位点	
pSLE1	?LDTINRRPSKKRGGTRSLGLAALIGLASS/LQLLSTYQG (SEQ ID NO:32)	0.702	0.292	0.352	imx2/部分
pJME		0.998	0.000	0.778	imx2/部分
	MWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:33)				
pCJEME	MNEGSIMWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:34)	0.985	0.012	0.785	imx 2/100%
pCBJE1-14	MGRKQNKRGGNEGSIMWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:35)	0.791	0.199	0.623	im x 1/100%
pcDNA3prM-E	MSKKRGGSSETSVLMVFMLIGFAAA/LKLSNFQGK (SEQ ID NO:36)	0.721	0.277	0.622	im x 4/部分
					gg x 2-4/100%
pCBWN	MGKRSAGSIMWLASLAVVIACAGA/VTL SNFQGK (SEQ ID NO:37)	0.976	0.024	0.526	im x 1/100%
pI012D2ME	MNVLRGFRKEIGRMLNILNRRRRRTAGMIIMLPITVMA/FILITTRNGE (SEQ ID NO:38)	0.165	0.778	0.164	id x 2/无
SV-PE	MVGLQKRGKRRSATDWMSWLLVITLLGMTLA/ATVRKERGD (SEQ ID NO:39)	0.943	0.056	0.899	im 或 gg x 2/100%
pWRG7077-RSSE	MGWLLV VVLLGVTLA/ATVRKERGD (SEQ ID NO:40)	1.000	0.000	0.912	ggx2/100%
pWRG7077-CEE	MSWLLVITLLGMTIA/ATVRKERGD (SEQ ID NO:41)	0.999	0.000	0.821	gg x 2/100%

^a 使用 SignalP HMM 程序计算信号肽(SP)。锚定肽(AP)及信号酶切割位点 (C 位点) 的概率。

使用单字母氨基酸密码子, 带电荷的氨基酸用下划线和粗体字进行强调。信号酶切割将 SP 和 prM 分开的位点以 “/” 表示。DNA 疫苗通过肌内 (im), 皮内 (id), 或者基因枪 (gg) 的方法接种。

表 8. 用不同剂量的联合 WN 和 JE 病毒 DNA 疫苗免疫小鼠产生的中和抗体(Nt)反应

	pCBWN + pCBJE1-14			pCB 对照
每种质粒的剂量(μ g)	100 +100	40 + 40	20 +20	10 +10 100
含有 Nt 的小鼠的百分数:				
WN 病毒/JE 病毒:	100 / 100	100 / 700	70 / 0	60 /0 0 / 0
PRNT ₉₀ 滴度的范围:				
WN 病毒:	1:320-1:80	1:80 – 1:20	1:80 - < 1:10	1:20 - < 1:10 < 1:10
JE 病毒:	1:40 - 1:10	1:10 - < 1:10	< 1:10	< 1:10 < 1:10

每组 10 只三周龄雌性 ICR 远交小鼠通过 i.m.注射单剂量所显示的联合质粒 DNA。免疫 12 周后收集的血清样品通过噬斑减少中和试验 (PRNT) 进行检测。使用 JE 病毒 (SA-14 毒株) 和西尼罗病毒 (NY-6480 毒株), 基于 90%噬斑减少的百分比分别计算抗 JE 和 WN 病毒的终点滴度。

表 9. 用来构建 DEN-2 病毒 prM-E 表达质粒的寡核苷酸, 以及显示的嵌合 DEN-2 和 JE E 的接合区域。

100% DEN-2 prM-E:	
D2KasI-438 ^a	5'TGTGCAG <u>GGCGCCTT</u> CCATTTAACACACGTAACG (SEQ ID NO:48)
CD2NotI-2402	5'TCGAG <u>CGGCGCGCT</u> CAACTAATTAGGCCTGCACCATGACTC (SEQ ID NO:49)
90% DEN-2 E & 10% JE E:	
T7	5'CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:50)
CD2BstXI-2244	5'ATAGATTGCT <u>CC</u> AAACACT <u>TGG</u> TGG (SEQ ID NO:51)
JE-2281	5'ACTCCATAGGAAAAGCCGTTTCACC (SEQ ID NO:52)
CSP6	5'GCGAGCTCTAGCATTTAGGTGACACTATAG (SEQ ID NO:53)
DEN-2 ⇌ ⇌ JE	
90-10 接合区	Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr (SEQ ID NO:55)
	CTC CAC CAA GTG TTT GGT GGT GCC TTC AGA ACA (SEQ ID NO:54)
80% DEN-2 E & 20% JE E:	
T7	5'CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:56)
CD2BsmBI-2097	5'GAATT <u>CGTCTC</u> ACTTCCTTTCTTAAACCAGTTGAGCTTC (SEQ ID NO:57)
JEBsmBI-2175	5'GGAATT <u>CGTCTC</u> CGGAAGCACGCTGGGCAAGG (SEQ ID NO:58)
CSP6	5'GCGAGCTCTAGCATTTAGGTGACACTATAG 3' (SEQ ID NO:59)
DEN-2 ⇌ ⇌ JE	
80-20 接合区	Asn Trp Lys Lys Gly Ser Thr Leu Gly Lys Ala (SEQ ID NO:61)
	AAC TGG TTT AAG AAA GGA AGC ACG CTG GGC GCC (SEQ ID NO:60)

^a 寡核苷酸中编码的限制性酶切位点用黑体, 斜体, 和下划线表示

表 10. 通过间接荧光抗体试验(IFA)测定的重组 DEN-2 质粒表达的 DEN-2E 糖蛋白表位的特性

抗体		对照 ^a		质粒构建体 ^a		
MAb(表位) ^b	抗原结构域 ^c	PRNT ^d	DEN-2 感染的细胞	正常 细胞	pCBD2-1J-4-3	pCB8D2-2J-2-9-1
4G2(A1)	2	+/-	4+	-	4+	4+
4E5(A2)	2	是	3+	-	3-4+	2-3+
1B7(A5)	2	是	3-4+	-	4+	2-3+
1B4C-2(C1)	1	否	3-4+(8000)	-	2-3+(4000)	2-3+(8000)
2B3A-1(C3)	1	否	3-4+(≥3200)	-	3+(100)	2-3+(≥3200)
9A4D-1(C4)	1	否	3-4+	-	2-3+(400)	3+(≥12800)
3H5(B2)	3	是	4+	-	4+	4+
10A4D-2(B3)	3	是	2-3+	-	3-4+	2-3+
1A1D-2(B4)	3	是	4+	-	4+	3-4+
9D12-6		是	2-4+	-	2-3+	3-4+
2H2	prM	否	4+	-	3-4+	3-4+
1A2A-1	衣壳	否	2-3+	-	1+	1-2+

^aIFA 底物为丙酮-固定的 COS-1 细胞，其用 DEN-2 16681 感染，或者不感染作为对照，或者用 DEN-2 重组质粒转化。

^b基于和 DEN-2 16681 病毒的反应，单克隆抗体以预定的最佳稀释比使用。对于一些 MAbs，用括号表示的终点滴度被报道，其它的只报道了定性的值，范围是从 1+到 4+，3-4+为阳性，2+模棱两可，以及 1+为阴性。

^c基于 TBE 病毒的 E-糖蛋白的抗原结构域(Mandl 等)，“使用壁虱性脑炎病毒作为模型的黄病毒包膜蛋白 E 分子水

平的抗原结构” (“Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model,”) *J. Virol.* 63:564-571 (1989); Rey 等, “2A 分辨率的来自蜱虱性脑炎病毒的包膜糖蛋白” (“The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution,”) *Nature* 375:291-298 (1995)。

^d腹水 1:100 稀释比的噬斑减少中和活性, 使用 90%的噬斑减少终点, 除了 4G2 和 9D12-6, 对其报道了 50%的中和终点 (Henchal 等, “使用单克隆抗体表位分析登革-2 病毒粒子表面上的抗原决定簇” (“Epitopic analysis of antigenic determinants on the surface of dengue-2 virions using monoclonal antibodies,”) *Am. J. Trop. Med. Hyg.* 34:162-169 (1985); Roehrig 等, “牙买加登革-2 病毒包膜糖蛋白的单克隆抗体图谱” (“Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica,”) *Virology* 246:317-328 (1998))。

表 11. 通过抗原捕捉 ELISA 检测分泌型和膜结合型 DEN-2 重组蛋白

质粒	样品类型	终点 ELISA 滴度
PCBD2-14-6	PEG-沉淀的培养液 ^a	<1:10
PCBD2-14-6	PEG-沉淀、乙醇提取 的培养液 ^b	<1:20
PCBD2-14-6	疏水膜蛋白制备物 ^c	1:160
pCB9D2-1J-4-3	PEG-沉淀的培养液 ^a	<1:10
pCB9D2-1J-4-3	PEG-沉淀、乙醇提取 的培养液 ^b	<1:20
pCB9D2-1J-4-3	疏水膜蛋白制备物 ^c	1:80
pCB8D2-2J-2-9-1	PEG-沉淀的培养液 ^a	1:640
pCB8D2-2J-2-9-1	PEG-沉淀、乙醇提取 的培养液 ^b	1:80
pCB8D2-2J-2-9-1	疏水膜蛋白制备物 ^c	1:80
PEGFP	PEG-沉淀的培养液 ^a	<1:10
PEGFP	PEG-沉淀、乙醇提取 的培养液 ^b	<1:10
PEGFP	疏水膜蛋白制备物 ^c	<1:10

^a 来自质粒转化细胞的培养上清液用 10% 聚乙二醇(PEG)沉淀并重悬浮在 1/100 的原始体积中。

^bPEG-沉淀的培养上清液用 4%乙醇提取除去 PEG，沉淀重悬浮在 1/5 提取的体积中。

^c 疏水膜部分按材料和方法所述进行制备。

表 12. ICR 小鼠中三种 DEN-2 重组质粒的免疫原性

质粒 DNA ^b	小鼠 #	DEN-2 病毒的 ELISA				DEN-2 病 毒的 PRNT ^a 终点滴度 9 周 p.v.	JE 病毒 的 PRNT ^a 终点滴度 9 周 p.v.	
		筛选 3 周,p.v. ^c		筛选 6 周,p.v. ^c				终点滴 度 9 周
		1:100	1:400	1:100	1:400			
pCB8D 2-2J-2- 9-1	合并 1,2,4- 10	ND ^d	ND	+	+	64,000	ND	ND
	1	+	+	+	+	64,000	>1000	<2
	2	+	+	+	+	32,000	>1000	<2
	4	+	+	+	+	16,000	200	<2
	5	+	+	+	+	4,000	<10	<2
	6	+	+	+	+	16,000	200	<2
	7	+	-	+	+	64,000	100	<2
	8	+	-	+	+	8,000	40	<2
	9	+	+	+	+	6,400	<2	<4
	10	+	+	+	+	64,000	>1000	<2
pCB9D 2-1J-4- 3	合并 1-10	ND	ND	+	+	1,000	ND	<2 ^e
	1	-	-	+	-	400	<10	ND
	2	+	-	+	+	200	<10	ND
	3	+	+	+	+	4,000	<2	≤4
	4	+	-	+	-	200	<10	ND
	5	-	-	+	+	400	<10	ND
	6	+	+	+	+	4,000	<2	2
	7	-	+/-	-	-	100	<10	ND
	8	-	-	-	-	200	<10	ND
	9	+	-	+	-	4,000	<2	<2
pCBD2 -14-6	合并 1-10	ND	ND	+	-	200	<2 ^f	<2 ^g
	1	-	-	-	-	400	<10	ND
	2,3,6-9	-	-	-	-	<100	ND	ND
	4	+	+	+	+	1,000	<2	<2
	5	-	-	+	-	2,000	8	<2
	10	+	-	-	-	<100	ND	ND
pEGFP	混合组 1-10	-	ND	-	ND	<100	<2	<2

^aPRNT,噬斑减少中和试验, 90%中和终点。^b小鼠用 100 μ 的质粒 DNA 在第 0 周和第 3 周进行肌肉免疫。^cELISA 筛选采用 1:100 和 1:400 的稀释比血清。^dND, 没做^e合并 1,2,4,5,7,8。^f合并 2,3,6-10。^g合并 1-3,6-10。

表 13. 三种 DEN-2 重组质粒的特征概述

质粒	IFA ^a		Ag-捕捉 ELISA 滴度		DEN-2 ^b 的 ELISA 滴度		DEN-2PRNT ^c
	+/-	球状/分散	分泌的抗原	疏水膜蛋白制备物	≥1:100 血清数目	合并血清滴度	≥1:10 血清数目
pCB8D2-2J-2-9-1	+	分散	1:640	1:80	9/9	1:6400 0	7/9 ^d
pCB9D2-1J-4-3	+	球状	<1:10	1:80	10/10	1:1000	0/10
pCBD2-14-6	+	球状	<1:10	1:160	3/10	1:200	0/10

^a 间接荧光抗体测定(IFA)的染色特征, +或者-, 分散或者球状。

^b 来自用重组质粒免疫小鼠的血清的抗-DEN-2 ELISA 滴度。血清收集自免疫后 9 周 (第 0 周和第 3 周)。显示了总小鼠数中滴度 ≥1:100 的小鼠数, 包括合并血清样品的终点 ELISA 滴度。

^c 噬斑减少中和滴度(PRNT, 减少 90%) ≥1:10 的小鼠数/小鼠总数。免疫 9 周后, 收集血清。

^d 在具有中和抗体的 7 只小鼠中, 有 3 只 PRNT 滴度 >1:1000, 3 只滴度 ≥1:100 但 <1:1000, 1 只滴度为 1:40。

序列表

<110> 美国政府健康及人类服务部, 疾病控制和预防中心

(The Government of the United States of America, as represented by the Secretary, Department of Health and Human Services, c/o Centers for Disease Control and Prevention)

张光正 (Chang, Gwong-Jen J)

<120> 预防黄病毒感染的核酸疫苗

(Nucleic Acid Vaccines for Prevention of Flavivirus Infection)

<130> SCT032688-47

<150> PCT/US/10764

<151> 2002-04-04

<150> 09/826, 115

<151> 2001-04-04

<150> 09/701, 536

<151> 2000-11-29

<150> PCT/US99/12298

<151> 1999-06-03

<150> 60/087, 908

<151> 1998-06-04

<160> 61

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 48

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> CDS

<222> (25)... (48)

<221> misc_特征

<222> 1-48

<223> 扩增引物 14DV389

<400> 1

```
cttggtacct ctagagccgc cgcc atg ggc aga aag caa aac aaa aga      48
                Met Gly Arg Lys Gln Asn Lys Arg
                  1             5
```

<210> 2

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注＝合成构建体

<400> 2

```
Met Gly Arg Lys Gln Asn Lys Arg
  1             5
```

<210> 3

<211> 50

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注＝合成构建体

<221> misc_特征

<222> 1-50

<223> 扩增引物 c14DV2453

<400> 3

```
ttttcttttg cggccgctca aacttaagca tgcacattgg tcgctaagaa      50
```

<210> 4

<211> 48

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注＝合成构建体

<221> CDS

<222> (25)... (48)

<221> misc_特征

<222> (1)... (48)

<223> 扩增引物 YFDV389

<400> 4

```
cttggtagcct ctagagccgc cgcc atg cgt tcc cat gat gtt ctg act      48
                Met Arg Ser His Asp Val Leu Thr
                  1                5
```

<210> 5

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 5

```
Met Arg Ser His Asp Val Leu Thr
  1                5
```

<210> 6

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-41

<223> 扩增引物 cYFDV2452

<400> 6

```
ttttcttttg cggccgctca cgccccaact cctagagaaa c      41
```

<210> 7

<211> 51

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<221> CDS

<222> (25)... (54)

<221> misc_特征

<222> 1-54

<223> 扩增引物 SLEDV410

<400> 7

cttgggtacct ctagagccgc cgcc atg tct aaa aaa aga gga ggg acc aga 51

Met Ser Lys Lys Arg Gly Gly Thr Arg

1

5

<210> 8

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 8

Met Ser Lys Lys Arg Gly Gly Thr Arg

1

5

<210> 9

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<221> misc_特征

<222> 1-38

<223> 扩增引物 cSLEDV2449

<400> 9

ttttcttttg cggccgctta ggcttgacag ctggttgc

38

<210> 10

<211> 7500

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 备注=合成构建体

<221> CDS

<222> (916)... (3009)

<221> misc_特征

<222> 1-7500

<223> pCDJE 2-7

<400> 10

```

gacggatcgg gagatctccc gatccccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
ccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420
attgacgtca atgggtggac tatttacggg aaactgcccc cttggcagta catcaagtgt      480
atcatatgcc aagtagcccc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgcccgagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca      600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg      660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc      720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg      780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca      840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttggtacc      900
gagctcgccg ccgcc atg ggc aga aag caa aac aaa aga gga gga aat gaa      951

```

Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu

1

5

10

```

ggc tca atc atg tgg ctc gcg agc ttg gca gtt gtc ata gct tgt gcg      999
Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala

```

15

20

25

```

gga gcc atg aag ttg tcg aat ttc cag ggg aag ctt ttg atg acc atc      1047
Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile

```

30

35

40

aac aac acg gac att gca gac gtt atc gtg att ccc acc tca aaa gga	1095
Asn Asn Thr Asp Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly	
45 50 55 60	
 gag aac aga tgc tgg gtc cgg gca atc gac gtc ggc tac atg tgt gag	1143
Glu Asn Arg Cys Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu	
65 70 75	
 gac act atc acg tac gaa tgt cct aag ctt acc atg ggc aat gat cca	1191
Asp Thr Ile Thr Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro	
80 85 90	
 gag gat gtg gat tgc tgg tgt gac aac caa gaa gtc tac gtc caa tat	1239
Glu Asp Val Asp Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr	
95 100 105	
 gga cgg tgc acg cgg acc agg cat tcc aag cga agc agg aga tcc gtg	1287
Gly Arg Cys Thr Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val	
110 115 120	
 tcg gtc caa aca cat ggg gag agt tca cta gtg aat aaa aaa gag gct	1335
Ser Val Gln Thr His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala	
125 130 135 140	
 tgg ctg gat tca acg aaa gcc aca cga tat ctc atg aaa act gag aac	1383
Trp Leu Asp Ser Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn	
145 150 155	
 tgg atc ata agg aat cct ggc tat gct ttc ctg gcg gcg gta ctt ggc	1431
Trp Ile Ile Arg Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly	
160 165 170	
 tgg atg ctt ggc agt aac aac ggt caa cgc gtg gta ttt acc atc ctc	1479
Trp Met Leu Gly Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu	
175 180 185	
 ctg ctg ttg gtc gct ccg gct tac agt ttt aat tgt ctg gga atg ggc	1527
Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly	
190 195 200	
 aat cgt gac ttc ata gaa gga gcc agt gga gcc act tgg gtg gac ttg	1575
Asn Arg Asp Phe Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu	
205 210 215 220	

gtg ctg gaa gga gat agc tgc ttg aca atc atg gca aac gac aaa cca Val Leu Glu Gly Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro 225 230 235	1623
aca ttg gac gtc cgc atg att aac atc gaa gct agc caa ctt gct gag Thr Leu Asp Val Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu 240 245 250	1671
gtc aga agt tac tgc tat cat gct tca gtc act gac atc tcg acg gtg Val Arg Ser Tyr Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val 255 260 265	1719
gct cgg tgc ccc acg act gga gaa gcc cac aac gag aag cga gct gat Ala Arg Cys Pro Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp 270 275 280	1767
agt agc tat gtg tgc aaa caa ggc ttc act gac cgt ggg tgg ggc aac Ser Ser Tyr Val Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn 285 290 295 300	1815
gga tgt gga ctt ttc ggg aag gga agc att gac aca tgt gca aaa ttc Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe 305 310 315	1863
tcc tgc acc agt aaa gcg att ggg aga aca atc cag cca gaa aac atc Ser Cys Thr Ser Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile 320 325 330	1911
aaa tac gaa gtt ggc att ttt gtg cat gga acc acc act tcg gaa aac Lys Tyr Glu Val Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn 335 340 345	1959
cat ggg aat tat tca gcg caa gtt ggg gcg tcc cag gcg gca aag ttt His Gly Asn Tyr Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe 350 355 360	2007
aca gta aca ccc aat gct cct tcg ata acc ctc aaa ctt ggt gac tac Thr Val Thr Pro Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr 365 370 375 380	2055
gga gaa gtc aca ctg gac tgt gag cca agg agt gga ctg aac act gaa Gly Glu Val Thr Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu 385 390 395	2103

gcg ttt tac gtc atg acc gtg ggg tca aag tca ttt ctg gtc cat agg Ala Phe Tyr Val Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg 400 405 410	2151
gag tgg ttt cat gac ctc gct ctc ccc tgg acg tcc cct tcg agc aca Glu Trp Phe His Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr 415 420 425	2199
gcg tgg aga aac aga gaa ctc ctc atg gaa ttt gaa gag gcg cac gcc Ala Trp Arg Asn Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala 430 435 440	2247
aca aaa cag tcc gtt gtt gct ctt ggg tca cag gaa gga ggc ctc cat Thr Lys Gln Ser Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His 445 450 455 460	2295
cag gcgttg gca gga gcc atc gtg gtg gag tac tca agc tca gtg aag Gln Ala Leu Ala Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys 465 470 475	2343
tta aca tca ggc cac ctg aaa tgt agg ctg aaa atg gac aaa ctg gct Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala 480 485 490	2391
ctg aaa ggc aca acc tat ggc atg tgt aca gaa aaa ttc tcg ttc gcg Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala 495 500 505	2439
aaa aat ccg gcg gac act ggt cac gga aca gtt gtc att gaa ctc tcc Lys Asn Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser 510 515 520	2487
tac tct ggg agt gat ggc ccc tgc aaa att ccg att gct tcc gtt gcg Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala 525 530 535 540	2535
agc ctc aat gac atg acc ccc gtt ggg cgg ctg gtg aca gtg aac ccc Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro 545 550 555	2583
ttc gtc gcg act tcc agt gcc agc tca aag gtg ctg gtc gag atg gaa Phe Val Ala Thr Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu 560 565 570	2631

ccc ccc ttc gga gac tcc tac atc gta gtt gga agg gga gac aag cag Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln 575 580 585	2679
atc aac cac cat tgg cac aaa gct gga agc acg ctg ggc aag gcc ttt Ile Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe 590 595 600	2727
tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc gac aca Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr 605 610 615 620	2775
gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata gga aaa Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys 625 630 635	2823
gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt ggg gga Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly 640 645 650	2871
atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc tgg atg Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met 655 660 665	2919
ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta gcc aca Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr 670 675 680	2967
ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct taa Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala *	3009
685 690 695	
ttagtttgag cggccgctcg agcatgcac tagagggcc tttctatag tgtcacctaa atgctagagc tcgctgatca gcctcgactg tgccttctag ttgccagcca tctgttggtt gcccccccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag gcggaagaa ccagctgggg ctctaggggg tatccccacg cgccctgtag cggcgcatca agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt tcgccggett tccccgtcaa gctctaaatc ggggcatccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attaggggtga tggttcacgt agtgggcat cgccctgata gacgggtttt cgcccttga cgttgagtc cacgttcttt aatagtggac tcttggtcca aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag	3069 3129 3189 3249 3309 3369 3429 3489 3549 3609 3669 3729

```

ggattttggg gatttcggcc tatttggttaa aaaatgagct gatttaacaa aaatttaacg 3789
cgaattaatt ctgtggaatg tgtgtcagtt aggggtgtgga aagtccccag gctccccagg 3849
caggcagaag tatgcaaagc atgcatctca attagtcagc aaccagggtgt ggaaagtccc 3909
caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccatag 3969
tccccccct aactccgcc atcccccccc taactccgcc cagttccgcc cattctccgc 4029
cccatggctg actaatTTTT tttatttatg cagaggccga ggccgcctct gcctctgagc 4089
tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctcccgg 4149
gagcttgat atccattttc ggatctgatc aagagacagg atgaggatcg tttcgcatga 4209
ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct 4269
atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc 4329
aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg 4389
acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg 4449
acgttgctac tgaagcgga agggactggc tgctattggg cgaagtccg gggcaggatc 4509
tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc 4569
ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaac catcgcatcg 4629
agcagacacg tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc 4689
atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggccgcgatg cccgacggcg 4749
aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc 4809
gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag 4869
cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgttctctcg 4929
tgctttacgg tatcgccgct cccgattcgc agcgcacgc cttctatcgc cttcttgacg 4989
agttcttctg agcgggactc tggggttcga aatgaccgac caagcgacgc ccaacctgcc 5049
atcacgagat ttcgattcca ccgccgcct ctatgaaagg ttgggcttcg gaatcgtttt 5109
ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt tcttcgcccc 5169
ccccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt 5229
cacaataaaa gcatTTTTTT cactgcattc tagttgtggt ttgtccaaac tcatcaatgt 5289
atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata 5349
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 5409
cataaagtgt aaagcctggg gtgcctaatt agtgagctaa ctcacattaa ttgcgttgcg 5469
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcgcca 5529
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttctctgc tctactgactc 5589
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtataacg 5649
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 5709
ggccaggaac cgtaaaaagg ccgcgttgct ggcgTTTTT cataggctcc gccccctga 5769
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 5829
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 5889
taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgctttctc aatgctcacg 5949
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 6009
ccccgttcag cccgaccgct gcgccttctc cggtaactat cgtcttgagt ccaaccgggt 6069
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 6129
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac 6189
agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 6249
ttgatccggc aaacaaacca ccgctggtag cgggtgtttt tttgtttgca agcagcagat 6309
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc tttctacgg ggtctgacgc 6369

```



```

tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 6429
cacctagatc cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta 6489
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 6549
atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg 6609
cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggtccaga 6669
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggc ctgcaacttt 6729
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt 6789
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt 6849
tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatccccc 6909
gttgtgcaaa aaagcgggta gtccttcgg tctccgacg gttgtcagaa gtaagttggc 6969
cgcagtgtta tcaactcatg ttatggcagc actgcataat tctcttactg tcatgccatc 7029
cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 7089
gcggcgaccg agttgctctt gcccgcgctc aatacgggat aataccgcg cacatagcag 7149
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 7209
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 7269
ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa 7329
gggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg 7389
aagcatttat cagggttatt gtctcatgag cggatacata ttggaatgta tttagaaaaa 7449
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg t 7500

```

<210> 11

<211> 697

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<223> pCDJE 2-7

<400> 11

```

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

```

100	105	110
Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val Ser Val Gln Thr		
115	120	125
His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala Trp Leu Asp Ser		
130	135	140
Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn Trp Ile Ile Arg		
145	150	155
Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly Trp Met Leu Gly		
165	170	175
Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu Leu Leu Leu Val		
180	185	190
Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly Asn Arg Asp Phe		
195	200	205
Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly		
210	215	220
Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro Thr Leu Asp Val		
225	230	235
Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu Val Arg Ser Tyr		
245	250	255
Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val Ala Arg Cys Pro		
260	265	270
Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp Ser Ser Tyr Val		
275	280	285
Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu		
290	295	300
Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe Ser Cys Thr Ser		
305	310	315
Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile Lys Tyr Glu Val		
325	330	335
Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn His Gly Asn Tyr		
340	345	350
Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe Thr Val Thr Pro		
355	360	365
Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr Gly Glu Val Thr		
370	375	380
Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu Ala Phe Tyr Val		
385	390	395
Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg Glu Trp Phe His		
405	410	415
Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr Ala Trp Arg Asn		
420	425	430
Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala Thr Lys Gln Ser		
435	440	445
Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His Gln Ala Leu Ala		

```

      450              455              460
Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys Leu Thr Ser Gly
465              470              475              480
His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala Leu Lys Gly Thr
      485              490              495
Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala
      500              505              510
Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser
      515              520              525
Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala Ser Leu Asn Asp
      530              535              540
Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr
545              550              555              560
Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly
      565              570              575
Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His
      580              585              590
Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe Ser Thr Thr Leu
      595              600              605
Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe
      610              615              620
Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln
625              630              635              640
Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile
      645              650              655
Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met Gly Val Asn Ala
      660              665              670
Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr Gly Gly Val Leu
      675              680              685
Val Phe Leu Ala Thr Asn Val His Ala
      690              695

```

<210> 12

<211> 46

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-46

<223> WN 466

<400> 12

cttgggtaccc gtctcggcgc cgtgaccctc tcgaacttcc agggca

46

<210> 13

<211> 43

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-43

<223> CWN2444

<400> 13

agaggcactt gcacgtgcgg acttccgccg gcgaaaaaga aaa

43

<210> 14

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<223> JE 信号

<400> 14

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala

1

5

10

15

Val Val Ile Ala Cys Ala Gly Ala

20

<210> 15

<211> 5308

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> CDS

<222> (911)... (2987)

<221> misc_特征

<222> (1)... (5308)

<223> pCBWN

<400> 15

```

gacggatcgg gagatctccc gatccccctat ggtgcactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420
attgacgtca atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt      480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgcccagta catgacctta tgggactttc ctacttgcca gtacatctac gtattagtca      600
tcgctattac catggtgatg cggtttttgg cagtacatca atgggcgtgg atagcggttt      660
gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac      720
caaaatcaac gggactttcc aaaatgtcgt aacaactccg cccattgac gcaaatgggc      780
ggtaggcgtg tacggtggga ggtctatata agcagagctc tctggctaac tagagaaccc      840
actgcttact ggcttatcga aattaatacg actcactata gggagaccca agcttggtac      900
cgccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg      949
      Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala
              1              5              10

```

```

agc ttg gca gtt gtc ata gct tgt gca ggc gcc gtg acc ctc tcg aac      997
Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn
      15              20              25

```

```

ttc cag ggc aag gtg atg atg acg gta aat gct act gac gtc aca gat      1045
Phe Gln Gly Lys Val Met Met Thr Val Asn Ala Thr Asp Val Thr Asp
      30              35              40              45

```

```

gtc atc acg att cca aca gct gct gga aag aac cta tgc att gtc aga      1093
Val Ile Thr Ile Pro Thr Ala Ala Gly Lys Asn Leu Cys Ile Val Arg
              50              55              60

```

```

gca atg gat gtg gga tac atg tgc gat gat act atc act tat gaa tgc      1141
Ala Met Asp Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys
              65              70              75

```

```

cca gtg ctg tcg gct ggt aat gat cca gaa gac atc gac tgt tgg tgc      1189
Pro Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys
              80              85              90

```

aca aag tca gca gtc tac gtc agg tat gga aga tgc acc aag aca cgc	1237
Thr Lys Ser Ala Val Tyr Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg	
95 100 105	
cac tca aga cgc agt cgg agg tca ctg aca gtg cag aca cac gga gaa	1285
His Ser Arg Arg Ser Arg Arg Ser Leu Thr Val Gln Thr His Gly Glu	
110 115 120 125	
agc act cta gcg aac aag aag ggg gct tgg atg gac agc acc aag gcc	1333
Ser Thr Leu Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala	
130 135 140	
aca agg tat ttg gta aaa aca gaa tca tgg atc ttg agg aac cct gga	1381
Thr Arg Tyr Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly	
145 150 155	
tat gcc ctg gtg gca gcc gtc att ggt tgg atg ctt ggg agc aac acc	1429
Tyr Ala Leu Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr	
160 165 170	
atg cag aga gtt gtg ttt gtc gtg cta ttg ctt ttg gtg gcc cca gct	1477
Met Gln Arg Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala	
175 180 185	
tac agc ttc aac tgc ctt gga atg agc aac aga gac ttc ttg gaa gga	1525
Tyr Ser Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly	
190 195 200 205	
gtg tct gga gca aca tgg gtg gat ttg gtt ctc gaa ggc gac agc tgc	1573
Val Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys	
210 215 220	
gtg act atc atg tct aag gac aag cct acc atc gat gtg aag atg atg	1621
Val Thr Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met	
225 230 235	
aat atg gag gcg gcc aac ctg gca gag gtc cgc agt tat tgc tat ttg	1669
Asn Met Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu	
240 245 250	
gct acc gtc agc gat ctc tcc acc aaa gct gcg tgc ccg acc atg gga	1717
Ala Thr Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly	
255 260 265	

gaa gct cac aat gac aaa cgt gct gac cca gct ttt gtg tgc aga caa	1765
Glu Ala His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln	
270 275 280 285	
gga gtg gtg gac agg ggc tgg ggc aac ggc tgc gga cta ttt ggc aaa	1813
Gly Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys	
290 295 300	
gga agc att gac aca tgc gcc aaa ttt gcc tgc tct acc aag gca ata	1861
Gly Ser Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile	
305 310 315	
gga aga acc atc ttg aaa gag aat atc aag tac gaa gtg gcc att ttt	1909
Gly Arg Thr Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe	
320 325 330	
gtc cat gga cca act act gtg gag tcg cac gga aac tac tcc aca cag	1957
Val His Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln	
335 340 345	
gtt gga gcc act cag gca ggg aga ttc agc atc act cct gcg gcg cct	2005
Val Gly Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro	
350 355 360 365	
tca tac aca cta aag ctt gga gaa tat gga gag gtg aca gtg gac tgt	2053
Ser Tyr Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys	
370 375 380	
gaa cca cgg tca ggg att gac acc aat gca tac tac gtg atg act gtt	2101
Glu Pro Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val	
385 390 395	
gga aca aag acg ttc ttg gtc cat cgt gag tgg ttc atg gac ctc aac	2149
Gly Thr Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn	
400 405 410	
ctc cct tgg agc agt gct gga agt act gtg tgg agg aac aga gag acg	2197
Leu Pro Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr	
415 420 425	
tta atg gag ttt gag gaa cca cac gcc acg aag cag tct gtg ata gca	2245
Leu Met Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala	
430 435 440 445	

ttg ggc tca caa gag gga gct ctg cat caa gct ttg gct gga gcc att Leu Gly Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile 450 455 460	2293
cct gtg gaa ttt tca agc aac act gtc aag ttg acg tcg ggt cat ttg Pro Val Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu 465 470 475	2341
aag tgt aga gtg aag atg gaa aaa ttg cag ttg aag gga aca acc tat Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr 480 485 490	2389
ggc gtc tgt tca aag gct ttc aag ttt ctt ggg act ccc gcg gac aca Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr 495 500 505	2437
ggt cac ggc act gtg gtg ttg gaa ttg cag tac act ggc acg gat gga Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly 510 515 520 525	2485
cct tgc aaa gtt cct atc tcg tca gtg gct tca ttg aac gac cta acg Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr 530 535 540	2533
cca gtg ggc aga ttg gtc act gtc aac cct ttt gtt tca gtg gcc acg Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr 545 550 555	2581
gcc aac gct aag gtc ctg att gaa ttg gaa cca ccc ttt gga gac tca Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser 560 565 570	2629
tac ata gtg gtg ggc aga gga gaa caa cag atc aat cac cat tgg cac Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His 575 580 585	2677
aag tct gga agc agc att ggc aaa gcc ttt aca acc acc ctc aaa gga Lys Ser Gly Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly 590 595 600 605	2725
gcg cag aga cta gcc gct cta gga gac aca gct tgg gac ttt gga tca Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser 610 615 620	2773

ggt gga ggg gtg ttc acc tca gtt ggg aag gct gtc cat caa gtg ttc	2821
Val Gly Gly Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe	
625 630 635	
gga gga gca ttc cgc tca ctg ttc gga ggc atg tcc tgg ata acg caa	2869
Gly Gly Ala Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln	
640 645 650	
gga ttg ctg ggg gct ctc ctg ttg tgg atg ggc atc aat gct cgt gat	2917
Gly Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp	
655 660 665	
agg tcc ata gct ctc acg ttt ctc gca gtt gga gga gtt ctg ctc ttc	2965
Arg Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe	
670 675 680 685	
ctc tcc gtg aac gtg cac gcc t gaaggcggcc gctcgagcat gcatctagag	3017
Leu Ser Val Asn Val His Ala	
690	
ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct	3077
tctagttgcc agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaaggt	3137
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgattg tctgagtagg	3197
tgctattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac	3257
aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgga aagaaccagc	3317
tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg	3377
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc	3437
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt	3497
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc	3557
ataggtcccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa	3617
acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc	3677
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg	3737
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcggt cgctccaagc	3797
tgggctgtgt gcacgaaccc ccggttcagc ccgaccgctg cgccttatcc ggtaactatc	3857
gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca	3917
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact	3977
acggctacac tagaagaaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg	4037
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt	4097
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct	4157
tttctacggg gtctgacgct cagtgaacg aaaactcacg ttaagggtatt ttggtcatga	4217
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa	4277
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac	4337
ctatctcagc gatctgtcta tttcggtcat ccatagttgc ctgactcccc gtcgtgtaga	4397

```

taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 4457
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 4517
gaagtgggcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 4577
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 4637
tggtgtcacg ctctcgtttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 4697
gagttacatg atcccccatt ttgtgcaaaa aagcggtttag ctcttcggt cctccgatcg 4757
ttgtcagaag taagtggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 4817
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 4877
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 4937
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 4997
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 5057
ccaactgate ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 5117
ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 5177
tcctttttca atattattga agcatttata agggttattg tctcatgagc ggatacatat 5237
ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 5297
cacctgacgt c 5308

```

<210> 16

<211> 692

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述: 备注=合成构建体

<223> pCBWN

<400> 16

```

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

```

Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala Thr Arg Tyr
 130 135 140
 Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly Tyr Ala Leu
 145 150 155 160
 Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr Met Gln Arg
 165 170 175
 Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe
 180 185 190
 Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly Val Ser Gly
 195 200 205
 Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val Thr Ile
 210 215 220
 Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn Met Glu
 225 230 235 240
 Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr Val
 245 250 255
 Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala His
 260 265 270
 Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val Val
 275 280 285
 Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile
 290 295 300
 Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg Thr
 305 310 315 320
 Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly
 325 330 335
 Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val Gly Ala
 340 345 350
 Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr Thr
 355 360 365
 Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro Arg
 370 375 380
 Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr Lys
 385 390 395 400
 Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro Trp
 405 410 415
 Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met Glu
 420 425 430
 Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly Ser
 435 440 445
 Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val Glu
 450 455 460
 Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg
 465 470 475 480

```

Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys
      485                      490                      495
Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly
      500                      505                      510
Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys
      515                      520                      525
Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly
      530                      535                      540
Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala
      545                      550                      555                      560
Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
      565                      570                      575
Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly
      580                      585                      590
Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln Arg
      595                      600                      605
Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly Gly
      610                      615                      620
Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly Ala
      625                      630                      635                      640
Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu
      645                      650                      655
Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile
      660                      665                      670
Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val
      675                      680                      685
Asn Val His Ala
      690

```

<210> 17

<211> 5334

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> CDS

<222> (916)... (3007)

<221> misc_特征

<222> (1)... (5334)

<223> pCBE 1-14

<400> 17

```

gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
ccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420
attgacgtca atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt      480
atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgcccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca      600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg      660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc      720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg      780
gtaggcgtgt acgggtggag gtctatataa gcagagctct ctggctaact agagaacca      840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttggtacc      900
tctagagccg ccgcc atg ggc aga aag caa aac aaa aga gga gga aat gaa      951
      Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu
              1              5              10

```

```

ggc tca atc atg tgg ctc gcg agc ttg gca gtt gtc ata gct tgt gcg      999
Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala
      15              20              25

```

```

gga gcc atg aag ttg tcg aat ttc cag ggg aag ctt ttg atg acc atc      1047
Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile
      30              35              40

```

```

aac aac acg gac att gca gac gtt atc gtg att ccc acc tca aaa gga      1095
Asn Asn Thr Asp Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly
      45              50              55              60

```

```

gag aac aga tgc tgg gtc cgg gca atc gac gtc ggc tac atg tgt gag      1143
Glu Asn Arg Cys Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu
      65              70              75

```

```

gac act atc acg tac gaa tgt cct aag ctt acc atg ggc aat gat cca      1191
Asp Thr Ile Thr Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro
      80              85              90

```

```

gag gat gtg gat tgc tgg tgt gac aac caa gaa gtc tac gtc caa tat      1239
Glu Asp Val Asp Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr
      95              100              105

```

gga cgg tgc acg cgg acc agg cat tcc aag cga agc agg aga tcc gtg	1287
Gly Arg Cys Thr Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val	
110 115 120	
tcg gtc caa aca cat ggg gag agt tca cta gtg aat aaa aaa gag gct	1335
Ser Val Gln Thr His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala	
125 130 135 140	
tgg ctg gat tca acg aaa gcc aca cga tat ctc atg aaa act gag aac	1383
Trp Leu Asp Ser Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn	
145 150 155	
tgg atc ata agg aat cct ggc tat gct ttc ctg gcg gcg gta ctt ggc	1431
Trp Ile Ile Arg Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly	
160 165 170	
tgg atg ctt ggc agt aac aac ggt caa cgc gtg gta ttt acc atc ctc	1479
Trp Met Leu Gly Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu	
175 180 185	
ctg ctg ttg gtc gct ccg gct tac agt ttt aat tgt ctg gga atg ggc	1527
Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly	
190 195 200	
aat cgt gac ttc ata gaa gga gcc agt gga gcc act tgg gtg gac ttg	1575
Asn Arg Asp Phe Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu	
205 210 215 220	
gtg ctg gaa gga gat agc tgc ttg aca atc atg gca aac gac aaa cca	1623
Val Leu Glu Gly Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro	
225 230 235	
aca ttg gac gtc cgc atg att aac atc gaa gct agc caa ctt gct gag	1671
Thr Leu Asp Val Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu	
240 245 250	
gtc aga agt tac tgc tat cat gct tca gtc act gac atc tcg acg gtg	1719
Val Arg Ser Tyr Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val	
255 260 265	
gct cgg tgc ccc acg act gga gaa gcc cac aac gag aag cga gct gat	1767
Ala Arg Cys Pro Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp	
270 275 280	

aca aaa cag tcc gtt gtt gct ctt ggg tca cag gaa gga ggc ctc cat 2295
Thr Lys Gln Ser Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His
445 450 455 460

cag gcg ttg gca gga gcc atc gtg gtg gag tac tca agc tca gtg aag	2343
Gln Ala Leu Ala Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys	
465 470 475	
tta aca tca ggc cac ctg aaa tgt agg ctg aaa atg gac aaa ctg gct	2391
Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala	
480 485 490	
ctg aaa ggc aca acc tat ggc atg tgt aca gaa aaa ttc tcg ttc gcg	2439
Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala	
495 500 505	
aaa aat ccg gcg gac act ggt cac gga aca gtt gtc att gaa ctc tcc	2487
Lys Asn Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser	
510 515 520	
tac tct ggg agt gat ggc ccc tgc aaa att ccg att gct tcc gtt gcg	2535
Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala	
525 530 535 540	
agc ctc aat gac atg acc ccc gtt ggg cgg ctg gtg aca gtg aac ccc	2583
Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro	
545 550 555	
ttc gtc gcg act tcc agt gcc agc tca aag gtg ctg gtc gag atg gaa	2631
Phe Val Ala Thr Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu	
560 565 570	
ccc ccc ttc gga gac tcc tac atc gta gtt gga agg gga gac aag cag	2679
Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln	
575 580 585	
atc aac cac cat tgg cac aaa gct gga agc acg ctg ggc aag gcc ttt	2727
Ile Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe	
590 595 600	
tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc gac aca	2775
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr	
605 610 615 620	
gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata gga aaa	2823
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys	
625 630 635	


```

gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt ggg gga      2871
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
      640                      645                      650

atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc tgg atg      2919
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
      655                      660                      665

ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta gcc aca      2967
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
      670                      675                      680

ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct t aattagtttg      3017
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
      685                      690                      695

agcggccgct cgagcatgca tctagagggc cctattctat agtgtcacct aaatgctaga      3077
gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgtttg ttgccctcc      3137
cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag      3197
gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtg ggtggggcag      3257
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct      3317
atggcttctg aggcggaaag aaccagctgc attaatgaat cggccaacgc gcggggagag      3377
gcggtttgcg tattgggcgc tcttcgcctt cctcgctcac tgactcgctg cgctcggtcg      3437
ttcggtcgcg gcgagcggtg tcagctcact caaaggcggg aatacggtta tccacagaat      3497
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta      3557
aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa      3617
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc      3677
cccttggaag ctccctcgtg cgctctcctg ttccgacctt gccgcttacc ggatacctgt      3737
ccgcctttct cccttcggga agegtggcgc tttctcatag ctcacgctgt aggtatctca      3797
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc      3857
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtgaaga cagcacttat      3917
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta      3977
cagagttctt gaagtgggtg cctaactacg gctacactag aagaacagta tttggtatct      4037
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac      4097
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa      4157
aaggatctca agaagatcct ttgatctttt ctacgggggc tgacgctcag tggaacgaaa      4217
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt      4277
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggctcgaca      4337
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca      4397
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc      4457
ccagtgtcgc aatgataccg cgagaccac gctcaccggc tccagattta tcagcaataa      4517
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc      4577
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca      4637

```

```

acgttggtgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 4697
tcagctccgg ttcccaacga tcaaggcgag ttacatgata ccccatgttg tgcaaaaaag 4757
cggtagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 4817
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 4877
ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt 4937
gctcttggcc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 4997
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 5057
ccagttcgat gtaaccactc cgtgcaccca actgatcttc agcatctttt actttcacca 5117
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca 5177
cacgaaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 5237
gttattgtct catgagcgga tacatatattg aatgtattta gaaaaataaa caaatagggg 5297
ttccgcgcac atttccccga aaagtccac ctgacgt 5334

```

<210> 18

<211> 697

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述: 备注=合成构建体

<223> pCBE 1-14

<400> 18

```

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

```

Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly Trp Met Leu Gly
 165 170 175
 Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu Leu Leu Leu Val
 180 185 190
 Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly Asn Arg Asp Phe
 195 200 205
 Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly
 210 215 220
 Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro Thr Leu Asp Val
 225 230 235 240
 Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu Val Arg Ser Tyr
 245 250 255
 Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val Ala Arg Cys Pro
 260 265 270
 Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp Ser Ser Tyr Val
 275 280 285
 Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu
 290 295 300
 Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe Ser Cys Thr Ser
 305 310 315 320
 Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile Lys Tyr Glu Val
 325 330 335
 Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn His Gly Asn Tyr
 340 345 350
 Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe Thr Val Thr Pro
 355 360 365
 Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr Gly Glu Val Thr
 370 375 380
 Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu Ala Phe Tyr Val
 385 390 395 400
 Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg Glu Trp Phe His
 405 410 415
 Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr Ala Trp Arg Asn
 420 425 430
 Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala Thr Lys Gln Ser
 435 440 445
 Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His Gln Ala Leu Ala
 450 455 460
 Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys Leu Thr Ser Gly
 465 470 475 480
 His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala Leu Lys Gly Thr
 485 490 495
 Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala
 500 505 510

```

Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser
    515                      520                      525
Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala Ser Leu Asn Asp
    530                      535                      540
Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr
    545                      550                      555                      560
Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly
                      565                      570                      575
Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His
    580                      585                      590
Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe Ser Thr Thr Leu
    595                      600                      605
Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe
    610                      615                      620
Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln
    625                      630                      635                      640
Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile
                      645                      650                      655
Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met Gly Val Asn Ala
                      660                      665                      670
Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr Gly Gly Val Leu
                      675                      680                      685
Val Phe Leu Ala Thr Asn Val His Ala
    690                      695

```

<210> 19

<211> 5283

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> CDS

<222> (910)... (2965)

<400> 19

```

gacggatcgg gagatctccc gatccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gtagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttaggggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata      300
tgaggttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420

```

attgacgtca atgggtggac tatttacggg aaactgcccc cttggcagta catcaagtgt	480
atcatatgcc aagtagcccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
tcgctattac catgggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg	780
gtaggcgtgt acgggtgggag gtctatataa gcagagctct ctggctaact agagaaccca	840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttgggtacc	900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc	951
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser	
1 5 10	
ttg gca gtt gtc ata gct ggt aca agc gct acc acc atc cac cgg gac	999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Thr Thr Ile His Arg Asp	
15 20 25 30	
agg gaa gga tac atg gtt atg cgg gcc agt gga agg gac gct gca agc	1047
Arg Glu Gly Tyr Met Val Met Arg Ala Ser Gly Arg Asp Ala Ala Ser	
35 40 45	
cag gtc agg gta caa aac gga acg tgc gtc atc ctg gca aca gac atg	1095
Gln Val Arg Val Gln Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met	
50 55 60	
gga gag tgg tgt gaa gat tca atc acc tac tct tgc gtc acg att gac	1143
Gly Glu Trp Cys Glu Asp Ser Ile Thr Tyr Ser Cys Val Thr Ile Asp	
65 70 75	
cag gag gaa gaa ccc gtt gac gtg gac tgc ttc tgc cga ggt gtt gat	1191
Gln Glu Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Gly Val Asp	
80 85 90	
agg gtt aag tta gag tat gga cgc tgt gga agg caa gct gga tct agg	1239
Arg Val Lys Leu Glu Tyr Gly Arg Cys Gly Arg Gln Ala Gly Ser Arg	
95 100 105 110	
ggg aaa agg tct gtg gtc att cca aca cat gca caa aaa gac atg gtc	1287
Gly Lys Arg Ser Val Val Ile Pro Thr His Ala Gln Lys Asp Met Val	
115 120 125	
ggg cga ggt cat gca tgg ctt aaa ggt gac aat att cga gat cat gtc	1335
Gly Arg Gly His Ala Trp Leu Lys Gly Asp Asn Ile Arg Asp His Val	
130 135 140	

acc cga gtc gag ggc tgg atg tgg aag aac aag ctt cta act gcc gcc	1383
Thr Arg Val Glu Gly Trp Met Trp Lys Asn Lys Leu Leu Thr Ala Ala	
145 150 155	
att gtg gcc ttg gct tgg ctc atg gtt gat agt tgg atg gcc aga gtg	1431
Ile Val Ala Leu Ala Trp Leu Met Val Asp Ser Trp Met Ala Arg Val	
160 165 170	
act gtc atc ctc ttg gcg ttg agt cta ggg cca gtg tac gcc acg agg	1479
Thr Val Ile Leu Leu Ala Leu Ser Leu Gly Pro Val Tyr Ala Thr Arg	
175 180 185 190	
tgc acg cat ctt gag aac aga gat ttt gtg aca gga act caa ggg acc	1527
Cys Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln Gly Thr	
195 200 205	
acc aga gtg tcc cta gtt ttg gaa ctt gga ggc tgc gtg acc atc aca	1575
Thr Arg Val Ser Leu Val Leu Glu Leu Gly Gly Cys Val Thr Ile Thr	
210 215 220	
gct gag ggc aag cca tcc att gat gta tgg ctc gaa gac att ttt cag	1623
Ala Glu Gly Lys Pro Ser Ile Asp Val Trp Leu Glu Asp Ile Phe Gln	
225 230 235	
gaa agc ccg gct gaa acc aga gaa tac tgc ctg cac gcc aaa ttg acc	1671
Glu Ser Pro Ala Glu Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Thr	
240 245 250	
aac aca aaa gtg gag gct cgc tgt cca acc act gga ccg gcg aca ctt	1719
Asn Thr Lys Val Glu Ala Arg Cys Pro Thr Thr Gly Pro Ala Thr Leu	
255 260 265 270	
ccg gag gag cat cag gct aat atg gtg tgc aag aga gac caa agc gac	1767
Pro Glu Glu His Gln Ala Asn Met Val Cys Lys Arg Asp Gln Ser Asp	
275 280 285	
cgt gga tgg gga aac cac tgc ggg ttt ttt ggg aag ggc agt ata gtg	1815
Arg Gly Trp Gly Asn His Cys Gly Phe Phe Gly Lys Gly Ser Ile Val	
290 295 300	
gct tgt gca aag ttt gaa tgc gag gaa gca aaa aaa gct gtg ggc cac	1863
Ala Cys Ala Lys Phe Glu Cys Glu Glu Ala Lys Lys Ala Val Gly His	
305 310 315	

gtc tat gac tcc aca aag atc acg tat gtt gtc aag gtt gag ccc cac Val Tyr Asp Ser Thr Lys Ile Thr Tyr Val Val Lys Val Glu Pro His 320 325 330	1911
aca ggg gat tac ttg gct gca aat gag acc aat tca aac agg aaa tca Thr Gly Asp Tyr Leu Ala Ala Asn Glu Thr Asn Ser Asn Arg Lys Ser 335 340 345 350	1959
gca cag ttt acg gtg gca tcc gag aaa gtg atc ctg cgg ctc ggc gac Ala Gln Phe Thr Val Ala Ser Glu Lys Val Ile Leu Arg Leu Gly Asp 355 360 365	2007
tat gga gat gtg tcg ctg acg tgt aaa gtg gca agt ggg att gat gtc Tyr Gly Asp Val Ser Leu Thr Cys Lys Val Ala Ser Gly Ile Asp Val 370 375 380	2055
gcc caa act gtg gtg atg tca ctc gac agc agc aag gac cac ctg cct Ala Gln Thr Val Val Met Ser Leu Asp Ser Ser Lys Asp His Leu Pro 385 390 395	2103
tct gca tgg caa gtg cac cgt gac tgg ttt gag gac ttg gcg ctg ccc Ser Ala Trp Gln Val His Arg Asp Trp Phe Glu Asp Leu Ala Leu Pro 400 405 410	2151
tgg aaa cac aag gac aac caa gat tgg aac agt gtg gag aaa ctt gtg Trp Lys His Lys Asp Asn Gln Asp Trp Asn Ser Val Glu Lys Leu Val 415 420 425 430	2199
gaa ttt gga cca cca cat gct gtg aaa atg gat gtt ttc aat ctg ggg Glu Phe Gly Pro Pro His Ala Val Lys Met Asp Val Phe Asn Leu Gly 435 440 445	2247
gac cag acg gct gtg ctg ctc aaa tca ctg gca gga gtt ccg ctg gcc Asp Gln Thr Ala Val Leu Leu Lys Ser Leu Ala Gly Val Pro Leu Ala 450 455 460	2295
agt gtg gag ggc cag aaa tac cac ctg aaa agc ggc cat gtt act tgt Ser Val Glu Gly Gln Lys Tyr His Leu Lys Ser Gly His Val Thr Cys 465 470 475	2343
gat gtg gga ctg gaa aag ctg aaa ctg aaa ggc aca acc tac tcc atg Asp Val Gly Leu Glu Lys Leu Lys Leu Lys Gly Thr Thr Tyr Ser Met 480 485 490	2391

tgt gac aaa gca aag ttc aaa tgg aag aga gtt cct gtg gac agc ggc	2439
Cys Asp Lys Ala Lys Phe Lys Trp Lys Arg Val Pro Val Asp Ser Gly	
495 500 505 510	
cat gac aca gta gtc atg gag gta tca tac aca gga agc gac aag cca	2487
His Asp Thr Val Val Met Glu Val Ser Tyr Thr Gly Ser Asp Lys Pro	
515 520 525	
tgt cgg atc ccg gtg cgg gct gtg gca cat ggt gtc cca gcg gtt aat	2535
Cys Arg Ile Pro Val Arg Ala Val Ala His Gly Val Pro Ala Val Asn	
530 535 540	
gta gcc atg ctc ata acc ccc aat cca acc att gaa aca aat ggt ggc	2583
Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile Glu Thr Asn Gly Gly	
545 550 555	
gga ttc ata gaa atg cag ctg cca cca ggg gat aac atc atc tat gtg	2631
Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val	
560 565 570	
gga gac ctt agc cag cag tgg ttt cag aaa ggc agt acc att ggt aga	2679
Gly Asp Leu Ser Gln Gln Trp Phe Gln Lys Gly Ser Thr Ile Gly Arg	
575 580 585 590	
atg ttt gaa aaa acc cgc agg gga ttg gaa agg ctc tct gtg gtt gga	2727
Met Phe Glu Lys Thr Arg Arg Gly Leu Glu Arg Leu Ser Val Val Gly	
595 600 605	
gaa cat gca tgg gac ttt ggc tca gta ggc ggg gta ctg tct tct gtg	2775
Glu His Ala Trp Asp Phe Gly Ser Val Gly Gly Val Leu Ser Ser Val	
610 615 620	
ggg aag gca atc cac acg gtg ctg ggg gga gct ttc aac acc ctt ttt	2823
Gly Lys Ala Ile His Thr Val Leu Gly Gly Ala Phe Asn Thr Leu Phe	
625 630 635	
ggg ggg gtt gga ttc atc cct aag atg ctg ctg ggg gtt gct ctg gtc	2871
Gly Gly Val Gly Phe Ile Pro Lys Met Leu Leu Gly Val Ala Leu Val	
640 645 650	
tgg ttg gga cta aat gcc agg aat cca acg atg tcc atg acg ttt ctt	2919
Trp Leu Gly Leu Asn Ala Arg Asn Pro Thr Met Ser Met Thr Phe Leu	
655 660 665 670	

gct gtg ggg gct ttg aca ctg atg atg aca atg gga gtt ggg gca t	2965
Ala Val Gly Ala Leu Thr Leu Met Met Thr Met Gly Val Gly Ala	
675 680 685	
gagcggccgc tcgagcatgc atctagaggg ccctattcta tagtgcacc taaatgctag	3025
agctcgtga tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttcccctc	3085
ccccgtgcct tccttgacct tggaaggtgc cactcccact gtcctttcct aataaaatga	3145
ggaaattgca tcgattgtc tgagtaggtg tcattctatt ctgggggggtg gggtagggca	3205
ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtgggctc	3265
tatggcttct gaggcggaaa gaacagctgc attaatgaat cggccaacgc gcggggagag	3325
gcggtttgcg tattgggcgc tcttcgcctt cctcgctcac tgactcgtg cgtcggctc	3385
ttcggtcgcg gcgagcggta tcagctcact caaaggcggg aatacggta tccacagaat	3445
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta	3505
aaaaggccgc gttgctggcg tttttccata ggctccgcc ccctgacgag catcacaaaa	3565
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	3625
cccttggaag ctccctcgtg cgtctcctg ttcgaccct gccgcttacc ggatacctgt	3685
ccgcctttct cccttcggga agcgtggcg tttctcaatg ctcacgctgt aggtatctca	3745
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc	3805
accgctgcgc cttatccggg aactatcgtc ttgagtccaa cccggttaaga cagcacttat	3865
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta	3925
cagagtctct gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct	3985
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac	4045
aaaccaccgc tggtagcggg ggtttttttg ttgcaagca gcagattacg cgcagaaaaa	4105
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa	4165
actcacgta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt	4225
taaattaaaa atgaagtttt aatcaatct aaagtatata tgagtaaact tggctctgaca	4285
gttaccaatg cttaatcagt gaggcaccta tctcagcgt ctgtctatct cgttcattca	4345
tagttgctg actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc	4405
ccagtgtgc aatgataccg cgagaccac gtcaccggc tccagattta tcagcaataa	4465
accagccagc cggaagggcc gagcgagaa gtggtcctgc aactttatcc gcctccatcc	4525
agtctattaa ttgttgccg gaagctagag taagtagttc gccagttaat agtttgcgca	4585
acgttggtgc cattgtaca ggcacgtg tgacacgctc gtcgttttgt atggcttcat	4645
tcagctccgg ttcccaacga tcaaggcgag ttacatgat ccccatgttg tgcaaaaaag	4705
cggtagctc cttcggctct ccgatcgtg tcagaagtaa gttggccgca gtgttatcac	4765
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt	4825
ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt	4885
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc	4945
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat	5005
ccagttcgat gtaacccact cgtgcacca actgatcttc agcatctttt actttacca	5065
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca	5125
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg	5185
gttattgtct catgagcgga tacatatgtt aatgtattta gaaaaataaa caaatagggg	5245
ttccgcgcac atttccccga aaagtccac ctgacgtc	5283

<210> 20

<211> 681

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 20

```

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1           5           10           15
Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys Asn Arg
      20           25           30
Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr Phe Ser
      35           40           45
Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys Tyr Trp
      50           55           60
Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro Arg Glu
65           70           75           80
Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn Val Arg
      85           90           95
Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg Ser Arg
      100          105          110
Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys Thr Arg
      115          120          125
Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu Gln Lys
      130          135          140
Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr Ala Leu
145          150          155          160
Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val Val Ile
      165          170          175
Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His Cys Ile
      180          185          190
Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly Thr Trp
      195          200          205
Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met Ala Pro
      210          215          220
Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile Asp Arg
225          230          235          240
Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr His Val
      245          250          255
Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu Ala Glu
      260          265          270
Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp Arg Gly

```

275	280	285
Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys		
290	295	300
Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val Asp Gln		
305	310	315
Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly Ala Lys		
325	330	335
Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp Ala Leu		
340	345	350
Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala Thr Leu		
355	360	365
Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr Ile Ala		
370	375	380
Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala Gln Asp		
385	390	395
Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg Glu Met		
405	410	415
His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile Arg Val		
420	425	430
Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu Thr Gly		
435	440	445
Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr Lys Leu		
450	455	460
His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu Thr Leu		
465	470	475
Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe Val Lys		
485	490	495
Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val		
500	505	510
Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp Asp Leu		
515	520	525
Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala		
530	535	540
Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly		
545	550	555
Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr Tyr Gln		
565	570	575
Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln Thr Met		
580	585	590
Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp Asp Phe		
595	600	605
Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile His Thr		
610	615	620
Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn Trp Ile		

625	630	635	640
Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr			
	645	650	655
Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met			
	660	665	670
Met Phe Leu Ser Leu Gly Val Gly Ala			
	675	680	

<210> 21

<211> 5304

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> CDS

<222> (910)... (2986)

<400> 21

gacggatcgg gagatctccc gatccccctat ggtcgactct cagtacaatc tgctctgatg	60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggctgct gagtagtgcg	120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc	180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt	240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata	300
tggagttccg cgttacataa cttacggtaa atggccccgcc tggctgaccg cccaacgacc	360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
attgacgtca atgggtggac tatttacggt aaactgcca cttggcagta catcaagtgt	480
atcatatgcc aagtagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca	600
tcgtattac catggtgatg cggttttgga agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatggcg	780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca	840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttggtacc	900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc	951

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser

1

5

10

ttg gca gtt gtc ata gct ggt aca agc gct ttg cag tta tca acc tat	999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Leu Gln Leu Ser Thr Tyr	
15	20
	25
	30

cag ggg aaa gtg tta atg tca atc aac aag act gac gct caa agc gcc	1047
---	------

Gln Gly Lys Val Leu Met Ser Ile Asn Lys Thr Asp Ala Gln Ser Ala	
35 40 45	
ata aac att cct agt gcc aac gga gca aac act tgc att gtg agg gct	1095
Ile Asn Ile Pro Ser Ala Asn Gly Ala Asn Thr Cys Ile Val Arg Ala	
50 55 60	
cta gat gtg ggg gtc atg tgc aaa gat gac atc aca tac ctg tgc cca	1143
Leu Asp Val Gly Val Met Cys Lys Asp Asp Ile Thr Tyr Leu Cys Pro	
65 70 75	
gtg ctt tca gcg gga aat gat ccc gag gac att gac tgt tgg tgt gac	1191
Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Asp	
80 85 90	
gtc gaa gag gtg tgg gtg cac tac ggc aga tgc acg cgc atg gga cat	1239
Val Glu Glu Val Trp Val His Tyr Gly Arg Cys Thr Arg Met Gly His	
95 100 105 110	
tcg agg cgt agc cga cgg tca atc tct gtg cag cat cat gga gat tcc	1287
Ser Arg Arg Ser Arg Arg Ser Ile Ser Val Gln His His Gly Asp Ser	
115 120 125	
aca ctg gca aca aag aac acg cca tgg ttg gac acc gtg aaa acc acc	1335
Thr Leu Ala Thr Lys Asn Thr Pro Trp Leu Asp Thr Val Lys Thr Thr	
130 135 140	
aaa tac ttg aca aaa gta gaa aac tgg gtt ttg cgc aat cct gga tat	1383
Lys Tyr Leu Thr Lys Val Glu Asn Trp Val Leu Arg Asn Pro Gly Tyr	
145 150 155	
gcc cta gtt gcg ctg gcg att gga tgg atg ctc ggt agc aac aac aca	1431
Ala Leu Val Ala Leu Ala Ile Gly Trp Met Leu Gly Ser Asn Asn Thr	
160 165 170	
cag aga gtg gtt ttt gtg atc atg ctg atg ctg att gct ccg gca tac	1479
Gln Arg Val Val Phe Val Ile Met Leu Met Leu Ile Ala Pro Ala Tyr	
175 180 185 190	
agc ttc aac tgt ctg gga aca tca aac agg gac ttt gtc gag gga gcc	1527
Ser Phe Asn Cys Leu Gly Thr Ser Asn Arg Asp Phe Val Glu Gly Ala	
195 200 205	
agt ggg gca aca tgg att gac ttg gta ctt gaa ggg gga agc tgt gtc	1575

Ser Gly Ala Thr Trp Ile Asp Leu Val Leu Glu Gly Gly Ser Cys Val	
210 215 220	
aca gtg atg gca cca gag aaa cca aca ctg gac ttc aaa gtg atg aag	1623
Thr Val Met Ala Pro Glu Lys Pro Thr Leu Asp Phe Lys Val Met Lys	
225 230 235	
atg gag gct acc gag tta gcc act gtg cgt gag tat tgt tac gaa gca	1671
Met Glu Ala Thr Glu Leu Ala Thr Val Arg Glu Tyr Cys Tyr Glu Ala	
240 245 250	
acc ttg gac acg ctg tca aca gtg gca agg tgc ccc aca aca gga gaa	1719
Thr Leu Asp Thr Leu Ser Thr Val Ala Arg Cys Pro Thr Thr Gly Glu	
255 260 265 270	
gct cac aac acc aaa agg agt gac cca aca ttt gtc tgc aaa aga gat	1767
Ala His Asn Thr Lys Arg Ser Asp Pro Thr Phe Val Cys Lys Arg Asp	
275 280 285	
gtt gtg gac cgc gga tgg ggt aac gga tgt ggt ctg ttt gga aaa ggg	1815
Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly	
290 295 300	
agc att gac aca tgc gct aag ttc aca tgc aaa aac aag gca aca ggg	1863
Ser Ile Asp Thr Cys Ala Lys Phe Thr Cys Lys Asn Lys Ala Thr Gly	
305 310 315	
aag acg atc ttg aga gaa aac atc aag tat gag gtt gca atc ttt gtg	1911
Lys Thr Ile Leu Arg Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val	
320 325 330	
cat ggt tca acg gac tct acg tca cat ggc aat tac tct gag cag att	1959
His Gly Ser Thr Asp Ser Thr Ser His Gly Asn Tyr Ser Glu Gln Ile	
335 340 345 350	
gga aaa aac caa gcg gct aga ttc acc ata agc ccg caa gca ccg tcc	2007
Gly Lys Asn Gln Ala Ala Arg Phe Thr Ile Ser Pro Gln Ala Pro Ser	
355 360 365	
ttt acg gcc aac atg ggc gag tat gga aca gtt acc att gat tgt gaa	2055
Phe Thr Ala Asn Met Gly Glu Tyr Gly Thr Val Thr Ile Asp Cys Glu	
370 375 380	
gca aga tca gga atc aac acg gag gat tat tat gtt ttc act gtc aag	2103

Ala Arg Ser Gly Ile Asn Thr Glu Asp Tyr Tyr Val Phe Thr Val Lys	
385 390 395	
gag aag tca tgg cta gtg aac agg gac tgg ttt cac gac ttg aac ctt	2151
Glu Lys Ser Trp Leu Val Asn Arg Asp Trp Phe His Asp Leu Asn Leu	
400 405 410	
cca tgg acg agc cct gcc aca act gat tgg cgc aac aga gaa aca ctg	2199
Pro Trp Thr Ser Pro Ala Thr Thr Asp Trp Arg Asn Arg Glu Thr Leu	
415 420 425 430	
gtg gaa ttt gag gaa ccg cat gcc acc aag caa act gta gta gcc cta	2247
Val Glu Phe Glu Glu Pro His Ala Thr Lys Gln Thr Val Val Ala Leu	
435 440 445	
gga tcg caa gaa ggt gcc ctg cac aca gca ttg gct gga gcc att cca	2295
Gly Ser Gln Glu Gly Ala Leu His Thr Ala Leu Ala Gly Ala Ile Pro	
450 455 460	
gcc act gtt agc agc tca acc cta acc ttg caa tca ggg cat ttg aaa	2343
Ala Thr Val Ser Ser Ser Thr Leu Thr Leu Gln Ser Gly His Leu Lys	
465 470 475	
tgc aga gct aag ctt gac aag gtc aaa atc aag gga acg aca tat ggc	2391
Cys Arg Ala Lys Leu Asp Lys Val Lys Ile Lys Gly Thr Thr Tyr Gly	
480 485 490	
atg tgt gac tct gcc ttc acc ttc agc aag aac cca act gac aca ggg	2439
Met Cys Asp Ser Ala Phe Thr Phe Ser Lys Asn Pro Thr Asp Thr Gly	
495 500 505 510	
cac ggg aca gtg att gtg gaa ctg cag tat act gga agc aac gga ccc	2487
His Gly Thr Val Ile Val Glu Leu Gln Tyr Thr Gly Ser Asn Gly Pro	
515 520 525	
tgc cga gtt ccc atc tcc gtg act gca aac ctc atg gat ttg aca ccg	2535
Cys Arg Val Pro Ile Ser Val Thr Ala Asn Leu Met Asp Leu Thr Pro	
530 535 540	
gtt gga aga ttg gtc acg gtc aat ccc ttt ata agc aca ggg gga gcg	2583
Val Gly Arg Leu Val Thr Val Asn Pro Phe Ile Ser Thr Gly Gly Ala	
545 550 555	
aac aac aag gtc atg atc gaa gtt gaa cca ccc ttt ggc gat tct tac	2631

Asn Asn Lys Val Met Ile Glu Val Glu Pro Pro Phe Gly Asp Ser Tyr	
560 565 570	
atc gtc gtc gga aga ggc acc acc cag att aac tac cac tgg cac aaa	2679
Ile Val Val Gly Arg Gly Thr Thr Gln Ile Asn Tyr His Trp His Lys	
575 580 585 590	
gag gga agc agc att ggg aag gct ttg gcg acc aca tgg aaa gga gcc	2727
Glu Gly Ser Ser Ile Gly Lys Ala Leu Ala Thr Thr Trp Lys Gly Ala	
595 600 605	
caa cgg cta gcc gtc tta ggg gac aca gcg tgg gac ttt gga tct att	2775
Gln Arg Leu Ala Val Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Ile	
610 615 620	
gga gga gtt ttc aat tca att ggc aaa gct gtc cac caa gtt ttc gga	2823
Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln Val Phe Gly	
625 630 635	
gga gcg ttc agg act ctg ttc ggg gga atg tcc tgg atc aca cag ggg	2871
Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly	
640 645 650	
cta ctt gga gct ctt ctc ctg tgg atg ggg ttg cag gcc cgc gac agg	2919
Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Leu Gln Ala Arg Asp Arg	
655 660 665 670	
agc atc tcg ctg act cta ctg gct gtc gga ggg att ctc atc ttt ctg	2967
Ser Ile Ser Leu Thr Leu Leu Ala Val Gly Gly Ile Leu Ile Phe Leu	
675 680 685	
gca acc agc gtg caa gcc t gagcggccgc tcgagcatgc atctagaggg	3016
Ala Thr Ser Val Gln Ala	
690	
ccctattcta tagtgtcacc taaatgctag agctcgctga tcagcctcga ctgtgccttc	3076
tagttgccag ccatctgttg ttgtccctc ccccgctgcct tccttgaccc tggaagggtgc	3136
cactcccact gtcttttctt aataaaatga ggaaattgca tcgcattgtc tgagtaggtg	3196
tcattctatt ctgggggggtg gggtggggca ggacagcaag ggggaggatt gggaagacaa	3256
tagcaggcat gctggggatg cggtgggctc tatggcttct gaggcggaaa gaacagctgc	3316
attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttcgctt	3376
cctcgctcac tgactcgctg cgctcggtcg ttcggtgcg gcgagcggta tcagctcact	3436
caaaggcggg aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag	3496
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata	3556


```

ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 3616
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg 3676
ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc 3736
tttctcaatg ctcacgctgt aggtatctca gttcgggtgta ggtcgttcgc tccaagctgg 3796
gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc 3856
ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga 3916
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg 3976
gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt accttcgaa 4036
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg 4096
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 4156
ctacggggtc tgacgctcag tggaacgaaa actcacgta agggattttg gtcagttagat 4216
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 4276
aaagtatata tgagtaaaact tggcttgaca gttaccaatg cttaatcagt gaggcaccta 4336
tctcagcgat ctgtctatct cgttcatcca tagttgcctg actccccgtc gtgtagataa 4396
ctacgatacg ggagggctta ccatctggcc ccagtgcctg aatgataccg cgagaccac 4456
gctcaccggc tccagattta tcagcaataa accagccagc cggaaggggc gagcgcagaa 4516
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttggtgccgg gaagctagag 4576
taagtagttc gccagttaat agtttgcgca acgttggtgc cattgctaca ggcatcgtgg 4636
tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag 4696
ttacatgata ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg 4756
tcagaagtaa gttggccgca gtgttatcac tcatggttat ggagcactg cataattctc 4816
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat 4876
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggctcaata cgggataata 4936
ccgcgccaca tagcagaact taaaagtgc tcatcattgg aaaacgttct tcggggcgaa 4996
aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaaccact cgtgcaccca 5056
actgatcttc agcatctttt actttacca gcgtttctgg gtgagcaaaa acaggaaggc 5116
aaaatgccgc aaaaaaggga ataaggcgca cacggaaatg ttgaatactc atactcttc 5176
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 5236
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 5296
ctgacgctc 5304

```

<210> 22

<211> 692

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 22

```

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1           5           10          15
Val Val Ile Ala Gly Thr Ser Ala Leu Gln Leu Ser Thr Tyr Gln Gly
          20           25           30

```

Lys Val Leu Met Ser Ile Asn Lys Thr Asp Ala Gln Ser Ala Ile Asn
 35 40 45
 Ile Pro Ser Ala Asn Gly Ala Asn Thr Cys Ile Val Arg Ala Leu Asp
 50 55 60
 Val Gly Val Met Cys Lys Asp Asp Ile Thr Tyr Leu Cys Pro Val Leu
 65 70 75 80
 Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Asp Val Glu
 85 90 95
 Glu Val Trp Val His Tyr Gly Arg Cys Thr Arg Met Gly His Ser Arg
 100 105 110
 Arg Ser Arg Arg Ser Ile Ser Val Gln His His Gly Asp Ser Thr Leu
 115 120 125
 Ala Thr Lys Asn Thr Pro Trp Leu Asp Thr Val Lys Thr Thr Lys Tyr
 130 135 140
 Leu Thr Lys Val Glu Asn Trp Val Leu Arg Asn Pro Gly Tyr Ala Leu
 145 150 155 160
 Val Ala Leu Ala Ile Gly Trp Met Leu Gly Ser Asn Asn Thr Gln Arg
 165 170 175
 Val Val Phe Val Ile Met Leu Met Leu Ile Ala Pro Ala Tyr Ser Phe
 180 185 190
 Asn Cys Leu Gly Thr Ser Asn Arg Asp Phe Val Glu Gly Ala Ser Gly
 195 200 205
 Ala Thr Trp Ile Asp Leu Val Leu Glu Gly Gly Ser Cys Val Thr Val
 210 215 220
 Met Ala Pro Glu Lys Pro Thr Leu Asp Phe Lys Val Met Lys Met Glu
 225 230 235 240
 Ala Thr Glu Leu Ala Thr Val Arg Glu Tyr Cys Tyr Glu Ala Thr Leu
 245 250 255
 Asp Thr Leu Ser Thr Val Ala Arg Cys Pro Thr Thr Gly Glu Ala His
 260 265 270
 Asn Thr Lys Arg Ser Asp Pro Thr Phe Val Cys Lys Arg Asp Val Val
 275 280 285
 Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile
 290 295 300
 Asp Thr Cys Ala Lys Phe Thr Cys Lys Asn Lys Ala Thr Gly Lys Thr
 305 310 315 320
 Ile Leu Arg Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly
 325 330 335
 Ser Thr Asp Ser Thr Ser His Gly Asn Tyr Ser Glu Gln Ile Gly Lys
 340 345 350
 Asn Gln Ala Ala Arg Phe Thr Ile Ser Pro Gln Ala Pro Ser Phe Thr
 355 360 365
 Ala Asn Met Gly Glu Tyr Gly Thr Val Thr Ile Asp Cys Glu Ala Arg
 370 375 380

Ser Gly Ile Asn Thr Glu Asp Tyr Tyr Val Phe Thr Val Lys Glu Lys
 385 390 395 400
 Ser Trp Leu Val Asn Arg Asp Trp Phe His Asp Leu Asn Leu Pro Trp
 405 410 415
 Thr Ser Pro Ala Thr Thr Asp Trp Arg Asn Arg Glu Thr Leu Val Glu
 420 425 430
 Phe Glu Glu Pro His Ala Thr Lys Gln Thr Val Val Ala Leu Gly Ser
 435 440 445
 Gln Glu Gly Ala Leu His Thr Ala Leu Ala Gly Ala Ile Pro Ala Thr
 450 455 460
 Val Ser Ser Ser Thr Leu Thr Leu Gln Ser Gly His Leu Lys Cys Arg
 465 470 475 480
 Ala Lys Leu Asp Lys Val Lys Ile Lys Gly Thr Thr Tyr Gly Met Cys
 485 490 495
 Asp Ser Ala Phe Thr Phe Ser Lys Asn Pro Thr Asp Thr Gly His Gly
 500 505 510
 Thr Val Ile Val Glu Leu Gln Tyr Thr Gly Ser Asn Gly Pro Cys Arg
 515 520 525
 Val Pro Ile Ser Val Thr Ala Asn Leu Met Asp Leu Thr Pro Val Gly
 530 535 540
 Arg Leu Val Thr Val Asn Pro Phe Ile Ser Thr Gly Gly Ala Asn Asn
 545 550 555 560
 Lys Val Met Ile Glu Val Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
 565 570 575
 Val Gly Arg Gly Thr Thr Gln Ile Asn Tyr His Trp His Lys Glu Gly
 580 585 590
 Ser Ser Ile Gly Lys Ala Leu Ala Thr Thr Trp Lys Gly Ala Gln Arg
 595 600 605
 Leu Ala Val Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly
 610 615 620
 Val Phe Asn Ser Ile Gly Lys Ala Val His Gln Val Phe Gly Gly Ala
 625 630 635 640
 Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu
 645 650 655
 Gly Ala Leu Leu Leu Trp Met Gly Leu Gln Ala Arg Asp Arg Ser Ile
 660 665 670
 Ser Leu Thr Leu Leu Ala Val Gly Gly Ile Leu Ile Phe Leu Ala Thr
 675 680 685
 Ser Val Gln Ala
 690

<210> 23

<211> 5271

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 备注=合成构建体

<221> CDS

<222> (910)... (2953)

<400> 23

```

gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420
attgacgtca atgggtggac tatttacggg aaactgcccc cttggcagta catcaagtgt      480
atcatatgcc aagtagcccc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca      600
tcgtattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg      660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc      720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg      780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca      840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttggtacc      900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc      951

```

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser

1

5

10

```

ttg gca gtt gtc ata gct ggt aca agc gct gtg acc ttg gtg cgg aaa      999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys
15                20                25                30

```

```

aac aga tgg ttg ctc cta aat gtg aca tct gag gac ctc ggg aaa aca      1047
Asn Arg Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr
35                40                45

```

```

ttc tct gtg ggc aca ggc aac tgc aca aca aac att ttg gaa gcc aag      1095
Phe Ser Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys
50                55                60

```

```

tac tgg tgc cca gac tca atg gaa tac aac tgt ccc aat ctc agt cca      1143
Tyr Trp Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro
65                70                75

```

aga gag gag cca gat gac att gat tgc tgg tgc tat ggg gtg gaa aac	1191
Arg Glu Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn	
80 85 90	
gtt aga gtc gca tat ggt aag tgt gac tca gca ggc agg tct agg agg	1239
Val Arg Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg	
95 100 105 110	
tca aga agg gcc att gac ttg cct acg cat gaa aac cat ggt ttg aag	1287
Ser Arg Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys	
115 120 125	
acc cgg caa gaa aaa tgg atg act gga aga atg ggt gaa agg caa ctc	1335
Thr Arg Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu	
130 135 140	
caa aag att gag aga tgg ttc gtg agg aac ccc ttt ttt gca gtg acg	1383
Gln Lys Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr	
145 150 155	
gct ctg acc att gcc tac ctt gtg gga agc aac atg acg caa cga gtc	1431
Ala Leu Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val	
160 165 170	
gtg att gcc cta ctg gtc ttg gct gtt ggt ccg gcc tac tca gct cac	1479
Val Ile Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His	
175 180 185 190	
tgc att gga att act gac agg gat ttc att gag ggg gtg cat gga gga	1527
Cys Ile Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly	
195 200 205	
act tgg gtt tca gct acc ctg gag caa gac aag tgt gtc act gtt atg	1575
Thr Trp Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met	
210 215 220	
gcc cct gac aag cct tca ttg gac atc tca cta gag aca gta gcc att	1623
Ala Pro Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile	
225 230 235	
gat aga cct gct gag gtg agg aaa gtg tgt tac aat gca gtt ctc act	1671
Asp Arg Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr	
240 245 250	

138

aga gta ctg gcc ctg gga aac cag gaa ggc tcc ttg aaa aca gct ctt	2247
Arg Val Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu	
435 440 445	
act ggc gca atg agg gtt aca aag gac aca aat gac aac aac ctt tac	2295
Thr Gly Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr	
450 455 460	
aaa cta cat ggt gga cat gtt tct tgc aga gtg aaa ttg tca gct ttg	2343
Lys Leu His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu	
465 470 475	
aca ctc aag ggg aca tcc tac aaa ata tgc act gac aaa atg ttt ttt	2391
Thr Leu Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe	
480 485 490	
gtc aag aac cca act gac act ggc cat ggc act gtt gtg atg cag gtg	2439
Val Lys Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val	
495 500 505 510	
aaa gtg tca aaa gga gcc ccc tgc agg att cca gtg ata gta gct gat	2487
Lys Val Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp	
515 520 525	
gat ctt aca gcg gca atc aat aaa ggc att ttg gtt aca gtt aac ccc	2535
Asp Leu Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro	
530 535 540	
atc gcc tca acc aat gat gat gaa gtg ctg att gag gtg aac cca cct	2583
Ile Ala Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro	
545 550 555	
ttt gga gac agc tac att atc gtt ggg aga gga gat tca cgt ctc act	2631
Phe Gly Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr	
560 565 570	
tac cag tgg cac aaa gag gga agc tca ata gga aag ttg ttc act cag	2679
Tyr Gln Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln	
575 580 585 590	
acc atg aaa ggc gtg gaa cgc ctg gcc gtc atg gga gac acc gcc tgg	2727
Thr Met Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp	
595 600 605	

gat ttc agc tcc gct gga ggg ttc ttc act tcg gtt ggg aaa gga att	2775
Asp Phe Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile	
610 615 620	
cat acg gtg ttt ggc tct gcc ttt cag ggg cta ttt ggc ggc ttg aac	2823
His Thr Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn	
625 630 635	
tgg ata aca aag gtc atc atg ggg gcg gta ctt ata tgg gtt ggc atc	2871
Trp Ile Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile	
640 645 650	
aac aca aga aac atg aca atg tcc atg agc atg atc ttg gta gga gtg	2919
Asn Thr Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val	
655 660 665 670	
atc atg atg ttt ttg tct cta gga gtt ggg gcg t gagcgccgc	2963
Ile Met Met Phe Leu Ser Leu Gly Val Gly Ala	
675 680	
tcgagcatgc atctagaggg ccctattcta tagtgcacc taaatgctag agctcgctga	3023
tcagcctcga ctgtgccttc tagttgccag ccactctgttg tttgcccctc ccccgctgcct	3083
tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca	3143
tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtaggggca ggacagcaag	3203
ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtagggctc tatggcttct	3263
gaggcgga aaacagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg	3323
tattgggcgc tcttcgcgtt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg	3383
gcgagcggtg tcagctcact caaaggcggg aatacgggta tccacagaat caggggataa	3443
cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc	3503
gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa atcgacgctc	3563
aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcggttc cccttgaag	3623
ctccctcgctg cgctctcctg ttccgacct gccgcttacc ggatacctgt ccgcctttct	3683
cccttcggga agcgtggcgc tttctcaatg ctacgctgt aggtatctca gttcggtgta	3743
ggctgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgctgcgc	3803
cttatccggt aactatcgtc ttgagtccaa cccgtaaga cagacttat cgccactggc	3863
agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt	3923
gaagtgggtg cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct	3983
gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc	4043
tggtagcggg ggtttttttg tttgcaagca gcagattacg gcgagaaaaa aaggatctca	4103
agaagatcct ttgatctttt ctacgggggc tgacgctcag tggaacgaaa actcacgtta	4163
agggattttg gtcattgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa	4223
atgaagtgtt aaatcaatct aaagtatata tgagtaaact tggctctgaca gttaccaatg	4283
cttaatcagt gaggcaccta tctcagcgat ctgtctatct cgttcaccca tagttgcctg	4343
actccccgct gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgtctg	4403


```

aatgataccg cgagacccac gtcaccggc tccagattta tcagcaataa accagccagc 4463
cggaagggcc gagcgagaa gtggctctgc aactttatcc gcctccatcc agtctattaa 4523
ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca acgttggtgc 4583
cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg 4643
ttccaacga tcaaggcgag ttacatgata ccccatgttg tgcaaaaaag cggttagctc 4703
cttcggtect ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat 4763
ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg 4823
tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc 4883
ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg 4943
aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat 5003
gtaaccact cgtgcacca actgatcttc agcatctttt actttcacca gcgtttctgg 5063
gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca cacggaatg 5123
ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct 5183
catgagcgga tacatatgtg aatgtattta gaaaaataaa caaatagggg ttccgcgcac 5243
atctccccga aaagtgcac ctgacgtc 5271

```

<210> 24

<211> 681

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述: 备注=合成构建体

<400> 24

```

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

```

145	150	155	160
Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val Val Ile			
	165	170	175
Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His Cys Ile			
	180	185	190
Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly Thr Trp			
	195	200	205
Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met Ala Pro			
	210	215	220
Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile Asp Arg			
225	230	235	240
Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr His Val			
	245	250	255
Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu Ala Glu			
	260	265	270
Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp Arg Gly			
	275	280	285
Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys			
	290	295	300
Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val Asp Gln			
305	310	315	320
Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly Ala Lys			
	325	330	335
Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp Ala Leu			
	340	345	350
Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala Thr Leu			
	355	360	365
Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr Ile Ala			
	370	375	380
Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala Gln Asp			
385	390	395	400
Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg Glu Met			
	405	410	415
His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile Arg Val			
	420	425	430
Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu Thr Gly			
	435	440	445
Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr Lys Leu			
	450	455	460
His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu Thr Leu			
465	470	475	480
Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe Val Lys			
	485	490	495
Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val			

```

          500          505          510
Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp Asp Leu
          515          520          525
Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala
          530          535          540
Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly
545          550          555          560
Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr Tyr Gln
          565          570          575
Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln Thr Met
          580          585          590
Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp Asp Phe
          595          600          605
Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile His Thr
          610          615          620
Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn Trp Ile
625          630          635          640
Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr
          645          650          655
Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met
          660          665          670
Met Phe Leu Ser Leu Gly Val Gly Ala
          675          680

```

<210> 25

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<221> misc_特征

<222> 1-35

<223> POW 454

<400> 25

aaaagaaaaa gcgctaccac catccaccgg gacag

35

<210> 26

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-41

<223> CPOW 2417

<400> 26

actgttaccc tcaaccccggt actcgccggc gaaaaagaaa a

41

<210> 27

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<223> 修饰的 JE 信号

<400> 27

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala

1

5

10

15

Val Val Ile Ala Gly Thr Ser Ala

20

<210> 28

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-36

<223> YF 482

<400> 28

aaaagaaaaa gcgctgtgac cttggtgcgg aaaaac

36

<210> 29

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-41

<223> CYF 2433

<400> 29

acagagatcc tcaaccccgcc actcgccggc gaaaaagaaa a

41

<210> 30

<211> 41

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-41

<223> SLE 463

<400> 30

aaaagaaaaa gcgctttgca gttatcaacc tatcagggga a

41

<210> 31

<211> 40

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> misc_特征

<222> 1-40

<223> CSLE 2477

<400> 31

accgttggtc gcacgttcgg actcgccggc gaaaaagaaa

40

<210> 32

<211> 39

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 32

```
Leu Asp Thr Ile Asn Arg Arg Pro Ser Lys Lys Arg Gly Gly Thr Arg
 1           5           10          15
Ser Leu Leu Gly Leu Ala Ala Leu Ile Gly Leu Ala Ser Ser Leu Gln
          20           25           30
Leu Leu Ser Thr Tyr Gln Gly
          35
```

<210> 33

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 33

```
Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met
 1           5           10          15
Lys Leu Ser Asn Phe Gln Gly Lys
          20
```

<210> 34

<211> 30

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 34

```
Met Asn Glu Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile
 1           5           10          15
Ala Cys Ala Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys
          20           25           30
```

<210> 35

<211> 39

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 35

```
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu Gly Ser Ile Met
 1           5           10           15
Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met Lys
           20           25           30
Leu Ser Asn Phe Gln Gly Lys
           35
```

<210> 36

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 36

```
Met Ser Lys Lys Arg Gly Gly Ser Glu Thr Ser Val Leu Met Val Ile
 1           5           10           15
Phe Met Leu Ile Gly Phe Ala Ala Ala Leu Lys Leu Ser Asn Phe Gln
           20           25           30
Gly Lys
```

<210> 37

<211> 33

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 37

```
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1           5           10           15
Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn Phe Gln Gly
           20           25           30
Lys
```

<210> 38

<211> 46

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 38

```
Met Asn Val Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn
 1           5           10          15
Ile Leu Asn Arg Arg Arg Arg Thr Ala Gly Met Ile Ile Met Leu Ile
      20           25          30
Pro Thr Val Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu
      35           40          45
```

<210> 39

<211> 40

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 39

```
Met Val Gly Leu Gln Lys Arg Gly Lys Arg Arg Ser Ala Thr Asp Trp
 1           5           10          15
Met Ser Trp Leu Leu Val Ile Thr Leu Leu Gly Met Thr Leu Ala Ala
      20           25          30
Thr Val Arg Lys Glu Arg Gly Asp
      35           40
```

<210> 40

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 40

```
Met Gly Trp Leu Leu Val Val Val Leu Leu Gly Val Thr Leu Ala Ala
```


1	5	10	15
Thr	Val	Arg	Lys
	Glu	Arg	Gly
	Asp		
20			

<210> 41

<211> 24

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 41

Met Ser Trp Leu Leu Val Ile Thr Leu Leu Gly Met Thr Ile Ala Ala

1	5	10	15
Thr	Val	Arg	Lys
	Glu	Arg	Gly
	Asp		
20			

<210> 42

<211> 5292

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<221> CDS

<222> (910)... (2964)

<400> 42

gacggatcgg gagatctccc gateccctat ggtgcactct cagtacaatc tgctctgatg	60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg	120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc	180
ttagggtttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt	240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata	300
tggagttccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc	360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
attgacgtca atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt	480
atcatatgcc aagtagcccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatggcg	780
gtaggcgtgt acgggtggag gtctatataa gcagagctct ctggctaact agagaacca	840

ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc	900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc	951
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser	
1 5 10	
ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt	999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg	
15 20 25 30	
aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt	1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser	
35 40 45	
ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc	1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala	
50 55 60	
atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc	1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro	
65 70 75	
ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct	1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser	
80 85 90	
acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat	1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His	
95 100 105 110	
aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga	1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly	
115 120 125	
ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa	1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys	
130 135 140	
cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc	1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr	
145 150 155	
atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa	1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln	
160 165 170	

aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca	1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr	
175 180 185 190	
atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca	1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser	
195 200 205	
gga gga agc tgg gtt gac ata gtc tta gaa cat gga agc tgt gtg acg	1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr	
210 215 220	
acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca	1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr	
225 230 235	
gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag	1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys	
240 245 250	
cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc	1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro	
255 260 265 270	
agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg	1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met	
275 280 285	
gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc	1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly	
290 295 300	
att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa	1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys	
305 310 315	
gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac	1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His	
320 325 330	
tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag	1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys	
335 340 345 350	

gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr 355 360 365	2007
ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp 370 375 380	2055
ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val 385 390 395	2103
cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala 400 405 410	2151
gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe 415 420 425 430	2199
aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln 435 440 445	2247
gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met 450 455 460	2295
tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg 465 470 475	2343
atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly 480 485 490	2391
aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile 495 500 505 510	2439
gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro 515 520 525	2487

ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att	2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile	
530 535 540	
aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa	2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu	
545 550 555	
gca gaa cct cca ttc gga gac agc tac atc atc ata gga gta gag ccg	2631
Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro	
560 565 570	
gga caa ctg aag ctc aac tgg ttt aag aaa gga agt tct atc ggc caa	2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln	
575 580 585 590	
atg ttt gag aca aca atg agg ggg gcg aag aga atg gcc att tta ggt	2727
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly	
595 600 605	
gac aca gcc tgg gat ttt gga tcc ttg gga gga gtg ttt aca tct ata	2775
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile	
610 615 620	
gga aag gct ctc cac caa gtc ttt gga gca atc tat gga gct gcc ttc	2823
Gly Lys Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe	
625 630 635	
agt ggg gtt tca tgg act atg aaa atc ctc ata gga gtc att atc aca	2871
Ser Gly Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr	
640 645 650	
tgg ata gga atg aat tca cgc agc acc tca ctg tct gtg aca cta gta	2919
Trp Ile Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val	
655 660 665 670	
ttg gtg gga att gtg aca ctg tat ttg gga gtc atg gtg cag gcc	2964
Leu Val Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala	
675 680 685	
taattagttg agcgccgct cgagcatgca tctagagggc cctattctat agtgtcacct	3024
aaatgctaga gctcgtgat cagcctcgac tgtgccttct agttgccagc catctgttgt	3084
ttgccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcttttccta	3144

```

ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg 3204
gggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 3264
gggtgggctct atggcttctg aggcggaaag aaccagctgc attaatgaat cggccaacgc 3324
gcggggagag gcggtttgcg tattgggcgc tcttcgcgtt cctcgcctac tgactcgtg 3384
cgctcggctcg ttgggctcgc gcgagcggta tcagctcact caaaggcggg aatacggtta 3444
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 3504
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 3564
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 3624
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 3684
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 3744
aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 3804
gttcagcccc accgctgcgc cttatccggg aactatcgtc ttgagtcaa cccggtaga 3864
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 3924
ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag aagaacagta 3984
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 4044
tccggcaaac aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg 4104
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 4164
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 4224
tagatccttt taaattaaat atgaagtttt aaatcaatct aaagtatata tgagtaaact 4284
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctatct 4344
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggccta 4404
ccatctggcc ccagtgtcgc aatgataccg cgagaccac gctcaccggc tccagattta 4464
tcagcaataa accagccagc cggaagggcc gagcgagaa gtggtcctgc aactttatcc 4524
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 4584
agtttgcgca acgttggttc cattgctaca ggcacgtggt tgtcacgctc gtcgtttggt 4644
atggcttcat tcagctccgg tteccaacga tcaaggcgag ttacatgatc ccccatgttg 4704
tgcaaaaaag cggtagctc cttcggctct ccgatcgttg tcagaagtaa gttggccgca 4764
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 4824
agatgctttt ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg 4884
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 4944
ttaaaagtc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 5004
ctgttgagat ccagttcgat gtaaccact cgtgcacca actgatcttc agcatctttt 5064
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 5124
ataaggcgca cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 5184
atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 5244
caaatagggg ttccgcgcac atttccccga aaagtgcac ctgacgtc 5292

```

<210> 43

<211> 685

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 43

```

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1             5             10             15
Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg Asn Gly
      20             25             30
Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu
      35             40             45
Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala Met Asp
      50             55             60
Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro Leu Leu
65             70             75             80
Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser
      85             90             95
Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His Arg Arg
      100            105            110
Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly Leu Glu
      115            120            125
Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
      130            135            140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
      145            150            155            160
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
      165            170            175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
      180            185            190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
      195            200            205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
      210            215            220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
      225            230            235            240
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
      245            250            255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu
      260            265            270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp
      275            280            285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val
      290            295            300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val
      305            310            315            320
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
      325            330            335

```

Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
 340 345 350
 Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
 355 360 365
 Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
 370 375 380
 Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
 385 390 395 400
 Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
 405 410 415
 Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
 420 425 430
 Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
 435 440 445
 Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
 450 455 460
 Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
 465 470 475 480
 Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
 485 490 495
 Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
 500 505 510
 Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
 515 520 525
 Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
 530 535 540
 Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
 545 550 555 560
 Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln
 565 570 575
 Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln Met Phe
 580 585 590
 Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Asp Thr
 595 600 605
 Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile Gly Lys
 610 615 620
 Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly
 625 630 635 640
 Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr Trp Ile
 645 650 655
 Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val Leu Val
 660 665 670
 Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
 675 680 685

<210> 44

<211> 5293

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 备注=合成构建体

<221> CDS

<222> (910)... (2964)

<400> 44

```

gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggctcgt gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggtttag gcgtttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt      240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtcat agcccatata      300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc      360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc      420
attgacgtca atgggtggag tatttacggt aaactgcccc cttggcagta catcaagtgt      480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt      540
atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca      600
tcgctattac catggtgatg cggtttttggc agtacatcaa tgggcgtgga tagcggtttg      660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc      720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg      780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca      840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttgggtacc      900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc      951
      Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
            1             5             10

```

```

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt      999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
    15             20             25             30

```

```

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt      1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
            35             40             45

```

```

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc      1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
            50             55             60

```

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc	1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro	
65 70 75	
ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct	1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser	
80 85 90	
acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat	1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His	
95 100 105 110	
aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga	1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly	
115 120 125	
ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa	1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys	
130 135 140	
cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc	1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr	
145 150 155	
atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa	1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln	
160 165 170	
aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca	1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr	
175 180 185 190	
atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca	1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser	
195 200 205	
gga gga agc tgg gtt gac ata gtc tta gaa cat gga agc tgt gtg acg	1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr	
210 215 220	
acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca	1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr	
225 230 235	

gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys 240 245 250	1671
cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro 255 260 265 270	1719
agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met 275 280 285	1767
gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly 290 295 300	1815
att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys 305 310 315	1863
gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His 320 325 330	1911
tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys 335 340 345 350	1959
gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr 355 360 365	2007
ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp 370 375 380	2055
ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val 385 390 395	2103
cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala 400 405 410	2151

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc	2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe	
415 420 425 430	
aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa	2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln	
435 440 445	
gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg	2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met	
450 455 460	
tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga	2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg	
465 470 475	
atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga	2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly	
480 485 490	
aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata	2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile	
495 500 505 510	
gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct	2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro	
515 520 525	
ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att	2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile	
530 535 540	
aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa	2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu	
545 550 555	
gca gaa cct cca ttc gga gac agc cac atc atc ata gga gta gag ccg	2631
Ala Glu Pro Pro Phe Gly Asp Ser His Ile Ile Ile Gly Val Glu Pro	
560 565 570	
gga caa ctg aag ctc aac tgg ttt aag aaa gga agt tct atc ggc caa	2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln	
575 580 585 590	

atg ttt gag aca aca atg agg ggg gcg aag aga atg gcc att tta ggt	2727
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly	
595 600 605	
gac aca gcc tgg gat ttt gga tcc ttg gga gga gtg ttt aca tct ata	2775
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile	
610 615 620	
gga aag gct ctc cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt	2823
Gly Lys Ala Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe	
625 630 635	
ggg gga atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc	2871
Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu	
640 645 650	
tgg atg ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta	2919
Trp Met Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu	
655 660 665 670	
gcc aca ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct	2964
Ala Thr Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala	
675 680 685	
taattagttt gggcgccgc tcgagcatgc atctagaggg ccctattcta tagtgtcacc	3024
taaatgctag agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg	3084
tttgccttc ccccgctgcct tccttgacct tggaagggtgc cactcccact gtcctttcct	3144
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg	3204
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg	3264
cgggtgggctc tatggcttct gaggcggaaa gaaccagctg cattaatgaa tcggccaacg	3324
cgcggggaga ggcggtttgc gtattgggcg ctcttcgct tcctcgctca ctgactcgct	3384
gcgctcggtc gttcggtgc ggcgagcgg atcagctcac tcaaaggcgg taatacgggt	3444
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	3504
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga	3564
gcatcacaaa aatcgacgct caagtccag gtggcgaaac ccgacaggac tataaagata	3624
ccaggcgttt ccccttgaa gctccctcgt gcgctctcct gttccgacct tgccgcttac	3684
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gtcacgctg	3744
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc	3804
cgttcagccc gaccgctgcg ctttatccgg taactatcgt cttgagtcca acccggtaa	3864
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatatgt	3924
aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaagaacagt	3984
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	4044
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac	4104
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca	4164

```

gtggaacgaa aactcacgtt aagggaatttt ggtcatgaga ttatcaaaaa ggatcttcac 4224
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 4284
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 4344
tcgttcatcc atagtgcct gactccccgt cgtgtagata actacgatac gggagggctt 4404
accatctggc cccagtgtg caatgatacc gcgagaccca cgctcaccgg ctccagattt 4464
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 4524
cgcttccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 4584
tagtttgccg aacgttggtt ccattgtctac aggcacgtg gtgtcacgct cgtcgtttgg 4644
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt 4704
gtgcaaaaaa gcggttagct cttcgggtcc tccgatcgtt gtcagaagta agttggccgc 4764
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 4824
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 4884
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 4944
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 5004
gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt 5064
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaagg 5124
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 5184
cattiatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 5244
acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtc 5293

```

<210> 45

<211> 685

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 45

```

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

```

115	120	125
Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val		
130	135	140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met		
145	150	155
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala		
165	170	175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg		
180	185	190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly		
195	200	205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met		
210	215	220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala		
225	230	235
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr		
245	250	255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu		
260	265	270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp		
275	280	285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val		
290	295	300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val		
305	310	315
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly		
325	330	335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile		
340	345	350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr		
355	360	365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn		
370	375	380
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg		
385	390	395
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr		
405	410	415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn		
420	425	430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly		
435	440	445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser		
450	455	460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp		

```

465          470          475          480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
          485          490          495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
          500          505          510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
          515          520          525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
          530          535          540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545          550          555          560
Pro Pro Phe Gly Asp Ser His Ile Ile Ile Gly Val Glu Pro Gly Gln
          565          570          575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln Met Phe
          580          585          590
Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Asp Thr
          595          600          605
Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile Gly Lys
          610          615          620
Ala Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
625          630          635          640
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
          645          650          655
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
          660          665          670
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
          675          680          685

```

<210> 46

<211> 5293

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<221> CDS

<222> (910)... (2964)

<400> 46

```

gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgt gagtagtgcg      120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc      180
ttagggtag gcgttttgcg ctgcttcgcg atgtacgggc cagatatagc cgttgacatt      240

```


gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata	300
tggagttccg cgttacataa cttacggtaa atggcccgc tggtgaccg cccaacgacc	360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	420
attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt	480
atcatatgcc aagtagccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca	600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg	660
actcacggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg	780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca	840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gcttggtagc	900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc	951
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser	
1 5 10	
ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt	999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg	
15 20 25 30	
aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt	1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser	
35 40 45	
ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc	1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala	
50 55 60	
atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc	1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro	
65 70 75	
ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct	1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser	
80 85 90	
acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat	1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His	
95 100 105 110	
aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga	1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly	
115 120 125	
ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa	1335

Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys	
130 135 140	
cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc	1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr	
145 150 155	
atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa	1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln	
160 165 170	
aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca	1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr	
175 180 185 190	
atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca	1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser	
195 200 205	
gga gga agc tgg gtt gac ata gtc tta gaa cat ggg agc tgt gtg acg	1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr	
210 215 220	
acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca	1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr	
225 230 235	
gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag	1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys	
240 245 250	
cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc	1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro	
255 260 265 270	
agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg	1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met	
275 280 285	
gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc	1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly	
290 295 300	
att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa	1863

Ile	Val	Thr	Cys	Ala	Met	Phe	Arg	Cys	Lys	Lys	Asn	Met	Glu	Gly	Lys		
305						310						315					
gtt	gtg	caa	cca	gaa	aac	ttg	gaa	tac	acc	att	gtg	ata	aca	cct	cac	1911	
Val	Val	Gln	Pro	Glu	Asn	Leu	Glu	Tyr	Thr	Ile	Val	Ile	Thr	Pro	His		
320						325						330					
tca	ggg	gaa	gag	cat	gca	gtc	gga	aat	gac	aca	gga	aaa	cat	ggc	aag	1959	
Ser	Gly	Glu	Glu	His	Ala	Val	Gly	Asn	Asp	Thr	Gly	Lys	His	Gly	Lys		
335						340						345			350		
gaa	atc	aaa	ata	aca	cca	cag	agt	tcc	atc	aca	gaa	gca	gaa	ttg	aca	2007	
Glu	Ile	Lys	Ile	Thr	Pro	Gln	Ser	Ser	Ile	Thr	Glu	Ala	Glu	Leu	Thr		
			355						360						365		
ggt	tat	ggc	act	gtc	aca	atg	gag	tgc	tct	cca	aga	acg	ggc	ctc	gac	2055	
Gly	Tyr	Gly	Thr	Val	Thr	Met	Glu	Cys	Ser	Pro	Arg	Thr	Gly	Leu	Asp		
			370						375						380		
ttc	aat	gag	atg	gtg	ttg	ttg	cag	atg	gaa	aat	aaa	gct	tgg	ctg	gtg	2103	
Phe	Asn	Glu	Met	Val	Leu	Leu	Gln	Met	Glu	Asn	Lys	Ala	Trp	Leu	Val		
			385						390						395		
cac	agg	caa	tgg	ttc	cta	gac	ctg	ccg	tta	cca	tgg	ttg	ccc	gga	gcg	2151	
His	Arg	Gln	Trp	Phe	Leu	Asp	Leu	Pro	Leu	Pro	Trp	Leu	Pro	Gly	Ala		
400						405						410					
gac	aca	caa	ggg	tca	aat	tgg	ata	cag	aaa	gag	aca	ttg	gtc	act	ttc	2199	
Asp	Thr	Gln	Gly	Ser	Asn	Trp	Ile	Gln	Lys	Glu	Thr	Leu	Val	Thr	Phe		
415						420						425			430		
aaa	aat	ccc	cat	gcg	aag	aaa	cag	gat	gtt	gtt	gtt	tta	gga	tcc	caa	2247	
Lys	Asn	Pro	His	Ala	Lys	Lys	Gln	Asp	Val	Val	Val	Leu	Gly	Ser	Gln		
			435						440						445		
gaa	ggg	gcc	atg	cac	aca	gca	ctt	aca	ggg	gcc	aca	gaa	atc	caa	atg	2295	
Glu	Gly	Ala	Met	His	Thr	Ala	Leu	Thr	Gly	Ala	Thr	Glu	Ile	Gln	Met		
			450						455						460		
tca	tca	gga	aac	tta	ctc	ttc	aca	gga	cat	ctc	aag	tgc	agg	ctg	aga	2343	
Ser	Ser	Gly	Asn	Leu	Leu	Phe	Thr	Gly	His	Leu	Lys	Cys	Arg	Leu	Arg		
465						470						475					
atg	gac	aag	cta	cag	ctc	aaa	gga	atg	tca	tac	tct	atg	tgc	aca	gga	2391	

Met	Asp	Lys	Leu	Gln	Leu	Lys	Gly	Met	Ser	Tyr	Ser	Met	Cys	Thr	Gly	
480						485					490					
aag	ttt	aaa	gtt	gtg	aag	gaa	ata	gca	gaa	aca	caa	cat	gga	aca	ata	2439
Lys	Phe	Lys	Val	Val	Lys	Glu	Ile	Ala	Glu	Thr	Gln	His	Gly	Thr	Ile	
495					500					505					510	
gtt	atc	aga	gtg	caa	tat	gaa	ggg	gac	ggc	tct	cca	tgc	aag	atc	cct	2487
Val	Ile	Arg	Val	Gln	Tyr	Glu	Gly	Asp	Gly	Ser	Pro	Cys	Lys	Ile	Pro	
				515					520					525		
ttt	gag	ata	atg	gat	ttg	gaa	aaa	aga	cat	gtc	tta	ggt	cgc	ctg	att	2535
Phe	Glu	Ile	Met	Asp	Leu	Glu	Lys	Arg	His	Val	Leu	Gly	Arg	Leu	Ile	
			530					535					540			
aca	gtc	aac	cca	att	gtg	aca	gaa	aaa	gat	agc	cca	gtc	aac	ata	gaa	2583
Thr	Val	Asn	Pro	Ile	Val	Thr	Glu	Lys	Asp	Ser	Pro	Val	Asn	Ile	Glu	
		545					550					555				
gca	gaa	cct	cca	ttc	gga	gac	agc	tac	atc	atc	ata	gga	gta	gag	ccg	2631
Ala	Glu	Pro	Pro	Phe	Gly	Asp	Ser	Tyr	Ile	Ile	Ile	Gly	Val	Glu	Pro	
560					565						570					
gga	caa	ctg	aag	ctc	aac	tgg	ttt	aag	aaa	gga	agc	acg	ctg	ggc	aag	2679
Gly	Gln	Leu	Lys	Leu	Asn	Trp	Phe	Lys	Lys	Gly	Ser	Thr	Leu	Gly	Lys	
575					580					585					590	
gcc	ttt	tca	aca	act	ttg	aag	gga	gct	caa	aga	ctg	gca	gcg	ttg	ggc	2727
Ala	Phe	Ser	Thr	Thr	Leu	Lys	Gly	Ala	Gln	Arg	Leu	Ala	Ala	Leu	Gly	
				595					600					605		
gac	aca	gcc	tgg	gac	ttt	ggc	tct	att	gga	ggg	gtc	ttc	aac	tcc	ata	2775
Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser	Ile	Gly	Gly	Val	Phe	Asn	Ser	Ile	
		610						615					620			
gga	aaa	gcc	gtt	cac	caa	gtg	ttt	ggt	ggt	gcc	ttc	aga	aca	ctc	ttt	2823
Gly	Lys	Ala	Val	His	Gln	Val	Phe	Gly	Gly	Ala	Phe	Arg	Thr	Leu	Phe	
		625				630					635					
ggg	gga	atg	tct	tgg	atc	aca	caa	ggg	cta	atg	ggt	gcc	cta	ctg	ctc	2871
Gly	Gly	Met	Ser	Trp	Ile	Thr	Gln	Gly	Leu	Met	Gly	Ala	Leu	Leu	Leu	
640					645						650					
tgg	atg	ggc	gtc	aac	gca	cga	gac	cga	tca	att	gct	ttg	gcc	ttc	tta	2919

Trp Met Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu	
655	660
	665
	670
gcc aca ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct	2964
Ala Thr Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala	
675	680
	685
taattagttt gagcgccgc tcgagcatgc atctagaggg ccctattcta tagtgcacc	3024
taaatgctag agctcgtga tcagcctcga ctgtgccttc tagttgccag ccatctgttg	3084
tttgccttc ccccggtcct tccttgacct tggaaggtgc cactcccact gtcctttcct	3144
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctggggggtg	3204
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg	3264
cgggtgggctc tatggcttct gaggcggaaa gaaccagctg cattaatgaa tcggccaacg	3324
cgcggggaga ggcggtttgc gtattgggcg ctcttcgct tcctcgtca ctgactcgt	3384
gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacggtt	3444
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	3504
caggaaccgt aaaaaggcgg cgttgctggc gtttttccat aggtccgcc cccctgacga	3564
gcatacaaaa aatcgacgt caagtcagag gtggcgaaac ccgacaggac tataaagata	3624
ccagcggtt cccctggaa gtcctcctgt gcgctctcct gttccgacct tgccgcttac	3684
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcaat gtcacgctg	3744
taggtatctc agttcgtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc	3804
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccgtaag	3864
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt	3924
aggcgggtgt acagagtct tgaagtgggt gcctaactac ggctacacta gaaggacagt	3984
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	4044
atccggcaaa caaacaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac	4104
gcgcagaaaa aaagatctc aagaagatcc tttgatctt tctacggggt ctgacgctca	4164
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac	4224
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	4284
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	4344
tcgttcaccc atagtgcct gactccccgt cgtgtagata actacgatac gggagggctt	4404
accatctggc ccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagattt	4464
atcagcaata aaccagccag ccggaaggcg cgagcgcaga agtggtcctg caactttatc	4524
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa	4584
tagtttgccg aacgttggtt ccattgctac aggcacgtg gtgtcacgct cgctgtttgg	4644
tatggcttca ttcagctcgg gttcccaacg atcaaggcga gttacatgat ccccatgtt	4704
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc	4764
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt	4824
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg	4884
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac	4944
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc	5004
gctgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt	5064
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaaggg	5124
aataaggcgg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag	5184

catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 5244
 acaaataggg gttccgcgca catttccccg aaaagtgccca cctgacgtc 5293

<210> 47

<211> 685

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 47

Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1 5 10 15
 Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg Asn Gly
 20 25 30
 Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu
 35 40 45
 Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala Met Asp
 50 55 60
 Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro Leu Leu
 65 70 75 80
 Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser
 85 90 95
 Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His Arg Arg
 100 105 110
 Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly Leu Glu
 115 120 125
 Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
 130 135 140
 Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
 145 150 155 160
 Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
 165 170 175
 Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
 180 185 190
 Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
 195 200 205
 Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
 210 215 220
 Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
 225 230 235 240
 Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
 245 250 255

Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu		
260	265	270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp		
275	280	285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val		
290	295	300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val		
305	310	315
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly		
325	330	335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile		
340	345	350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr		
355	360	365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn		
370	375	380
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg		
385	390	395
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr		
405	410	415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn		
420	425	430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly		
435	440	445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser		
450	455	460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp		
465	470	475
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe		
485	490	495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile		
500	505	510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu		
515	520	525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val		
530	535	540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu		
545	550	555
Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln		
565	570	575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Thr Leu Gly Lys Ala Phe		
580	585	590
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr		
595	600	605

Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
 610 615 620
 Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
 625 630 635 640
 Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
 645 650 655
 Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
 660 665 670
 Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
 675 680 685

<210> 48

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 48

tgtgcaggcg ccttcattt aaccacacgt aacg

34

<210> 49

<211> 40

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 49

tcgagcggcc gctcaactaa ttaggcctgc accatgactc

40

<210> 50

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述; 备注=合成构建体

<400> 50

cttatcgaaa ttaatacgac tcactatagg

30

<210> 51
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 51
atagattgct ccaaacactt ggtgg 25

<210> 52
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 52
actccatagg aaaagccgtt cacc 24

<210> 53
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 53
gcgagctcta gcatttaggt gacactatag 30

<210> 54
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 54
ctccaccaag tgtttggtgg tgccttcaga aca 33

<210> 55

<211> 11

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 55

Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr

1

5

10

<210> 56

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 56

cttatcgaaa ttaatacgac tcactatagg

30

<210> 57

<211> 39

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 57

gaattcgtct cacttccttt cttaaaccag ttgagcttc

39

<210> 58

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述：备注=合成构建体

<400> 58

ggaattcgtc tcggaagcac gctgggcaag g

31

<210> 59
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 59
gcgagctcta gcatttaggt gacactatag 30

<210> 60
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 60
aactggttta agaaaggaag cacgctgggc gcc 33

<210> 61
<211> 11
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述：备注=合成构建体

<400> 61
Asn Trp Lys Lys Gly Ser Thr Leu Gly Lys Ala
1 5 10

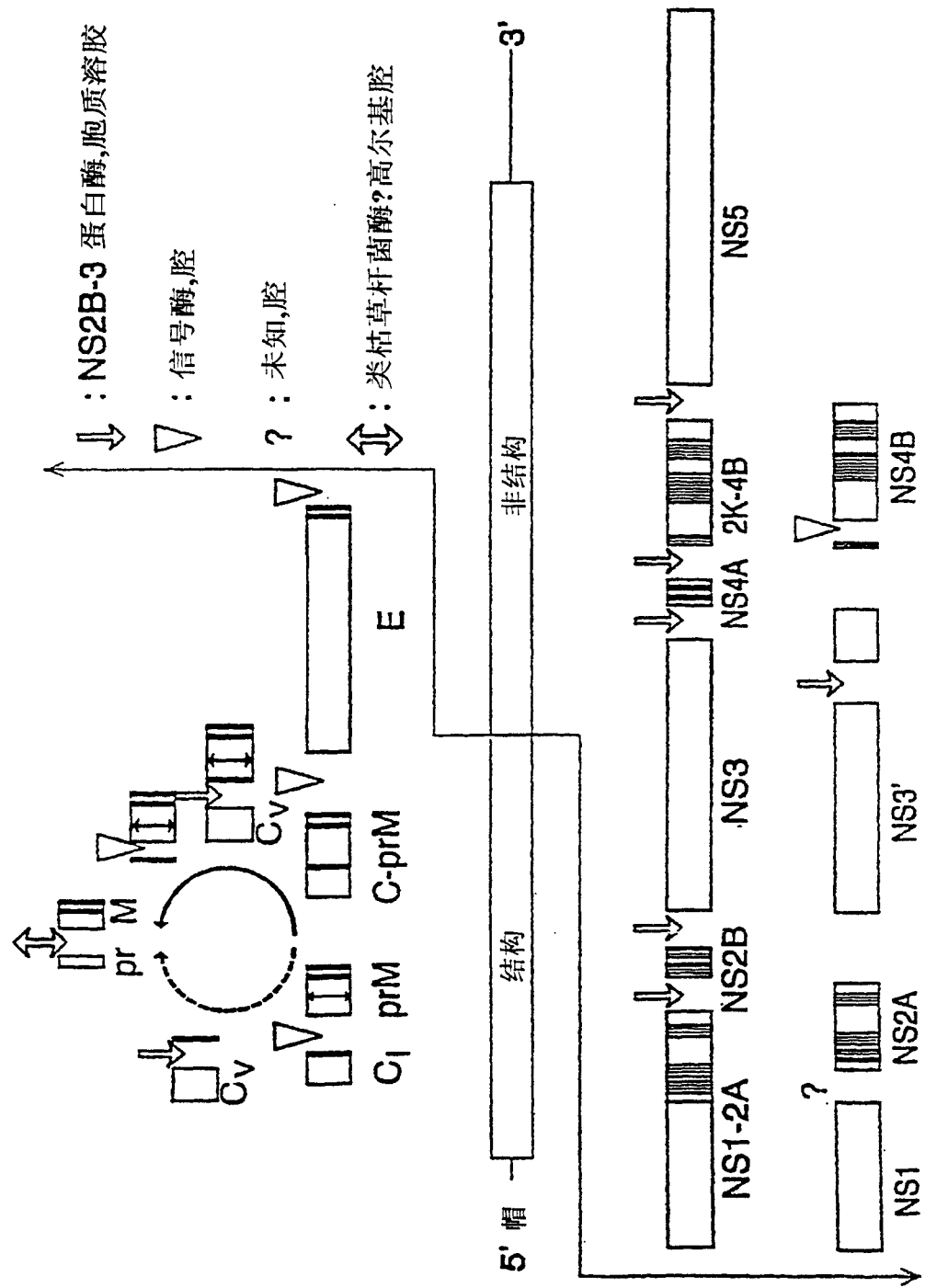


图1

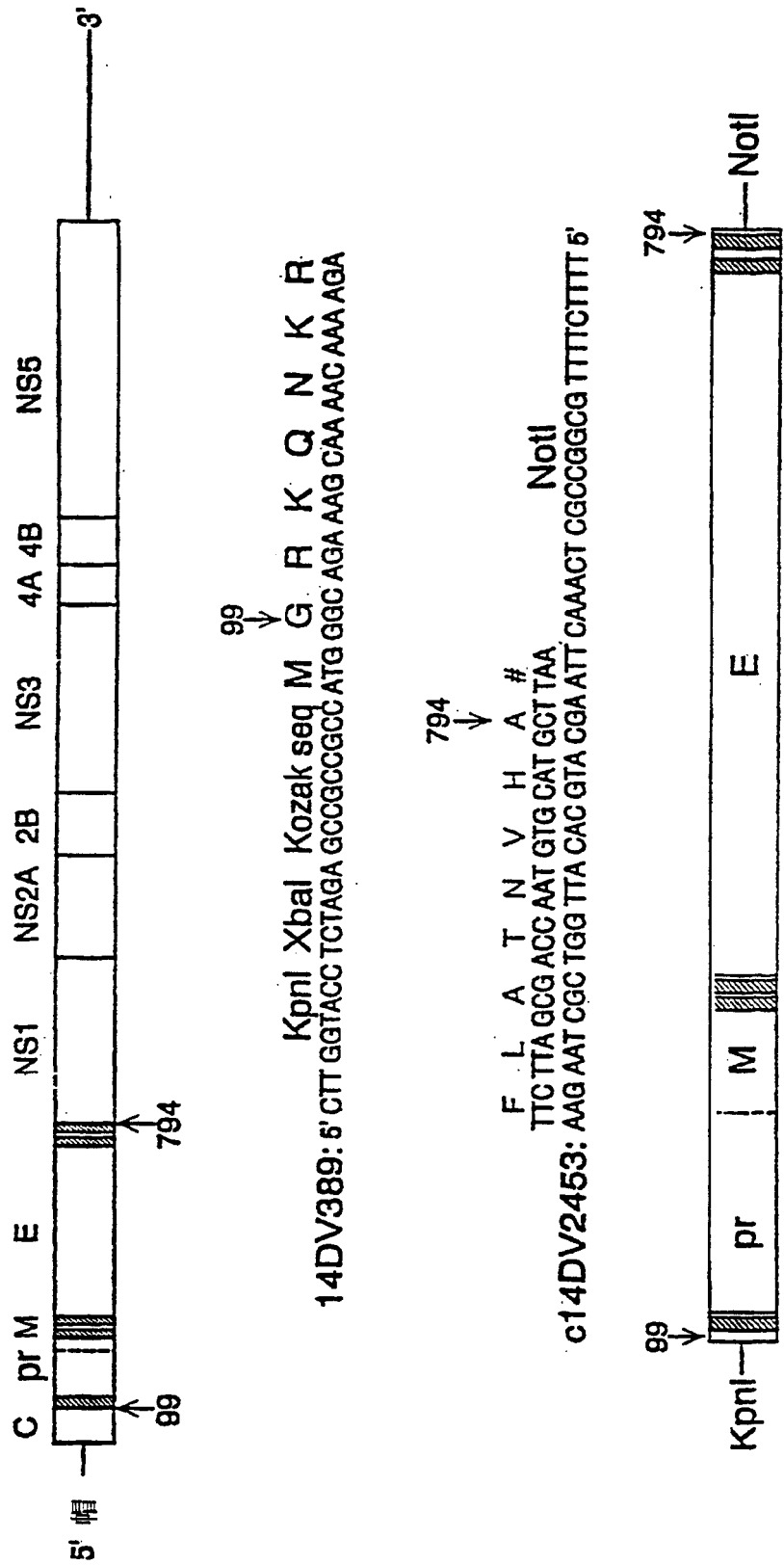


图2

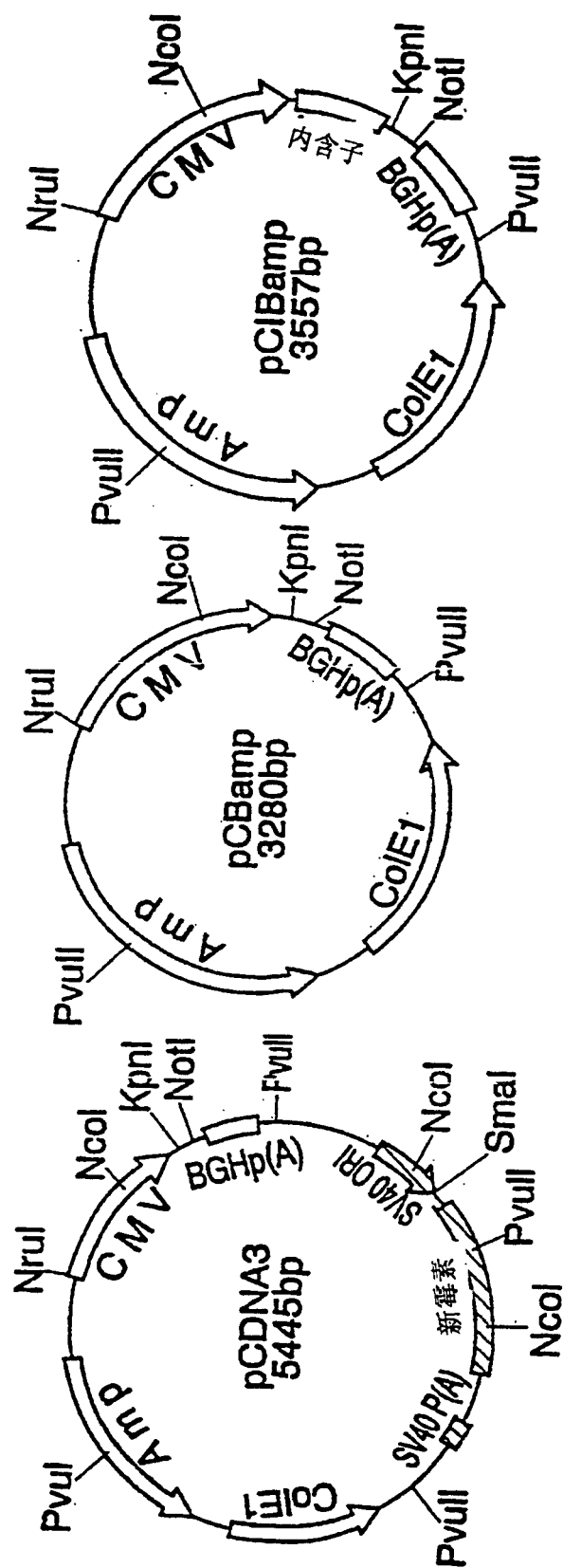


图3

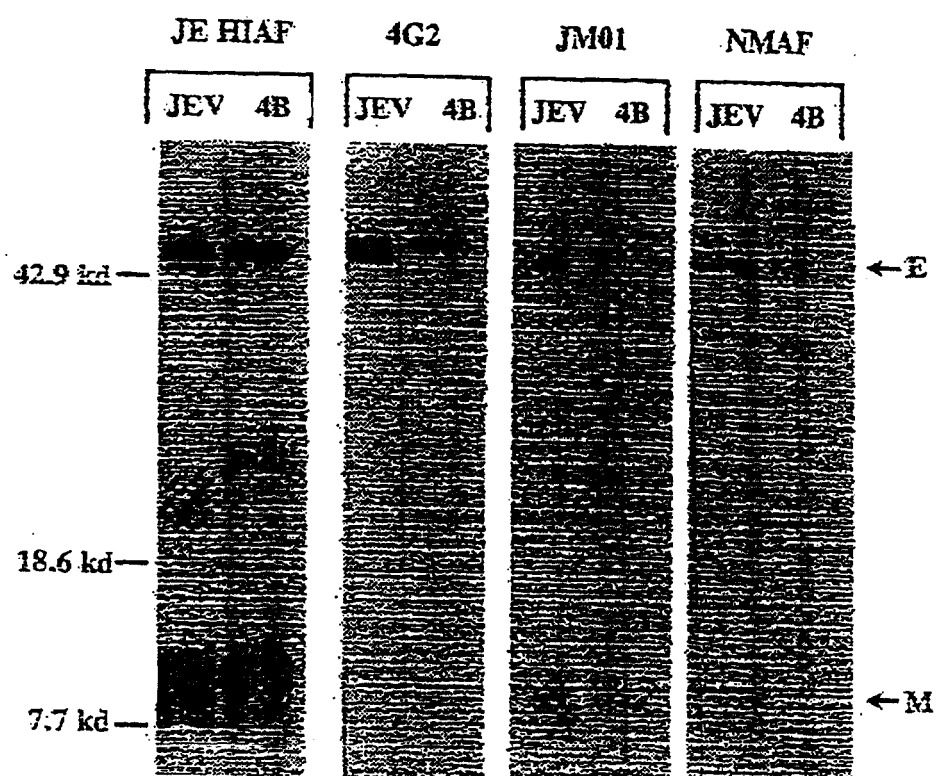


图4

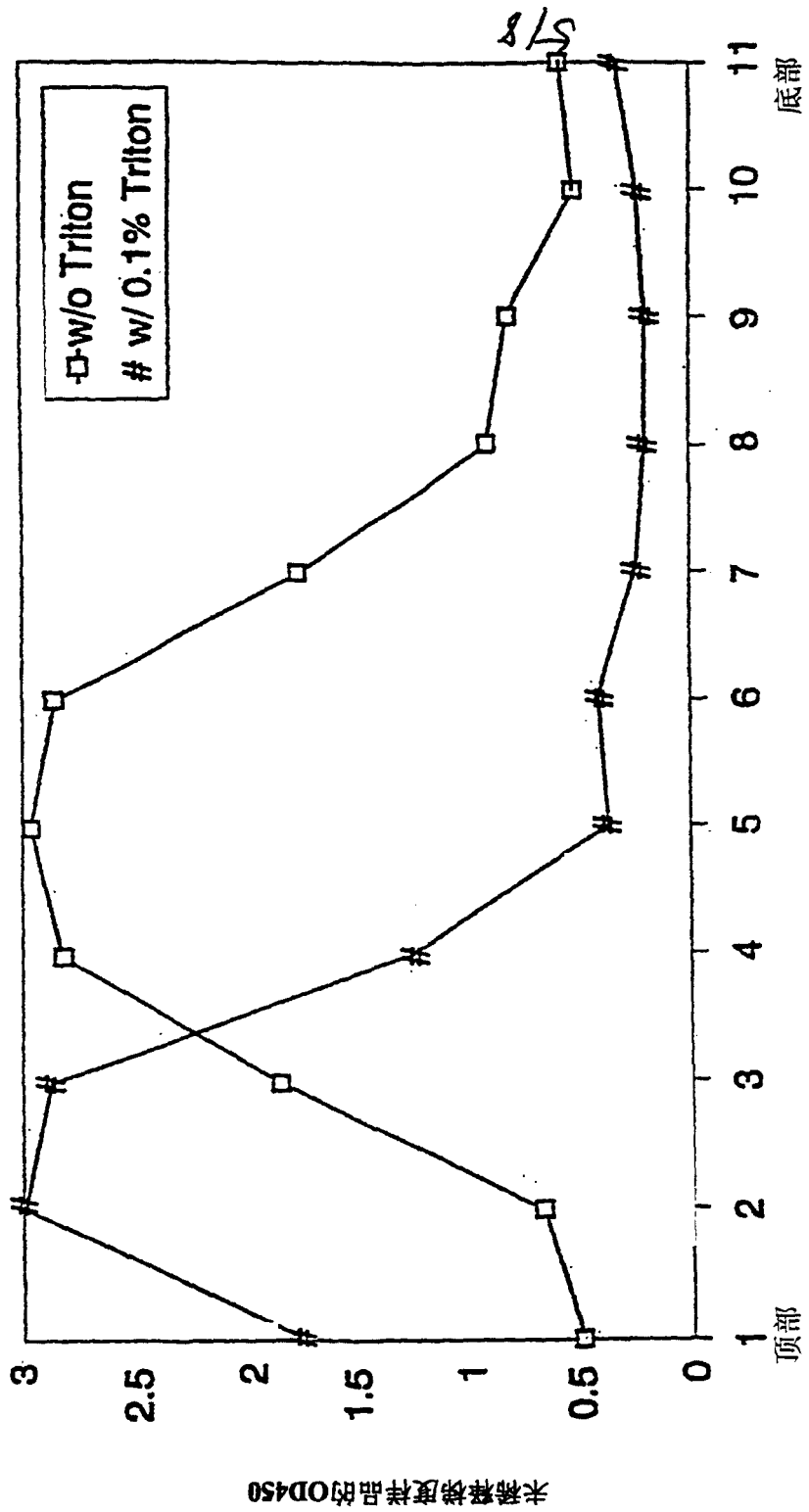


图5

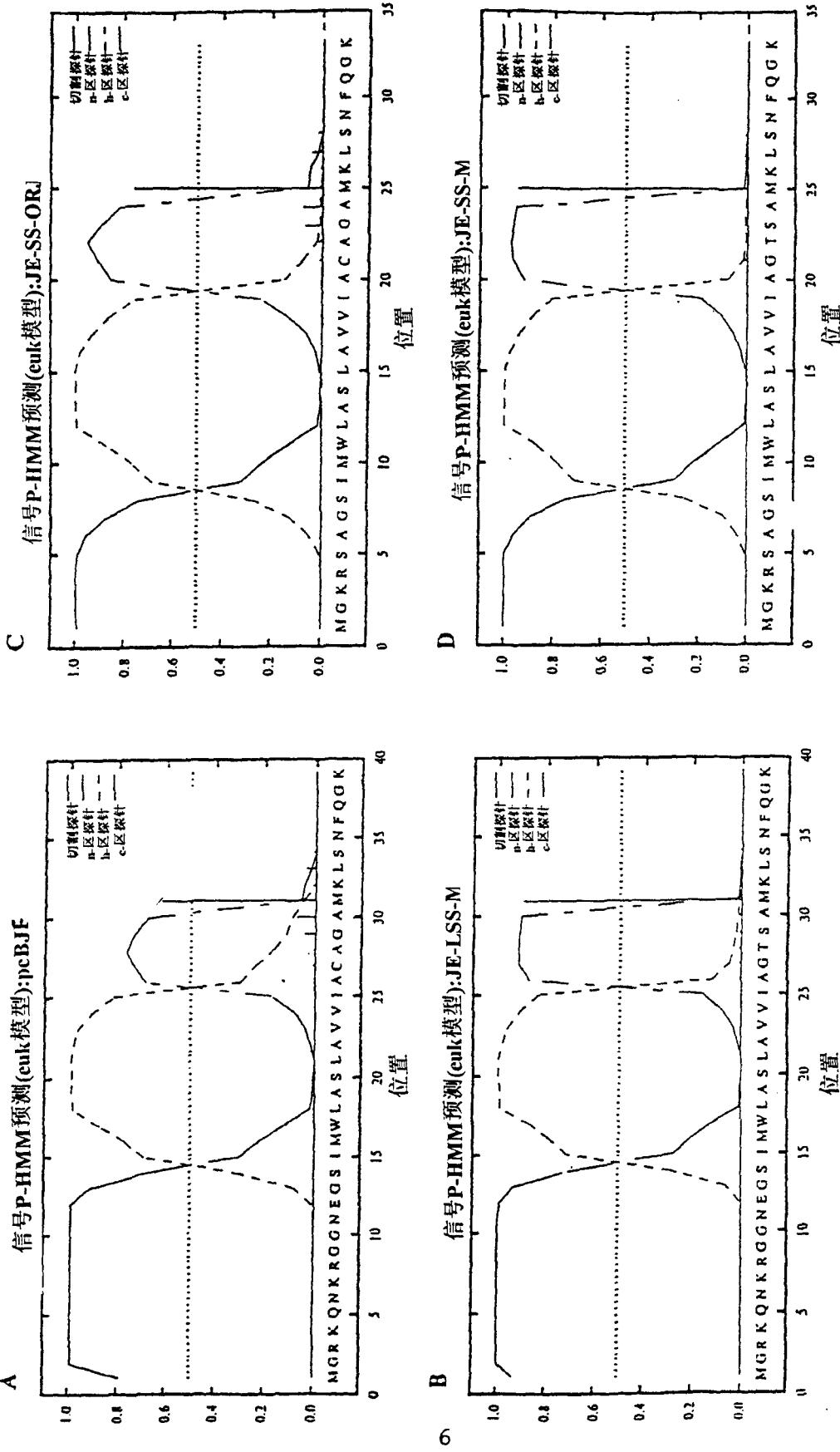


图6

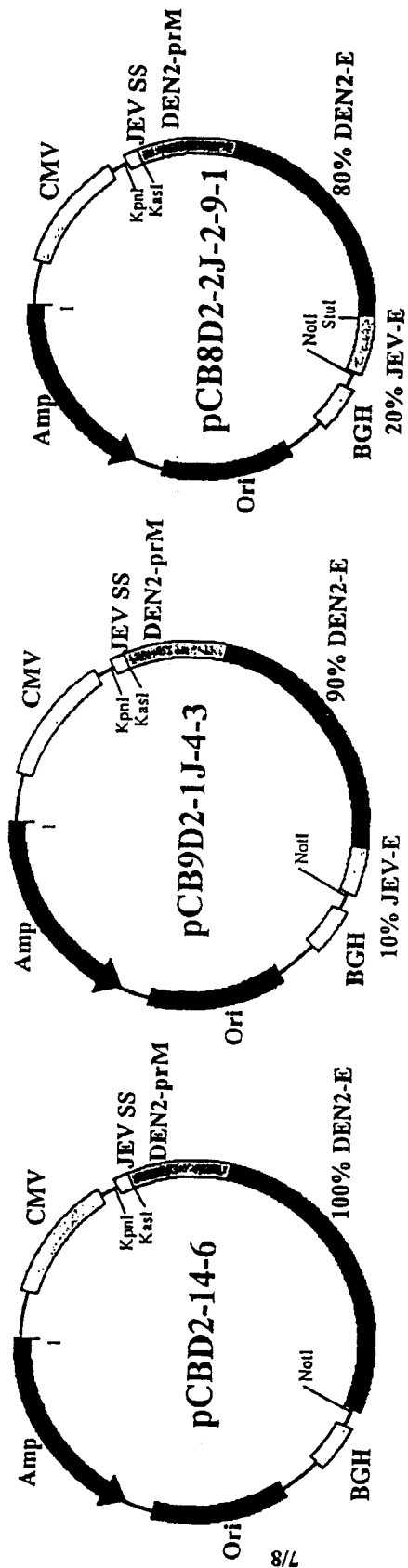


图7

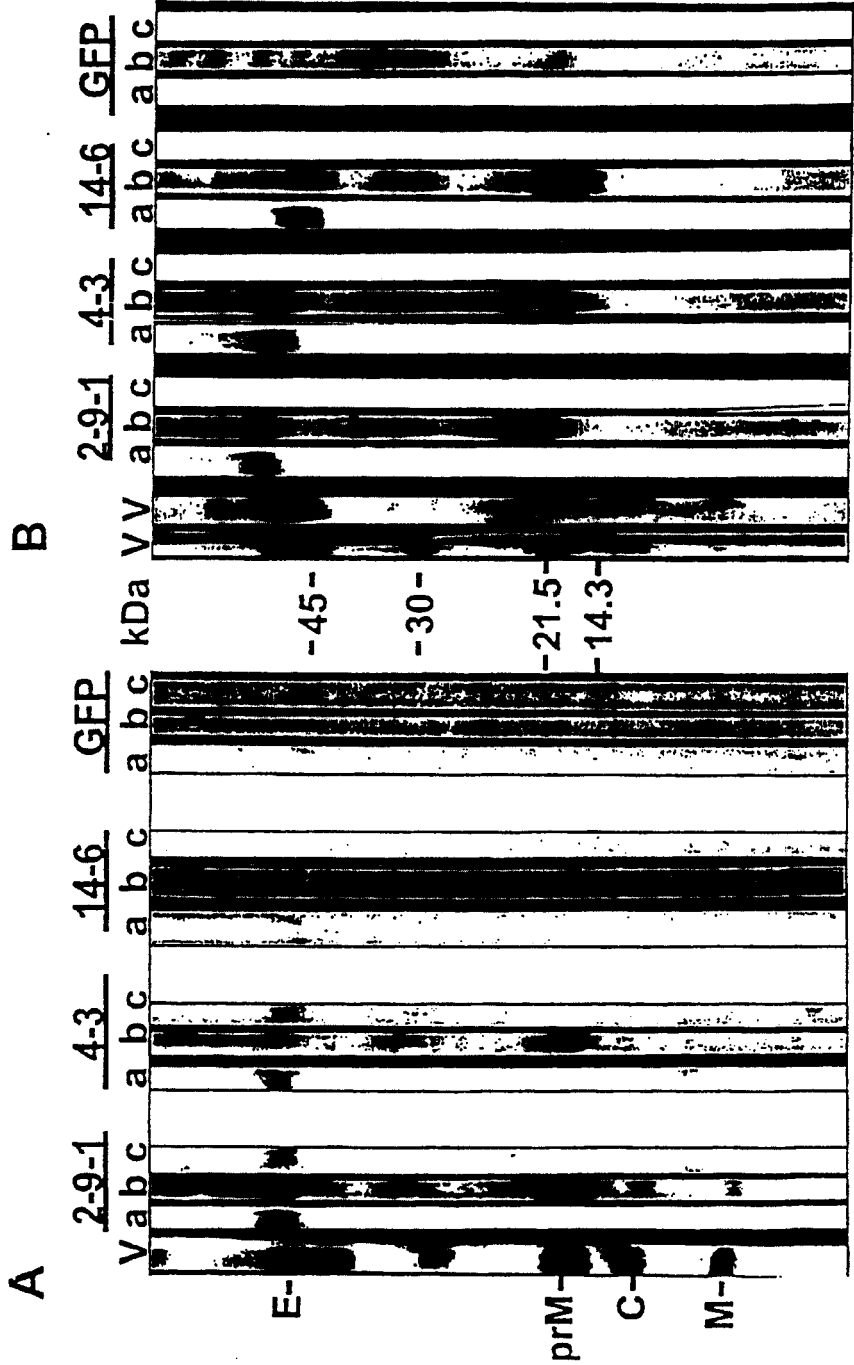


图8

专利名称(译)	预防黄病毒感染的核酸疫苗		
公开(公告)号	CN1304575C	公开(公告)日	2007-03-14
申请号	CN02807758.X	申请日	2002-04-04
[标]申请(专利权)人(译)	美国政府健康及人类服务部，疾病控制和预防中心		
当前申请(专利权)人(译)	美国政府健康及人类服务部 疾病控制和预防中心		
[标]发明人	张光正		
发明人	张光正		
IPC分类号	C12N15/33 A61K45/00 A61P31/12 C12N15/09 A61K31/711 A61K39/00 A61K39/12 A61K48/00 A61P31/14 C07K14/18 C07K16/10 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/40 C12Q1/70 G01N33/53 G01N33/569		
CPC分类号	C12N2770/24122 C07K2319/02 C07K16/1081 A61K39/12 A61K2039/53 A61K48/00 G01N33/56983 G01N2333/18 C12N2770/24022 C12Q1/701 C07K14/005 A61K2039/70 C12N2770/24134 Y02A50/386 Y02A50/388 Y02A50/39 Y02A50/394 Y02A50/53 Y02A50/60		
代理人(译)	丁业平 王维玉		
审查员(译)	吴文英		
优先权	09/826115 2001-04-04 US		
其他公开文献	CN1500152A		
外部链接	Espacenet SIPO		

摘要(译)

本发明包括含有转录单位的分离核酸，该转录单位编码第一种黄病毒的信号序列和第二种黄病毒的免疫原性黄病毒抗原或者含有来自多种黄病毒序列的嵌合免疫原性黄病毒抗原。本发明进一步包括核酸和蛋白疫苗以及该疫苗免疫受试者抗黄病毒感染的应用。本发明还提供了由本发明的核酸编码的抗原，响应于抗原引发产生的抗体以及抗原和/或抗体在检测黄病毒或诊断黄病毒感染中的应用。

100% DEN-2 prM-E:	
D2KasI-438*	5'TGTGCA GGCGC CTTCCATTTAACACACGTAACG (SEQ ID NO:48)
CD2BsmBI-2402	5'TCGA GGCGCGGCTC AACTAATTAGGCTGCACCATGACTC (SEQ ID NO:49)
90% DEN-2 E & 10% JE E:	
T7	5'CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:50)
CD2BsmBI-2244	5'ATAGATTGCT CC AAACACT TGG TGG (SEQ ID NO:51)
JE-2281	5'ACTCCATAGGAAAAGCCGTTCCACC (SEQ ID NO:52)
CSP6	5'GCGAGCTCTAGCAATTAGGTGACACTATAG (SEQ ID NO:53)
90-10 接合区	Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr (SEQ ID NO:55) CTC CAC CAA GTG TTT GGT GGT GCC TTC AGA ACA (SEQ ID NO:54)
80% DEN-2 E & 20% JE E:	
T7	5'CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:56)
CD2BsmBI-2097	5'GAATT CGTCTC ACTTCTTTCTTAAACCAAGTTGAGCTTC (SEQ ID NO:57)
JEbsmBI-2175	5'GGAATT CGTCTC GGAAGCACGCTGGCAAGG (SEQ ID NO:58)
CSP6	5'GCGAGCTCTAGCAATTAGGTGACACTATAG 3' (SEQ ID NO:59)
80-20 接合区	Asn Trp Lys Lys Gly Ser Thr Leu Gly Lys Ala (SEQ ID NO:61) AAC TGG TTT AAG AAA GGA AGC ACG CTG GGC GCC (SEQ ID NO:60)