



(12) 发明专利

(10) 授权公告号 CN 111995664 B

(45) 授权公告日 2024. 08. 02

(21) 申请号 202010272504.4  
(22) 申请日 2012.11.09  
(65) 同一申请的已公布的文献号  
    申请公布号 CN 111995664 A  
(43) 申请公布日 2020.11.27  
(30) 优先权数据  
    61/558,800 2011.11.11 US  
    61/654,157 2012.06.01 US  
(62) 分案原申请数据  
    201280066355.X 2012.11.09  
(73) 专利权人 变异生物技术公司  
    地址 加拿大安大略省  
    专利权人 索邦大学  
(72) 发明人 大卫·E·安德森  
    安妮-凯瑟琳·弗卢基格  
    夏洛特·弗里伯特  
    大卫·克拉茨曼  
(74) 专利代理机构 北京集佳知识产权代理有限公司 11227  
    专利代理师 郑斌 尹玉峰

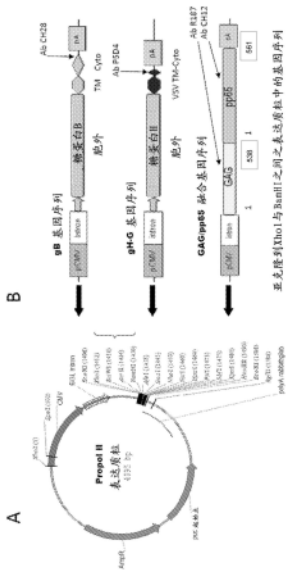
(51) Int.Cl.  
    C07K 14/045 (2006.01)  
    A61K 39/12 (2006.01)  
    A61K 39/245 (2006.01)  
    A61P 31/22 (2006.01)  
(56) 对比文件  
    CN 106102769 A, 2016.11.09  
    US 2009123494 A1, 2009.05.14  
    US 6440730 B1, 2002.08.27  
    WO 2010128338 A2, 2010.11.11  
    GenBank.G protein precursor  
    [Vesicular stomatitis Indiana virus].  
    《GenBank》.2011, ADX53329.1.  
    GenBank.envelope glycoprotein B  
    [Human herpesvirus 5].《GenBank》.2009,  
    ACM48044.1.

审查员 周奋进

权利要求书1页 说明书60页  
序列表67页 附图17页

(54) 发明名称  
    用于治疗巨细胞病毒的组合物和方法  
(57) 摘要

本发明涉及用于治疗巨细胞病毒的组合物和方法。本公开提供了可用于治疗HCMV感染的组合物和方法。如本文所述,所述组合物和方法基于开发包含病毒样颗粒(VLP)的免疫原性组合物,所述病毒样颗粒包含一个或更多个莫洛尼鼠白血病病毒(MMLV)核心蛋白并且包含一个或更多个HCMV表位,例如来自HCMV的包膜糖蛋白gB和/或gH和/或被膜蛋白pp65。除了其他方面以外,本发明涵盖了可导致有益免疫应答的抗原(例如,包膜糖蛋白和结构蛋白)之组合的识别,所述免疫应答包括例如体液应答(例如中和抗体的产生)和细胞应答(例如T细胞活化)二者。



CN 111995664 B

1. 病毒样颗粒 (VLP), 其包含:

第一多肽, 其由SEQ ID NO: 1的氨基酸序列组成; 以及

第二多肽, 其由SEQ ID NO: 10的氨基酸序列组成。

2. 药物组合物, 其包含权利要求1所述的VLP和可药用载体。

3. 权利要求2所述的药物组合物, 其还包含细胞因子。

4. 权利要求2所述的药物组合物, 其还包含凝胶型佐剂。

5. 权利要求2所述的药物组合物, 其还包含微生物佐剂。

6. 权利要求2所述的药物组合物, 其还包含油乳剂佐剂。

7. 权利要求2所述的药物组合物, 其还包含基于乳化剂的佐剂。

8. 权利要求2所述的药物组合物, 其还包含颗粒佐剂。

9. 权利要求2所述的药物组合物, 其还包含合成佐剂。

10. 权利要求2所述的药物组合物, 其还包含聚合物佐剂。

11. 用于生产VLP的方法, 所述方法包括:

用包含编码第一多肽之核苷酸序列的载体和包含编码第二多肽之核苷酸序列的载体共转染宿主细胞, 所述第一多肽由SEQ ID NO: 1的氨基酸序列组成, 所述第二多肽由SEQ ID NO: 10的氨基酸序列组成; 以及

在合适的培养基中于允许表达由所述载体编码之蛋白的条件下培养所述宿主细胞。

12. 权利要求2所述的药物组合物在制造用于在对象中减少HCMV感染之症状的频率、严重程度或延迟所述症状发作之药物中的用途。

## 用于治疗巨细胞病毒的组合物和方法

[0001] 本申请是申请号为201280066355.X的中国专利申请的分案申请,原申请是2012年11月9日提交的PCT国际申请PCT/IB2012/002854于2014年7月4日进入中国国家阶段的申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求于2011年11月11日提交的美国临时申请号61/558,800和于2012年6月1日提交的美国临时申请号61/654,157的优先权,二者的内容通过引用整体并入本文。

### 技术领域

[0004] 本申请涉及用于治疗巨细胞病毒的组合物和方法。

### 背景技术

[0005] 人巨细胞病毒(HCMV)( $\beta$ -疱疹病毒)是一种无所不在的病原体。在免疫正常的人中HCMV感染通常被忽视,其具有最温和的且非特异性的症状。相比之下,某些高危人群,例如在免疫抑制患者(例如AIDS患者或移植受者)中,以及产前感染后,HCMV感染具有严重的表现(Staras SA等,2006 Clin Infect Dis 43(9):1143-51;Hebart H等,2004 Hum Immunol 65(5):432-6;Rowshani AT等,2005 Transplantation 79(4):381-6)。现有的治疗方法包括使用免疫球蛋白和抗病毒制剂(例如更昔洛韦及其衍生物),其在高危人群中预防性使用或在非常早期的感染过程中使用时是最有效的。然而,现有疗法的特征为显著的毒性和有限的疗效,尤其是对晚发性疾病(Boeckh M.,2004 Pediatr Transplant 8(增刊5):19-27;Limaye AP.,2004 Transplantation 78(9):1390-6),而且他们对先天性HCMV疾病尚无影响。开发有效的疫苗来防止HCMV疾病被认为是重要的公共卫生优先事项(Arvin AM等,2004 Clin Infect Dis 39(2):233-9)。

### 发明内容

[0006] 除了其他方面以外,本发明提供了可用于预防、治疗和/或研究人巨细胞病毒(HCMV)感染的方法和组合物。在一些实施方案中,本发明提供了病毒样颗粒(virus-like particles,VLP),其包含一个或多个莫洛尼鼠白血病病毒(Moloney Murine leukemia virus,MMLV)核心蛋白并且包含一个或多个HCMV表位,例如来自HCMV的包膜糖蛋白gB和/或gH和/或被膜蛋白pp65。除了其他方面以外,本发明涵盖了抗原(例如,包膜糖蛋白和结构蛋白)之组合可导致有益的免疫应答之诱导改善,所述免疫应答包括例如体液应答(例如中和抗体的产生)和细胞应答(例如T细胞活化)二者。可提供的VLP特征在于他们不含有病毒DNA并且无感染性。

[0007] 在一些实施方案中,提供的VLP被脂质膜包围,任选地含有来自病毒包膜糖蛋白(例如gB和/或gH)的一个或多个表位,其是在病毒中和抗体的诱导中发挥作用的抗原。

[0008] 在一些实施方案中,提供的VLP含有来自病毒结构蛋白(例如pp65)的一个或多个表位,其是在诱导细胞免疫应答(例如T细胞应答)中发挥作用的抗原。在一些实施方案

中,利用的病毒结构蛋白(例如pp65)不仅刺激T辅助细胞的形成还诱导抗HCMV的细胞毒性T淋巴细胞(CTL)。

[0009] 在一些实施方案中,本发明提供了病毒包膜糖蛋白(例如gB和/或gH)的变体。在一些实施方案中,变体病毒包膜糖蛋白为融合蛋白或包含融合蛋白。在一些实施方案中,病毒糖蛋白的变体包含异源蛋白结构域(例如,来自不同蛋白的跨膜和/或胞质结构域)。在一些实施方案中,病毒结构蛋白的变体包含异源抗原或表位。在一些实施方案中,本发明提供了包含病毒结构蛋白之变体的VLP。在一些实施方案中,病毒结构蛋白的变体为融合蛋白或包含融合蛋白。

[0010] 本申请中所用的术语“约”和“大约”等同使用。在本申请中使用的带有或不带有约/大约的任意数值意味着涵盖被相关领域的普通技术人员理解的任何正常波动。

[0011] 本发明的其他特性、目的和益处将在以下的详细描述中将变得明显。然而,应当理解,尽管所述详细描述显示了本发明的实施方案,其仅通过举例说明的方式给出而非限制。从该详细描述,本发明范围内的多种改变和修饰对本领域技术人员将变得显而易见。

## 附图说明

[0012] 附图仅用于举例说明目的而非限制。

[0013] 图1示出DNA表达质粒图谱(A)和示例性重组表达质粒的构建(B)。

[0014] 图2示出HEK293包装细胞上示例性异源表面抗原的FACS分析(A)和示例性VLP组合物中异源抗原表达的蛋白质印迹分析(B)。

[0015] 图3的(A)和(B)示出两种示例性VLP组合物的颗粒大小测定和多分散指数。

[0016] 图4示出在用示例性gB/pp65(A)、gH-G/pp65(B)或gB/gH-G/pp65(C)VLP组合物处理之小鼠中的ELISA效价。

[0017] 图5示出用示例性VLP组合物处理的小鼠中的中和抗体活性(在人成纤维细胞中测定)。

[0018] 图6示出用示例性VLP组合物处理的小鼠中的中和抗体活性(在人上皮细胞中测定)。

[0019] 图7示出用示例性VLP组合物相对于重组gB蛋白处理的小鼠中的中和抗体活性(在人成纤维细胞中测定)。

[0020] 图8示出用HEK293细胞中表达的示例性VLP组合物处理之小鼠中的抗原特异性CTL应答,其描述为基于CFSE衰减、选通CD3<sup>+</sup>CD8<sup>+</sup>T细胞的CTL频率(A)或pp65-特异性CTL增殖的频率(B)。

[0021] 图9示出用HEK293细胞中表达的示例性VLP组合物处理的兔中的抗-gB和中和抗体效价(在人成纤维细胞中测定)。

[0022] 图10示出用HEK293细胞中表达的示例性VLP组合物处理的兔中的中和抗体效价(在人上皮细胞中测定)。

[0023] 图11示出通过(A)切向流过滤(TFF)或(B)阴离子交换(AEX)色谱纯化的CHO细胞中表达的示例性VLP组合物之负染电子显微镜(EM)图像。

[0024] 图12示出用在CHO细胞中表达并通过切向流过滤(TFF)和阴离子交换(AEX)色谱纯化的示例性VLP组合物处理之兔中的中和抗体效价(在人成纤维细胞中测定)。

[0025] 图13示出用通过切向流过滤 (TFF) 和阴离子交换 (AEX) 色谱纯化的CHO细胞中表达的示例性VLP组合物处理之兔中的中和抗体效价 (在人上皮细胞中测定)。

[0026] 图14示出用在CHO细胞中表达并通过切向流过滤 (TFF) 和阴离子交换 (AEX) 色谱纯化的示例性VLP组合物处理之兔中产生的抗体之亲和力指数。

## 具体实施方式

[0027] 定义

[0028] 为了本发明变得更易于理解,下文中首先对某些术语进行定义。以下术语和其他术语的另外定义列在整个说明书中。

[0029] 氨基酸:本文所用的术语“氨基酸”以其最广泛的含义,指可被并入多肽链的任何化合物和/或物质。在一些实施方案中,氨基酸具有通式结构 $\text{H}_2\text{N}-\text{C}(\text{H})(\text{R})-\text{COOH}$ 。在一些实施方案中,氨基酸是天然存在的氨基酸。在一些实施方案中,氨基酸是合成的氨基酸;在一些实施方案中,氨基酸是d-氨基酸;在一些实施方案中,氨基酸是l-氨基酸。“标准氨基酸”指任何在天然存在之肽中常规存在的20种标准l-氨基酸。“非标准氨基酸”指除了标准氨基酸之外的任何氨基酸,无论是合成制备的还是从天然来源获得的。本文所用的“合成氨基酸”涵盖化学修饰的氨基酸,包括但不限于盐、氨基酸衍生物(例如酰胺)和/或取代。氨基酸(包括肽中羧基和/或氨基末端氨基酸)可通过甲基化、酰胺化、乙酰化、保护基团和/或用其他化学基团取代进行修饰,这可改变肽的循环半衰期且不负面影响其活性。氨基酸可以参与二硫键。氨基酸可以包含一种或更多种翻译后修饰,例如与一种或更多种化学实体(例如甲基基团、乙酸盐/酯基团、乙酰基基团、磷酸盐/酯基团、甲酰基部分、类异戊二烯基团、硫酸盐/酯基团、聚乙二醇部分、脂质部分、碳水化合物部分、生物素部分等)结合。术语“氨基酸”与“氨基酸残基”可互换使用,并且可以指游离的氨基酸和/或肽的氨基酸残基。从使用该术语的上下文中,指游离氨基酸还是指肽的残基将会是显而易见的。

[0030] 抗原:本文所使用的术语“抗原”是指包含有一种或更多种被抗体识别的表位(线性表位、构象表位或两者)的物质。在某些实施方案中,抗原是或包含病毒或病毒多肽。在一些实施方案中,术语“抗原”指亚单位抗原(即从完整病毒分开或分离的抗原,所述抗原与所述完整病毒本质上相关;例如与病毒样颗粒相关的抗原)。或者或此外,在一些实施方案中,术语“抗原”指杀死的、减毒或灭活的病毒。在某些实施方案中,抗原是“免疫原”。

[0031] 大约或约:本文所用的应用至一个或更多个目的值的术语“大约”或“约”,指类似于所述参考值的值。在某些实施方案中,术语“大约”或“约”指落入所述参考值之任意方向(大于或小于)的25%、20%、19%、18%、17%、16%、15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更小之内(除了这样的数字将超过可能值的100%的情况)的一系列值。

[0032] 改善:本文所用的术语“改善”意味着预防、减少或减轻状态,或改进对象的状态。改善包括但不需要完全恢复或完全预防疾病、病症或病状(例如HCMV感染)。术语“预防”指延缓疾病、病症或病状的发作。当疾病、病症或病状的发作已经被延迟了预定的时间段时,可以视为完成了预防。

[0033] 特征部分:本文所用的术语物质的“特征部分”,以最广泛的含义,是享有完整物质的指定程度之结构同一性的一部分。在某些实施方案中,特征部分享有完整物质的至少一

个功能特征。例如,蛋白质或多肽的“特征部分”是含有氨基酸之连续延伸,或者氨基酸连续延伸之集合的一部分,其一起构成了蛋白质或多肽的特征。在一些实施方案中,每个这种连续延伸通常含有至少2、5、10、15、20、50或更多个氨基酸。一般而言,物质(例如蛋白质或抗体等)的特征部分是(除以上详细说明的序列和/或结构同一性)享有相关的完整物质之至少一个功能特征的部分。在一些实施方案中,特征部分可以为生物学活性。

[0034] 特征序列:“特征序列”是指存在于多肽或核酸家族的所有成员中的序列,并且由此可被本领域的普通技术人员用来定义家族的成员。

[0035] 胞质结构域:如本领域已知的,多肽有时具有跨膜、胞质和/或胞外结构域。一般而言,本文所用的“胞质结构域”指具有存在于胞质中的性质的结构域。应当理解,不需要胞质结构域中的每一个氨基酸都存在于胞质中。例如,在一些实施方案中,胞质结构域的特征在于蛋白质的指定延伸或部分基本上位于胞质中。如本领域中公知的,可以使用多种算法分析氨基酸或核酸序列以预测蛋白亚细胞定位(例如胞质定位)。示例性的这样的程序包括psort(PSORT.org)、Prosites(prosite.expasy.org)等等。

[0036] 剂型:本文所用的术语“剂型”和“单位剂型”指用于待治疗之患者的治疗剂的物理不连续单位。各个单位含有预定量的经计算活性物质以产生期望的治疗效果。然而,应当理解,组合物的总剂量将由主治医师在合理的医学判断范围内决定。

[0037] 给药方案:本文所用的术语“给药方案”(或“治疗方案”)是单独施用至对象的通常被时间段隔开的一组单位剂量(通常多于一个)。在一些实施方案中,给定的治疗剂具有推荐的给药方案,其可涉及一个或更多次给药。在一些实施方案中,给药方案包括多次给药,各次之间被相同长度的时间段相互隔开;在一些实施方案中,给药方案包括多次给药以及将各次给药分开的至少两种不同的时间段。

[0038] 表达:如本文使用的核酸序列的“表达”指一个或更多个以下事件:(1)从DNA序列产生RNA模板(例如通过转录);(2)加工RNA转录物(例如通过剪切、编辑、5'帽形成和/或3'尾形成);(3)将RNA翻译成多肽或蛋白;和/或(4)多肽或蛋白质的翻译后修饰。

[0039] 胞外结构域:如本领域已知的,多肽有时具有跨膜、胞质和/或胞外结构域。一般而言,本文所用的“胞外结构域”指具有存在于细胞之外部的性质的结构域。应当理解,不需要胞外结构域中的每个氨基酸都存在于细胞的外部。例如,在一些实施方案中,胞外结构域的特征在于蛋白的指定延伸或部分基本上位于细胞外。如本领域中公知的,可以使用多种算法分析氨基酸或核酸序列以预测蛋白亚细胞定位(例如细胞外定位)。示例性的这样的程序包括psort(PSORT.org)、Prosites(prosite.expasy.org)等等。

[0040] 融合蛋白:本文所用的术语“融合蛋白”一般指包含至少两个片段的多肽,所述片段各自显示出与(1)天然存在的和/或(2)代表多肽的功能结构域的肽部分具有高度的氨基酸同一性。通常,含有至少两个这样的片段的多肽被认为是融合蛋白,如果所述两个片段是这样的部分:(1)不包括在天然的同肽中,和/或(2)之前在单个多肽中未相互连接,和/或(3)已经通过人手工作用相互连接。

[0041] 基因:术语“基因”具有其本领域理解的含义。本领域的普通技术人员应当理解,术语“基因”可以包括基因调控序列(例如启动子、增强子等)和/或内含子序列。还应当理解,基因的定义包括涉及不编码蛋白而编码功能RNA分子(例如tRNA、RNAi-诱导剂等)的核酸。出于澄清的目的,我们注释本说明书中使用的术语“基因”一般指编码蛋白的核酸部分;本

领域的普通技术人员将从上下文中明白,该术语可任选地涵盖调控序列。这一定义不旨在排除将术语“基因”应用于非蛋白编码表达单位,而是为了澄清在大多数情况下,在本文中使用的该术语指编码蛋白的核酸。

[0042] 基因产物或表达产物:本文所用的术语“基因产物”或“表达产物”一般指从基因(加工前和/或加工后)转录的RNA或者由从该基因转录的RNA编码的多肽(修饰前或修饰后)。

[0043] 异源:对于蛋白或多肽,本文所用的术语“异源”指非天然存在于特定有机体中的蛋白或多肽,例如反转录病毒或VLP。在一些实施方案中,异源蛋白或多肽在特定的反转录病毒病毒体中是非天然存在的。对于蛋白结构域,本文所用的术语“异源”一般指在特定蛋白中非天然存在的蛋白结构域。

[0044] 免疫原性:本文所用的术语“免疫原性”意味着能够在宿主动物中针对非宿主实体(例如HCMV抗原)产生免疫应答。在某些实施方案中,这种免疫应答形成了通过针对特定感染性生物(例如HCMV)的疫苗引起之保护性免疫的基础。

[0045] 免疫应答:本文所用的术语“免疫应答”指在动物中引起的应答。免疫应答可以指细胞免疫、体液免疫或者可以涉及二者。还可免疫应答限制在免疫系统的一部分。例如,在某些实施方案中,免疫原性组合物可以诱导IFN  $\gamma$  应答增加。在某些实施方案中,免疫原性组合物可以诱导黏膜IgA应答(例如,在鼻和/或直肠清洗液中测量的)。在某些实施方案中,免疫原性组合物可以诱导全身的IgG应答(例如,在血清中测量的)。在某些实施方案中,免疫原性组合物可以诱导病毒中和抗体或者中和抗体应答。

[0046] 改善、增加或减少:本文所用的术语“改善”、“增加”或“减少”或语法等价物,指示相对于基线测量的值,所述基线测量例如在本文所述的处理开始之前在相同的个体中的测量或在缺乏本文所述之处理的情况下在对照个体(或多个对照个体)中的测量。

[0047] 个体、对象、患者:本文所用的术语“对象”、“个体”或“患者”指人或非人哺乳动物对象。在一些实施方案中,被治疗的个体(也称为“患者”或“对象”)是患有疾病(例如HCMV感染)的个体(胎儿、婴儿、儿童、青少年或成人)。在一些实施方案中,对象有HCMV感染的风险。在一些实施方案中,所述对象是免疫抑制的对象。例如,在一些实施方案中,免疫抑制的对象选自HIV感染的对象、AIDS患者、移植受者、儿童对象以及妊娠对象。在一些实施方案中,所述对象已经暴露于HCMV感染。在一些实施方案中,所述对象是人。

[0048] 分离的:本文所用的术语“分离的”指如下物质和/或实体,其已经(1)从在初始产生(无论在自然和/或在实验环境中)时与其相关联的至少一些组分分离,和/或(2)通过人手工产生、制备和/或制造。分离的物质和/或实体可以从初始与其相关的其他组分的约10%、约20%、约30%、约40%、约50%、约60%、约70%、约80%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或大于约99%分离。在一些实施方案中,分离的药剂为约80%、约85%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或者大于约99%纯的。如果物质基本上没有其他组分,那么该物质是“纯的”。如本文所述的,分离的物质和/或实体之百分比纯度的计算不应当包括赋形剂(例如,缓冲液、溶剂、水等)。

[0049] 接头:本文所用的术语“接头”指(例如在融合蛋白中)非在天然蛋白的特定位置出现的适当长度的氨基酸序列,并且一般被设计为柔性和/或插入结构(例如两个蛋白部分之

间的 $\alpha$ -螺旋)。一般而言,接头允许融合蛋白的两个或更多个结构域保留各自结构域之50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或更多的生物活性。接头还可以指间隔子。

[0050] 核酸:本文所用的术语“核酸”,以其最广泛的含义,指被并入或能够并入寡核苷酸链中的任何化合物和/或物质。在一些实施方案中,核酸是通过磷酸二酯键被并入或能够并入寡核苷酸链的化合物和/或物质。在一些实施方案中,“核酸”指单个核苷酸残基(例如核苷酸和/或核苷)。在一些实施方案中,“核酸”指包含单个核苷酸残基的寡核苷酸链。本文所用的术语“寡核苷酸”和“多核苷酸”可互换使用。在一些实施方案中,“核酸”涵盖RNA以及单链和/或双链DNA和/或cDNA。此外,术语“核酸”、“DNA”、“RNA”和/或类似的术语包括核酸类似物,例如具有非磷酸二酯骨架的类似物。例如,在本领域已知并且在骨架中具有肽键而非磷酸二酯键的所谓的“肽核酸”被视为在本发明的范围内。术语“编码氨基酸序列的核苷酸序列”包括为彼此间的兼并版本(degenerate version)和/或编码相同氨基酸序列的所有核苷酸序列。编码蛋白和/或RNA的核苷酸序列可以包括内含子。核酸可以从天然来源纯化,使用重组表达系统产生并任选地纯化,化学地合成等。其中,例如在化学合成分子的情况下,合适的核酸分子可包括核苷酸类似物,例如具有经化学修饰的碱基或糖、骨架修饰等的类似物。除非另有说明,核酸序列以5'至3'方向给出。本文所用的术语“核酸片段”是指为较长核酸序列之一部分的核酸序列。在许多实施方案中,核酸片段包含至少3、4、5、6、7、8、9、10或更多个残基。在一些实施方案中,核酸是或包含天然核苷(例如,腺苷、胸苷、鸟苷、胞苷、尿苷、脱氧腺苷、脱氧胸苷、脱氧鸟苷和脱氧胞苷);核苷类似物(2-氨基腺苷、2-硫代胸苷、肌苷、吡咯并嘧啶、3-甲基腺苷、5-甲基胞苷、C-5丙炔基-胞苷、C-5丙炔基-尿苷、2-氨基腺苷、C5-溴尿苷、C5-氟尿苷、C5-碘尿苷、C5-丙炔基-尿苷、C5-丙炔基-胞苷、C5-甲基胞苷、2-氨基腺苷、7-脱氮腺苷、7-脱氮鸟苷、8-氧代腺苷、8-氧代鸟苷、0(6)-甲基鸟嘌呤和2-硫代胞苷);化学修饰的碱基;生物修饰的碱基(例如甲基化的碱基);插入的碱基;修饰的糖(例如2'-氟核糖、核糖、2'-脱氧核糖、阿拉伯糖和己糖);和/或经修饰的磷酸盐/酯基团(例如硫代磷酸盐/酯和5'-N-亚磷酰胺键)。在一些实施方案中,本发明具体地涉及“未经修饰的核酸”,意指核酸(例如多核苷酸和残基,包括核苷酸和/或核苷)尚未被化学修饰以利于或实现递送。

[0051] 可药用:本文所用的术语“可药用”指在合理的医学判断范围内,适合用于与人类和动物的组织接触而没有过度的毒性、刺激性、过敏反应或其它问题或并发症,与合理的利益/风险比相当的物质。

[0052] 多肽:本文所用的“多肽”一般而言是至少两个氨基酸通过肽键彼此连接的串。在一些实施方案中,多肽可以包含至少3至5个氨基酸,每一个通过至少一个肽键与其他相连接。本领域的普通技术人员应当理解,任选地,多肽有时包括“非天然”氨基酸或者仍能够并入多肽链的其他部分。

[0053] 多蛋白:本文所用的术语“多蛋白”一般指在合成后可以被切割以产生几个功能不同的多肽的蛋白。多蛋白通常由单个氨基酸序列编码。在一些实施方案中,未切割的多蛋白保留其组成部分的生物活性。一些病毒产生这样的多蛋白,例如,Gag多蛋白,其可作为功能多蛋白保留或者可被加工成几个功能不同的多肽。功能上,Gag多肽被分成3个结构域:膜结合结构域,其使Gag多蛋白靶向细胞膜;相互作用结构域,其促进Gag聚合;以及晚期结构域,



其利于新生病毒体从宿主细胞释放。一般而言,介导病毒颗粒组装的Gag蛋白的形式是多蛋白。

[0054] 自组装部分:一般而言,本文所用的“自组装部分”指采取限定的排列而无需来自外源的指导或管理之实体的相关延伸。在一些实施方案中,实体是蛋白。在一些实施方案中,实体是多蛋白。在一些这样的实施方案中,相关延伸为至少10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500或更多个残基。自组装可以在例如在细胞的内容物内(例如体内)展示。或者或此外,自组装可以在细胞内容物外部(例如体外)展示。自组装可以为分子内(例如折叠)和/或分子间。在一些实施方案中,自组装可以是巨大分子,实体由此自组装成复合物和/或延展的巨大分子结构。自组装实体可以表现出广泛的结构模体(motif),包括但不限于颗粒、纤维、片层和带。在一些实施方案中,实体的自组装对于实体的生物学功能是重要的。例如,在一些实施方案中,脂质的自组装导致细胞膜结构的形成。在一些实施方案中,蛋白(例如病毒结构蛋白)的自组装导致颗粒结构(例如,病毒颗粒结构)的形成。例如,病毒结构多蛋白可以含有靶向序列,其能够引导其定位至其宿主细胞的细胞膜(例如质膜、内体等),病毒结构多蛋白可以从所述宿主细胞出芽形成VLP,其含有包围病毒结构多蛋白的宿主细胞膜材料。

[0055] 基本同源性(Substantial homology):本文所用的短语“基本同源性”指氨基酸或核酸序列之间的比较。本领域的技术人员应当理解,如果两个序列在对应的位置含有同源残基,那么他们一般被认为是“基本同源”。同源残基可以为相同残基。或者,同源残基可以为具有大致相似的结构和/或功能特征的非相同残基。例如,本领域的普通技术人员公知,某些氨基酸通常分类为“疏水的”或“亲水的”氨基酸,和/或具有“极性”或“非极性”侧链。相同类型的一个氨基酸对另一个的替代经常可以认为是同源替代。

[0056] 如本领域所公知的,可以用多种算法的任一种进行氨基酸或核酸序列的比较,包括那些市售的计算机程序,例如用于核苷酸序列的BLASTN和用于氨基酸序列的BLASTP、缺口的BLAST和PSI-BLAST。示例性的这样的程序描述于Altschul,等,Basic local alignment search tool,J.Mol.Biol.,215(3):403-410,1990;Altschul,等,Methods in Enzymology266:460-480(1996);Altschul,等,"Gapped BLAST and PSI-BLAST:a new generation of protein database search programs",Nucleic Acids Res.25:3389-3402,1997;Baxevanis,等,Bioinformatics:A Practical Guide to the Analysis of Genes and Proteins,Wiley,1998;and Misener,等,(编辑),Bioinformatics Methods and Protocols(Methods in Molecular Biology,132卷),Humana Press,1999。除了鉴别同源序列,上述程序通常提供同一性程度的指示。在一些实施方案中,如果在残基的相关延伸中其对应残基的至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多为同源,那么两个序列被认为是基本同源。在一些实施方案中,相关延伸是整个序列。在一些实施方案中,相应延伸为至少10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500或更多个残基。

[0057] 基本同一性:本文所用的短语“基本同一性”指氨基酸或核酸序列之间的比较。本领域的普通技术人员应当理解,如果两个序列在对应位置含有相同的残基,那么他们一般

被认为是“基本同一”。本领域公知,可以用多种算法中的任一种进行氨基酸或核酸序列的比较,包括那些市售的计算机程序,例如用于核苷酸序列的BLASTN和用于氨基酸序列的BLASTP、缺口的BLAST和PSI-BLAST。示例性的这样的程序描述于Altschul,等,Basic local alignment search tool,J.Mol.Biol.,215(3):403-410,1990;Altschul,等,Methods in Enzymology 266:460-480(1996);Altschul,等,Nucleic Acids Res.25:3389-3402,1997;Baxevanis,等,Bioinformatics:A Practical Guide to the Analysis of Genes and Proteins,Wiley,1998;and Misener,等,(编辑),Bioinformatics Methods and Protocols(Methods in Molecular Biology,132卷),Humana Press,1999。除了鉴别相同序列,上述程序通常提供同一性程度的指示。在一些实施方案中,如果在残基的相关延长中其对应残基的至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多为同一,两个序列被认为是基本同一的。在一些实施方案中,相关延伸是整个序列。在一些实施方案中,相关延伸为至少10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500或更多个残基。

[0058] 患有:“患有”疾病、病症或病况(例如HCMV感染)的个体已经被诊断为具有和/或表现出该疾病、病症或病况的一种或更多种症状。

[0059] 易感:“易感”疾病、病症或病况(例如HCMV感染)的个体具有发生疾病、病症或病况的风险。在一些实施方案中,易感疾病、病症或病况的个体不表现出疾病、病症或病况的任何症状。在一些实施方案中,易感疾病、病症或病况的个体已经被诊断为具有疾病、病症和/或病况。在一些实施方案中,易感疾病、病症或病况的个体是已经暴露于与该疾病、病症或病况的发生相关之条件的个体(例如个体已经暴露于HCMV)。

[0060] 症状减轻:根据本发明,当特定疾病、病症或病况的一种或更多种症状的程度(强度、严重程度等)或频率降低时,则“症状减轻”。出于澄清的目的,特定症状之发作的延迟被认为是减少所述症状之频率的一种形式。不旨在将本发明仅限制为症状消除的情况。本发明具体地考虑到使得虽然未完全消除,但一种或更多种症状减轻(并且对象的病症由此“改善”)的治疗。

[0061] 治疗有效量:本文所用的术语“治疗有效量”指以可应用于任何医疗治疗的合理的益处/风险比,足以对所治疗的对象产生治疗效果的量。治疗效果可为客观的(例如可通过某种测试或标志物测量)或主观的(例如对象给出效果的指示或感觉到效果)。特别地,“治疗有效量”指有效地治疗、改善或预防期望的疾病或病症或者表现出可检测的治疗或预防效果(例如通过改善与疾病相关的症状、预防或延迟疾病的发作、和/或还减轻疾病症状之严重程度或频率)的治疗蛋白或组合物的量。治疗有效量通常以可以包含多种单位剂量的给药方案施用。对于任何特定的免疫原性组合物,治疗有效量(和/或有效给药方案内的适当单位剂量)可以根据(例如)施用途径或与其他药剂的组合变化。此外,对于任何特定患者,具体的治疗有效量(和/或单位剂量)可以取决于多种因素,包括:被治疗的病症和病症的严重程度;所采用的具体药剂的活性;所采用的具体组合物;患者的年龄、体重、一般健康状况、性别和饮食;施用时间、施用途径、和/或所采用的具体免疫原性组合物的排泄或代谢速率;治疗的持续时间;以及医药领域中公知的类似因素。

[0062] 跨膜结构域:本领域中已知,多肽有时具有跨膜、胞质和/或胞外结构域。一般而

言,本文所用的“跨膜结构域”指具有存在于膜(例如跨越细胞膜的一部分或全部)中之性质的结构域。应当理解,不需要跨膜结构域中的每个氨基酸都存在于膜中。例如,在一些实施方案中,跨膜结构域的特征在于蛋白的指定延伸或部分基本上定位于膜中。本领域中公知,可以使用多种算法分析氨基酸或核酸序列以预测蛋白亚细胞定位(例如跨膜定位)。示例性的这样的程序包括psort(PSORT.org)、Prosites(prosite.expasy.org)等等。

[0063] 治疗:本文所用的术语“治疗”指对于特定疾病、病症和/或病况(例如HCMV感染)的一种或更多种症状或特征或者对该疾病的易感性予以部分或完全地减轻、改善、缓解、抑制、延迟其发作、减轻其严重程度、和/或减轻其发生率的免疫原性组合物的任何施用。这样的治疗可以对不表现出相关疾病、病症和/或病况之病征的对象进行,和/或仅表现出该疾病、病症和/或病况之早期病征的对象进行。或者或此外,这样的治疗可以对表现出相关疾病、病症和/或病况之一种或更多种已确定病征的对象进行。在某些实施方案中,术语“治疗”指患者的接种疫苗。

[0064] 接种疫苗:本文所用的术语“接种疫苗”指旨在对例如致病剂(例如,HCMV)产生免疫应答的组合物的施用。出于本发明的目的,接种疫苗可以在暴露于致病剂之前、期间和/或之后施用,并且在某些实施方案中,在暴露于致病剂之前、期间和/或不久之后施用。在一些实施方案中,接种疫苗包括疫苗组合物以适当间隔的时间多次施用。

[0065] 载体:本文所用的“载体”指能够转运与其缔合的另一个核酸的核酸分子。在一些实施方案中,载体能够染色体外复制和/或表达在宿主细胞(例如真核和/或原核细胞)中与其相连接的核酸。能够引导有效连接的基因之表达的载体在本文中称作“表达载体”。

[0066] 某些实施方案的详细描述

[0067] 除了其他方面以外,本发明提供了用于预防、治疗和/或研究人巨细胞病毒(HCMV)感染的方法和组合物。在一些实施方案中,本发明提供了病毒样颗粒(virus-like particles,VLP),其包含一个或更多个莫洛尼鼠白血病毒(Moloney Murine leukemia virus,MMLV)核心蛋白并且包含一个或更多个HCMV表位,例如来自HCMV的包膜糖蛋白gB和/或gH和/或被膜蛋白(tegument protein)pp65。除了其他方面以外,本发明涵盖了以下认识:抗原(例如,包膜糖蛋白和结构蛋白)之组合可导致改善的有益免疫应答之诱导,所述有益免疫应答包括例如体液应答(例如中和抗体的产生)和细胞应答(例如T细胞活化)二者。所提供的VLP特征可在于他们不含有病毒RNA或DNA并且无感染性。在一些实施方案中,提供的VLP确实含有病毒RNA或DNA并且有感染性。在一些这样的实施方案中,提供的VLP可用作DNA疫苗。

[0068] 在一些实施方案中,对象中的体液免疫应答持续至少约1个月、至少约2个月、至少约3个月、至少约4个月、至少约5个月、至少约6个月、至少约7个月、至少约8个月、至少约9个月、至少约10个月、至少约11个月、至少约12个月、至少约13个月、至少约14个月、至少约15个月、至少约16个月、至少约17个月、至少约18个月、至少约19个月、至少约20个月、至少约21个月、至少约22个月、至少约23个月、至少约24个月、至少约28个月、至少约32个月、至少约36个月、至少约40个月、至少约44个月、至少约48个月或更长。在一些实施方案中,对象中的细胞免疫应答持续至少约1个月、至少约2个月、至少约3个月、至少约4个月、至少约5个月、至少约6个月、至少约7个月、至少约8个月、至少约9个月、至少约10个月、至少约11个月或至少12个月。

[0069] 在一些实施方案中,所提供的VLP被脂质膜包围,所述脂质膜任选地含有来自病毒包膜糖蛋白(例如gB和/或gH)的一种或更多种表位,其是在病毒中和抗体的诱导中发挥作用的抗原。

[0070] 在一些实施方案中,所提供的VLP含有来自病毒结构蛋白(例如pp65)的一个或更多个表位,其是在细胞免疫应答(例如,T细胞应答)的诱导中发挥作用的抗原。在一些实施方案中,所利用的病毒结构蛋白(例如pp65)不仅刺激T辅助细胞( $T_H$ )的形成而且还诱导针对HCMV的细胞毒性T淋巴细胞(CTL)。

[0071] 在一些实施方案中,本发明提供了多种病毒包膜糖蛋白(例如gB和/或gH)。在一些实施方案中,变体病毒包膜糖蛋白是或者包含融合蛋白。在一些实施方案中,病毒糖蛋白的变体包含异源蛋白结构域(例如,来自不同蛋白的跨膜和/或胞质结构域)。在一些实施方案中,病毒结构蛋白的变体包含异源抗原或表位。在一些实施方案中,本发明提供了包含病毒结构蛋白之变体的VLP。在一些实施方案中,病毒结构蛋白的变体是或者包含融合蛋白。

[0072] I. 病毒样颗粒(VLP)

[0073] 反转录病毒是包膜的RNA病毒,其属于反转录病毒科(Retroviridae)。宿主细胞被反转录病毒感染后,RNA经酶反转录酶转录为DNA。然后DNA通过整合酶并入宿主细胞的基因组并且之后作为宿主细胞DNA的一部分复制。反转录病毒科包括以下属: $\alpha$ 反转录病毒属(Alpharetrovirus)、 $\beta$ 反转录病毒属(Betaretrovirus)、 $\gamma$ 反转录病毒属(Gammaretrovirus)、 $\delta$ 反转录病毒属(Deltaretrovirus)、 $\epsilon$ 反转录病毒属(Epsilonretrovirus)、慢病毒属(Lentivirus)和泡沫病毒属(Spumavirus)。反转录病毒科的宿主一般为脊椎动物。反转录病毒产生感染性病毒体,其含有被来自宿主细胞膜的脂质双层包围的球形核壳(与病毒结构蛋白复合的病毒基因组)。

[0074] 反转录病毒载体可用于产生包膜的病毒体,其是感染性的并且有复制能力或者复制缺陷的。有复制能力的感染性反转录病毒载体含有用于病毒体合成的所有必需基因并且一旦发生宿主细胞的感染后持续增殖自身。复制缺陷的感染性反转录病毒载体在初始感染后不传播。这通过用待转移的基因或核苷酸序列替换反转录病毒的大多数编码区完成;使得载体能够制造另外的复制循环所述的蛋白。

[0075] 或者或此外,反转录病毒载体可用于产生病毒样颗粒(VLP),其缺乏来自反转录病毒的基因组并且是非感染性和非复制的。由于VLP有益的性质,VLP可以用作抗原递送系统。此外,因为VLP是非感染性的,它们可作为免疫原性组合物(例如,疫苗)安全地施用。VLP通常结构类似于上述的包膜的病毒体,但缺乏来自反转录病毒的基因组,使得病毒复制将不太可能发生。一些病毒(例如鼠白血病病毒,例如莫洛尼鼠白血病病毒(MMLV))的衣壳蛋白(例如Gag)的表达导致自组装成类似于相应天然病毒的颗粒,该颗粒没有病毒遗传物质。

[0076] 已经制备了各种各样的VLP。例如,已经制备了包含单种或多种衣壳蛋白,有或者没有包膜蛋白和/或表面糖蛋白的VLP。在一些情况下,VLP是非包膜的并且通过仅表达一种主要衣壳蛋白装配,如所示的从嗜肝DNA病毒(例如Engerix<sup>TM</sup>,GSK和Recombivax HB<sup>TM</sup>,Merck)、乳头状瘤病毒(例如Cervarix<sup>TM</sup>,GSK和Gardasil<sup>TM</sup>,Merck)、细小病毒(parovirus)或多瘤病毒制备的VLP。在一些实施方案中,VLP是包膜的并且可包含存在于相应的天然病毒中的多个抗原蛋白。VLP通常类似于其对应的天然病毒并且可为多价的颗粒结构。在一些实施方案中,抗原蛋白可以存在于VLP内作为VLP结构的组分和/或在VLP的表面上。本发明

涵盖了以下认识:与其他形式的抗原呈递(例如不与VLP缔合的可溶性抗原)相比,在VLP的情形中的抗原呈递有利于诱导针对该抗原的中和抗体。中和抗体大多数经常识别三级或四级结构,这经常需要以其天然病毒构象呈递抗原蛋白(如包膜糖蛋白)。或者或此外,VLP可用于在诱导细胞免疫(例如T细胞应答)的情形下呈递抗原。本发明还涵盖了这样的观点,在VLP系统中使用抗原组合可产生改进的免疫应答。

[0077] A. 结构蛋白

[0078] 在一些实施方案中,本发明利用由一种或更多种反转录病毒结构蛋白(例如Gag)组成的VLP。在一些实施方案中,根据本发明所使用的结构蛋白为 $\alpha$ 反转录病毒属(例如禽白血病病毒)、 $\beta$ 反转录病毒属(小鼠乳腺瘤病毒)、 $\gamma$ 反转录病毒属(鼠白血病病毒)、 $\delta$ 反转录病毒属(牛白血病病毒)、 $\epsilon$ 反转录病毒属(大眼梭鲈皮肤肉瘤病毒,walley dermal sarcoma virus)、慢病毒属(人免疫缺陷病毒1)和泡沫病毒属(黑猩猩泡沫病毒)结构蛋白。在某些实施方案中,结构蛋白为鼠白血病病毒(MLV)结构蛋白。这些反转录病毒的基因组在数据库中可容易地得到。所有这些反转录病毒的Gag基因具有整体上类似的结构并且在各个反转录病毒群中在氨基酸水平是保守的。反转录病毒Gag蛋白主要功能为病毒组装。多蛋白形式的Gag基因产生VLP的核心结构蛋白。MLV Gag基因编码65kDa多蛋白前体,其在成熟的病毒体中由MLV蛋白酶蛋白水解切割为4个结构蛋白(基质(MA);p12;衣壳(CA)和核壳(NC))。反转录病毒组装不成熟衣壳,其由从Gag多肽形成的Gag多蛋白组成但缺乏其他病毒元件(如病毒蛋白酶),Gag作为该未成熟病毒颗粒的结构蛋白。功能上,Gag多蛋白分为3个结构域:膜结合结构域,其将Gag多蛋白靶向细胞膜;相互作用结构域,其促进Gag聚合;以及晚期结构域,其利于新生病毒体从宿主细胞的释放。介导病毒颗粒组装的Gag蛋白之形式为多蛋白。

[0079] 在一些实施方案中,根据本发明使用的反转录病毒结构蛋白为Gag多肽。本文所用的术语“Gag多肽”为来自反转录病毒的结构多肽,其负责本文所述的VLP的形成并且指其氨基酸序列包括Gag之至少一个特征序列的多肽序列。本领域和本领域普通技术人员已知来自多种反转录病毒的各种各样的Gag序列,参考这样的序列,可容易地鉴别一般Gag蛋白和/或特定Gag多肽的特征性序列。

[0080] 根据本发明使用的示例性Gag多肽如以下SEQ ID NO:1所示。在一些实施方案中,合适的Gag多肽与已知的反转录病毒Gag多肽基本同源。例如,Gag多肽可以为经修饰的反转录病毒Gag多肽,其与野生型或天然存在的Gag多肽(例如SEQ ID NO:1)相比含有一个或更多个氨基酸替代、缺失和/或插入,同时基本上保留自组装活性。因此,在一些实施方案中,适合用于本发明的Gag多肽基本上与MMLV Gag多肽(SEQ ID NO:1)同源。在一些实施方案中,适合用于本发明的Gag多肽具有与SEQ ID NO:1有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中,适合用于本发明的Gag多肽与MMLV Gag多肽(SEQ ID NO:1)基本同一。在一些实施方案中,适合用于本发明的Gag多肽具有与SEQ ID NO:1有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。

[0081] MMLV Gag氨基酸序列(SEQ ID NO:1)

[0082]

MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGTFN  
RDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLPL  
EPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLLEDPPPYRDP RPPPSDRDGNNG  
EATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFSDDL YNW  
KNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLGTLTGEEKQRVLL EARKAVRGD  
DGRPTQLPNEVDAAFLERP DWYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNLAKV

[0083]

KGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRKLERLED  
LKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRRHREM  
SKLLATVVSQGKQDRQGGERRRSQ LDRDQCAYCKEKGHWAKDCPKKPRGPRGPRPQT  
SLLTLDD (SEQ ID NO:1)

MMLV Gag核苷酸序列 (SEQ ID NO:2)

[0084]

ATGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACCTGGAAAGATGTC  
GAGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGTTGGGTACCTT  
CTGCTCTGCAGAAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACCTTTAA  
CCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTTTTACCTGGCCCGCATGGACA  
CCCAGACCAGGTCCCCTACATCGTGACCTGGGAAGCCTTGCTTTTGACCCCCCTCC  
CTGGGTCAAGCCCTTTGTACACCCTAAGCCTCCGCTCCTCTTCTCCATCCGCCCCG  
TCTCTCCCCCTTGAACCTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCC  
TCACTCCTTCTCTAGGCGCCAAACCTAAACCTCAAGTTCTTTCTGACAGTGGGGGGC  
CGCTCATCGACCTACTTACAGAAGACCCCCCGCTTATAGGGACCCAAGACCACCCC  
CTTCCGACAGGGACGGAATGGTGGAGAAGCGACCCCTGCGGGAGAGGCACCGGA  
CCCCTCCCCAATGGCATCTCGCCTACGTGGGAGACGGGAGCCCCCTGTGGCCGACTC  
CACTACCTCGCAGGCATTCCCCCTCCGCGCAGGAGGAAACGGACAGCTTCAATACT  
GGCCGTTCTCCTCTTCTGACCTTTACAACCTGGAAAAATAATAACCTTCTTTTTCTGA  
AGATCCAGGTAAACTGACAGCTCTGATCGAGTCTGTTCTCATCACCCATCAGCCCAC  
CTGGGACGACTGTCAGCAGCTGTTGGGGACTCTGCTGACCGGAGAAGAAAAACAAC  
GGGTGCTCTTAGAGGCTAGAAAGGCGGTGCGGGCGATGATGGGCGCCCCACTCAA  
CTGCCCAATGAAGTCGATGCCGCTTTTCCCCCTCGAGCGCCAGACTGGGATTACACC  
ACCCAGGCAGGTAGGAACCACTAGTCCACTATCGCCAGTTGCTCCTAGCGGGTCTC  
CAAAACGCGGGCAGAAGCCCCACCAATTTGGCCAAGGTAAAAGGAATAACACAAG  
GGCCCAATGAGTCTCCCTCGGCCTTCTAGAGAGACTTAAGGAAGCCTATCGCAGGT  
ACACTCCTTATGACCCTGAGGACCCAGGGCAAGAACTAATGTGTCTATGTCTTTCA  
TTTGGCAGTCTGCCCCAGACATTGGGAGAAAAGTTAGAGAGGTTAGAAGATTTAAAA  
AACAAGACGCTTGGAGATTTGGTTAGAGAGGCAGAAAAGATCTTTAATAAACGAGA  
AACCCCGGAAGAAAGAGAGGAACGTATCAGGAGAGAAACAGAGGAAAAAGAAGA  
ACGCCGTAGGACAGAGGATGAGCAGAAAGAGAAAGAAAGAGATCGTAGGAGACAT  
AGAGAGATGAGCAAGCTATTGGCCACTGTCTGTTAGTGGACAGAAACAGGATAGACA  
GGGAGGAGAACGAAGGAGGTCCCAACTCGATCGCGACCAGTGTGCCTACTGCAAAG  
AAAAGGGGCACTGGGCTAAAGATTGTCCCAAGAAACCACGAGGACCTCGGGGACC  
AAGACCCAGACCTCCCTCCTGACCCTAGATGAC (SEQ ID NO:2)

[0085]

密码子优化的MMLV Gag核苷酸序列 (SEQ ID NO:3)

ATGGGACAGACAGTCACTACACCCCTGAGCCTGACACTGGGACATTGGAAAGACGT  
GGAGAGGATTGCACATAACCAGAGCGTGGACGTGAAGAAACGGAGATGGGTACCT  
TTTGCTCCGCCGAGTGGCCAACATTCAATGTGGGATGGCCCCGAGATGGCACCTTCA  
ACCGGGACCTGATCACTCAGGTGAAGATCAAGGTCTTCTCTCCAGGACCCACGGCC  
ATCCAGATCAGGTGCCCTACATCGTCACCTGGGAGGCTCTGGCATTGACCCCCCTC  
CATGGGTGAAGCCTTTCGTCCACCCAAAACCACTCCACCACTGCCTCCATCTGCCC  
CTAGTCTGCCACTGGAACCCCTCGGTCAACCCACCCAGAAGCTCCCTGTATCCCG  
CACTGACACCTAGCCTGGGGGCCAAGCCTAAACCACAGGTGCTGTCTGATAGTGGC  
GGCCTCTGATCGATCTGCTGACCGAGGACCCTCCACCATAACGCGACCCACGACCT

[0086]

CCACCAAGCGACCGGGACGGAAACGGAGGAGAGGCTACACCCGCAGGCGAAGCCC  
CCGATCCTAGTCCAATGGCATCAAGGCTGCGCGGGAGGCGCAACCTCCAGTGGCC  
GACTCAACCACAAGCCAGGCATTTCCACTGAGGGCCGGGGGAAATGGACAGCTCCA  
GTATTGGCCCTTCTCTAGTTCAGATCTGTACAACTGGAAGAACAATAACCCTAGCTT  
CAGCGAGGACCCAGGCAAACTGACCGCCCTGATCGAATCCGTGCTGATTACCCACC  
AGCCACATGGGACGATTGTCAGCAGCTCCTGGGCACCCTGCTGACCGGAGAGGAA  
AAGCAGAGAGTGCTGCTGGAGGCTAGGAAAGCAGTCCGCGGGGACGATGGAAGGC  
CAACACAGCTCCCCAATGAGGTGGATGCCGCTTTCCTCTGGAACGGCCAGATTGGG  
ACTATACTACCCAGGCTGGACGCAACCACCTGGTGCAATTACCGGCAGCTCCTGCTGG  
CTGGACTGCAGAATGCAGGGCGCAGCCCCACTAACCTGGCCAAGGTGAAAGGAATC  
ACCCAGGGCCCCAATGAGTCCCCTTCTGCATTCCCTGGAGCGGCTGAAGGAAGCCTAC  
CGACGGTATACTCCCTACGATCCTGAGGACCCAGGCCAGGAAACCAACGTGAGTAT  
GAGCTTCATCTGGCAGTCCGCTCCTGACATTGGCCGAAAACCTGGAGCGGCTGGAAG  
ATCTGAAGAACAAGACCCTGGGCGACCTGGTGCGGGAGGCAGAAAAAGATCTTCAAC  
AAAAGGGAGACTCCAGAGGAACGGGAGGAAAGAATTAGAAGGGAAACAGAGGAA  
AAGGAGGAACGCCGACGGACTGAGGATGAACAGAAGGAGAAAGAAAGAGACCGGC  
GGCGGCACCGGGAGATGTCTAAGCTGCTGGCCACCGTGGTCAGTGGCCAGAAACAG  
GATCGACAGGGAGGAGAGCGACGGAGAAGCCAGCTCGATCGGGACCAGTGC GCCT  
ATTGTAAGGAAAAAGGGCATTGGGCTAAGGACTGCCCCAAGAAACCCAGAGGCCCA  
CGCGGGCCCCGACCTCAGACTTCCCTGCTGACCCTGGACGAT (SEQ ID NO:3)

[0087]

密码子优化的MMLV Gag核苷酸序列 (SEQ ID NO:21)

[0088]

ATGGGACAGACCGTCACAACACCCCTGAGCCTGACCCCTGGGACATTGGAAAGACGT  
GGAGAGGATCGCACATAACCAGAGCGTGGACGTGAAGAAACGGAGATGGGTACACA  
TTCTGCAGTGCTGAGTGGCCAACTTTTAATGTGGGATGGCCCCGAGACGGCACTTTC  
AACAGGGATCTGATCACCCAGGTGAAGATCAAGGTCTTTAGCCCAGGACCTCACGG  
ACATCCAGACCAGGTGCCTTATATCGTCACTGAGGAGGCACTGGCCTTCGATCCCCC  
TCCATGGGTGAAGCCATTTGTCCACCCAAAACCACCTCCACCACTGCCTCCAAGTGC  
CCCTTCACTGCCACTGGAACCACCCCGGAGCACACCACCCCGCAGCTCCCTGTATCC  
TGCTCTGACTCCATCTCTGGGCGCAAAGCCAAAACCACAGGTGCTGAGCGACTCCG  
GAGGACCACTGATTGACCTGCTGACAGAGGACCCCCACCATAACCGAGATCCTCGG  
CCTCCACCAAGCGACCGCGATGGAAATGGAGGAGAGGCTACTCCTGCCGGCGAAGC  
CCCTGACCCATCTCCAATGGCTAGTAGGCTGCGCGGACAGGCGCGAGCCTCCAGTGG  
CAGATAGCACCATCCAGGCCTTCCCTCTGAGGGCTGGGGGAAATGGGCAGCTC  
CAGTATTGGCCATTTTCTAGTTTACAGACCTGTACAACTGGAAGAACAATAACCCCTCT  
TTCAGTGAGGACCCCGGCAAACCTGACCGCCCTGATCGAATCCGTGCTGATTACCCAT  
CAGCCCACATGGGACGATTGTGTCAGCAGCTCCTGGGCACCCTGCTGACCGGAGAGGA  
AAAGCAGCGCGTGCTGCTGGAGGCTCGCAAAGCAGTCCGAGGGGACGATGGACGG  
CCCACACAGCTCCCTAATGAGGTGGACGCCGCTTTTCCACTGGAAAGACCCGACTGG  
GATTATACTACCCAGGCAGGGAGAAACCACCTGGTCCATTACAGGCAGCTCCTGCT  
GGCAGGCCTGCAGAATGCCGGGAGATCCCCACCAACCTGGCCAAGGTGAAAGGCA  
TCACACAGGGGCCTAATGAGTCACCAAGCGCCTTTCTGGAGAGGCTGAAGGAAGCT  
TACCGACGGTATACCCCATACGACCCTGAGGACCCCGGACAGGAAACAAACGTCTC  
CATGTCTTTCATCTGGCAGTCTGCCCCAGACATTGGGCGGAAGCTGGAGAGACTGGA  
AGACCTGAAGAACAAGACCCTGGGCGACCTGGTGCGGGAGGCTGAAAAGATCTTCA

ACAAACGGGAGACCCCGAGGAAAGAGAGGAAAGGATTAGAAGGGAAACTGAGGA  
AAAGGAGGAACGCCGACGGACCGAGGACGAACAGAAGGAGAAAGAACGAGATCG  
GCGGCGGCACCGGAGATGTCAAAGCTGCTGGCCACCGTGGTCAGCGGACAGAAAC  
AGGACAGACAGGGAGGAGAGCGACGGAGAAGCCAGCTCGACAGGGATCAGTGCGC  
ATACTGTAAGGAAAAAGGCCATTGGGCCAAGGATTGCCCCAAAAAGCCAAGAGGAC  
CAAGAGGACCAAGACCACAGACATCACTGCTGACCCTGGACGAC (SEQ ID NO:21)

[0089] 通常自然情况下,Gag蛋白包含大C端延伸,其可以含有反转录病毒蛋白酶、反转录酶和整合酶酶促活性。然而,VLP的组装通常不需要这样的组分存在。在一些情况下,反转录病毒Gag蛋白可单独(例如,缺乏C端延伸,缺乏一种或更多种基因组RNA、反转录酶、病毒蛋白酶或包膜蛋白)自组装以形成VLP,无论在体外还是在体内(Sharma S等,1997 Proc.Natl.Acad.Sci.USA 94:10803-8)。反转录病毒Gag多蛋白可单独低聚并组装成VLP。

[0090] 在一些实施方案中,根据本发明使用的Gag多肽缺乏C端延伸和/或含有经修饰的C端延伸。Gag多肽可以任选地包含一个或更多个另外的多肽(例如,异源抗原)。在一些实施方案中,Gag多肽与异源抗原共表达(例如在独立的启动子下和/或作为独立的蛋白)。在一些实施方案中,Gag多肽表达为具有异源抗原的融合蛋白。Gag多肽可连接至异源抗原以产生不改变Gag功能的融合蛋白。例如,可将异源抗原的编码序列剪接成Gag多肽编码序列,例如在Gag多肽编码序列的3'端。在一些实施方案中,可将异源抗原的编码序列框内(in frame)剪接成Gag多肽编码序列。在一些实施方案中,Gag多肽编码序列和异源抗原可以由单个启动子表达。在一些实施方案中,在Gag多肽的C端插入(例如融合到)异源抗原。不希望受任何理论的限制,认为自组装的Gag多肽与异源抗原的融合产生了融合蛋白,其充当未修改的Gag并且因此将允许所述抗原并入到所产生的VLP之结构组分中。在一些实施方案中,VLP结构组分充当有效的免疫原(例如用于诱导细胞免疫应答)。例如,所提供的VLP可以包括反转录病毒Gag多肽(例如MMLV Gag)和HCMV的结构组分(例如pp65)。在一些这样的实施方案中,pp65并入到VLP中并且充当用于引发针对HCMV之免疫应答的抗原。



[0091] 根据本发明所用的示例性Gag-pp65融合多肽如以下SEQ ID NO:4所示。在一些实施方案中,合适的Gag多肽融合蛋白包含与已知的反转录病毒Gag多肽基本同源的Gag多肽之全部或部分,和与已知的pp65多肽基本同源的pp65多肽之全部或部分。例如,与野生型或天然存在的Gag多肽和/或pp65多肽相比,Gag-pp65多肽融合蛋白可以含有一个或更多个氨基酸替代、缺失和/或插入,同时保留基本的自组装活性。因此,在一些实施方案中,适合用于本发明的Gag-pp65多肽融合蛋白与SEQ ID NO:4基本同源。在一些实施方案中,适合用于本发明的Gag-pp65多肽融合蛋白具有与SEQ ID NO:4有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中,适合用于本发明的Gag-pp65多肽融合蛋白与SEQ ID NO:4基本同一。在一些实施方案中,适合用于本发明的Gag-pp65多肽融合蛋白具有与SEQ ID NO:4有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。

[0092] MLLV Gag-CMV pp65氨基酸序列 (SEQ ID NO:4)

**MGQTVTTPLSLTLGHWKDV**ERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDG  
TFNRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPP  
SAPSLPLEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLTTEDPPPYRDP  
PPSDRDGNGGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQ  
YWPFSDDLYNWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLGTLLTGEE  
KQRVLLLEARKAVRGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQL  
LLAGLQNAGRSPTNLAKVKGITQGPNEPSAFLERLKEAYRRYTPYDPEDPGQETN  
VSMFSIWQSAPDIGRKLERLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETE  
EKEERRRTEDEQKEKERDRRRHREMSKLLATVVSGQKQDRQGGERRRSQLDRDQ

[0093] **CAYCKEKGHWAKDCPKKPRGPRGPRPQTSL**TLDDCESRGRRCPMISVLGPISGHV  
LKAVFSRGDTPVLPHETRLQTGIHVRVSQPSLILVSQYTPDSTPCHRGDNQLQVQHTYF  
TGSEVENVSVNVHNPTGRSICPSQEPMSIYVYALPLKMLNIPSINVHHYPSAAERKHRHL  
PVADAVIHASGKQMWQARLTVSGLAWTRQQNQWKEPDVYYTSAFVFPTKDVALRHV  
VCAHELVCSEMENTRATKMQVIGDQYVKVYLESFCEDEVPSGKLFMHVTLGSDVEEDLT  
MTRNPQPFMRPHERNGFTVLCPKNMIKPGKISHIMLDVAFTSHEHFGLLCPKSIPGLSIS  
GNLLMNGQQIFLEVQAIRETVELRQYDPVAALFFFDIDLLLQRGPQYSEHPTFTSQYRIQ

**GKLEYRHTWDRHDEGAAQGGDDVWTS**GSDSDEELVTTERKTPRVTGGGAMAGASTSA  
GRKRKSASSATACTAGVMTRGRLKAESTVAPEEDTDESDNEIHNPVFTWPPWQAGI  
LARNLVPMVATVQGQNLKYQEFFWDANDIYRIFAELEGVWQPAAQPKRRRRHRQDALP

**GPCIASTPKKHRG\* (SEQ ID NO:4)** (MLLV氨基酸序列加粗)

[0094] MLLV Gag-CMV pp65核苷酸序列 (SEQ ID NO:5)

[0095]

ATGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCAGTGGAAAAGAT  
GTCGAGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGTTGGGT  
TACCTTCTGCTCTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACG  
GCACCTTTAACCGAGACCTCATCACCCAGGTAAAGATCAAGGTCTTTTACCTG  
GCCCCGATGGACACCCAGACCAGGTCCCCTACATCGTGACCTGGGAAGCCTTG  
GCTTTTGACCCCCCTCCCTGGGTCAAGCCCTTTGTACACCCTAAGCCTCCGCCT  
CCTCTTCTCCATCCGCCCGTCTCTCCCCCTTGAACCTCCTCGTTCGACCCCCG  
CCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGCGCCAAACCTAAA  
CCTCAAGTTCTTTCTGACAGTGGGGGGCGGCTCATCGACCTACTTACAGAAGA  
CCCCCGCCTTATAGGGACCCAAGACCACCCCTTCCGACAGGGACGGAATG  
GTGGAGAAGCGACCCCTGCGGGAGAGGACCGGACCCCTCCCCAATGGCATCT  
CGCCTACGTGGGAGACGGGAGCCCCCTGTGGCCGACTCCACTACCTCGCAGGC  
ATTCCCCCTCCGCGCAGGAGGAAACGGACAGCTTCAATACTGGCCGTTCTCCT  
CTTCTGACCTTTACAACCTGGAAAAATAATAACCCCTTCTTTTTCTGAAGATCCAG  
GTAAACTGACAGCTCTGATCGAGTCTGTTCTCATCACCCATCAGCCCACCTGGG  
ACGACTGTCAGCAGCTGTTGGGGACTCTGCTGACCGGAGAAGAAAAACAACGG  
GTGCTCTTAGAGGCTAGAAAGGCGGTGCGGGGCGATGATGGGCGCCCCACTCA  
ACTGCCCAATGAAGTCGATGCCGCTTTTCCCCTCGAGCGCCCAGACTGGGATT  
ACACCACCCAGGCAGGTAGGAACCACCTAGTCCACTATCGCCAGTTGCTCCTA  
GCGGGTCTCCAAAACGCGGGCAGAAGCCCCACCAATTGGCCAAGGTAAAAGG  
AATAACACAAGGGCCCAATGAGTCTCCCTCGGCCTTCTAGAGAGACTTAAGG  
AAGCCTATCGCAGGTACACTCCTTATGACCCTGAGGACCCAGGGCAAGAACT  
AATGTGTCTATGTCTTTCATTTGGCAGTCTGCCCCAGACATTGGGAGAAAGTTA  
GAGAGGTTAGAAGATTTAAAAACAAGACGCTTGGAGATTGGTTAGAGAGGC  
AGAAAAGATCTTTAATAAACGAGAAACCCCGGAAGAAAGAGAGGAACGTATCA  
GGAGAGAAACAGAGGAAAAAGAAACGCCGTAGGACAGAGGATGAGCAGAA  
AGAGAAAGAAAGAGATCGTAGGAGACATAGAGAGATGAGCAAGCTATTGGCCA  
CTGTCTAGTGGACAGAAACAGGATAGACAGGGAGGAGAACGAAGGAGGTC  
CCAACCTCGATCGCGACCACTGTGCCTACTGCAAAGAAAAGGGGCACTGGGCTA  
AAGATTGTCCCAAGAAACCACGAGGACCTCGGGGACCAAGACCCCAGACCTCC  
CTCCTGACCCTAGATGACTGTGAGTCGCGCGGTGCGCGTTGTCCCGAAATGATATC  
CGTACTGGGTCCCATTTTCGGGGCACGTGCTGAAAGCCGTGTTAGTTCGCGGCGACAC  
GCCGGTGCTGCCGCACGAGACGCGACTCCTGCAGACGGGTATCCACGTGCGCGTGA  
GCCAGCCCTCGCTGATCCTGGTGTGCGAGTACACGCCGACTCGACGCCATGCCACC  
GCGGGCACAATCAGCTGCAGGTGCAGCACACGTACTTTACGGGCAGCGAGGTGGAG  
AACGTGTCGGTCAACGTGCACAACCCACGGGCCGGAGCATCTGCCCCAGCCAAGA  
GCCCATGTCGATCTATGTGTACGCGCTGCCGCTCAAGATGCTGAACATCCCCAGCAT  
CAACGTGCACCACTACCCGTGCGCGGCCGAGCGCAAACACCGACACCTGCCCGTAG  
CTGACGCTGTGATTCACGCGTCGGGCAAGCAGATGTGGCAGGCGCGTCTCACGGTCT  
CGGGACTGGCCTGGACGCGTCAGCAGAACCAGTGGAAGAGCCCGACGTCTACTAC

[0096]

ACGTCAGCGTTTCGTGTTTCCCACCAAGGACGTGGCACTGCGGCACGTGGTGTGCGCG  
CACGAGCTGGTTTGCTCCATGGAGAACACGCGCGCAACCAAGATGCAGGTGATAGG  
TGACCAGTACGTCAAGGTGTACCTGGAGTCCTTCTGCGAGGACGTGCCCTCCGGCAA  
GCTCTTTATGCACGTACGCTGGGCTCTGACGTGGAAGAGGACCTGACGATGACCCG  
CAACCCGCAACCCTTCATGCGCCCCACGAGCGCAACGGCTTTACGGTGTGTGTCC  
CAAAAATATGATAATCAAACCGGGCAAGATCTCGCACATCATGCTGGATGTGGCTTT  
TACCTCACACGAGCATTTTGGGCTGCTGTGTCCCAAGAGCATCCCGGGCCTGAGCAT  
CTCAGGTAACCTATTGATGAACGGGCAGCAGATCTTCCTGGAGGTGCAAGCGATAC  
GCGAGACCGTGGAACCTGCGTCAGTACGATCCCGTGGCTGCGCTCTTCTTTTTCGATA  
TCGACTTGCTGCTGCAGCGCGGGCCTCAGTACAGCGAACACCCACCTTCACCAGCC  
AGTATCGCATCCAGGGCAAGCTTGAGTACCGACACACCTGGGACCGGCACGACGAG  
GGTGCCGCCCAGGGCGACGACGACGTCTGGACCAGCGGATCGGACTCCGACGAGGA  
ACTCGTAACCAACCGAGCGCAAGACGCCCCGCGTTACCGGCGGCGGCGCCATGGCGG  
GCGCCTCCACTTCCGCGGGGCCGCAAACGCAAATCAGCATCCTCGGCGACGGCGTGC  
ACGGCGGGCGTTATGACACGCGGCCGCCTTAAGGCCGAGTCCACCGTCGCGCCCGA  
AGAGGACACCGACGAGGATTCCGACAACGAAATCCACAATCCGGCCGTGTTACCT  
GGCCGCCCTGGCAGGCCGGCATCTGGCCCGCAACCTGGTGCCCATGGTGGCTACG  
GTTACAGGGTCAGAATCTGAAGTACCAGGAGTTCTTCTGGGACGCCAACGACATCTAC  
CGCATCTTCGCCGAATTGGAAGGCGTATGGCAGCCCGCTGCGCAACCCAAACGTCG  
CCGCCACCGGCAAGACGCCTTGCCCGGGCCATGCATCGCCTCGACGCCCAAAAGC  
ACCGAGGTTAG (SEQ ID NO:5) (MMLV Gag核苷酸序列加粗)

[0097]

密码子优化的MMLV Gag-CMV pp65核苷酸序列 (SEQ ID NO:6)

[0098]

ATGGGACAGACAGTCACTACACCCCTGAGCCTGACACTGGGACATTGGAAAGA  
CGTGGAGAGGATTGCACATAACCAGAGCGTGGACGTGAAGAAACGGAGATGG  
GTCACCTTTTGCTCCGCCGAGTGGCCAACATTCAATGTGGGATGGCCCCGAGA  
TGGCACCTTCAACCGGGACCTGATCACTCAGGTGAAGATCAAGGTCTTCTCTCC  
AGGACCCACGGCCATCCAGATCAGGTGCCCTACATCGTCACCTGGGAGGCTC  
TGGCATTTGACCCCCCTCCATGGGTGAAGCCTTTCGTCCACCCAAAACACCTC  
CACCCTGCCTCCATCTGCCCTAGTCTGCCACTGGAACCCCTCGGTCAACCC  
CACCCAGAAGCTCCCTGTATCCCGCACTGACACCTAGCCTGGGGGGCCAAGCCT  
AAACCACAGGTGCTGTCTGATAGTGGCGGGCCTCTGATCGATCTGCTGACCGA  
GGACCCTCCACCATAACCGCGACCCACGACCTCCACCAAGCGACCGGGACGGAA  
ACGGAGGAGAGGCTACACCCGCGAGGCGAAGCCCCCGATCCTAGTCCAATGGCA  
TCAAGGCTGCGCGGGAGGCGCGAACCTCCAGTGGCCGACTCAACCACAAGCCA  
GGCATTTCCACTGAGGGCCGGGGGAAATGGACAGCTCCAGTATTGGCCCTTCT  
CTAGTTCAGATCTGTACAACTGGAAGAACAATAACCCCTAGCTTCAGCGAGGAC  
CCAGGCAAACTGACCGCCCTGATCGAATCCGTGCTGATTACCCACCAGCCAC  
ATGGGACGATTGTGACGAGCTCCTGGGCACCCTGCTGACCGGAGAGGAAAAGC  
AGAGAGTGCTGCTGGAGGCTAGGAAAGCAGTCCGCGGGGACGATGGAAGGCC  
AACACAGCTCCCCAATGAGGTGGATGCCGCTTTCCTCTGGAACGGCCAGATT  
GGGACTATACTACCCAGGCTGGACGCAACCACCTGGTGCATTACCGGCAGCTC  
CTGCTGGCTGGACTGCAGAATGCAGGGCGCAGCCCCACTAACCTGGCCAAGGT  
GAAAGGAATCACCCAGGGCCCCAATGAGTCCCCTTCTGCATTCTGGAGCGGC  
TGAAGGAAGCCTACCGACGGTATACTCCCTACGATCCTGAGGACCCAGGCCAG  
GAAACCAACGTGAGTATGAGCTTCATCTGGCAGTCCGCTCCTGACATTGGCCG

[0099]

AAAACCTGGAGCGGCTGGAAGATCTGAAGAACAAGACCCTGGGCGACCTGGTGC  
GGGAGGCAGAAAAAGATCTTCAACAAAAGGGAGACTCCAGAGGAACGGGAGGA  
AAGAATTAGAAGGGAAACAGAGGAAAAAGGAGGAACGCCGACGGACTGAGGAT  
GAACAGAAGGAGAAAGAAAGAGACCGGCGGCGGCACCGGGAGATGTCTAAGC  
TGCTGGCCACCGTGGTCACTGGCCAGAAACAGGATCGACAGGGAGGAGAGCG  
ACGGAGAAGCCAGCTCGATCGGGACCAGTGCGCCTATTGTAAGGAAAAAGGGC  
ATTGGGCTAAGGACTGCCCCAAGAAACCCAGAGGCCACGCGGGCCCCGACCT  
CAGACTTCCCTGCTGACCCTGGACGATTGCGAGAGCCGGGCGGCGGTGCCCA  
GAAATGATCTCTGTGCTGGGGCCCATTAGTGGACATGTGCTGAAGGCCGTCTTCTCC  
AGGGGAGACACCCCCGTGCTGCCTCACGAGACTCGACTGCTGCAGACCGGCATCCA  
TGTGCGGGTCTCCAGCCCTCTCTGATTCTGGTGTACAGTATACACCAGATAGCAC  
TCCCTGCCACAGAGGAGACAATCAGCTCCAGGTGCAGCATACCTACTTTACAGGCTC  
CGAGGTCGAAAACGTGTCTGTCAATGTGCACAACCCTACCGGCAGGAGCATCTGTC  
CTAGCCAGGAGCCAATGAGCATCTACGTGTACGCCCTGCCTCTGAAGATGCTGAATA  
TCCCATCAATTAACGTCCACCATTACCCTAGCGCAGCCGAACGGAAGCACAGACAT  
CTGCCAGTGGCCGACGCTGTCATCCATGCCAGCGGCAAACAGATGTGGCAGGCAAG  
ACTGACCGTGTCCGGGCTGGCCTGGACAAGGCAGCAGAATCAGTGGAAGGAGCCCG  
ACGTGTACTATAACAGCGCCTTCGTGTTCCCTACCAAAGACGTGGCCCTGAGACATG  
TGGTGTGCGCACATGAGCTGGTGTGCAGCATGGAAAAACTAGGGCCACCAAGATG  
CAGGTCATCGGCGATCAGTATGTCAAAGTGTACCTGGAGAGTTTTTTCGAAGACGTG  
CCATCAGGGAAGCTGTTTCATGCATGTGACCCTGGGCAGCGATGTGAGGAAGACCT  
GACCATGACAAGAAATCCACAGCCCTTTATGAGACCCACGAGAGGAATGGGTTCA  
CTGTGCTGTGCCCCAAGAACATGATCATTAAAGCCTGGAAAAATCAGTCATATTATGC  
TGGATGTGGCCTTTACATCACACGAGCATTTCGGACTGCTGTGCCCCAAATCCATCC  
CTGGACTGAGCATTTCGGCAATCTGCTGATGAACGGCCAGCAGATCTTCCTGGAAG  
TGCAGGCCATCCGGGAGACCGTCGAACTGCGACAGTATGACCCAGTGGCTGCACTG  
TTCTTTTTTCGACATCGACCTGCTGCTGCAGCGAGGACCACAGTACAGCGAGCACCT  
ACTTTTACCTCCCAGTATCGGATTCAGGGGAAGCTGGAGTACAGGCACACCTGGGAT  
CGCCATGACGAAGGAGCCGCTCAGGGGGACGATGACGTGTGGACATCTGGCAGTGA  
TTCAGACGAGGAACTGGTGACAACTGAGCGAAAAACCCCCGGGTGACAGGAGGA  
GGGGCAATGGCAGGGGGCCAGCACCAGCGCAGGGCGGAAGCGAAAAAGCGCCAGCA  
GCGCCACAGCATGTACCGCCGGCGTGATGACTAGAGGAAGGCTGAAGGCCGAGTCT  
ACAGTCGCTCCCGAGGAAGATACTGACGAGGATAGTGACAATGAAATCCACAACCC  
CGCCGTGTTACCTGGCCACCTTGGCAGGCAGGGATTCTGGCTCGCAACCTGGTCCC  
CATGGTGGCAACCGTCCAGGGACAGAATCTGAAGTATCAGGAGTTTTTCTGGGATGC  
TAACGACATCTACCGGATTTTTGCAGAGCTGGAAGGCGTGTGGCAGCCAGCAGCCC  
AGCCCAAACGACGGAGACATCGACAGGACGCTCTGCCAGGACCTTGTATCGCCAGC  
ACACCAAAGAAGCACAGGGGGCTAA (SEQ ID NO:6) (MMLV Gag核苷酸序列加粗)

[0100]

密码子优化的MMLV Gag-CMV pp65核苷酸序列 (SEQ ID NO:22)

[0101]

ATGGGACAGACCGTCACAACACCCCTGAGCCTGACCCTGGGACATTGGAAAGA  
CGTGGAGAGGATCGCACATAACCAGAGCGTGGACGTGAAGAAACGGAGATGG  
GTCACATTCTGCAGTGCTGAGTGGCCAACCTTTAATGTGGGATGGCCCCGAGA  
CGGCACCTTCAACAGGGATCTGATCACCCAGGTGAAGATCAAGGTCTTTAGCC

[0102]

CAGGACCTCACGGACATCCAGACCAGGTGCCTTATATCGTCACCTGGGAGGCA  
CTGGCCTTCGATCCCCCTCCATGGGTGAAGCCATTTGTCCACCCAAAACCACCT  
CCACCCTGCCTCCAAGTGCCCTTCACTGCCACTGGAACCACCCCGGAGCAC  
ACCACCCCGCAGCTCCCTGTATCCTGCTCTGACTCCATCTCTGGGCGCAAAGCC  
AAAACCACAGGTGCTGAGCGACTCCGGAGGACCCTGATTGACCTGCTGACAG  
AGGACCCCCACCATAACGAGATCCTCGGCCTCCACCAAGCGACCGCGATGGA  
AATGGAGGAGAGGCTACTCCTGCCGGCGAAGCCCCCTGACCCATCTCCAATGGC  
TAGTAGGCTGCGCGGCAGGCGCGAGCCTCCAGTGGCAGATAGCACCATCCC  
AGGCCTTCCCTCTGAGGGCTGGGGGAAATGGGCAGCTCCAGTATTGGCCATTT  
TCTAGTTCAGACCTGTACAACCTGGAAGAACAATAACCCCTCTTTCAGTGAGGAC  
CCCGGCAAACCTGACCGCCCTGATCGAATCCGTGCTGATTACCCATCAGCCAC  
ATGGGACGATTGTGACGAGCTCCTGGGCACCCTGCTGACCGGAGAGGAAAAGC  
AGCGCGTGCTGCTGGAGGCTCGCAAAGCAGTCCGAGGGGACGATGGACGGCC  
CACACAGCTCCCTAATGAGGTGGACGCGCTTTTCCACTGGAAAGACCCGACT  
GGGATTATACTACCCAGGCAGGGAGAAACCACCTGGTCCATTACAGGCAGCTC  
CTGCTGGCAGGCCTGCAGAAATGCCGGGAGATCCCCCACCACCTGGCCAAGGT  
GAAAGGCATCACACAGGGGCCCTAATGAGTCACCAAGCGCCTTTCTGGAGAGGC  
TGAAGGAAGCTTACCGACGGTATACCCCATACGACCCTGAGGACCCCGGACAG  
GAAACAAACGTCTCCATGTCTTTCATCTGGCAGTCTGCCCCAGACATTGGGCG  
GAAGCTGGAGAGACTGGAAGACCTGAAGAACAAGACCCTGGGCGACCTGGTG  
CGGGAGGCTGAAAAGATCTTCAACAAACGGGAGACCCCCGAGGAAAGAGAGG  
AAAGGATTAGAAGGGAAACTGAGGAAAAGGAGGAACGCCGACGGACCGAGGA  
CGAACAGAAGGAGAAAGAACGAGATCGGCGGGCGGCACCGGGAGATGTCAAAG  
CTGCTGGCCACCGTGGTCAGCGGACAGAAACAGGACAGACAGGGAGGAGAGC  
GACGGAGAAGCCAGCTCGACAGGGATCAGTGCGCATACTGTAAGGAAAAAGGC  
CATTGGGCCAAGGATTGCCCCAAAAAGCCAAGAGGACCAAGAGGACCAAGACC  
ACAGACATCACTGCTGACCCTGGACGACTGCGAGAGCCGGGGCCGGCGGTGCC  
AGAAATGATCTCTGTGCTGGGGCCCATAGTGGACATGTGCTGAAGGCCGTCTTCTC  
CAGGGGAGACACCCCGTGCTGCCTCACGAGACTCGACTGCTGCAGACCGGCATCC  
ATGTGCGGGTCTCCAGCCCTCTCTGATTCTGGTGTACAGTATACACCAGATAGCA  
CTCCCTGCCACAGAGGAGACAATCAGCTCCAGGTGCAGCATACCTACTTTACAGGCT  
CCGAGGTGCAAAACGTGTCTGTCAATGTGCACAACCCTACCGGCAGGAGCATCTGT  
CCTAGCCAGGAGCCAATGAGCATCTACGTGTACGCCCTGCCTCTGAAGATGCTGAAT  
ATCCCATCAATTAACGTCCACCATTACCCTAGCGCAGCCGAACGGAAGCACAGACA  
TCTGCCAGTGGCCGACGCTGTATCCATGCCAGCGGCAAACAGATGTGGCAGGCAA  
GACTGACCGTGTCCGGGCTGGCCTGGACAAGGCAGCAGAATCAGTGGAAGGAGCCC  
GACGTGTACTATACCAGCGCCTTCGTGTTCCCTACCAAAGACGTGGCCCTGAGACAT  
GTGGTGTGCGCACATGAGCTGGTGTGCAGCATGGAAAACACTAGGGCCACCAAGAT  
GCAGGTATCGGCGATCAGTATGTCAAAGTGTACCTGGAGAGTTTTTGCGAAGACGT  
GCCATCAGGGAAGCTGTTTCATGCATGTGACCCTGGGCAGCGATGTCGAGGAAGACC  
TGACCATGACAAGAAATCCACAGCCCTTTATGAGACCCACGAGAGGAATGGGTTC  
ACTGTGCTGTGCCCCAAGAACATGATCATTAAGCCTGGAAAAATCAGTCATATTATG  
CTGGATGTGGCCTTTACATCACACGAGCATTTTCGGACTGCTGTGCCCCAAATCCATC  
CCTGGACTGAGCATTTCCGGCAATCTGCTGATGAACGGCCAGCAGATCTTCTGGAA  
GTGCAGGCCATCCGGGAGACCGTCGAACTGCGACAGTATGACCCAGTGGCTGCACT  
GTTCTTTTTTCGACATCGACCTGCTGCTGCAGCGAGGACCACAGTACAGCGAGCACC  
TACTTTTACCTCCAGTATCGGATTACAGGGGAAGCTGGAGTACAGGCACACCTGGGA  
TCGCCATGACGAAGGAGCCGCTCAGGGGGACGATGACGTGTGGACATCTGGCAGTG  
ATTCAGACGAGGAACTGGTGACAACCTGAGCGAAAAACCCCCGGGTGACAGGAGG  
AGGGGCAATGGCAGGGGCCAGCACACGCGCAGGGCGGAAGCGAAAAAGCGCCAGC  
AGCGCCACAGCATGTACCGCCGGCGTGATGACTAGAGGAAGGCTGAAGGCCGAGTC  
TACAGTCGCTCCCGAGGAAGATACTGACGAGGATAGTGACAATGAAATCCACAACC  
CCGCCGTGTTACCTGGCCACCTTGGCAGGCAGGGATTCTGGCTCGCAACCTGGTCC  
CCATGGTGGCAACCGTCCAGGGACAGAATCTGAAGTATCAGGAGTTTTTCTGGGATG  
CTAACGACATCTACCGGATTTTTCAGAGCTGGAAGGCCGTGTGGCAGCCAGCAGCC  
CAGCCCAAACGACGGAGACATCGACAGGACGCTCTGCCAGGACCTTGTATCGCCAG

[0103]

**CACACCAAAGAAGCACAGGGGCTAA (SEQ ID NO:22)** (MMLV Gag核苷酸序列加粗)

[0104] 在一些实施方案中,本发明提供了核酸,其编码Gag多肽或Gag多肽的特征部分。在某些实施方案中,核酸可为DNA或RNA,并且可为单链或双链。在一些实施方案中,本发明的核酸可包含一种或更多种非天然的核苷酸;在另一些实施方案中,本发明的核酸仅包含天然核苷酸。

#### [0105] B. 包膜蛋白

[0106] 在一些实施方案中,本发明使用由来自HCMV的一种或更多种包膜多肽(例如gB和/或gH)组成的VLP。本文所用的术语“包膜多肽”指其氨基酸序列包含包膜糖蛋白的至少一个特征序列的多肽序列。本领域和本领域技术人员已知来自多种病毒(包括但不限于HCMV)的各种各样的包膜糖蛋白序列,参考这样的序列,可容易地鉴别一般包膜糖蛋白和/或特定包膜糖蛋白的特征性序列。在一些实施方案中,包膜多肽包含胞质、跨膜和/或胞外部分或结构域。

[0107] 在一些实施方案中,来自HCMV的包膜多肽包含不天然地存在于HCMV蛋白中的跨膜和胞质结构域。例如,在一些实施方案中,来自HCMV的包膜多肽包含来自另一个HCMV蛋白(例如gB或gH)的跨膜和胞质结构域。在一些实施方案中,来自HCMV的包膜多肽包含天然地存在于水泡性口炎病毒(VSV)中的跨膜结构域和胞质结构域。本领域已知,多肽有时具有跨膜、胞质和/或胞外结构域。一般而言,本文所用的“跨膜结构域”指具有存在于膜中(例如跨越细胞膜的一部分或全部)的性质的结构域。应当理解,不需要跨膜结构域中的每个氨基酸都存在于膜中。例如,在一些实施方案中,跨膜结构域的特征在于蛋白的指定延伸或部分定位于膜中。本领域中公知,可以使用多种算法分析氨基酸或核酸序列以预测蛋白亚细胞定位(例如跨膜定位)。示例性的这样的程序包括psort (PSORT.org)、Prosit (prosite.expasy.org)等等。一般而言,本文所用的“胞质结构域”指具有存在于胞质中的性质的结构域。应当理解,不需要胞质结构域中的每一个氨基酸都存在于胞质中。例如,在一些实施方案中,胞质结构域的特征在于蛋白的指定延伸或部分基本上位于胞质中。本领域中公知,可以使用多种算法分析氨基酸或核酸序列以预测蛋白亚细胞定位(例如胞质定位)。示例性的这样的程序包括psort (PSORT.org)、Prosit (prosite.expasy.org)等等。

[0108] VSV-G的跨膜结构域的功能为将病毒糖蛋白靶向细胞膜(Compton T等,1989 Proc Natl Acad Sci USA 86:4112-4116)。将VSV-G的跨膜和胞质结构域与另一个蛋白的跨膜和胞质结构域交换已经用于将蛋白导向细胞膜,当天然的蛋白质不天然的发生这一情况或者需要辅助共表达的蛋白以完成这一情况时(Garrone P等,2011 Sci Transl Med 94:;)。

[0109] 除了其他方面以外,本发明涵盖以下认识:包含病毒(例如MLV)之结构组分和一个或更多个异源表面抗原(例如包膜蛋白)的VLP对于抗原递送和诱发针对异源抗原的免疫应答特别有效。

#### [0110] C. 异源抗原

[0111] 由于中和抗体通常能够预防感染,HCMV的包膜蛋白(例如糖蛋白gB和gH)是用于产生针对HCMV之中和抗体的重要靶标。已经开发了用于HCMV感染的治疗(例如gB亚单位疫苗)并且在试验动物和临床研究中测试。然而,来自这样的研究的结果已经证明了在人中,抗体应答不是终生的并且不足以在所有情况下有效地治疗HCMV。对于仅基于HCMV之gB的亚单位疫苗的受限效果,已经提出了原因,依次是免疫应答中毒株特异性变化、细胞免疫应答的不

充分诱导、和所用抗原的结构限制,所述抗原的表位被认为是构象依赖的。本发明人认识到,开发包含在VLP表面上以其天然构象存在之一种或更多种包膜多肽抗原的HCMV疫苗导致中和抗体的诱导(例如,经体液免疫应答),并且开发包含一种或更多种结构蛋白抗原(例如被膜蛋白pp65)的HCMV疫苗导致辅助T细胞( $T_H$ 淋巴细胞)和细胞毒性T细胞(CTL)的诱导(例如经细胞介导的免疫应答)。中和抗体通常针对病毒包膜蛋白形成,并且特别地针对HCMV糖蛋白gB和gH。 $T_H$ 细胞被病毒的被膜结构蛋白(例如HCMV pp65(ppUL83))刺激。此外,pp65在针对HCMV的CTL应答的诱导中发挥重要作用。

[0112] 应当理解,所提供的VLP可以包含任何异源的抗原,包括来自HCMV的异源抗原。例如,在一些实施方案中,根据本发明的VLP包含一种或更多种HCMV包膜多肽。在一些实施方案中,根据本发明的VLP包含一种或更多种HCMV结构多肽。在一些实施方案中,根据本发明的VLP包含一种或更多种HCMV包膜多肽和一种或更多种HCMV结构多肽。以下提供了示例性而非限制性的HCMV抗原列表。

[0113] gB-糖蛋白复合物(gC)I

[0114] 表征得最完全的HCMV的糖蛋白复合物是gB复合物(gB;UL55)。已经证明来自CMV血清反应阳性之个体的血清含有对gB的抗体,并且在康复期血清中,中和抗体应答的至多70%为是gB特异性的(Marshall GS等,1994 J Med Virol 43:77-83)。

[0115] 示例性的HCMVgB多肽氨基酸和核酸序列分别如以下SEQ ID NO:7和SEQ ID NO:8所示。在一些实施方案中,合适的gB多肽与已知的HCMV gB多肽基本同源。例如,gB多肽可以为经修饰的HCMV gB多肽,其与野生型或天然存在的gB多肽(例如SEQ ID NO:7)相比含有一个或更多个氨基酸替换、缺失和/或插入。因此,在一些实施方案中,适合用于本发明的gB多肽与HCMV gB多肽(SEQ ID NO:7)基本同源。在一些实施方案中,适合用于本发明的HCMV多肽具有与SEQ ID NO:7有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中,适合用于本发明的gB多肽与HCMV gB多肽(SEQ ID NO:7)基本同一。在一些实施方案中,适合用于本发明的gB多肽具有与SEQ ID NO:7有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。

[0116] HCMV gB氨基酸序列(SEQ ID NO:7)

[0117]

MESRIWCLVVCVNL CIVCLGAAVSSSSSTRGTSATHSHSSHTTSAAHSRSGSVSQRVTSS  
QTVSHGVNETIYNTTLKYGDVVGVNTTKYPYRVCSMAQGTDLIRFERNIVCTSMKPINE  
DLDEGIMVVYKRNIVAHTFKVRVYQKVLTFRRSYAYIHTTYLLGSNTEYVAPPMWEIH  
HINSHSQCYSSYSRVIAGTVFVAYHRDSYENKTMQLMPDDYSNTHSTRYVTVKDQWHS  
RGSTWLYRETCNLNCMVTITTARSKYPYHFFATSTGDVVDISPFYNGTNRNASYFGENA  
DKFFIFPNYITIVSDFGRPNSALETHRLVAFLERADSVISWDIQDEKNVTCQLTFWEASERT  
IRSEAEDSYHFSSAKMTATFLSKKQEVNMSDSALDCVRDEAINKLQQIFNTSYNQTYEK  
YGNVSVFETTGGLVVFWQGIKQKSLVELERLANRSSLNLTHNRTKRSTDGNNATHLSN  
MESVHNLVYAQLQFTYDTLRGYINRALAQIAEAWCVDQRRTLEVFKELSKINPSAILS  
AIYNKPIAARFMGDVGLASCVTINQTSVKVLRDMNVKESPGRCYSRPVVIFNFANSSYVQ  
YQQLGEDNEILLGNHRTEECQLPSLKIFIAGNSAYEYVDYLFKRMIDLSSISTVDSMIALD  
IDPLENTDFRVLELYSQKELRSINVFDLEEIMREFNSYKQRVKYVEDKVVDPLPPYLKGL  
DDLMSGLGAAGKAVGVAIGAVGGAVASVVEGVATFLKNPFGAFTIILVAIAVVIIIIYLIY  
TRQRRLCMOPLQNLFPYLVSADGTTVTSGNTKDTSLQAPPSYEESVYNSSGRKGPGPPSS

DASTAAPPYTNEQAYQMLLALVRLDAEQRAQONGTDSLDGQTGTQDKGQKPNLLDRL  
RHRKNGYRHLKDSDEEENV\* (SEQ ID NO:7) (TM和CD加下划线)

[0118]

HCMV gB核苷酸序列 (SEQ ID NO:8)



[0119]

ATGGAATCCAGGATCTGGTGCCTGGTAGTCTGCGTTAACTTGTGTATCGTCTGTCTG  
GGTGCTGCGGTTTCCTCATCTTCTACTCGTGGAACCTTCTGCTACTCACAGTCACCATT  
CCTCTCATACGACGTCTGCTGCTCATTCTCGATCCGGTTCAGTCTCTCAACGCGTAAC  
TTCTTCCCAAACGGTCAGCCATGGTGTAAACGAGACCATCTACAACACTACCCTCAA  
GTACGGAGATGTGGTGGGGGTCAACACCACCAAGTACCCCTATCGCGTGTGTTCTAT  
GGCACAGGGTACGGATCTTATTCGCTTTGAACGTAATATCGTCTGCACCTCGATGAA  
GCCCATCAATGAAGACCTGGACGAGGGCATCATGGTGGTCTACAAACGCAACATCG  
TCGCGCACACCTTTAAGGTACGAGTCTACCAGAAGGTTTTGACGTTTCGTCTAGCT  
ACGCTTACATCCACACCCTTATCTGCTGGGACGCAACACGGAATACGTGGCGCCTC  
CTATGTGGGAGATTCATCATATCAACAGTCACAGTCAGTGCTACAGTTCCTACAGCC  
GCGTTATAGCAGGCACGGTTTTCTGTGGCTTATCATAGGGACAGCTATGAAAACAAA  
ACCATGCAATTAATGCCCAGCATTATTCCAACACCCACAGTACCCGTTACGTGACG  
GTCAAGGATCAATGGCACAGCCGCGGCAGCACCTGGCTCTATCGTGAGACCTGTAA  
TCTGAATTGTATGGTGACCATCACTACTGCGCGCTCCAAGTATCCCTATCATTTTTTC  
GCAACTTCCACGGGTGATGTGGTTGACATTTCTCCTTTCTACAACGGAACATAATCGC  
AATGCCAGCTATTTTGGAGAAAACGCCGACAAGTTTTTTCATTTTTCCGAACCTACACT  
ATCGTCTCCGACTTTGGAAGACCGAATTCTGCGTTAGAGACCCACAGGTTGGTGGCT  
TTTCTTGAACGTGCGGACTCAGTGATCTCCTGGGATATACAGGACGAGAAGAATGTT  
ACTTGTCAACTCACTTTCTGGGAAGCCTCGGAACGCACCATTTCGTTCCGAAGCCGAG  
GACTCGTATCACTTTTCTTCTGCCAAAATGACCGCCACTTTCTTATCTAAGAAGCAAG  
AGGTGAACATGTCCGACTCTGCGCTGGACTGTGTACGTGATGAGGCCATAAATAAGT  
TACAGCAGATTTTCAATACTTCATACAATCAAACATATGAAAAATATGGAAACGTGT  
CCGTCTTTGAAACCACTGGTGGTTTGGTGGTGTCTGGCAAGGTATCAAGCAAAAAT  
CTCTGGTGGAACCTCGAACGTTTGGCCAACCGCTCCAGTCTGAATCTTACTCATAATA  
GAACCAAAAGAAGTACAGATGGCAACAATGCAACTCATTATCCAACATGGAGTCG  
GTGCACAATCTGGTCTACGCCCAGCTGCAGTTCACCTATGACACGTTGCGCGGTTAC  
ATCAACCGGGCGCTGGCGCAAATCGCAGAAGCCTGGTGTGTGGATCAACGGCGCAC  
CCTAGAGGTCTTCAAGGAACTTAGCAAGATCAACCCGTCAGCTATTCTCTCGGCCAT  
CTACAACAAACCGATTGCCGCGCGTTTCATGGGTGATGTCTGGGTCTGGCCAGCTG  
CGTGACCATTAACCAAAACAGCGTCAAGGTGCTGCGTGATATGAATGTGAAGGAAT  
CGCCAGGACGCTGCTACTCACGACCAGTGGTCATCTTTAATTTGCCAACAGCTCGT  
ACGTGCAGTACGGTCAACTGGGCGAGGATAACGAAATCCTGTTGGGCAACCAACCGC  
ACTGAGGAATGTCAGCTTCCCAGCCTCAAGATCTTCATCGCCGGCAACTCGGCCTAC  
GAGTACGTGGACTACCTCTTCAAACGCATGATTGACCTCAGCAGCATCTCCACCGTC  
GACAGCATGATCGCCCTAGACATCGACCCGCTGGAAAACACCGACTTCAGGGTACT  
GGAACCTTACTCGCAGAAAGAATTGCGTTCATCAACGTTTTTGATCTCGAGGAGAT  
CATGCGCGAGTTCAATTTCGTATAAGCAGCGGGTAAAGTACGTGGAGGACAAGGTAG  
TCGACCCGCTGCCGCCCTACCTCAAGGGTCTGGACGACCTCATGAGCGGCCTGGGCG  
CCGCGGGAAGGCCGTTGGCGTAGCCATTGGGGCCGTGGGTGGCGCGGTGGCCTCC  
GTGGTCAAGGCGTTGCCACCTTCCTCAAAAACCCCTTCGGAGCCTTCACCATCATC  
CTCGTGGCCATAGCCGTCGTCAATTATTTATTTGATCTATACTCGACAGCGGCGTC

TCTGCATGCAGCCGCTGCAGAACCTCTTTCCCTATCTGGTGTCCGCCGACGGGACCA  
CCGTGACGTGCGGCAACACCAAGACACGTCGTTACAGGCTCCGCCTTCCTACGAG  
GAAAGTGTTTATAATTCTGGTCGCAAAGGACCGGGACCAACCGTCGTCTGATGCATCC  
ACGGCGGCTCCGCCTTACACCAACGAGCAGGCTTACCAGATGCTTCTGGCCCTGGTC  
CGTCTGGACGCAGAGCAGCGAGCGCAGCAGAACGGTACAGATTCTTTGGACGGACA  
GACTGGCACGCAGGACAAGGGACAGAAGCCCAACCTGCTAGACCGACTGCGACACC  
GCAAAAACGGCTACCGACACTTGAAAGACTCCGACGAAGAAGAGAACGTCTGA

(SEQ ID NO:8) (TM和CD加下划线)

[0120]

密码子优化的HCMV gB核苷酸序列 (SEQ ID NO:9)

[0121]

ATGGAGTCAAGGATTTGGTGCCTGGTTCGTGTGCGTCAATCTGTGCATCGTCTGTCTG  
GGGGCTGCCGTGTCATCAAGTTCTACAAGAGGCACCAGCGCCACCCACTCACACCA  
TAGCTCCCATAACCACATCCGCCGCTCACTCCCGGTCTGGCAGCGTGAGCCAGAGAGT  
CACATCTAGTCAGACCGTGAGCCACGGGGTCAACGAGACCATCTACAATACTACCC  
TGAAGTATGGCGACGTGGTTCGGGGTGAACACAACATAAATACCCATATAGGGTCTGC  
AGTATGGCCCAGGGCACTGATCTGATTAGATTTCGAAAGGAACATCGTGTGCACCAG  
CATGAAGCCCATTAATGAGGACCTGGATGAAGGGATCATGGTGGTCTACAAACGCA  
ATATTGTGGCCATACCTTCAAGGTGCGAGTCTATCAGAAAGTGCTGACATTTTCGGA  
GATCTTACGCATATATCCACACCACATACCTGCTGGGGAGTAACACCGAGTATGTGG  
CTCCCCCTATGTGGGAAATTCACCATATCAATAGCCATTCCAGTGCTACTCAAGCT  
ACAGCAGAGTGATCGCTGGAACAGTGTTCGTGCGATACCACAGAGACTCTTATGAG  
AACAAAGACTATGCAGCTCATGCCCCGACGATTACAGCAATACACATTCCACTAGATAT  
GTGACAGTCAAAGATCAGTGGCACTCAAGGGGCAGCACCTGGCTGTACCGCGAGAC  
ATGCAACCTGAATTGTATGGTGACTATCACTACCGCTAGATCCAAGTACCCCTATCA  
CTTCTTTGCAACTTCCACCGGGGACGTGGTTCGATATTTCTCCTTTCTACAACGGCACA  
AACCGGAATGCATCTTATTTTGGGGAGAACGCCGACAAGTTCTTTATTTTCCCAAAT  
TACACCATCGTGTCTGATTTTGGCAGACCCAACAGTGCCTTGGAGACACATCGACTG  
GTGGCATTCCTGGAACGGGCCGACTCCGTCATTTCTTGGGACATCCAGGATGAGAAG  
AATGTGACCTGCCAGCTCACCTTCTGGGAGGCCAGCGAACGCACCATCCGATCCGA  
GGCTGAAGATTCTTACCACTTCTCCTCTGCCAAAATGACAGCTACTTTTCTGAGCAA  
GAAACAGGAGGTGAACATGTCTGACAGTGCTCTGGATTGCGTGCGGGACGAAGCAA  
TTAATAAGCTGCAGCAGATCTTCAACACATCATAACAACAGACTTACGAGAAGTAC  
GGAAACGTGAGCGTCTTCGAAACAACCTGGCGGGCTGGTGGTCTTTTGGCAGGGCAT  
CAAGCAGAAATCCCTGGTGGAGCTGGAAAGGCTGGCCAATCGCAGTTCAGTGAACC  
TGACTCATAATCGGACCAAGAGATCTACAGACGGAAACAATGCCACACATCTGTCT  
AACATGGAGAGTGTGCACAATCTGGTCTACGCTCAGCTCCAGTTTACCTACGACACA  
CTGAGAGGCTATATTAACAGGGCACTGGCCAGATCGCTGAAGCATGGTGCGTGGA  
TCAGAGGCGCACCCCTGGAGGTCTTCAAGGAACTGTCCAAAATCAACCCTTCAGCAA  
TTCTGAGCGCCATCTACAATAAGCCAATTGCAGCCAGGTTTATGGGAGACGTGCTGG  
GCCTGGCCAGTTGCGTCACTATCAACCAGACCTCAGTGAAGGTCCTGCGCGATATGA  
ATGTGAAAGAGAGTCCCGGCAGATGCTATTACAGGCCTGTGGTCATCTTCAACTTTG  
CTAATAGCTCCTACGTGCAGTATGGACAGCTCGGCGAGGACAACGAAATTCTGCTG  
GGGAATCACAGGACCGAGGAATGTCAGCTCCCTAGCCTGAAGATTTTCATCGCTGG  
AAACTCCGCATACGAGTATGTGGATTACCTGTTCAAGCGGATGATTGACCTGTCTAG  
TATCTCCACTGTGGATTCTATGATTGCCCTGGACATCGATCCACTGGAAAAATACCGA  
CTTCAGGGTGCTGGAGCTGTATAGCCAGAAGGAACTGCGCTCCATCAACGTGTTTCGA  
TCTGGAGGAAATTATGAGAGAGTTTAATAGCTACAAGCAGAGGGTGAAATATGTGCG  
AAGATAAGGTGGTCGACCCCCTGCCACCCTACCTGAAAGGCCTGGACGATCTGATG  
AGCGGGCTGGGAGCTGCAGGGAAGGCAGTGGGAGTCGCTATCGGCGCAGTGGGAG  
GAGCCGTGGCCAGCGTGGTCGAGGGAGTGGCAACATTCTGAAAAACCCCTTCGGG  
GCCTTCACCATCATTCTGGTGGCAATCGCCGTGGTCATCATTATCTACCTGATCTACA  
CAAGGCAGCGGCGGCTGTGCATGCAGCCTCTGCAGAACCTGTTTCCATACCTGGTGA  
GCGCCGACGGGACCACAGTCACCTCAGGAAATACTAAGGATACCTCTCTGCAGGCC  
CCCCAAGTTACGAGGAATCAGTGTATAACAGCGGCAGAAAAGGACCAGGACCACC  
TTCAAGCGACGCCAGCACTGCCGCTCCACCCTACACCAATGAGCAGGCCTATCAGAT  
GCTGCTGGCTCTGGTGCCTGGATGCCGAACAGCGAGCTCAGCAGAACGGGACCG  
ACTCCCTGGATGGACAGACCGGAACACAGGACAAGGGACAGAAACCTAATCTGCTG  
GATCGGCTGCGGCACAGAAAAAACGGGTATAGGCACCTGAAGGACTCCGACGAAG  
AAGAAAATGTCTAA (SEQ ID NO:9) (TM和CD加下划线)

[0122]

[0123] 在一些实施方案中,根据本发明使用的gB多肽缺乏跨膜结构域和/或胞质结构域和/或包含经修饰的跨膜结构域和/或胞质结构域。gB多肽可以任选地包含一种或更多种另外的多肽(例如异源跨膜结构域和/或胞质结构域多肽)。在一些实施方案中,gB多肽表达为与异源多肽的融合蛋白。gB多肽可连接至异源多肽以产生融合蛋白而不改变gB功能和/或

抗原性。例如,可将异源多肽的编码序列剪接成gB多肽编码序列中,例如在gB多肽编码序列的3'端。在一些实施方案中,可将异源多肽的编码序列框内剪接成gB多肽编码序列。在一些实施方案中,gB多肽编码序列和异源多肽可以通过单一启动子表达。在一些实施方案中,在gB多肽的C末端插入(例如融合至)异源多肽。

[0124] 在一些实施方案中,异源多肽是或包含存在于水泡性口炎病毒(VSV)中的跨膜结构域和/或胞质结构域。在一些实施方案中,缺乏跨膜结构域和/或胞质结构域的gB与来自VSV的跨膜结构域和/或胞质结构域融合。根据本发明使用的示例性gB-VSV融合多肽如下SEQ ID NO:10所示。在一些实施方案中,合适的gB-VSV多肽融合蛋白包含与已知gB多肽基本同源的gB多肽的全部或一部分和与已知VSV多肽基本同源的VSV多肽的全部或一部分。例如,与野生型或天然存在的gB和/或VSV多肽相比,gB-VSV多肽融合蛋白可以含有一个或多个氨基酸替换、缺失和/或插入。因此,在一些实施方案中,适合用于本发明的gB-VSV多肽融合蛋白与SEQ ID NO:10基本同源。在一些实施方案中,适合用于本发明的gB-VSV多肽融合蛋白具有与SEQ ID NO:10有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中,适合用于本发明的gB-VSV多肽融合蛋白与SEQ ID NO:10基本同一。在一些实施方案中,适合用于本发明的gB-VSV多肽融合蛋白具有与SEQ ID NO:10有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。本文所用的“gB-G”指HCMV gB-VSV TM/CTD融合蛋白。

[0125] HCMV gB-G氨基酸序列(SEQ ID NO:10)

MESRIWCLVVCVNLCTVCLGAAVSSSSSTRGTSATHSHSSHTTSAAHSRSGSVSQRVTS  
QTVSHGVNETIYNTTLKYGDVVGVTTKYPYRVCSMAQGTDLIRFERNIVCTSMKPINE  
DLDEGIMVVYKRNIVAHTFKVRVYQKVLTFRRSYAYIHTTYLLGSNTEYVAPPMWEIH  
HINSHSQCYSSYSRVIAGTVFVAYHRDSYENKTMQLMPDDYSNTHSTRYVTVKDQWHS  
RGSTWLYRETCNLNCMVTITTARSKYPYHFFATSTGDDVDDISPFYNGTNRNASYFGENA  
DKFFIFPNYTIVSDFGRPNSALETHRLVAFLERADSVISWDIQDEKNVTCQLTFWEASERT  
[0126] IRSEAEDSYHFSSAKMTATFLSKKQEVNMSDSALDCVRDEAINKLQIFNTSYNQTYEK  
YGNVSVFETTGGLVVFWQGIKQKSLVELERLANRSSLNLTHNRTKRSTDGNNATHLSN  
MESVHNLVYAQLQFTYDTRLRGYINRALAQIAEAWCVDQRRRTLEVFKELSKINPSAILS  
YNKPIAARFMGDVGLGLASCVTINQTSVKVLRDMNVKESPGRCYSRPVVIFNFANSSYVQ  
YGQLGEDNEILLGNHRTTEECQLPSLKIFIAGNSAYEYVDYLFKRMIDLSSISTVDSMIALD  
IDPLENTDFRVLELYSQKELRSINVDLEEIMREFNSYKQRVKYVEDKVVDPLPPYLKGL  
DDLMSGLGAAGKAVGVAIGAVGGAVASVVEGVATFLKNPFFFFIIGLIIGLFLVLRVGIHL  
CIKLKHTKKRQIYTDIEMNRLGK\* (SEQ ID NO:10) (TM和CTD加下划线)

[0127] HCMV gB-G核苷酸序列(SEQ ID NO:11)

ATGGAATCCAGGATCTGGTGCCTGGTAGTCTGCGTTAACTTGTGTATCGTCTGTCTG  
GGTGCTGCGGTTTCCTCATCTTCTACTCGTGGAACCTTCTGCTACTCACAGTCACCAT  
[0128] CCTCTCATACGACGTCTGCTGCTCATTCTCGATCCGGTTCAGTCTCTCAACGCGTAAC  
TTCTTCCCAAACGGTCAGCCATGGTGTTAACGAGACCATCTACAACACTACCCTCAA  
GTACGGAGATGTGGTGGGGGTCAACACCACCAAGTACCCCTATCGCGTGTGTTCTAT

[0129]

GGCACAGGGTACGGATCTTATTCGCTTTGAACGTAATATCGTCTGCACCTCGATGAA  
GCCCATCAATGAAGACCTGGACGAGGGCATCATGGTGGTCTACAAACGCAACATCG  
TCGCGCACACCTTTAAGGTACGAGTCTACCAGAAGGTTTTGACGTTTTCGTCGTAGCT  
ACGCTTACATCCACACCACTTATCTGCTGGGCAGCAACACGGAATACGTGGCGCCTC  
CTATGTGGGAGATTTCATCATATCAACAGTCACAGTCAGTGCTACAGTTCCTACAGCC  
GCGTTATAGCAGGCACGGTTTTCTGTGGCTTATCATAGGGACAGCTATGAAAACAAA  
ACCATGCAATTAATGCCCAGCATTATTCCAACACCCACAGTACCCGTTACGTGACG  
GTCAAGGATCAATGGCACAGCCGCGGCAGCACCTGGCTCTATCGTGAGACCTGTAA  
TCTGAATTGTATGGTGACCATCACTACTGCGCGCTCCAAGTATCCCTATCATTTTTTC  
GCAACTTCCACGGGTGATGTGGTTGACATTTCTCCTTTCTACAACGGAACATAATCGC  
AATGCCAGCTATTTTTGGAGAAAACGCCGACAAGTTTTTCATTTTTCCGAACCTACACT  
ATCGTCTCCGACTTTGGAAGACCGAATTCTGCGTTAGAGACCCACAGGTTGGTGGCT  
TTTTCTGAACGTGCGGACTCAGTGATCTCCTGGGATATACAGGACGAGAAGAATGTT  
ACTTGTCAACTCACTTTCTGGGAAGCCTCGGAACGCACCATTTCGTTCCGAAGCCGAG  
GACTCGTATCACTTTTCTTCTGCCAAAATGACCGCCACTTTCTTATCTAAGAAGCAAG  
AGGTGAACATGTCCGACTCTGCGCTGGACTGTGTACGTGATGAGGCCATAAATAAGT  
TACAGCAGATTTTCAATACTTCATACAATCAAACATATGAAAAATATGGAAACGTGT  
CCGTCTTTGAAACCACTGGTGGTTTGGTGGTGTCTGGCAAGGTATCAAGCAAAAAT  
CTCTGGTGGAACCTCGAACGTTTGGCCAAACCGCTCCAGTCTGAATCTTACTCATAATA  
GAACCAAAAGAAGTACAGATGGCAACAATGCAACTCATTTATCCAACATGGAGTCG  
GTGCACAATCTGGTCTACGCCAGCTGCAGTTCACCTATGACACGTTGCGCGGTTAC  
ATCAACCGGGCGCTGGCGCAAATCGCAGAAGCCTGGTGTGTGGATCAACGGCGCAC  
CCTAGAGGTCTTCAAGGAACTTAGCAAGATCAACCCGTCAGCTATTCTCTCGGCCAT  
CTACAACAAACCGATTGCCGCGCTTTCATGGGTGATGTCCTGGGTCTGGCCAGCTG  
CGTGACCATTAACCAAACAGCGTCAAGGTGCTGCGTGATATGAATGTGAAGGAAT  
CGCCAGGACGCTGCTACTCACGACCAGTGGTCATCTTTAATTTCCGCAACAGCTCGT  
ACGTGCAGTACGGTCAACTGGGCGAGGATAACGAAATCCTGTTGGGCAACCACCGC  
ACTGAGGAATGTCAGCTTCCCAGCCTCAAGATCTTCATCGCCGGCAACTCGGCCTAC  
GAGTACGTGGACTACCTCTTCAAACGCATGATTGACCTCAGCAGCATCTCCACCGTC  
GACAGCATGATCGCCCTAGACATCGACCCGCTGGAAAACACCGACTTCAGGGTACT  
GGAACCTTACTCGCAGAAAGAATTGCGTTCCATCAACGTTTTTGTATCTCGAGGAGAT  
CATGCGCGAGTTCAATTCGTATAAGCAGCGGGTAAAGTACGTGGAGGACAAGGTAG  
TCGACCCGCTGCCGCCCTACCTCAAGGGTCTGGACGACCTCATGAGCGGCCTGGGCG  
CCGCGGGAAGGGCCGTTGGCGTAGCCATTGGGGCCGTGGGTGGCGCGGTGGCCTCC  
GTGGTCAAGGCGTTGCCACCTTCCTCAAAAACCCCTTTTTCTTTATCATAGGGTTAA  
TCATTGGACTATTCTTGGTTCTCCGAGTTGGTATCCATCTTGCATTAAATTAAAGCA  
CACCAAGAAAAGACAGATTTATACAGACATAGAGATGAACCGACTTGAAAGTAA

(SEQ ID NO:11) (TM和CTD加下划线)

[0130]

密码子优化的HCMA gB-G核苷酸序列 (SEQ ID NO:12)

[0131]

ATGGAGTCAAGGATTTGGTGTCTGGTCTGCGTCAACCTGTGCATTGTCTGCCTG  
GGAGCCGCCGTCTCATCATCTACCCGAGGCACATCCGCCACTCACTCTCACCAT  
AGTCCCATAACACATCCGCCGCTCACTCAAGAAGCGGGTCCGTGTCTCAGAGGGTC  
ACATCTAGTCAGACCGTGAGCCATGGAGTCAACGAGACAATCTACAATACTACCT  
GAAGTATGGAGACGTGGTCGGCGTGAACACAATAACCCCTATAGGGTCTGCT  
CTATGGCCCAGGGGACAGATCTGATCCGATTTGAACGGAACATCGTGTGCACTAGC

[0132]

ATGAAGCCTATCAATGAGGACCTGGATGAAGGAATTATGGTGGTCTACAAACGAAA  
TATCGTGGCCCATACTTTTAAGGTGAGAGTCTATCAGAAAGTGCTGACCTTCCGGAG  
AAGCTACGCTTATATTACACCACATACCTGCTGGGGTCCAACACCGAGTATGTGGC  
ACCCCTATGTGGGAAATCCACCATATTAATAGTCATTACAGTGCTACTCAAGCTA  
CAGCAGAGTGATCGCTGGAACCGTGTTCTGCTCGCATACCACAGAGACAGTTATGAGA  
ACAAGACAATGCAGCTCATGCCGACGATTACAGTAATACCCATTCAACAAGATAT  
GTGACCGTCAAAGATCAGTGGCACTCTCGCGGCAGTACCTGGCTGTACCGAGAGAC  
ATGCAACCTGAATTGTATGGTGACAATTACTACCGCCAGAAGCAAGTACCCTTATCA  
CTTCTTTGCTACCTCAACAGGGGACGTGGTTCGACATCAGCCCCTTCTACAACGGAAC  
AAACCGGAATGCCTCCTATTTTCGGCGAGAACGCTGACAAATTCTTTATCTTCCCCAA  
CTACACTATCGTGAGCGATTTTCGGCAGACCTAACAGTGCCCTGGAGACCCATCGGCT  
GGTGGCATTTCTGGAAAGAGCCGACAGCGTGATCTCCTGGGACATTACAGGATGAGA  
AGAATGTGACCTGCCAGCTCACCTTCTGGGAGGCCAGCGAAAGAACCATCAGGTCC  
GAGGCAGAAGATTCTTACCACTTTTCCTCTGCAAAAATGACTGCCACCTTCTGTCC  
AAGAAACAGGAGGTGAACATGAGCGACTCCGCACTGGATTGCGTGCGGGACGAAGC  
CATCAATAAGCTGCAGCAGATCTTCAACACATCTTACAACCAGACTTACGAGAAGTA  
CGGCAACGTGAGTGTCTTTGAAACAACTGGCGGGCTGGTGGTCTTCTGGCAGGGGAT  
CAAGCAGAAATCTCTGGTGGAGCTGGAACGGCTGGCCAATAGAAGTTCACCTGAACC  
TGACTCATAATCGCACCAAGCGATCCACAGACGGAAACAATGCAACTCATCTGAGC  
AACATGGAGTCCGTGCACAATCTGGTCTACGCCAGCTCCAGTTCACCTACGACACC  
CTGCGAGGCTATATCAACCGGGCCCTGGCTCAGATTGCAGAAGCCTGGTGCCTGGAT  
CAGAGGCGCACCCCTGGAGGTCTTTAAGGAACTGAGCAAAATTAACCCATCTGCTATC  
CTGAGTGCAATCTACAATAAGCCCATCGCAGCCAGGTTCATGGGGGACGTGCTGGG  
ACTGGCCTCCTGCGTCACTATCAACCAGACCTCTGTGAAGGTCTGCGCGATATGAA  
TGTGAAAGAGAGTCCTGGCAGGTGTTATTCACGCCAGTGGTCATCTTCAACTTCGC  
TAATAGCTCCTACGTGCAGTATGGCCAGCTCGGGGAGGACAACGAAATCCTGCTGG  
GAAATCACAGGACCGAGGAATGTCAGCTCCCAAGTCTGAAGATCTTTATTGCCGGC  
AACTCAGCTTACGAGTATGTGGATTACCTGTTCAAACGCATGATCGACCTGTCTAGT  
ATTTCAACAGTGGATAGCATGATCGCCCTGGACATTGATCCCCTGGAAAATACTGAC  
TTCAGGGTGCTGGAGCTGTATAGCCAGAAGGAACTGCGCTCCATTAACGTGTTTGAT  
CTGGAGGAAATCATGAGGGAGTTCAATTCCTACAAGCAGCGCGTGAAATATGTGCA  
AGATAAGGTGGTCGACCCTCTGCCACCCTACCTGAAAGGCCTGGACGATCTGATGA  
GCGGGCTGGGAGCTGCAGGCAAGGCAGTGGGAGTCGCCATCGGAGCTGTGGGAGGC  
GCTGTGCGATCCGTGGTCGAGGGAGTGGCTACCTTTCTGAAGAACCCATTCTTTTC  
ATCATCGGCCTGATCATTGGGCTGTTCTGGTGTGAGAGTCGGCATCCACCTGTGC  
ATTAAGCTGAAGCACACCAAGAAGAGGCAGATCTACACCGATATTGAAATGAACAG  
ACTGGGCAAGTGA (SEQ ID NO:12) (TM和CTD加下划线)

[0133] gH-糖蛋白复合物(gC) III

[0134] gc III复合物包含糖蛋白gH(UL75)、gL(UL115)和gO(UL74) (Urban M等,1996 J Gen Virol 77:1537-47)。像gB一样,gH在人致病性疱疹病毒中是保守的并且在HCMV复制期间的多个步骤中发挥重要作用。HCMV编码两个gH/gL复合物:gH/gL/gO和gH/gL/UL128/UL130/UL131复合物(Wang D和Shenk T 2005 Proc Natl Acad USA 102:18153-8)。含有gO的复合物通常足以使成纤维细胞感染HCMV,而含有UL128/UL130/UL131的复合物对于HCMV感染内皮细胞和上皮细胞是需要的(Wang D和Shenk T 2005 J Virol 79 10330-8)。天然感染HCMV通常引发对表皮细胞进入有特异性的高效价中和抗体并且已经证明针对gH/gL/UL128/UL130/UL131表位的抗体可以包含这一活性的主要组分(Macagno A et al.,2010 J Virol 84:1005-13)。对gH的免疫学研究已经证明在哺乳动物细胞中,所述蛋白需要另外的多肽(如gL)用于正确地加工和转运至细胞膜(Urban M等,1996 J Gen Virol 77:1537-1547)。如果单独表达,仅在胞质和/或核膜中发现gH(Cranage MP等,1988 J Virol 62:1416-1422)。

[0135] 示例性HCMV gH多肽氨基酸和核酸序列分别如以下SEQ ID NO:13和SEQ ID NO:14所示。在一些实施方案中,合适的gH多肽与已知HCMV gH多肽基本同源。例如,gH多肽可以为经修饰的HCMV gH多肽,其与野生型或天然存在的gH多肽(例如SEQ ID NO:13)相比含有一个或多个氨基酸替换、缺失和/或插入。因此,在一些实施方案中,适合用于本发明的gH多肽与HCMV gH多肽(SEQ ID NO:13)基本同源。在一些实施方案中,适合用于本发明的HCMV gH多肽具有与SEQ ID NO:13有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中,适合用于本发明的gH多肽与HCMV gH多肽(SEQ ID NO:13)基本同一。在一些实施方案中,适合用于本发明的gH多肽具有与SEQ ID NO:13有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。

[0136] HCMV gH氨基酸序列(SEQ ID NO:13)

MRPGLPSYLIVLAVCLLSHLLSSRYGAEASEPLDKAFHLLLNTYGRPIRFLRENTTQCT  
YNSSLRNSTVVRENAISFNFFQSYNQYYVFHMPRCLFAGPLAEQFLNQVDLTETLERYQ  
QRLNTYALVSKDLASYRSFSQQLKAQDSLGEQPTTVPPPIDLSIPHVWMPPTTPHGW  
ESHTTSGLHRPHFNQTCILFDGHDLLFSTVTPCLHQGFYLIDELRYVKITLTEDFFVVTVSI

[0137]

DDDTPMLLIFGHLPRVLFKAPYQRDNFILRQTEKHELLVLVKKDQLNRHSYKDPDFLD  
AALDFNYLDLSALLRNSFHRYAVDVLKSGRCQMLDRRTVEMAFAYALALFAAARQEE  
AGAQVSVPRALDRQAALLQIQEFMITCLSQTTPRTTLLLYPTAVDLAKRALWTPNQITDI  
TSLVRLVYILSKQNQQHLIPQWALRQIADFALKLHKTHLASFLSAFARQELYLMGSLVH  
SMLVHTTERREIFIVETGLCSLAELSHFTQLLAHPHHEYLSDLYTPCSSSGRRDHSLERLT  
RLFPDATVPTTVPAALSILSTMQPSTLETFPDLFCLPLGESFSALTVEHVSYYVTNQYLI  
KGISYPVSTTVVGQSLITQDSQTKCELTRNMHTTHSITAALNISLENCAFCQSALLEYD  
DTQGVINIMYMHDSDDVLFALDPYNEVVVSSPRTHYLMLLKNGTVLEVTDVVVDATDS  
RLLMMSVYALSAIIGIYLLYRMLKTC\* (SEQ ID NO:13) (TM和CTD加下划线)

[0138] HCMV gH核苷酸序列(SEQ ID NO:14)

[0139]

ATGCGGCCAGGCCTCCCCTCCTACCTCATCGTCCTCGCCGTCTGTCTCCTCAGCCACC  
TACTTTTCGTACGATATGGCGCAGAAGCCATATCCGAACCGCTGGACAAAGCGTTTC  
ACCTACTGCTCAACACCTACGGGAGACCCATCCGCTTCCTGCGTGAAAAACACCACC  
AGTGTACCTACAATAGCAGCCTCCGTAACAGCACGGTCGTAGGGAAAAACGCCATC  
AGTTTCAACTTTTTCCAAAGCTATAATCAATACTATGTATTCCATATGCCTCGATGTC  
TTTTTGCGGGTCTCTGGCGGAGCAGTTTCTGAACCAGGTAGATCTGACCGAAACCC  
TGGAAAGATACCAACAGAGACTTAACACTTACGCGCTGGTATCCAAAGACCTGGCC  
AGCTACCGATCTTTTTCGCAGCAGCTAAAGGCACAGGACAGCCTAGGTGAACAGCC  
CACCCTGTGCCACCACCCATTGACCTGTCAATACCTCACGTTTGGATGCCACCGCA  
AACCCTCCACACGGCTGGACAGAATCACATACCACCTCAGGACTACACCGACCAC  
ACTTTAACCAGACCTGTATCCTCTTTGATGGACACGATCTACTATTCAGCACCGTCAC  
ACCTTGTTTGCACCAAGGCTTTTACCTCATCGACGAACACTACGTTACGTTAAAAATAAC  
ACTGACCGAGGACTTCTTCGTAGTTACGGTGTCCATAGACGACGACACACCCATGCT  
GCTTATCTTTCGGCCATCTTCCACGCGTACTCTTTAAAGCGCCCTATCAACGCGACAA  
CTTTATACTACGACAAACTGAAAAACACGAGCTCCTGGTGCTAGTTAAGAAAGATC  
AACTGAACCGTCACTCTTATCTCAAAGACCCGACTTTCTTGACGCCGCACTTGACT  
TCAACTACCTGGACCTCAGCGCACTACTACGTAACAGCTTTCACCGTTACGCCGTGG  
ATGTACTCAAAGCGGTTCGATGTGATGCTGGACCGCCGCACGGTAGAAATGGCC  
TTCGCCTACGCATTAGCACTGTTTCGACGAGCCCGACAAGAAGAGGCCGGCGCCCA  
AGTCTCCGTCCCACGGGCCCTAGACCGCCAGGCCGCACTCTTACAAATACAAGAATT  
TATGATCACCTGCCTCTCACAAACACCACCACGCACCACGTTGCTGCTGTATCCAC  
GGCCGTGGACCTGGCCAAACGAGCCCTTTGGACACCGAATCAGATCACCGACATCA  
CCAGCCTCGTACGCCTGGTCTACATACTCTCTAAACAGAATCAGCAACATCTCATCC  
CCCAGTGGGCACTACGACAGATCGCCGACTTTGCCCTAAAACACACAAACGCAC  
CTGGCCTCTTTTCTTTCAGCCTTCGCGCGTCAAGAACTCTACCTCATGGGCAGCCTCG  
TCCACTCCATGCTAGTACATACGACGGAGAGACGCGAAATCTTCATCGTAGAAACG  
GGCCTCTGTTTATTAGCCGAGCTATCACACTTTACGCAGTTGCTAGCTCATCCGCAC  
CACGAATACCTCAGCGACCTGTACACACCCTGTTCCAGTAGCGGGCGACGCGATCA  
CTCGCTCGAACGCCTCACACGTCTCTTCCCCGATGCCACCGTCCCCACTACCGTTCCC  
GCCGCCCTCTCCATCCTATCTACCATGCAACCAAGCACGCTAGAAACCTTCCCCGAC  
CTGTTTTGTCTGCCGCTCGGCGAATCCTTCTCCGCGCTGACCGTCTCCGAACACGTCA  
GTTATGTGCTAACAAACCAGTACCTGATCAAAGGTATCTCCTACCCTGTCTCCACCA  
CCGTCGTAGGCCAGAGCCTCATCATCACCCAGACGGACAGTCAAATAAATGCGAA  
CTGACGCGCAACATGCATACCACACACAGCATCACAGCGGCGCTCAACATTTCCCTA  
GAAAACTGCGCCTTTTGCCAAAGCGCCCTACTAGAATACGACGACACGCAAGGCGT

CATCAACATCATGTACATGCACGACTCGGACGACGTCCTTTTTCGCCCTGGATCCCTA  
CAACGAAGTGGTGGTCTCATCTCCGCGAACTCACTACCTCATGCTTTTGAAAAACGG  
TACGGTCCTAGAAGTAACTGACGTCGTGCTGGACGCTACCGACAGTCGTCTCCTCAT  
GATGTCCGTCTACGCGCTATCGGCCATCATCGGCATCTATCTGCTCTACCGCATGCTC

AAGACATGCTGA (SEQ ID NO:14) (TM和CTD加下划线)

[0140]

密码子优化的HCMV gH核苷酸序列 (SEQ ID NO:15)

[0141]

ATGAGACCTGGACTGCCTTCTTATCTGATTGTGCTGGCCGCTCTGCCTGCTGTCACATC  
TGCTGAGTTCACGCTATGGGGCTGAGGCTATCTCCGAGCCACTGGACAAGGCTTTTC  
ACCTGCTGCTGAACACCTACGGGAGACCCATTAGGTTCTGCGGAGAATACCACA  
CAGTGCACATATAACAGCTCCCTGCGGAACAGCACTGTGGTCCGCGAAAACGCCAT  
CTCTTTTAATTTCTTTCAGAGTTACAACCAGTACTACGTGTTCCATATGCCACGCTGT  
CTGTTTGCAGGACCCCTGGCCGAGCAGTTCCTGAACCAGGTGGACCTGACCGAGAC  
ACTGGAAGATAACCAGCAGAGGCTGAATACCTATGCCCTGGTGAGTAAGGATCTGG  
CTTCATATCGGTCTTTTCAGTCAGCAGCTCAAGGCCAGGACTCACTGGGCGAGCAGC  
CTACTACCGTGCCCCCTCCAATCGATCTGAGCATTCCACACGTCTGGATGCCCCCTC  
AGACAACTCCCCACGGCTGGACCGAAAGCCATACCACATCCGGGCTGCACAGACCC  
CATTTCAACCAGACATGCATCCTGTTTGATGGGCACGACCTGCTGTTTCAGCACTGTG  
ACCCCTTGCTGCATCAGGGATTCTACCTGATCGATGAGCTGAGATATGTGAAAATT  
ACACTGACTGAAGACTTCTTTGTGGTCACCGTGAGCATCGACGATGACACACCAATG  
CTGCTGATTTTTGGACACCTGCCCCGGGTGCTGTTCAAGGCCCCCTACCAGCGAGAC  
AACTTTATTCTGCGGCAGACCGAGAAACACGAACTGCTGGTGCTGGTCAAGAAAGA  
TCAGCTCAACAGGCATAGCTATCTGAAGGACCCCGACTTCTGGATGCCGCTCTGGA  
CTTCAACTACCTGGACCTGTGAGCACTGCTGCGGAATAGCTTCCACAGATATGCCGT  
GGATGTCCTGAAATCCGGAAGATGCCAGATGCTGGACCGGAGAACCCTGGAGATG  
GCATTTGCCTACGCTCTGGCACTGTTTCGACCCGCTAGGCAGGAGGAAGCAGGCGC  
TCAGGTGTCCGTCCCTCGCGCACTGGATCGACAGGCAGCCCTGCTGCAGATCCAGGA  
GTTTCATGATTACCTGTCTGTCTCAGACACCACCCAGAACTACCCTGCTGCTGTACCC  
CACTGCCGTGGACCTGGCTAAGAGGGCACTGTGGACCCCTAACCAGATCACTGATA  
TTACCTCTCTGGTGCGCTGGTCTATATCTGAGTAAACAGAATCAGCAGCACCTGA  
TCCCACAGTGGGCCCTGCGACAGATTGCCGACTTCGCTCTGAAGCTGCACAAAACCC  
ATCTGGCTTCCTTCTGTCTGCATTTGCCCGCCAGGAGCTGTACCTGATGGGCTCTCT  
GGTGACAGTATGCTGGTCCATACAACTGAGAGGCGCGAAATCTTTATTGTGGAGAC  
AGGGCTGTGCAGCCTGGCTGAACTGTCCCACTTCACTCAGCTCCTGGCCCATCCTCA  
CCATGAGTACCTGTCCGATCTGTATACCCCATGTTCTAGTTTCAGGCCGACGGGACCA  
CTCTCTGGAACGACTGACTCGGCTGTTTCCTGATGCAACCGTGCTACCACCGTGCC  
CGCCGCCCTGAGTATCCTGTCAACAATGCAGCCAAGCACACTGGAGACTTTCCCCGA  
CCTGTTTTGCCTGCCCTCTGGGGGAGTCATTTCAGCGCCCTGACCGTGTCAGAACATGT  
CAGCTACGTGGTCAAAACCAGTATCTGATCAAGGGAATTTCTACCCCGTGTCTAC  
TACCGTGGTCCGCCAGAGTCTGATCAATTACCAGACAGATTACAGACTAAATGTGA  
GCTGACCCGGAATATGCACACAACTCATAGCATCACCGCCGCTCTGAACATTTCCCT  
GGAGAATTGCGCTTTTTGTGTCAGAGTGCAGTCTGGAATACGATGACACACAGGGCGT  
GATCAACATTATGTATATGCACGATAGCGATGACGTGCTGTTTCGCTCTGGACCCCTA  
CAACGAGGTGGTCTGTGAGCTCCCCCTCGCACTCATTATCTGATGCTGCTGAAGAATGG  
AACAGTGCTGGAAGTCACTGATGTCGTGGTTCGATGCCACAGACTCCCGGCTGCTGAT

GATGTCTGTGTACGCACTGTCCGCCATCATCGGCATCTATCTGCTGTATCGAATGCTG  
AAAACCTGTTGA (SEQ ID NO:15) (TM和CTD加下划线)

[0142] 在一些实施方案中,根据本发明使用的gH多肽缺乏跨膜结构域和/或胞质结构域和/或包含经修饰的跨膜结构域和/或胞质结构域。gH多肽可以任选地包含一种或更多种另外的多肽(例如异源跨膜结构域和/或胞质结构域多肽)。在一些实施方案中,gH多肽表达为与异源多肽的融合蛋白。gH多肽可连接至异源多肽以产生融合蛋白而不改变gH功能和/或抗原性。例如,可将异源多肽的编码序列剪接成gH多肽编码序列中,例如在gH多肽编码序列的3'端。在一些实施方案中,可将异源多肽的编码序列框内剪接成gH多肽编码序列。在一些实施方案中,gH多肽编码序列和异源多肽可以通过单一启动子表达。在一些实施方案中,在gH多肽的C末端插入(例如融合至)异源多肽。

[0143] 在一些实施方案中,异源多肽是或包含存在于水泡性口炎病毒(VSV)中的跨膜结构域和/或胞质结构域。在一些实施方案中,缺乏跨膜结构域和/或胞质结构域的gH与来自



VSV的跨膜结构域和/或胞质结构域融合。根据本发明使用的示例性gH-VSV融合多肽如下SEQ ID NO:16所示。在一些实施方案中,合适的gH-VSV多肽融合蛋白包含与已知HCMV gH多肽基本同源的gH多肽的全部或一部分和与已知VSV多肽基本同源的VSV多肽的全部或一部分。例如,与野生型或天然存在的gH和/或VSV多肽相比,gH-VSV多肽融合蛋白可以含有一个或更多个氨基酸替换、缺失和/或插入。因此,在一些实施方案中,适合用于本发明的gH-VSV多肽融合蛋白与SEQ ID NO:16基本同源。在一些实施方案中,适合用于本发明的gH-VSV多肽融合蛋白具有与SEQ ID NO:16有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中,适合用于本发明的gH-VSV多肽融合蛋白与SEQ ID NO:16基本同一。在一些实施方案中,适合用于本发明的gH-VSV多肽融合蛋白具有与SEQ ID NO:16有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。本文所用的“gH-G”指HCMV gH-VSV TM/CTD融合蛋白。

[0144] HCMV gH-G氨基酸序列 (SEQ ID NO:16)

MRPGLPSYLIVLAVCLLSHLLSSRYGAEASEPLDKAFHLLLNTYGRPIRFLRENTTQCT  
YNSSLRNSTVVRENAISFNFFQSYNQYYVFHMPRCLFAGPLAEQFLNQVDLTETLERYQ  
QRLNTYALVSKDLASYRSFSQQLKAQDSLGEQPTTVPPIDLSIPHVWMPPTTPHGW  
ESHTTSGLHRPHFNQTCILFDGHDLLFSTVTPCLHQGFYLIDELRYVKITLTEDFFVVT  
DDDTPMLLIFGHLPRVLFKAPYQRDNFILRQTEKHELLVLVKKDQLNRHSYKDPDFLD  
AALDFNYLDLSALLRNSFHRYAVDVLKSGRCQMLDRRTVEMAFAYALALFAAARQEE  
AGAQVSVPRALDRQAALLQIQEFMITCLSQTTPRTTLLLYPTAVDLAKRALWTPNQITDI  
TSLVRLVYILSKQNQQHLIPQWALRQIADFALKLHKTHLASFLSAFARQELYLMGSLVH  
SMLVHTTERREIFIVETGLCSLAELSHFTQLLAHPHHEYLSDLTPCSSSGRRDHSRLRLT  
RLFPDATVPTTTPAALSILSTMQPSTLETFPDLFCLPLGESFSALTVEHVSYYVTNQYLI  
KGISYPVSTTVVGQSLITQTDSTQTKCELTRNMHTTHSITAALNISLENCAFCQSALLEYD  
DTQGVINIMYMHDSDDVLFALDPYNEVVSSPRTHYLMLLKNGTVLEVTDVVVDATDS  
RFFFIIGLIIGLFLVLRVGIHLCKLKHKKRQIYTDIEMNRLGK\* (SEQ ID NO:16)

(TM和CTD加下划线)

[0146] HCMV gH-G核苷酸序列 (SEQ ID NO:17)

ATGCGGCCAGGCCTCCCCTCCTACCTCATCGTCCTCGCCGTCTGTCTCCTCAGCCACC  
TACTTTCGTACGATATGGCGCAGAAGCCATATCCGAACCGCTGGACAAAGCGTTTC  
ACCTACTGCTCAACACCTACGGGAGACCCATCCGCTTCCTGCGTGAAAACACCAACC  
AGTGTACCTACAATAGCAGCCTCCGTAACAGCACGGTCGTCAGGGAAAACGCCATC  
AGTTTCAACTTTTTCCAAAGCTATAATCAATACTATGTATTCCATATGCCTCGATGTC  
TTTTTGCGGGTCTCTGGCGGAGCAGTTTCTGAACCAGGTAGATCTGACCGAAACCC  
TGGAAGATACCAACAGAGACTTAACACTTACGCGCTGGTATCCAAAGACCTGGCC  
AGCTACCGATCTTTTTTCGCAGCAGCTAAAGGCACAGGACAGCCTAGGTGAACAGCC  
CACCCTGTGCCACCACCCATTGACCTGTCAATACCTCACGTTTGGATGCCACCGCA  
AACCCTCCACACGGCTGGACAGAATCACATACCACCTCAGGACTACACCGACCAC  
ACTTTAACCAGACCTGTATCCTCTTTGATGGACACGATCTACTATTCAGCACCGTCAC  
ACCTTGTTTGCACCAAGGCTTTTACCTCATCGACGAACCTACGTTACGTTAAATAAC  
ACTGACCGAGGACTTCTTCGTAGTTACGGTGTCCATAGACGACGACACACCCATGCT  
GCTTATCTTCGGCCATCTTCCACGCGTACTCTTTAAAGCGCCCTATCAACGCGACAA  
CTTTATACTACGACAACTGAAAAACACGAGCTCCTGGTGCTAGTTAAGAAAGATC  
AACTGAACCGTCACTCTTATCTCAAAGACCCGGACTTTCTTGACGCCGCACTTGACT  
TCAACTACCTGGACCTCAGCGCACTACTACGTAACAGCTTTCACCGTTACGCCGTGG  
ATGTACTCAAAGCGGTCGATGTCAGATGCTGGACCGCCGCACGGTAGAAATGGCC  
TTCGCCTACGCATTAGCACTGTTTCGCAGCAGCCCGACAAGAAGAGGCCGCGCCCA

[0147]

[0148]

AGTCTCCGTCCACGGGCCCTAGACCGCCAGGCCGCACTCTTACAAATACAAGAATT  
TATGATCACCTGCCTCTCACAAACACCACCACGACCGTTCGCTGCTGTATCCCAC  
GGCCGTGGACCTGGCCAAACGAGCCCTTTGGACACCGAATCAGATACCGACATCA  
CCAGCCTCGTACGCCTGGTCTACATACTCTCTAAACAGAATCAGCAACATCTCATCC  
CCCAGTGGGCACTACGACAGATCGCCGACTTTGCCCTAAAACTACACAAAACGCAC  
CTGGCCTCTTTTCTTTCAGCCTTCGCGCGTCAAGAACTCTACCTCATGGGCAGCCTCG  
TCCACTCCATGCTAGTACATACGACGGAGAGACGCGAAATCTTCATCGTAGAAACG  
GGCCTCTGTTTCATTAGCCGAGCTATCACACTTTACGCAGTTGCTAGCTCATCCGCAC  
CACGAATACCTCAGCGACCTGTACACACCCTGTTCCAGTAGCGGGCGACGCGATCA  
CTCGCTCGAACGCCTCACACGTCTCTTCCCCGATGCCACCGTCCCCACTACCGTTCCC  
GCCGCCCTCTCCATCCTATCTACCATGCAACCAAGCACGCTAGAAACCTTCCCCGAC  
CTGTTTTGTCTGCCGCTCGGCGAATCCTTCTCCGCGCTGACCGTCTCCGAACACGTCA  
GTTATGTCGTAACAAACCAGTACCTGATCAAAGGTATCTCCTACCCTGTCTCCACCA  
CCGTCGTAGGCCAGAGCCTCATCATCACCCAGACGGACAGTCAAATAAATGCGAA  
CTGACGCGCAACATGCATACCACACACAGCATCACAGCGGCGCTCAACATTTCCCTA  
GAAAACTGCGCCTTTTGCCAAAGCGCCCTACTAGAATACGACGACACGCAAGGCGT  
CATCAACATCATGTACATGCACGACTCGGACGACGTCCTTTTCGCCCTGGATCCCTA  
CAACGAAGTGGTGGTCTCATCTCCGCGAACTCACTACCTCATGCTTTTGAAAAACGG  
TACGGTCTAGAAAGTAACTGACGTCGTCTGGACGCTACCGACAGTCGTTTTTCTTT  
ATCATAGGGTTAATCATTGGACTATTCTTGGTTCTCCGAGTTGGTATCCATCTTTGCA  
TAAATTAAAGCACACCAAGAAAAGACAGATTTATACAGACATAGAGATGAACCGA

CTTGGAAAGTAA (SEQ ID NO:17) (TM和CTD加下划线)

[0149]

密码子优化的HCMV gH-G核苷酸序列 (SEQ ID NO:18)

[0150]

ATGCGACCCGGAAGTACCTGATTGTCTGGCTGTCTGTCTGCTGTACAC  
CTGCTGAGTTCAAGATATGGGGCCGAAGCCATCAGCGAGCCACTGGACAAGGCTTT  
CCACCTGCTGTGAACACCTACGGCAGACCCATTAGGTTTCTGCGCGAGAATACCAC  
ACAGTGCACATATAACAGCTCCCTGAGGAATAGCACTGTGGTCCGCGAAAACGCCA  
TCTCTTTCAATTTCTTTCAGAGTTACAACCAGTACTACGTGTTCCATATGCCACGCTG  
TCTGTTTCGACAGGACCCCTGGCCGAGCAGTTTCTGAACCAGGTGGACCTGACCGAGAC  
ACTGGAAAGATAACCAGCAGAGGCTGAATACCTATGCCCTGGTGAAGGATCTGG  
CTTCATATCGGTCTTTCAGTCAGCAGCTCAAGGCCCAGGACTCTCTGGGAGAGCAGC  
CTACTACCGTGCCCCCTCCAATCGATCTGAGTATTCCACACGTCTGGATGCCCCCTC  
AGACAACTCCCCACGGATGGACCGAAAGCCATAACCACATCCGGCCTGCACAGACCC  
CACTTCAACCAGACATGCATCCTGTTTCGATGGCCACGACCTGCTGTTTTCCACTGTG  
ACCCCTTGTCTGCATCAGGGGTTCTACCTGATCGATGAGCTGAGATATGTGAAGATT  
ACACTGACTGAAGACTTCTTTGTGGTCAACCGTGTCTATCGACGATGACACACCAATG  
CTGCTGATTTTCGGACACCTGCCCCGGGTGCTGTTCAAGGCCCCCTACCAGCGAGAC  
AACTTCATCCTGCGGCAGACCGAGAAACACGAAGTGTGGTGGTCAAGAAAGA  
TCAGCTCAACCGGCATTCCCTATCTGAAGGACCCCGACTTCCTGGATGCCGCTCTGGA  
CTTTAACTACCTGGACCTGTCAGCACTGCTGCGGAATAGCTTTCACAGATATGCCGT  
GGATGTCCTGAAATCTGGGCGCTGCCAGATGCTGGACCGGAGAACCGTGGAGATGG  
CATTCGCCTACGCTCTGGCACTGTTTGCAGCCGCTCGGCAGGAGGAAGCAGGAGCTC  
AGGTGAGTGTCCCTCGCGCACTGGATCGACAGGCAGCCCTGCTGCAGATCCAGGAG  
TTCATGATTACCTGTCTGAGCCAGACACCACCCAGAACTACCCTGCTGCTGTACCCC  
ACTGCCGTGGACCTGGCTAAGAGGGCACTGTGGACCCCTAACCAGATCACTGATATT

[0151]

ACCAGCCTGGTGAGACTGGTCTATATCCTGTCCAAACAGAATCAGCAGCACCTGATC  
CCACAGTGGGCCCTGAGGCAGATTGCCGACTTTGCTCTGAAGCTGCACAAAACCCAT  
CTGGCTTCCTTTCTGTCTGCATTCGCCAGACAGGAGCTGTACCTGATGGGATCTCTGG  
TGCACAGTATGCTGGTCCATACTGAGAGGCGCGAAATCTTCATTGTGGAGACA  
GGCCTGTGCAGCCTGGCTGAACTGTCCCACTTTACTCAGCTCCTGGCCCATCCTCAC  
CATGAGTACCTGTCAGATCTGTATACCCCATGTTCTAGTTCAGGACGACGGGACCAC  
AGCCTGGAACGACTGACTCGGCTGTTCCCTGATGCAACCGTGCCTACCACCGTGCCC  
GCCGCCCTGAGTATCCTGTCAACAATGCAGCCAAGCACACTGGAGACTTTTCCCGAC  
CTGTTCTGCCTGCCTCTGGGCGAGAGCTTCAGCGCCCTGACCGTGAGCGAACATGTC  
AGCTACGTGGTCACAAACCAGTATCTGATCAAGGGGATTTCCTACCCCGTGTCTACT  
ACCGTGGTTCGGACAGTCCCTGATCATTACCCAGACAGATTCTCAGACTAAATGTGAG  
CTGACCAGAAATATGCACACAACCTCATAGTATCACCGCCGCTCTGAACATTTCACTG  
GAGAAATTGCGCTTTCTGTCACTCCGCACTGCTGGAATACGATGACACACAGGGCGTG  
ATCAACATTATGTATATGCACGATTCTGATGACGTGCTGTTTGCTCTGGACCCCTACA  
ACGAGGTGGTCGTGAGCTCCCCAGAACTCATTATCTGATGCTGCTGAAGAATGGCA  
CAGTGCTGGAAGTCACTGATGTCTGCTGCTGATGCCACAGACTCCCGCTTCTTTTTCAT  
CATTGGCCTGATCATTGGGCTGTTCCCTGGTGTCTGCGAGTCGGCATCCACCTGTGCAT  
CAAGCTGAAGCATACAAAGAAGAGACAGATCTACACCGATATTGAAATGAACAGGC  
TGGGCAAATGA (SEQ ID NO:18) (TM和CTD加下划线)

[0152] 在一些实施方案中, gH多肽包含存在于gB中的跨膜结构域和/或胞质结构域。根据本发明使用的编码HCMV gH-HCMV gB TM/CTD融合多肽的示例性核苷酸如下SEQ ID NO: 20所示。在一些实施方案中, 适合用于本发明的HCMV gH-HCMV gB TM/CTD多肽与SEQ ID NO: 20编码的多肽基本同源。在一些实施方案中, 适合用于本发明的HCMV gH-HCMV gB TM/CTD多肽具有与SEQ ID NO: 20编码的多肽有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同源性的氨基酸序列。在一些实施方案中, 适合用于本发明的HCMV gH-HCMV gB TM/CTD多肽与SEQ ID NO: 20编码的多肽基本同一。在一些实施方案中, 适合用于本发明的HCMV gH-HCMV gB TM/CTD多肽具有与SEQ ID NO: 20编码的多肽有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高同一性的氨基酸序列。

[0153] HCMV gH-HCMVgB TM/CTD核苷酸序列 (SEQ ID NO: 20)

[0154]

ATGCGGCCAGGCCTCCCCTCCTACCTCATCGTCCTCGCCGTCTGTCTCCTCAGCCACC  
TACTTTCGTACGATATGGCGCAGAAGCCATATCCGAACCGCTGGACAAAGCGTTTC  
ACCTACTGCTCAACACCTACGGGAGACCCATCCGCTTCTCTGCGTGAAAACACCACCC  
AGTGTACCTACAATAGCAGCCTCCGTAACAGCACGGTCGTCAGGGAAAACGCCATC  
AGTTTCAACTTTTTCCAAAGCTATAATCAATACTATGTATTCCATATGCCTCGATGTC  
TTTTTGCGGGTCTCTGCGGAGCAGTTTCTGAACCAGGTAGATCTGACCGAAACCC  
TGAAAGATACCAACAGAGACTTAACACTTACGCGCTGGTATCCAAAGACCTGGCC  
AGCTACCGATCTTTTTCGCAGCAGCTAAAGGCACAGGACAGCCTAGGTGAACAGCC  
CACCCTGTGCCACCACCCATTGACCTGTCAATACCTCACGTTTGGATGCCACCGCA  
AACCCTCCACACGGCTGGACAGAATCACATACCACCTCAGGACTACACCGACCAC  
ACTTTAACCAGACCTGTATCCTCTTTGATGGACACGATCTACTATTCAGCACCGTCAC  
ACCTTGTTTGCACCAAGGCTTTTACCTCATCGACGAACTACGTTACGTTAAAATAAC  
ACTGACCGAGGACTTCTTCGTAGTTACGGTGTCCATAGACGACGACACACCCATGCT  
GCTTATCTTCGGCCATCTTCCACGCGTACTCTTTAAAGCGCCCTATCAACGCGACAA  
CTTTATACTACGACAACTGAAAAACACGAGCTCCTGGTGCTAGTTAAGAAAGATC  
AACTGAACCGTCACTCTTATCTCAAAGACCCGGACTTTCTTGACGCCGCACTTGACT  
TCAACTACCTGGACCTCAGCGCACTACTACGTAACAGCTTTCACCGTTACGCCGTGG  
ATGTACTCAAAAGCGGTTCGATGTCAGATGCTGGACCGCCGCACGGTAGAAATGGCC  
TTCGCTACGCATTAGCACTGTTTCGACGAGCCCGACAAGAAGAGGCCGGCGCCCA  
AGTCTCCGTCCCACGGGCCCTAGACCGCCAGGCCGCACTCTTACAAATACAAGAATT  
TATGATCACCTGCCTCTCACAAACACCACCACGCACCACGTTGCTGCTGTATCCAC  
GGCCGTGGACCTGGCCAAACGAGCCCTTTGGACACCGAATCAGATCACCGACATCA  
CCAGCCTCGTACGCCTGGTCTACATACTCTCTAAACAGAATCAGCAACATCTCATCC  
CCCAGTGGGCACTACGACAGATCGCCGACTTTGCCCTAAACTACACAAAACGCAC  
CTGGCCTCTTTTTCTTTCAGCCTTCGCGCGTCAAGAACTCTACCTCATGGGCAGCCTCG  
TCCACTCCATGCTAGTACATACGACGGAGAGACGCGAAATCTTCATCGTAGAAACG  
GGCCTCTGTTTCATTAGCCGAGCTATCACACTTTACGCAGTTGCTAGCTCATCCGCAC  
CACGAATACCTCAGCGACCTGTACACACCCTGTTCCAGTAGCGGGCGACGCGATCA  
CTCGCTCGAACGCCTCACACGTCTCTTCCCCGATGCCACCGTCCCCACTACCGTTCCC  
GCCGCCCTCTCCATCCTATCTACCATGCAACCAAGCACGCTAGAAACCTTCCCCGAC  
CTGTTTTGTCTGCCGCTCGGCGAATCCTTCTCCGCGCTGACCGTCTCCGAACACGTCA  
GTTATGTCGTAACAAACCAGTACCTGATCAAAGGTATCTCCTACCCTGTCTCCACCA  
CCGTCGTAGGCCAGAGCCTCATCATCACCCAGACGGACAGTCAAATAAATGCGAA  
CTGACGCGCAACATGCATACCACACACAGCATCACAGCGGCGCTCAACATTTCCCTA  
GAAACTGCGCCTTTTGCCAAAGCGCCCTACTAGAATACGACGACACGCAAGGCGT  
CATCAACATCATGTACATGCACGACTCGGACGACGTCCTTTTCGCCCTGGATCCCTA  
CAACGAAGTGGTGGTCTCATCTCCGCGAACTCACTACCTCATGCTTTTGAAAAACGG  
TACGGTCTTAGAAGTAACTGACGTCGTCTGTGGACGCTACCGACAGTCGTTTCGGAG  
CCTTCACCATCATCCTCGTGGCCATAGCCGTCGTCATTATCATTTATTTGATCT  
ATACTCGACAGCGGCGTCTCTGCATGCAGCCGCTGCAGAACCTCTTTCCTATC  
TGGTGTCCGCGACGGGACCACCGTGACGTCGGGCAACACCAAAGACACGTCG  
TTACAGGCTCCGCCTTCCTACGAGGAAAGTGTTTATAATTCTGGTTCGCAAAGGA  
CCGGGACCACCGTCGTCTGATGCATCCACGGCGGCTCCGCCTTACACCAACGA  
GCAGGCTTACCAGATGCTTCTGGCCCTGGTCCGTCTGGACGCAGAGCAGCGAG  
CGCAGCAGAACGGTACAGATTCTTTGGACGGACAGACTGGCACGCAGGACAAG  
  
GGACAGAAGCCCAACCTGCTAGACCGACTGCGACACCGCAAAAACGGCTACCG  
ACACTTGAAAGACTCCGACGAAGAAGAGAACGTCTGA (SEQ ID NO:20)

(gH肽信号加下划线;gB TM-CTD加粗)

[0155] 其他(包括gN和gM-糖蛋白复合物(gC) II)

[0156] 除了gB和gH,其他的包膜糖蛋白可以用于疫苗开发。例如,含有gN(UL73)和gM(UL100)的gcH复合物是特别感兴趣的。HCMV病毒体蛋白质组分析已经证明了gcII是病毒颗粒中表达最丰富的糖蛋白,强调了其在保护性免疫中的潜在重要性(Varnum SM等,2005

Human Gene Ther 16:1143-50)。HCMV感染在大多数血清反应阳性的个体中引发gC特异性抗体应答(Shimamura M等,2006 J Virol 80:4591-600),并且含有gCII抗原gM和gN的DNA疫苗已经示出在兔和小鼠中引发中和抗体应答(Shen S等,2007 Vaccine 25:3319-27)。

[0157] pp65

[0158] 对HCMV的细胞免疫应答包括MHC II类限制的CD4<sup>+</sup>和MHC I类限制的细胞毒性CD8<sup>+</sup>对多种病毒抗原的T-淋巴细胞应答,许多所述病毒抗原存在于病毒被膜(位于包膜和核壳之间的病毒颗粒区域)中。例如,HCMV pp65蛋白(UL83)已经示出在HCMV感染后引发显著的CD8<sup>+</sup>T-淋巴细胞应答(McLaughlin-Taylor E 1994 J Med Virol 43:103-10)。这一65kDa病毒蛋白是表达最丰富的HCMV的结构蛋白之一。示例性pp65序列在本文中描述。

[0159] 其他(包括IE1,pp150)

[0160] 引发T淋巴细胞应答的其他蛋白包括即刻早期-1(immediate early-1)蛋白(UL123)和pp150(UL32)(Gibson L等,2004 J Immunol 172:2256-64;La Rosa C等,2005 Human Immunol 66:116-26)。

[0161] 如上所述,Gag多肽可以任选地包含一种或更多种另外的多肽(例如异源抗原)。在一些实施方案中,Gag多肽与异源抗原共表达(例如在独立的启动子下和/或作为独立的蛋白)。Gag多肽可与异源抗原共表达而不改变Gag功能。不希望受任何理论的限制,认为自组装Gag多肽与异源包膜抗原的共表达将允许该抗原并入所产生的VLP之包膜或脂质双层。在一些实施方案中,VLP包膜组分充当有效的免疫原(例如用于诱导体液免疫应答)。在一些实施方案中,Gag多肽表达为具有异源抗原的融合蛋白。例如,可将异源抗原的编码序列剪接成Gag多肽编码序列,例如在Gag多肽编码序列的3'端。在一些实施方案中,可将异源抗原的编码序列框内成Gag多肽编码序列。在一些实施方案中,Gag多肽编码序列和异源抗原可以通过单个启动子表达。在一些实施方案中,在Gag多肽的C端插入(例如融合至)异源抗原。不希望受任何理论的限制,认为自组装Gag多肽与异源抗原的融合将允许所述抗原并入所产生的VLP之结构组分中。在一些实施方案中,VLP结构组分充当有效的免疫原(例如用于诱导细胞免疫应答)。例如,所提供的VLP可以包含反转录病毒gag多肽(例如MLV gag)和HCMV的结构组分(例如,pp65)。在一些这样的实施方案中,pp65并入VLP中并且充当用于引发针对HCMV之免疫应答的抗原。

[0162] 所提供的VLP可以含有结构反转录病毒蛋白(例如Gag多肽),其设置和构造为使得它自组装以形成VLP并且位于VLP内部。在一些实施方案中,所提供的VLP含有包膜蛋白(例如gB和/或gH),其设置和构造为使得包膜蛋白(例如gB和/或gH)的一个或更多个表位位于VLP表面上。在一些实施方案中,所提供的VLP含有融合结构蛋白(例如Gag/pp65),其设置和构造为使得结构蛋白(例如pp65)的一个或更多个表位位于VLP内部。

[0163] II.VLP的生产

[0164] 应当理解,包含VLP的组合物通常将包含具有一系列尺寸之VLP的混合物。应当理解,以下所列的直径值对应于该混合物中最频繁的直径。在一些实施方案中,组合物中大于90%的小泡将具有位于最频繁值之50%以内的直径(例如1000±500nm)。在一些实施方案中,分布可以更窄,例如组合物中大于90%的小泡可以具有位于最频繁值之40%、30%、20%、10%或5%以内的直径。在一些实施方案中,可使用超声处理或超-超声处理(ultra-sonication)以促进VLP形成和/或改变VLP大小。在一些实施方案中,可以使用过滤、透析

和/或离心来调整VLP尺寸分布。

[0165] 一般而言,根据本公开内容的方法产生的VLP可以为任何尺寸。在一些实施方案中,组合物可以包含直径为约20nm至约300nm的VLP。在一些实施方案中,VLP的特征在于它直径范围的下限被20、30、40、50、60、70、80、90或100nm界定并且上限被300、290、280、270、260、250、240、230、220、210、200、190、180或170nm界定。在一些实施方案中,群体内的VLP示出平均直径,其范围的下限被20、30、40、50、60、70、80、90或100nm界定并且上限被300、290、280、270、260、250、240、230、220、210、200、190、180或170nm界定。在一些实施方案中,群体中的VLP的多分散指数小于0.5,例如小于0.45、小于0.4或小于0.3。在一些实施方案中,VLP直径通过纳米化(nanosizing)测定。在一些实施方案中,VLP直径通过电子显微镜测定。

[0166] A. 体外/离体VLP生产

[0167] 根据本发明提供的VLP可以根据技术人员已知的一般方法学制备。例如,可以使用已知病毒之序列制备多种核酸分子、基因组或重组载体或质粒。这样的序列可从库(bank)获得,并且材料可以从多种保藏中心(collection)、发表的质粒等获得。可使用技术人员公知的技术分离并操作这些元件,或者从本领域可得的质粒分离。还可以从计算机分析或通过(高通量)筛选文库产生多种合成或人工序列。重组表达VLP的多肽需要构建含有多核苷酸的表达载体,所述多核苷酸编码一种或更多种多肽。一旦已经获得了编码一种或更多种多肽的多核苷酸,可以使用本领域已知的技术通过重组DNA技术产生用于产生所述多肽的载体。根据本发明可以利用的表达载体包括但不限于哺乳动物表达载体、杆状病毒表达载体、植物表达载体(例如,花椰菜花叶病毒(CaMV)、烟草花叶病毒(TMV))、质粒表达载体(例如,Ti质粒)等等。

[0168] 根据本发明,可以使用的示例性VLP表达质粒如下SEQ ID NO:19所示:

[0169] Propol II表达质粒(SEP ID NO:19)

[0170]

CTAGAGAGCTTGGCCCATTCGCATACGTTGTATCCATATCATAATATGTACATTTATAT  
TGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAG  
TAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAA  
CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCCATTTGACGTCA  
ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGG  
GTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA  
AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG  
TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTA  
TTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT  
CACGGGGATTTCGAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTGTTTGGCACC  
AAAATCAACGGGACTTTCCAAAATGTGCTAACAACCTCCGCCCCATTGACGCAAATG  
GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG  
TCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGA  
CCGATCCAGCCTCCGGTCGACCGATCCTGAGAACTTCAGGGTGAGTTTGGGGACCCT  
TGATTGTTCTTTCTTTTTCGCTATTGTAAAATTCATGTTATATGGAGGGGGCAAAGTT  
TTCAGGGTGTTGTTTAGAATGGGAAGATGTCCCTTGATCACCATGGACCCTCATGA  
TAATTTTGTTCCTTTCACTTTCTACTCTGTTGACAACCATTTGTCTCCTCTTATTTTCTTT  
TCATTTTCTTGTAACTTTTTCGTTAACTTTAGCTTGCATTTGTAACGAATTTTAAAT  
TCACTTTTGTGTTATTTGTCAGATTGTAAGTACTTTCTCTAATCACTTTTTTTTCAAGGC  
AATCAGGGTATATTATATTGTACTTCAGCACAGTTTTAGAGAACAAATTGTTATAATT  
AAATGATAAGGTAGAATATTTCTGCATATAAATTCTGGCTGGCGTGGAAATATTCTT  
ATTGGTAGAAACAACCTACATCCTGGTCATCATCCTGCCTTTCTCTTTATGGTTACAAT  
GATATACACTGTTTGAGATGAGGATAAAATACTCTGAGTCCAAACCGGGCCCCCTCTG  
CTAACCATGTTTCATGCCTTCTTTTTCCTACAGCTCCTGGGCAACGTGCTGGTTAT  
TGTGCTGTCTCATCATTTTGGCAAAGAATTCTCGAGCGTACGCCTAGGGGATCCAG  
CGCTATTTAAATGCTAGCATGCATGTTAACCCTGCAGGGGTACCGCGGCCGCAAGCT  
TAGATCCGTCGAGGAATTCATCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGC  
TGGTGTGGCCAATGCCCTGGCTCAAAATACCACTGAGATCTTTTCCCTCTGCCAA  
AATTATGGGGACATCATGAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAA

[0171]

ATTTATTTTCATTGCAATAGTGTGTTGGAATTTTTTGTGTCTCTCACTCGGAAGGACA  
TATGGGAGGGCAAATCATTTAAAACATCAGAATGAGTATTTGGTTTAGAGTTTGGA  
ACATATGCCCATATGCTGGCTGCCATGAACAAAGGTTGGCTATAAAGAGGTCATCA  
GTATATGAAACAGCCCCCTGCTGTCCATTCCCTATTCCATAGAAAAGCCTTGACTTG  
AGGTTAGATTTTTTTTATATTTTGTGTTTGTGTTATTTTTTCTTTAACATCCCTAAAAT  
TTTCCTTACATGTTTTACTAGCCAGATTTTTCTCTCTCTCTGACTACTCCAGTCATA  
GCTGTCCCTCTTCTCTTATGGAGATCCCTCGACGGATCGGCCGCAATTTCGTAATCATG  
TCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA  
GCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATT  
AATTGCGTTGCGCTCACTGCCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCA  
TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCCGC  
TTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGC  
TCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA  
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCT  
GGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAG  
TCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAA  
GCTCCCTCGTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCCTT  
TCTCCCTTCGGGAAGCGTGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCTG  
GTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGAC  
CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA  
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG  
TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT  
TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG  
ATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTTGCAAGCAGCAGAT  
TACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGA  
CGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAA  
GGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTA  
TATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCT  
CAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCCTGACTCCCCGTGCTGTAGATAAC  
TACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACC  
CACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG  
CGCAGAAGTGGTCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGG  
GAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT  
ACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTACAGTCCGGTTCCC  
AACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGGTTAGCTC  
CTTCGGTCCCTCCGATCGTTGTGAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTT  
ATGGCAGCACTGCATAATTCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGA  
CTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT  
CTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG  
CTCATCATTTGAAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTG  
AGATCCAGTTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTT  
TCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGG  
AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTG  
AAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAA  
AAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAAATTGT  
AAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTTGTAAATCAGCTCATTTTTT



[0172]

AACCAATAGGCCGAAATCGGCCAAAATCCCTTATAAAATCAAAAGAATAGACCGAGAT  
AGGGTTGAGTGTTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC  
CAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCAT  
CACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTA  
AAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAA  
GGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTC  
ACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC  
CCATTGCGCCATTGAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC  
GCTATTACGCCAGCTGGCGAAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAA  
CGCCAGGGTTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGAGCGCGCGTAA  
TACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCT (SEQ ID  
NO:19)

[0173] 所提供的VLP可以根据本领域已知的技术制备。例如,在一些实施方案中,所提供的VLP可以在任何可获得的蛋白表达系统中产生。通常,通过常规技术将表达载体转移至宿主细胞中并且随后通过常规技术培养经转染的细胞以产生VLP。在一些实施方案中,使用细胞的瞬时转染产生VLP。在一些实施方案中,使用稳定转染的细胞产生VLP。可用于VLP产生的典型细胞系包括但不限于哺乳动物细胞系,例如人胚肾(HEK) 293、WI 38、中国仓鼠卵巢(CHO)、猴肾(COS)、HT1080、C10、HeLa、幼仓鼠肾(BHK)、3T3、C127、CV-1、HaK、NS/O和L-929细胞。具体的非限制性实例包括但不限于,BALB/c小鼠骨髓瘤系(NS0/1, ECACC No: 85110503);人成视网膜细胞(PER.C6 (CruCell, Leiden, The Netherlands));用SV40(COS-7, ATCC CRL 1651)转化的猴肾CV1系;人胚肾系(293或亚克隆的用于在悬浮培养物中生长的293细胞, Graham等, J. Gen Virol., 36:59 (1977));幼仓鼠肾细胞(BHK, ATCC CCL 10);中国仓鼠卵巢细胞+/-DHFR(CHO, Urlaub和Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980));小鼠支持细胞(TM4, Mather, Biol. Reprod., 23:243-251 (1980));猴肾细胞(CV1, ATCC CCL 70);非洲绿猴肾细胞(VERO-76, ATCC CRL-1587);人宫颈癌细胞(HeLa, ATCC CCL 2);犬肾细胞(MDCK, ATCC CCL 34);布法罗大鼠肝细胞(buffalo rat liver cell) (BRL 3A, ATCC CRL 1442);人肺细胞(W138, ATCC CCL75);人肝细胞(Hep G2, HB 8065);小鼠乳腺肿瘤(MMT 060562, ATCC CCL51);TRI细胞(Mather等, Annals N.Y. Acad. Sci., 383:44-68 (1982));MRC 5细胞;FS4细胞;和人肝细胞瘤系(Hep G2)。在一些实施方案中,可用于VLP生产的细胞系包括昆虫(例如Sf-9、Sf-21、Tn-368、Hi5)或植物(例如豆科、谷类(cereal)或烟草)细胞。应当理解,在一些实施方案中,特别是当糖基化对蛋白功能重要时,优选哺乳动物细胞用于蛋白表达和/或VLP生产(参见,例如Roldao A等, 2010 Expt Rev Vaccines 9: 1149-76)。

[0174] 应当理解,可以选择调节插入序列的表达或以特定方式修饰并加工基因产物的细胞株。蛋白产物的这样的修饰(例如糖基化)和加工(例如切割或转运至膜)对于VLP的产生或者VLP多肽或另外多肽(例如佐剂或另外的抗原)之功能可以是重要的。对于蛋白和基因产物的转录后加工和修饰,不同的细胞具有特有的或特异性的机制。可选择适当的细胞系或宿主系统以确保表达的异种蛋白的正确修饰和加工。一般而言,根据本发明,可使用真核宿主细胞(也称为包装细胞(packaging cell) (例如293T人胚肾细胞)),其具有用于适当加工基因产物之初级转录、糖基化和磷酸化的细胞机制(cellular machinery)。

[0175] 可以根据已知的技术纯化VLP,例如离心、梯度、蔗糖梯度超速离心、切向流过滤和色谱(例如,离子交换(阴离子和阳离子)、亲和和尺寸柱色谱)或差别溶解度(differential

solubility),等等。或者或此外,可以直接使用细胞上清液,无需纯化步骤。可以向纯化的VLP添加另外的实体(例如另外的抗原或佐剂)。

[0176] 在一些实施方案中,所提供的多核苷酸序列是密码子优化的。密码子优化是本领域公知的并且涉及使用密码子的改造(modification of codon usage)使得产生更高水平的蛋白。

#### [0177] B. 体内VLP生产

[0178] 根据本发明,所提供的VLP可以根据本领域公知的方法制备成DNA疫苗。例如,在一些实施方案中,将一种或更多种载体或质粒(例如,如上所述的那些)施用至对象使得受者细胞表达由载体或质粒编码的多肽。在一些实施方案中,表达这样的多肽的受者细胞产生包含所述多肽的VLP。

#### [0179] C. 单价、二价、三价eVLP

[0180] 根据本发明,可以用本文所述的单个表达载体转染细胞。在一些实施方案中,单个表达载体编码一个以上的VLP成分(element)(例如,一个以上的结构多蛋白、CMV被膜多肽、CMV糖蛋白等)。例如,在一些实施方案中,单个表达载体编码两个或更多个VLP成分。在一些实施方案中,单个表达载体编码三个或更多个VLP成分。

[0181] 在一些实施方案中,用两个或更多个表达载体转染细胞。例如,在一些实施方案中,用编码Gag多肽的第一载体和编码HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的第二载体转染细胞。在一些这样的实施方案中,生产包含HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的“单价”VLP。在一些实施方案中,用编码Gag多肽的第一载体、编码HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的第二载体以及编码另一种HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的第三载体转染细胞。在一些这样的实施方案中,生产包含两个HCMV包膜糖蛋白(例如gB和gH-G或者gB-G和gH-G)的“二价”VLP。在一些实施方案中,用编码Gag-pp65融合多肽的第一载体和编码HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的第二载体转染细胞。在一些这样的实施方案中,生产包含HCMV结构蛋白和HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的“二价”VLP。在一些实施方案中,用编码Gag-pp65融合多肽的第一载体、编码HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的第二载体以及编码另一种HCMV包膜糖蛋白(例如gB或gB-G或gH-G)的第三载体转染细胞。在一些这样的实施方案中,生产包含HCMV结构蛋白和2种HCMV包膜糖蛋白(例如gB和gH-G或者gB-G和gH-G)的“三价”VLP。

[0182] 在一些实施方案中,混合单价、二价或三价VLP。例如,在一些实施方案中,混合单价和二价VLP以形成三价VLP混合物。在一些实施方案中,混合两个二价VLP以形成三价VLP混合物。

#### [0183] III. HCMV感染和治疗

[0184] 人巨细胞病毒(HCMV)( $\beta$ -疱疹病毒)是一种广泛存在的病原体。一般而言,疱疹病毒进入细胞是复杂的过程,其通过吸附和受体结合起始并且随后是病毒包膜与细胞膜的融合。融合一般在质膜或内体膜处发生。HCMV体内感染多种细胞类型,包括上皮细胞、内皮细胞和成纤维细胞(Plachter B等,1996 Adv Virus Res 46:195-261)。它与成纤维细胞的质膜融合(Compton T等,1992 Virology 191:387-395),但通过胞吞作用进入视网膜色素上皮细胞和脐静脉内皮细胞(Bodaghi B等,1999 J Immunol 162:957-964;Ryckman BJ等,2006 J Virol 80:710-722)。疱疹病毒通过何种机制选择其进入细胞的途径仍不清楚。通常假定

进入途径主要由宿主细胞确定,但有证据表明病毒体糖蛋白的向性作用(tropic role)(Wang X等,1998 J Virol 72:5552-5558)。如之前提到的,HCMV编码两个gH/gL复合物:gH/gL/gO和gH/gL/UL128/UL130/UL131。含有gO的复合物足以使成纤维细胞感染,而含有pUL128/UL130/UL131的复合物对于HCMV感染内皮细胞和上皮细胞是重要的。

[0185] HCMV感染50%至85%的40岁以下的成年人(Gershon AA等,1997in Viral Infections of Humans,第四版,New York;Plenum Press:229-251)。出生后获得HCMV的大多数健康个体很少发生症状。然而,HCMV疾病在免疫低下的个体(例如造血细胞移植(HCT)和实体器官移植(SOT)的受者)中导致显著的发病率和死亡率(Pass RF 2001 Cytomegalovirus.In Fields Virology.第四版,Philadelphia;Lippincott Williams&Wilkins:2675-2705)。在SOT或HCT人群中,HCMV疾病可由于由供体器官或HCT传输的新感染而发生,或者可由于受者中潜伏病毒的再激活而复发。在HIV感染的个体中,尽管可获得抗反转录病毒治疗,但HCMV感染仍加速AIDS和死亡的进程(Deayton JR等,2004 Lancet 363:2116-2121)。除了在US,HCMV是最常见的子宫内感染并在每年大约8000例婴儿中导致引起死亡或严重出生缺陷的先天性畸形,包括耳聋和智力迟钝(mental retardation),(Stagon S等,1986 JAMA 256:1904-1908)。

[0186] 控制HCMV的免疫应答未完全理解。通过与其他人疱疹病毒的类比,可以推测细胞和体液免疫应答都发挥了重要作用(Kohl S 1992 Current topics in Microbiology and Immunology 179:75-88)。对于鼠CMV,显示无论细胞毒性T细胞应答还是中和抗体的被动转运均足以防止致命挑战(Rapp M等,1993 Multidisciplinary Approach to Understanding Cytomegalovirus Disease:327-332;Reddehase MJ等,198 J Virology 61:3102-3108)。

[0187] 在免疫低下的人中,CMV的控制主要与细胞免疫应答相关;CD8<sup>+</sup>和CD4<sup>+</sup>T淋巴细胞二者似乎对防止CMV疾病均是重要的(Gamadia LE等,2003 Blood 101:2686-2692;Cobbold M等,2005 J Exp Med 202:379-386)。对CMV的细胞免疫应答包括对存在于病毒被膜(包膜与衣壳之间的病毒颗粒区域)中的多种抗原的CD4<sup>+</sup>辅助T淋巴细胞和CD8<sup>+</sup>细胞毒性T淋巴细胞应答。最近对来自健康供体的CMV特异性CD4<sup>+</sup>和CD8<sup>+</sup>T细胞的研究使用了来自一系列CMV开放阅读框的重叠肽(overlapping peptide)来鉴别CMV感染后被识别的抗原(Sylwester AW等,2005 J Exp Med 202:673-685)。CMV被膜磷蛋白65(pp65)和表面糖蛋白gB是最经常被CD4<sup>+</sup>T细胞识别的抗原,并且pp65也是最经常被CD8<sup>+</sup>T细胞识别的抗原之一。

[0188] 与移植情况(transplant setting)相反,母体针对病毒的体液免疫应答对于在新生儿中预防HCMV疾病似乎是重要的。对表面糖蛋白(特别是gB)的抗体似乎对于防止CMV的母体-新生儿转移至关重要(Fowler KB等,2003 JAMA 289:1008-1011)。此外,在更早的接种疫苗研究中,显示出防止再感染与中和抗体相关(Adler SP等,1995 J Infectious Diseases 171:26-32)。对CMV的体液免疫应答以对存在于病毒颗粒之外部包膜中的病毒包膜糖蛋白(例如gB和gH)的应答为主。

[0189] 在HVMC的情况下,难以直接评估免疫学效应物(effector)功能,因为病毒是严格物种特异性的并且没有可用的动物模型系统。然而,鼠CMV和豚鼠CMV已经被用于评估这些宿主物种中的疫苗策略。

[0190] 诱导保护性T细胞和中和抗体应答二者的CMV疫苗具有防止感染或改善由先天性

感染或移植导致的CMV疾病的潜力。

[0191] 在人中测试的第一个活的减毒HCMV疫苗候选者基于实验室改造的AD169毒株。用另一个实验室改造的临床分离株(isolate) (Towne毒株) 进行的随后实验证实了活的减毒疫苗能够引发中和抗体以及CD4+和CD8+T淋巴细胞应答。在肾移植受者中的一系列研究中评估了Towne疫苗的效力。尽管Towne疫苗不提供对HCMV疾病的保护性效果,它无法防止移植后的HCMV感染(Plotkin SA等,1984 Lancet 1:528-530)。还在血清反应阴性之母的安慰剂对照研究中评估了Towne疫苗,这些母亲的孩子参加政府或社区所办的幼儿园(group daycare),这种情况下无法防止这些母亲从她们的被HCMV感染的儿童获得感染(Adler SP等,1995 J Infectious Diseases 171:26-32)。对这些研究的解释为Towne被过度减毒(overattenuated)了。为了探究这种可能性,一系列的遗传重组体,其中CMV的未减毒的“Toledo”毒株的区域替代Towne基因组的对应区域,导致Towne/Toledo“嵌合体”的构造,其含有部分而不是全部使得Towne疫苗弱化的突变(Heineman TC等2006 J Infect Disease 193:1350-1360)。正在I期临床试验中测试4种Towne/Toledo“嵌合体”的安全性和耐受性。关于建立潜伏HCMV感染之潜在风险的长期安全性关注已经阻碍了活的、减毒疫苗的进一步开发。

[0192] 重要的亚单位CMV疫苗候选者基于包膜糖蛋白gB(纯化的重组gB疫苗由Sanofi-Pasteur Vaccines制造),因为这种蛋白能够在自然感染期间引发高效价的病毒中和抗体。重组gB疫苗引发中和抗体应答并且具有优异的安全谱,然而它排除了中和抗体应答的其他糖蛋白靶标和更重要地T淋巴细胞靶标。该疫苗需要MF59佐剂以优化免疫原性。在最近的实验中,这种疫苗在年轻妇女中的2期临床试验中对预防CMV感染提供了总共50%的效力(Pass RF等,2009 N Engl J Med 360:1191-1199)。正作为亚单位疫苗候选者被评估的其他病毒蛋白包括pp65和IE1,它们都引发T细胞应答。

[0193] DNA疫苗在动物中引发强大的细胞和体液免疫应答并且很好地匹配疫苗设计中的特异性和精确度(precision)。已经开发了用于CMV的DNA疫苗,并且聚焦于gB、IE1和pp65蛋白作为候选的靶免疫原。使用编码pp65和gB之质粒DNA的二价CMV DNA疫苗候选者(Wloch MK 2008 J Infectious Diseases 297:1634-1642),和还包括编码IE1基因产物之第三质粒的三价疫苗候选者(Jacobson MA 2009 Vaccine 27:1540-1548),已经由Vical Vaccines开发(专利US 7,410,795)。无论施用途径如何,单独的三价DNA疫苗的免疫原性最小。然而,在施用活的减毒CMV(Towne)后,CMV DNA疫苗确实看起来安全地引发了所观察到之CMV抗原的记忆应答。

[0194] 在载体疫苗的途径中,所关注的基因产物在非复制(通常为病毒)载体的环境中表达。这种的一个实例是由Virogenetics and Sanofi-Pasteur Vaccines开发的称作ALVAC的金丝雀痘载体(canarypox vector),其是减毒的无法在哺乳动物细胞中复制的痘病毒。已经在临床试验中测试了表达CMV gB的ALVAC和表达pp65的ALVAC(US 6,267,965)。ALVAC-CMV(gB)不诱导中和抗体,但在用Towne毒株CMV后续感染后确实引发了更高的中和抗体效价(Adler SP等1999 J Infectious Diseases 180:843-846),尽管在用gB亚单位/MF59疫苗后续免疫后,它似乎没有促进中和抗体效价(Bernstein DI等2002 J Infectious Diseases 185:686-690)。表达pp65、ALVAC-CMV(pp64)的金丝雀痘载体在所有初始血清反应阴性的志愿者中诱导了持久的CTL应答,其频率与天然血清反应阳性的个体相当

(Berencsi K等2001 J Infectious Diseases 183:1171-1179)。用于表达gB作为载体疫苗的另一种途径是使用AlphaVax Inc的 $\alpha$ 病毒复制子系统(US 7,419,674)。这一途径涉及源自甲病毒的减毒株(委内瑞拉马脑炎(Venezuelan Equine Encephalitis,VEE)病毒)的繁殖缺陷的单周期(single-cycle)RNA复制子载体系统以产生表达pp65、IE1或gB蛋白的病毒样复制子颗粒(VRP)(Berstein等,2010 Vaccine 28:484-493)。使用了两组分甲病毒复制子疫苗来表达三种CMV蛋白作为CMV gB(Towne毒株)和pp65/IE1融合蛋白的可溶形式(Reap EA等2007 Vaccine 25:7441-7449),该疫苗被发现是安全的并且诱导高水平的中和抗体和多功能的CD4<sup>+</sup>和CD8<sup>+</sup>抗原特异性T细胞应答。高剂量组的几何平均滴度(GMT,Geometric Mean Titre)为在试验中所测试的12例自然感染的CMV血清反应阳性个体中之GMT的一半左右。

[0195] 目前处于临床前开发中的用于针对HCMV接种疫苗的新颖候选者为“密体”疫苗。密体(dense body,DB)是细胞培养物中在CMV的复制期间形成的包膜的、复制缺陷的颗粒。它们含有包膜糖蛋白和大量的pp65蛋白。DB是非感染性的且是免疫原性的,但不能在疫苗受者中建立潜在的HCMV感染。DB已经示出能够在小鼠中,在没有病毒基因表达的情况下诱导病毒中和抗体和T细胞应答(Pepperl S等,2000 J Virol 74:6132-6146-WO 00/53729和US 6,713,070)。

#### [0196] IV. 药物组合物

[0197] 本发明提供了包含所提供的VLP和/或所提供的糖蛋白变体的药物组合物。在一些实施方案中,本发明提供了VLP和至少一种可药用的赋形剂。这样的药物组合物可以任选地包含一种或更多种另外的治疗活性物质和/或与其结合施用。在一些实施方案中,所提供的药物组合物可用于药物中。在一些实施方案中,所提供的药物组合物可在HCMV或者与HCMV感染相关或有关之负面结果的治疗或预防中用作预防剂(即疫苗)。在一些实施方案中,所提供的药物组合物可用于治疗性施用,例如在患有或者易患HCMV感染的个体中。在一些实施方案中,药物组合物配制成用于对人施用。

[0198] 例如,本文提供的药物组合物可以以无菌可注射的形式(例如,适用于皮下注射或静脉内输注的形式)提供。例如,在一些实施方案中,以适合于注射的液体剂型提供药物组合物。在一些实施方案中,提供的药物组合物为粉剂(例如冻干和/或灭菌的)(任选地在真空下),其在注射前用水性稀释剂(例如水、缓冲液、盐溶液等)重构。在一些实施方案中,药物组合物在水、氯化钠溶液、乙酸钠溶液、苄醇溶液、磷酸盐缓冲盐水等中稀释和/或重构。在一些实施方案中,应当将粉剂与水性稀释剂温和地混合(例如,不要震荡)。

[0199] 在一些实施方案中,所提供的药物组合物包含一种或更多种可药用赋形剂(例如防腐剂、惰性稀释剂、分散剂、表面活性剂和/或乳化剂、缓冲剂等)。合适的赋形剂包括例如水、盐水、葡萄糖、蔗糖、海藻糖、甘油、乙醇或类似物,及其组合。此外,如果需要,疫苗可以含有辅助物质,例如润湿剂或乳化剂、pH缓冲剂或增强疫苗有效性的佐剂。在一些实施方案中,药物组合物包含一种或更多种防腐剂。在一些实施方案中,药物组合物不包含防腐剂。

[0200] 在一些实施方案中,药物组合物以可以被冷却和/或被冷冻的形式提供。在一些实施方案中,药物组合物以不能被冷却和/或被冷冻的形式提供。在一些实施方案中,重构的溶液和/或液体剂型可以在重构后贮存一段时间(例如2小时、12小时、24小时、2天、5天、7天、10天、2周、1个月、2个月或更长)。在一些实施方案中,VLP制剂长于指定时间的贮存导致

VLP的降解。

[0201] 液体剂型和/或重构的溶液在施用前可以包含颗粒物质和/或变色。在一些实施方案中,如果溶液变色或浑浊和/或如果颗粒物质在过滤后保留,则该溶液不应被使用。

[0202] 本文所述的药物组合物制剂可以通过药学领域中已知的或以后开发的任何方法制备。在一些实施方案中,这样的制备(preparatory)方法包括使活性成分与一种或更多种赋形剂和/或一种或更多种辅助物质结合,然后如果需要和/或期望,将产品成型和/或包装成期望的单一或多次剂量单位。

[0203] 根据本发明的药物组合物可以以散装、作为单一单位剂量和/或作为多个单一单位剂量制备、包装和/或销售。本文所用的“单位剂量”是包含预定量活性成分的药物组合物的离散量。活性成分的量一般等于会施用至对象的剂量和/或这种剂量的方便分数,例如这种剂量的二分之一或三分之一。

[0204] 根据本发明的药物组合物中活性成分、可药用赋形剂和/或任何另外的成分的相对量可以根据所治疗对象的身份、体型(size)和/或病症、和/或根据组合物待施用的途径而变化。例如,组合物可以包含0.1%至100%(重量/重量)的活性成分。

[0205] 本发明的药物组合物可以另外地包含可药用赋形剂,本文所用的赋形剂可以为或包含适合于期望的特定剂型的溶剂、分散介质、稀释剂或其它液体载剂、分散或悬浮助剂、表面活性剂、等张剂、增稠剂或乳化剂、防腐剂、固体粘合剂、润滑剂等。Remington的The Science and Practice of Pharmacy,第21版,A.R.Gennaro,(Lippincott,Williams&Wilkins,Baltimore,MD,2006)公开了用于配制药组合物的多种赋形剂以及用于其制备的多种已知技术。任何常规的赋形剂介质,除非与物质或其衍生物不相容,例如通过产生任何不期望的生物学作用或者以有害的方式与药物组合物的任何其他组分相互作用,其使用涵盖在本发明的范围内。

[0206] 在一些实施方案中,药物组合物具有充分的免疫原性作为疫苗(例如,在缺乏佐剂的情况下)。在一些实施方案中,药物组合物通过包含佐剂增强免疫原性。根据本发明可以使用任何佐剂。大量的佐剂是已知的;国立卫生研究院(National Institutes of Health)准备了许多这类化合物的有用的汇编并且可以找到([www.niaid.nih.gov/daids/vaccine/pdf/compendium.pdf](http://www.niaid.nih.gov/daids/vaccine/pdf/compendium.pdf))。还参见Allison(1998,Dev.Biol.Stand.,92:3-11;通过引用并入本文),Unkeless等(1998,Annu.Rev.Immunol.,6:251-281;通过引用并入本文)和Phillips等(1992,Vaccine,10:151-158;通过引用并入本文)。本领域已知数百种不同的佐剂并且可以在本发明的实践中采用。根据本发明可利用的示例性佐剂包括但不限于细胞因子、凝胶型佐剂(例如氢氧化铝、磷酸铝、磷酸钙等);微生物佐剂(例如,包含CpG模体的免疫调节性DNA序列;内毒素,如单磷酰脂质A;外毒素,例如霍乱毒素、大肠杆菌热不稳定毒素和百日咳毒素;胞壁酰二肽等);油乳剂和基于乳化剂的佐剂(如弗氏佐剂、MF59[Novartis]、SAF等);颗粒佐剂(例如,脂质体、生物可降解的微球体、皂苷类等);合成佐剂(例如,非离子型嵌段共聚物、胞壁酰肽类似物、聚磷腈、合成的多聚核苷酸等);和/或其组合。其它的示例性佐剂包括一些聚合物(例如聚磷腈;描述于美国专利5,500,161中,其通过引用并入本文)、Q57、QS21、角鲨烯、四氯十烷氧化物(tetrachlorodecaoxide),等。可药用赋形剂之前已经在以上题为“药物组合物”的部分中进一步详细描述。

[0207] V.施用

[0208] 所提供的组合物和本公开内容的方法可用于在对象(包括成人和儿童)中预防和/或治疗HCMV感染。然而,一般而言,它们可以用于任何动物。在某些实施方案中,本文的方法可以用于兽医应用,例如犬科或猫科应用。如果需要的话,本文的方法也可以用于农场动物,例如绵羊、禽、牛、猪和马种类(breed)。

[0209] 本文所用的术语“对象”、“个体”或“患者”指人或非人哺乳动物对象。被治疗的个体(也称为“患者”或“对象”)是患有疾病(例如HCMV感染)的个体(胎儿、婴儿、儿童、青少年或成年人)。在一些实施方案中,对象有HCMV感染的风险。在一些实施方案中,对象为免疫抑制的对象。例如,在一些实施方案中,免疫抑制的对象选自:HIV感染的对象、AIDS患者、移植受者、儿童对象和妊娠对象。在一些实施方案中,对象已经暴露于HCMV感染。在一些实施方案中,对象为人。

[0210] 本文所述的组合物通常会以诱导免疫应答所需或足以诱导免疫应答的量和时间施用。给药方案可以由一段时间内的单一剂量或多次剂量组成。待施用的免疫原性组合物(例如VLP)的确切量可以随对象而变化并且可以取决于几种因素。因此,将会理解,一般而言,所使用的确切剂量会由开处方的医生决定并且会不仅仅依赖于对象的体重和施用途径,而且还依赖于对象的年龄和症状的严重程度和/或感染之风险。在某些实施方案中,特定量的VLP组合物作为单一剂量施用。在某些实施方案中,特定量的VLP组合物作为多于一次的剂量施用(例如相隔1至12个月的1至3次剂量)。在某些实施方案中,特定量的VLP组合物作为单一剂量在几个时机施用(例如相隔1至12个月的1至3次剂量)。

[0211] 在一些实施方案中,所提供的组合物以初始剂量和至少一次加强剂量施用。在一些实施方案中,所提供的组合物以初始剂量和至少两次加强剂量施用。在一些实施方案中,所提供的组合物以初始剂量和三次加强剂量施用。在一些实施方案中,所提供的组合物以初始剂量和四次加强剂量施用。在一些实施方案中,提供的组合物以初始剂量和在所述初始剂量后约1个月、约2个月、约3个月、约4个月、约5个月或约6个月的至少一次加强剂量施用。在一些实施方案中,所提供的组合物在初始剂量后约6个月、约7个月、约8个月、约9个月、约10个月、约11个月或约1年以第二加强剂量施用。在一些实施方案中,所提供的组合物每1年、每2年、每3年、每4年、每5年、每6年、每7年、每8年、每9年或每10年以加强剂量施用。

[0212] 在某些实施方案中,所提供的组合物可以配制成用于肠胃外递送(例如通过注射)。在这样的实施方案中,例如,可以通过输注或无针注射技术静脉内、肌内、皮内、或皮下施用。在某些实施方案中,组合物可以配制成用于口服递送、经口递送、鼻内递送、口腔递送、舌下递送、透皮递送、经皮递送、腹膜内递送、阴道内递送、直肠递送或颅内递送。

[0213] 实施例

[0214] 以下实施例描述了制造和实施本文所述之某些组合物的一些示例性方式。应当理解,这些实施例仅用于举例说明目的并且不意味着限制本文所述的组合物和方法的范围。

[0215] 实施例1:DNA表达质粒的构建

[0216] 该实施例描述了用于表达重组HCMV基因序列(例如gB、gB-G、gH-G和Gag/pp65融合基因序列)之表达质粒和构建体(construct)的开发。标准的表达质粒通常由以下组成:哺乳动物来源的启动子序列、内含子序列、多聚腺苷酸化信号序列(PolyA)、复制序列的pUC起始点(pUC-pBR322是复制起始的colE1起始点/位点并且用于在细菌(如大肠杆菌)(DH5 $\alpha$ )中复制质粒)以及作为用于质粒菌斑(plasmid plaque)选择之可选择标志物的抗生素抗性基

因。质粒中内含子后是多个限制酶位点,其可用于剪接成感兴趣的基因或部分基因序列。

[0217] Propol II表达质粒含有pHCMV(用于HCMV的早期启动子)、 $\beta$ -球蛋白内含子(BGL内含子),兔球蛋白多聚腺苷酸化信号序列(PolyA),复制序列的pUC起始点(pUC-pBR322是复制起始的colE1起始点/位点并且用于在细菌(如大肠杆菌)(DH5 $\alpha$ )中复制质粒)以及氨苄青霉素抗性基因 $\beta$ -内酰胺酶(Amp<sup>R</sup>-质粒的可选择标志物赋予对氨苄青霉素的抗性)(100 $\mu$ g/ml)(图1的A)。

[0218] 图1的B描绘了示例性重组表达质粒。例如,为开发pHCMV-Gag MMLV表达构建体(“MLV-Gag”),将编码MMLV之Gag多蛋白(没有其C端聚合酶(Pol)序列的Gag)的互补DNA(cDNA)序列克隆到Propol II(pHCMV)表达载体中。为开发gB表达构建体(“gB”),gB的全长序列被密码子优化以用于人表达(GenScript)并且被克隆到Propol II表达载体中,所述载体包含gB的胞外部分、跨膜结构域(TM)和胞质(Cyto)部分。为开发gH-G表达构建体(“gH-G”),将仅编码胞外部分的gH截短序列与VSV-G的TM和Cyto部分融合在一起并且将其密码子优化以用于人表达(GenScript)并克隆到Propol II表达载体中。类似地,为开发gB-G表达构建体(“gB-G”),将仅编码胞外部分的gB截短序列与VSV-G的TM和Cyto部分融合在一起并且将其密码子优化以用于人表达(GenScript)并克隆到Propol II表达载体中。为开发Gag/pp65表达构建体(“Gag/pp65”),将编码MMLV之Gag多蛋白(没有其C端聚合酶序列的Gag)的序列与pp65的全长序列融合并且将其密码子优化以用于人表达(GenScript)并克隆到Propol II表达载体中。

[0219] DNA质粒在感受态大肠杆菌(DH5 $\alpha$ )中扩增并且用无内毒素制备试剂盒根据标准方案纯化。

[0220] 实施例2:病毒样颗粒(VLP)的生产

[0221] 这一实施例描述了用于生产病毒样颗粒(VLP)的方法,所述病毒样颗粒含有实施例1中所述的多种重组HCMV抗原。

[0222] 使用磷酸钙法用MMLV-Gag DNA表达质粒瞬时转染HEK 293T细胞(ATCC,CRL-11268)并且与gB或gB-G(数据未示出)或gH-G DNA表达质粒共转染。或者,将细胞用Gag/pp65 DNA表达质粒转染并用gB或gB-G(数据未示出)或gH-G DNA表达质粒共转染。应当理解,细胞可用MMLV-Gag DNA表达质粒转染并且用gB和gH-G或者gB-G和gH-G DNA表达质粒二者共转染。由HEK293细胞表达的多种HCMV抗原通过流式细胞术确认(图2的A)。转染后48至72小时,收获含有VLP的上清液并且通过0.45 $\mu$ m孔径膜过滤,并通过SW32 Beckman转子(25,000rpm,2小时,4 $^{\circ}$ C)中的20%蔗糖垫,经超速离心进一步浓缩和纯化。将沉淀重悬在无内毒素PBS(GIBCO)中以获得浓缩500倍的VLP原液。通过Bradford测定定量试剂盒(BioRad)在等分试样上测定总蛋白。将纯化的VLP贮存在-80 $^{\circ}$ C直到使用。对于MMLV-Gag、gB、gH-G和/或MMLV-Gag/pp65融合蛋白的表达,通过带有对gB之特异性抗体(对gB的CH 28小鼠单克隆抗体;Virusys Corporation;Pereira,L等1984 Virology 139:73-86)、gH-G之特异性抗体(对抗VSV-G标签抗体P5D4的小鼠单克隆抗体;Abcam plc)和pp65之特异性抗体(对UL83/pp65的CH12小鼠单克隆抗体;Virusys Corporation;Pereira,L.等1982 Infect Immun 36:924-932)的SDS-Page和蛋白质印迹分析了各个批次的经纯化VLP(图2的B)。使用增强的化学发光(ECL)检测了抗体。

[0223] 实施例3:病毒样颗粒(VLP)的物理化学表征



[0224] VLP的物理化学分析包括粒径测定和使用Malvern Instrument Zeta-Sizer Nano series (ZEN3600) 评估多分散指数。从纳米化分析获得的示例性结果在图3的A和图3的B中示出。在两个不同实验室中使用相同的重组表达载体产生了示例性VLP组合物(gH-G/pp65二价VLP组合物)并且两种VLP制剂均给出了直径为150nm至180nm的平均粒径。这与CMV病毒体的尺寸一致,其报道为150nm至200nm的尺寸(1997 J Pathol 182:273-281)。0.214至0.240的低多分散指数(PdI)指示产品同质性和窄的尺寸分布。

[0225] 实施例4:VLP在小鼠中的免疫原性和中和活性

[0226] 在6至8周龄的雌性BALB/C小鼠中测试了如实施例2中所述制备的VLP组合物(每个测试组最少6只动物)。用200 $\mu$ l的VLP组合物腹膜内免疫小鼠2次,一次在0天(引发(Prime))并且一次在第56天(第8周强化)。用10 $\mu$ g、25 $\mu$ g或50 $\mu$ g(总蛋白)的二价gB/Gag/pp65、二价gH-G/Gag/pp65或三价gB/gH-G/Gag/pp65 VLP组合物处理小鼠。为评估小鼠中的体液免疫应答,免疫接种前以及第一次和第二次免疫接种后在0、2、3、4、6、8、9、10、12和14周从研究组中的所有小鼠收集血液。研究设计总结在表1中。

[0227] 表1

[0228]

测试编号#	剂量	测试编号描述	免疫接种时间表(周)
1	50 $\mu$ g	pp65/gB二价VLP	0, 8
2	25 $\mu$ g	pp65/gB二价VLP	0, 8
3	10 $\mu$ g	pp65/gB二价VLP	0, 8
4	50 $\mu$ g	pp65/gH-G二价VLP	0, 8
5	25 $\mu$ g	pp65/gH-G二价VLP	0, 8
6	10 $\mu$ g	pp65/gH-G二价VLP	0, 8
7	50 $\mu$ g	三价pp65/gB/gH-G VLP	0, 8
8	25 $\mu$ g	三价pp65/gB/gH-G VLP	0, 8
9	10 $\mu$ g	三价pp65/gB/gH-G VLP	0, 8

[0229] 使用用来自被HCMV毒株AD169感染之MRC-5细胞的裂解物包被的市售ELISA板(IBM International)进行酶联免疫吸附测定(ELISA)。未加工的市售HCMV裂解物(其含有所有CMV相关的抗原并且可用于检测IgG免疫应答)被用作通过ELISA测定小鼠血清HCMV IgG含量的阳性对照。将小鼠血清的连续稀释(在TBS-T/BSA/DMEM 10%FCS中稀释)与包被的板在室温孵育2小时。洗涤板后,以1/10,000的稀释度添加抗小鼠辣根过氧化物酶(HRP)缀合的第二抗体并孵育1小时,随后添加四甲基联苯胺(TMB)底物溶液。通过添加HCL 1N终止反应并且在ELISA微孔板阅读器中在450nm下读出吸光度。图4示出在8周第二次免疫接种后,对于各个二价和三价VLP组合物,小鼠之持久抗体和强强化的证据。

[0230] 使用微量中和测定确定了对HCMV的中和抗体应答。用感染培养基(含有1%加热灭活的FBS、1%氨基酸混合物、1%青霉素-链霉素以及无Ca<sup>+2</sup>和Mg<sup>+2</sup>之PBS的DME)稀释标准量的HCMV毒株VR1814(内皮细胞向性的CMV毒株-Revello MG等J Gen Virol 82:1429-1438),并添加至等体积的经加热灭活之测试血清的系列稀释并在转子上在37°C下孵育1小时。将血清/HCMV混合物添加到在24孔组织培养板中之盖玻片上生长的人包皮成纤维细胞(HFF)或视网膜色素上皮细胞(ARPE-19细胞)并且在37°C下,5%CO<sub>2</sub>中孵育2小时。将板用PBS清洗两次然后在37°C下,5%CO<sub>2</sub>中培养48小时。用4%多聚甲醛固定细胞,使其与抗IE1单克隆抗

体(对IE 1/2的CH 160小鼠单克隆抗体或对IE1的CH443单克隆抗体;Virusys Corporation)在室温下反应1小时,随后与FITC标记的山羊抗小鼠抗体在室温下反应45分钟。通过荧光显微术确定了表达IE1的细胞数目。在各个放血时间点(0、3、6、8和9周)测试混合小鼠血清(1:6稀释)的中和活性,一式两份。图5示出在8周第二次免疫接种后,各个二价和三价VLP组合物对小鼠之中和抗体(在成纤维细胞中测定)的诱导。图6示出在8周第二次免疫接种后,各个二价VLP组合物对小鼠之中和抗体的诱导(在成纤维细胞中测定)。

[0231] 在另一个研究中,在6至8周龄的雌性BALB/C小鼠中测试了如实施例2所述制备的单价gB-G VLP组合物(每个测试组最少8只动物)。用200 $\mu$ l的VLP组合物腹膜内免疫小鼠2次,一次在0天(引发)并且一次在第56天(第8周强化)。用等量的每次注射20 $\mu$ g的gB含量(通过ELISA确定)处理小鼠。为评估小鼠中的体液免疫应答,第一次免疫接种前、第一次免疫接种后在第3和6周、第二次免疫接种前在第8周以及第二次免疫接种后在第9周(从研究开始起)从研究组中的所有小鼠收集血液。研究设计总结在表2中。

[0232] 表2

测试编号#	剂量	测试编号描述	免疫接种时间表(周)
1	20 $\mu$ g	gB-G单价VLP	0, 8
2	20 $\mu$ g	重组gB	0, 8

[0234] 在成纤维细胞中基于表达绿色荧光蛋白的HCMV病毒(TB40)使用微量中和测定以及用流式细胞术分析感染的(GFP+)HFF细胞,确定了对HCMV的中和抗体应答。为了评估血清样品中中和抗体的存在,将血清与表达GFP的HCMV预孵育足够中和抗体减少HCMV之感染性的一段时间。血清和HCMV之预孵育混合物的连续稀释用于接触易受HCMV感染的宿主细胞(成纤维细胞或上皮细胞)。通过流式细胞术确定表达报告基因构建体(GFP)的细胞数以评估病毒制剂的感染效价。图7描述了相对于重组gB,在0和8周用单价gB-G VLP免疫2次的小鼠的中和抗体效价。将如所述的免疫接种前和免疫接种后收集的血清混合,并在10%豚鼠补体存在的情况下相对于阳性对照CMV超球蛋白(hyperglobulin)(Cytogam<sup>TM</sup>)测试。如图7所示,相比于由重组gB蛋白引发的,单价gB-G VLP组合物引发成纤维细胞感染的更快且更有效的中和。

[0235] 可以看出(或者已经阅读本说明书的本领域技术人员会理解),除了其他方面以外,该数据证明了与重组gB相比,VLP(例如包含gB-G的VLP)出乎意料好的活性。

[0236] 在另一个研究中,在6至8周龄的雌性BALB/C小鼠中测试了如实施例2所述制备的二价gB/Gag/pp65 VLP组合物(每个测试组最少4只动物)。用200 $\mu$ l的VLP组合物腹膜内免疫小鼠2次,一次在0天(引发)并且一次在第56天(第8周强化)。用二价gB/Gag/pp65 VLP处理小鼠并在14天后收集脾细胞。使用脾细胞刺激(1:5比率)富集的CD8<sup>+</sup>T细胞,以确定针对pp65或Gag之CTL的频率,所述脾细胞用编码Gag/pp65或Gag的质粒转染24小时。基于CFSE衰减、选通CD3<sup>+</sup>CD8<sup>+</sup>T细胞确定CTL频率(图8的A)。该散点图显示了pp65特异性CTL减去针对Gag的应答后,增殖的频率(图8的B)。如在图8中描述的,二价gB/Gag/pp65 VLP在免疫的小鼠中引发了pp65特异性的CTL。

[0237] 实施例5:VLP在兔中的免疫原性和中和活性

[0238] 在6至8周龄新西兰白兔中测试了如实施例2所述制备的二价gB/Gag/pp65和gH-G/Gag/pp65 VLP组合物(每个测试组最少6只动物)。兔用0.5ml的VLP组合物肌肉免疫3次,一

次在0天(引发),一次在第56天(第8周强化),一次在168天(24周强化)。用25 $\mu$ g或50 $\mu$ g或100 $\mu$ g(总蛋白)的二价gB/Gag/pp65、二价gH-G/Gag/pp65或三价gB/gH-G/Gag/pp65(以1:1比率混合在一起的二价gB/Gag/pp65和gH-G/Gag/pp65)VLP组合物处理兔。为评估兔中的体液免疫应答,第一次免疫接种前和第一次免疫接种后在第2、4、6、8周和第二次免疫接种后在第10、13、16、20和24周(从研究开始起)和第三次免疫接种后在第26和28周(从研究开始起)从研究组中的所有兔收集血液。研究设计总结在表3中。

[0239] 表3

测试编号#	剂量	测试编号描述	免疫接种时间表(周)
1	100 $\mu$ g	gB/pp65 二价 VLP	0, 8, 24
2	50 $\mu$ g	gB/pp65 二价 VLP	0, 8, 24
3	25 $\mu$ g	gB/pp65 二价 VLP	0, 8, 24
4	100 $\mu$ g	gH-G/pp65 二价 VLP	0, 8, 24
5	25 $\mu$ g	gH-G/pp65 二价 VLP	0, 8, 24
6	各 100 $\mu$ g	gB/pp65 二价 VLP + gH-G/pp65 二价 VLP (1:1 比率)	0, 8, 24
7	各 25 $\mu$ g	gB/pp65 二价 VLP + gH-G/pp65 二价 VLP (1:1 比率)	0, 8, 24

[0241] 通过ELISA测试单独兔血清针对重组gB抗原的反应性(在图9的A左轴上绘出)。如实施例4所述,在成纤维细胞中基于表达GFP的CMV病毒(TB40)使用微量中和测定以及用流式细胞术分析感染的(GFP+)HFF细胞确定了对HCMV的中和抗体应答(在图9的B右轴上绘出)。将如所述的免疫接种前和免疫接种后收集的兔血清混合并在HFF成纤维细胞中在对表达GFP的HCMV的补体存在的条件下,测试了相对于阳性对照CMV超球蛋白(Cytogam<sup>TM</sup>)的中和活性。端点效价(Endpoint titer)被绘制并代表在CMV感染的细胞中相对于匹配的免疫前血清减少50%(在图9的A左轴上绘出)。在流式细胞术分析感染的(GFP<sup>+</sup>)细胞期间收集50,000个细胞(在图9的B右轴上绘出)。如图9所示,二价gB/Gag/pp65 VLP组合物在兔中针对成纤维细胞感染引发高效价结合(A)和高效价中和抗体(B)应答。

[0242] 还在上皮细胞中基于表达GFP的HCMV病毒(Towne TS15-rR)使用微量中和测定以及用流式细胞术分析感染的(GFP+)ARPE-19细胞测试了混合的兔血清对HCMV的中和抗体应答。将如所述的免疫接种前和免疫接种后收集的兔血清混合并在ARPE-19上皮细胞中在针对表达GFP之HCMV的补体存在的条件下测试了相对于阳性对照CMV超球蛋白(Cytogam<sup>TM</sup>)的中和活性。如图10中所示,出乎意料地,二价gB/Gag/pp65和二价gH-G/Gag/pp65 VLP组合物的组合表现出协同作用,相对于gB/Gag/pp65或gH/Gag/pp65,在兔中针对上皮细胞感染的中和抗体应答更快速。

[0243] 可以看出(或者已经阅读本说明书的本领域技术人员会理解),除了其他方面以

外,该数据证明了与包含gB/Gag/pp65的VLP或包含gH-G/Gag/pp65的VLP相比,VLP(例如通过包含gB/Gag/pp65之VLP和包含gH-G/Gag/pp65之VLP的组合)出乎意料地好的活性。

[0244] 实施例6:病毒样颗粒(VLP)之按比例增加的生产和纯化

[0245] 如下生产和纯化VLP。用选择的质粒(如实施例1所述制备)以预定浓度和比率转染细胞密度为 $1.5 \times 10^6$ 至 $2.0 \times 10^6$ 细胞/mL的CHO细胞。添加填充物DNA(Stuffer DNA)使得总DNA浓度升至 $1 \mu\text{g/mL}$ 细胞培养物。将用于转染的质粒首先通过MaxiPrep或GigaPrep质粒纯化试剂盒(Qiagen)纯化。用于转染以将DNA递送至细胞的PEIMAX以6:1(PEI:DNA重量/重量)的比率提供。将细胞培养72小时,然后使用转子JS-4.2A(Beckman Coulter)在1L瓶中,将培养物以4000rpm离心20分钟。将上清液通过0.8/0.45 $\mu\text{m}$ 滤器(AcroPak 500 Capsule, Pall)过滤。然后通过切向流过滤(TFF)浓缩经过滤的上清液并对含组氨酸的缓冲液渗滤。将经渗滤的物质装载至阴离子交换层析柱(AEX)上,其中收集流出液。然后将流出液通过0.45 $\mu\text{m}$ 无菌过滤以分装成不同的体积。

[0246] TFF程序涉及两个TFF膜(Pellicon 2 Mini 500kDa截止,0.1 $\text{m}^2$ 表面积)通过将它们平行固定在不锈钢外壳中在0.5M NaOH中的过夜消毒(sanitization)。在渗余物(retentate)以及在渗透物(permeate)侧上运行水至中性pH后,使用磷酸盐缓冲液(PBS)平衡膜。然后将经过滤的上清液装载至TFF渗余物再循环环中。开始的入口压力为10.5psi,渗透物流速为400mL/分钟,其在浓缩的末期降低至约200mL/分钟。浓缩到约10至20倍以后,用5倍体积的20mM组氨酸+150mM NaCl,pH 5.5完成渗滤。透过流关闭再循环另一个5分钟后,收集渗滤物质并用等体积的组氨酸缓冲液(his-buffer)漂洗。然后收集渗余物并与之前收集的渗余物混合。为保持膜的功能性,用PBS漂洗膜10分钟;用0.5M NaOH漂洗40分钟(通过渗透物和渗余物向废物冲刷200ml后);并且最后在渗余物和渗透物中用水漂洗至中性pH。然后将膜贮存在冷藏室中的0.1M NaOH溶液中。

[0247] 使用AEX柱色谱以减少DNA含量,使用20mL HiLoad 16/10琼脂糖HP柱,以1.6mL/分钟的流速来用平衡缓冲液(20mM组氨酸+150mM NaCl,pH 5.5)平衡,并且以3.2mL/分钟的流速进行装载、洗涤和剥离(stripping)步骤。在0.8mL/分钟下进行清洁步骤。使用图表记录器监测280nm下的紫外吸收。使用Gradifrac色谱系统(GE Healthcare),其在使用前被消毒。使用5柱体积的超Q水(低内毒素)从柱移除第一乙醇(在柱中作为贮藏缓冲液以20%体积/体积存在)。通过5柱体积的平衡缓冲液,随后5柱体积的剥离缓冲液(20mM组氨酸+1000mM NaCl,pH 5.5)以调节柱。再次通过5柱体积的平衡缓冲液以准备用于装载的柱。以3.2mL/分钟进行装载,之后柱用5柱体积的平衡缓冲液清洗或者直到观察到基线。在材料的上样过程中,从UV峰值开始直到UV峰值跌落至UV吸光度之最大峰值高度的约10%,收集流出物。然后通过5柱体积的剥离缓冲液(20mM组氨酸+1000mM NaCl,pH 5.5)剥离柱并收集峰值。之后是5柱体积的另外剥离缓冲液(20mM组氨酸+2000mM NaCl,pH 5.5)以除去任何强力键合的蛋白和核酸并且收集峰值。然后用1M NaOH(4柱体积,0.8mL/分钟)清洗柱以移除任何沉淀的蛋白。然后用水(0.8mL/分钟)将柱漂洗至中性pH(通常4至5柱体积)。然后20%酒精(4柱体积,0.8mL/分钟)通过柱以除去任何脂蛋白或脂质(2柱体积)。在这一阶段,柱或者贮存或者用水漂洗(4柱体积)以重新开始第二批次的循环。

[0248] 图11描绘了经纯化的gB-G单价VLP的图像,其由CHO细胞产生并且随后经过TFF(图11的A)和AEX(图11的B)纯化方法,接着是负染电子显微镜分析方法。如图11所示,TFF(图11

的A)和AEX纯化(图11的B)后呈现了完整gB-G单价VLP。

[0249] 实施例7:纯化的VLP在兔中的免疫原性和中和活性

[0250] 在6至8周龄新西兰白兔中测试了如实施例6所述制备的单价gB-G VLP组合物(每个测试组最少6只动物)。将兔用0.5ml (50 $\mu$ g gB含量)的VLP组合物肌肉免疫3次,一次在0天(引发),一次在第56天(第8周强化),一次在第168天(第24周强化)。为评估兔中的体液免疫应答,第一次免疫接种前和第一次免疫接种后在第2、4、6、8周和第二次免疫接种后在第10、13、16、20和24周(从研究开始起)从研究组中的所有兔收集血液。研究设计总结在表4中。

[0251] 表4

	测试编号#	剂量	测试编号描述	免疫接种时间表(周)
[0252]	1	50 $\mu$ g gB 含量	gB-G 单价 VLP (TFF&AEX 纯化的)	0, 8, 24

[0253] 图12描绘了成纤维细胞感染的有效中和,其通过来自用TFF/AEX纯化的CHO细胞产生的gB单价VLP(图12中“gB eVLPs”)免疫之兔的血清引发。这一中和作用超过用阳性对照CMV超球蛋白(Cytogam<sup>TM</sup>)达到的中和作用。图12还包括公开的Towne疫苗和佐剂的gB亚单位疫苗(gB+MF59<sup>TM</sup>)的中和效价数据(Cui X等2008 Vaccine 26:5760-5766)。

[0254] 图13举例说明了上皮细胞感染的有效中和,其通过来自用TFF/AEX纯化的CHO细胞产生的gB-G单价VLP(图13中“gB eVLPs”)免疫之兔的血清引发。这一中和作用比得上用阳性对照CMV超球蛋白(Cytogam<sup>TM</sup>)达到的中和作用。图13还包括公开的Towne疫苗和佐剂的gB亚单位疫苗(gB+MF59<sup>TM</sup>)的中和效价数据(Cui X等2008 Vaccine 26:5760-5766)。

[0255] 图14描绘了在用TFF/AEX纯化的CHO细胞产生的gB-G单价VLP组合物免疫之兔中引发的抗体之亲和力指数。混合的兔血清和阳性对照CMV超球蛋白(Cytogam<sup>TM</sup>)以1:600,000稀释并且在存在或不存在5M尿素的情况下通过ELISA针对全长重组gB抗原测试。如前所述(Marshau BC和Adler S 2003 Viral Immunol 16:491-500)确定了抗体亲和力。如图14所示,通过用TFF/AEX纯化的CHO细胞产生的gB-G单价VLP的免疫接种在兔中引起了高亲和力中和抗体的快速诱导。两次gB-G VLP免疫接种后实现了最大的抗体亲和力。

[0256] 其他实施方案

[0257] 从本文公开的公开内容之实践或说明书考虑,本公开内容的其他实施方案对本领域技术人员将是显而易见的。本说明书和实施例旨在仅被视为示例性的,本公开内容的真实范围由以下权利要求书指出。本文涉及的任何参考文献的内容通过引用整体并入本文。

[0258] 序列表说明

[0259] SEQ ID NO:1描述MMLV Gag氨基酸序列。

[0260] SEQ ID NO:2描述MMLV Gag核苷酸序列。

[0261] SEQ ID NO:3描述密码子优化的MMLV Gag核苷酸序列。

[0262] SEQ ID NO:4描述MMLV Gag-CMV pp65氨基酸序列。

[0263] SEQ ID NO:5描述MMLV Gag-CMV pp65核苷酸序列。

[0264] SEQ ID NO:6描述密码子优化的MMLV Gag-CMV pp65核苷酸序列。

[0265] SEQ ID NO:7描述HCMV gB氨基酸序列。

[0266] SEQ ID NO:8描述HCMV gB核苷酸序列。

- [0267] SEQ ID NO:9描述密码子优化的HCMV gB核苷酸序列。
- [0268] SEQ ID NO:10描述HCMV gB-G氨基酸序列。
- [0269] SEQ ID NO:11描述HCMV gB-G核苷酸序列。
- [0270] SEQ ID NO:12描述密码子优化的HCMV gB-G核苷酸序列。
- [0271] SEQ ID NO:13描述HCMV gH氨基酸序列。
- [0272] SEQ ID NO:14描述HCMV gH核苷酸序列。
- [0273] SEQ ID NO:15描述密码子优化的HCMV gH核苷酸序列。
- [0274] SEQ ID NO:16描述HCMV gH-G氨基酸序列。
- [0275] SEQ ID NO:17描述HCMV gH-G核苷酸序列。
- [0276] SEQ ID NO:18描述密码子优化的HCMV gH-G核苷酸序列。
- [0277] SEQ ID NO:19描述Propol II表达质粒核苷酸序列。
- [0278] SEQ ID NO:20描述HCMV gH-HCMV gB TM/CTD核苷酸序列。
- [0279] SEQ ID NO:21描述密码子优化的MMLV Gag核苷酸序列。
- [0280] SEQ ID NO:22描述密码子优化的MMLV Gag-CMV pp65核苷酸序列。
- [0281] 以下内容对应于母案申请中的原始权利要求书,现作为说明书的一部分并入此处:
- [0282] 1.病毒样颗粒(VLP),其包含:
- [0283] 第一多肽,所述第一多肽为鼠白血病病毒(MLV) gag多肽,其中其氨基酸序列包含与具有SEQ ID NO:1氨基酸序列的参考MLV gag蛋白之自组装部分显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分,以及
- [0284] 第二多肽,其氨基酸序列包含被结合人巨细胞病毒(HCMV)蛋白的一种或更多种抗体识别的表位。
- [0285] 2.项1所述的VLP,其中所述第二多肽具有包含被HCMV糖蛋白B(gB)蛋白的一种或更多种抗体识别之表位的氨基酸序列。
- [0286] 3.项1所述的VLP,其中所述第二多肽具有包含被HCMV糖蛋白H(gH)蛋白的一种或更多种抗体识别之表位的氨基酸序列。
- [0287] 4.项1至3中任一项所述的VLP,其中所述第二多肽具有包含跨膜结构域、胞质结构域或二者的氨基酸序列。
- [0288] 5.项4所述的VLP,其中所述跨膜结构域不天然地存在于HCMV蛋白中。
- [0289] 6.项4所述的VLP,其中所述跨膜结构域天然地存在于水泡性口炎病毒(VSV)蛋白中。
- [0290] 7.项4所述的VLP,其中所述跨膜结构域具有不同于SEQ ID NO:7和SEQ ID NO:13中所存在氨基酸序列的氨基酸序列。
- [0291] 8.项4所述的VLP,其中所述跨膜结构域具有选自以下的氨基酸序列:存在于SEQ ID NO:7中的跨膜结构域,存在于SEQ ID NO:13中的跨膜结构域和存在于SEQ ID NO:16中的跨膜结构域。
- [0292] 9.项4所述的VLP,其中所述跨膜结构域具有SEQ ID NO:10中752至795位残基的氨基酸序列。

[0293] 10. 项4所述的VLP, 其中所述跨膜结构域具有由SEQ ID NO:20中2152至2622位核苷酸编码的氨基酸序列。

[0294] 11. 项4至10中任一项所述的VLP, 其中所述第二多肽具有包含以下的氨基酸序列: 1) 存在于SEQ ID NO:13中的表位; 以及2) 包含SEQ ID NO:10中752至795位残基之跨膜结构域的氨基酸序列。

[0295] 12. 项4至10中任一项所述的VLP, 其中所述第二多肽具有包含以下的氨基酸序列: 1) 存在于SEQ ID NO:13中的表位; 以及2) 包含由SEQ ID NO:20中2152至2622位核苷酸所编码氨基酸序列之跨膜结构域的氨基酸序列。

[0296] 13. 项11或项12所述的VLP, 其中所述第二多肽具有氨基酸序列, 其包含存在于SEQ ID NO:13中的除所述表位外的一个或更多个序列, 其中所述一个或更多个序列不包含完整的跨膜结构域。

[0297] 14. 项1至13中任一项所述的VLP, 其中所述第一多肽为融合蛋白或包含融合蛋白。

[0298] 15. 项1至13中任一项所述的VLP, 其中所述第二多肽为融合蛋白或包含融合蛋白。

[0299] 16. 跨膜的经修饰的gH多肽。

[0300] 17. 项16所述的多肽, 其包含存在于HCMV中的gH蛋白之N端胞外结构域。

[0301] 18. 项16或项17所述的多肽, 其包含跨膜结构域、胞质结构域或二者。

[0302] 19. 项18所述的多肽, 其中所述跨膜结构域、胞质结构域或二者不天然地存在于所述gH蛋白中。

[0303] 20. 项18所述的多肽, 其中所述跨膜结构域、胞质结构域或二者天然地存在于VSV蛋白中。

[0304] 21. 项19所述的多肽, 其中所述跨膜结构域、胞质结构域或二者天然地存在于gB蛋白中。

[0305] 22. 项16所述的多肽, 其包含SEQ ID NO:16。

[0306] 23. 项16至22中任一项所述的多肽, 其中所述多肽存在于VLP中。

[0307] 24. 多核苷酸, 其编码项16至22中任一项所述的多肽。

[0308] 25. 载体, 其包含项24所述的多核苷酸。

[0309] 26. 病毒样颗粒(VLP), 其包含多肽, 所述多肽包含存在于HCMV中的gB蛋白之表位。

[0310] 27. 项26所述的VLP, 其中所述多肽包含存在于HCMV中的gB蛋白之N端胞外结构域。

[0311] 28. 项26或27所述的VLP, 其中所述多肽包含跨膜结构域、胞质结构域或者二者。

[0312] 29. 项28所述的VLP, 其中所述跨膜结构域、胞质结构域或者二者不天然地存在于gB蛋白中。

[0313] 30. 项28所述的VLP, 其中所述跨膜结构域、胞质结构域或者二者天然地存在于VSV蛋白中。

[0314] 31. 项26所述的VLP, 其中所述多肽包含SEQ ID NO:7。

[0315] 32. 项30所述的VLP, 其中所述多肽包含SEQ ID NO:10。

[0316] 33. 融合蛋白, 其包含:

[0317] a) 第一多肽, 所述第一多肽为MLV gag多肽, 其中其氨基酸序列包含与具有SEQ ID NO:1氨基酸序列的参考MLV gag蛋白之自组装部分显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分, 以及

[0318] b) 第二多肽,所述第二多肽为包含pp65表位的HCMV pp65多肽,其中其氨基酸序列包含与具有SEQ ID NO:4中539至1099位残基之氨基酸序列的参考pp65蛋白显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分。

[0319] 34. 包含融合蛋白的病毒样颗粒(VLP),其包含:

[0320] a) 第一多肽,所述第一多肽为MLV gag多肽,其中其氨基酸序列包含与具有SEQ ID NO:1氨基酸序列的参考MLV gag蛋白之自组装部分显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分,以及

[0321] b) 第二多肽,所述第二多肽为包含pp65表位的HCMV pp65多肽,其中其氨基酸序列包含与具有SEQ ID NO:4中539至1099位残基之氨基酸序列的参考pp65蛋白显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分。

[0322] 35. 病毒样颗粒(VLP),其包含:

[0323] 第一多肽,其包含存在于HCMV中的gB蛋白之表位;和

[0324] 第二多肽,其包含跨膜的经修饰的gH多肽。

[0325] 36. 项35所述的VLP,其中所述第一多肽包含存在于HCMV中的gB蛋白之N端胞外结构域。

[0326] 37. 项36所述的VLP,其中所述第一多肽包含跨膜结构域、胞质结构域或者二者。

[0327] 38. 项37所述的VLP,其中所述跨膜结构域、胞质结构域或者二者不天然地存在于gB蛋白中。

[0328] 39. 项38所述的VLP,其中所述跨膜结构域、胞质结构域或者二者天然地存在于VSV蛋白中。

[0329] 40. 项35所述的VLP,其中所述第一多肽包含SEQ ID NO:7。

[0330] 41. 项39所述的VLP,其中所述第一多肽包含SEQ ID NO:10。

[0331] 42. 项35所述的VLP,其中所述第二多肽包含天然地存在于HCMV中的gH蛋白之N端胞外结构域。

[0332] 43. 项35或42所述的VLP,其中所述第二多肽包含跨膜结构域、胞质结构域或者二者。

[0333] 44. 项43所述的VLP,其中所述跨膜结构域、胞质结构域或者二者不天然地存在于gH蛋白中。

[0334] 45. 项43所述的VLP,其中所述跨膜结构域、胞质结构域或者二者天然地存在于VSV蛋白中。

[0335] 46. 项35所述的VLP,其中所述第二多肽包含SEQ ID NO:16。

[0336] 47. 病毒样颗粒(VLP),其包含:

[0337] 融合蛋白,其包含:

[0338] a) 第一多肽,所述第一多肽为MLV gag多肽,其中其氨基酸序列包含与具有SEQ ID NO:1氨基酸序列的参考MLV gag蛋白之自组装部分显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分,以及

[0339] b) 第二多肽,所述第二多肽为包含pp65表位的HCMV pp65多肽,其中其氨基酸序列



包含与具有SEQ ID NO:4中539至1099位残基之氨基酸序列的参考pp65蛋白显示至少70%、75%、80%、85%、90%、95%、98%、99%或100%同一性的长度为至少100、200、300或更多个氨基酸的部分;以及

[0340] 存在于HCMV中的第二蛋白之表位。

[0341] 48.项47所述的VLP,其中所述融合蛋白设置并构建为使得其自组装以形成VLP,从而所述pp65表位位于所述VLP内部。

[0342] 49.项47或项48所述的VLP,其设置和构建为使得第二HCMV蛋白表位位于VLP表面上。

[0343] 50.VLP,其包含:

[0344] 融合蛋白,其包含融合在HCMV中所存在pp65蛋白之表位上游的MLV中所存在gag蛋白之N端部分;以及

[0345] 多肽,其包含存在于HCMV中的gB蛋白之表位。

[0346] 51.项50所述的VLP,其中所述gag蛋白之N端部分不包含存在于MLV中的C端聚合酶序列。

[0347] 52.项50或项51所述的VLP,其中所述gag蛋白之N端部分包含全长蛋白的1至538位残基。

[0348] 53.项50至52中任一项所述的VLP,其中所述gB蛋白的表位存在于全长gB蛋白中。

[0349] 54.项50至52中任一项所述的VLP,其中所述gB蛋白的表位被结合全长gB中表位的抗体识别。

[0350] 55.项50至52中任一项所述的VLP,其中所述gB蛋白的表位被结合SEQ ID NO:7中表位的抗体识别。

[0351] 56.项50至52中任一项所述的VLP,其中包含gB表位的所述多肽具有包含全长gB蛋白中1至751位残基的氨基酸序列。

[0352] 57.项50至52中任一项所述的VLP,其中包含gB表位的所述多肽具有包含SEQ ID NO:7之至少15个连续残基的氨基酸序列。

[0353] 58.项50至52中任一项所述的VLP,其中包含gB表位的所述多肽具有包含SEQ ID NO:7的氨基酸序列。

[0354] 59.项50至52中任一项所述的VLP,其中包含gB表位的所述多肽包含融合蛋白。

[0355] 60.项59所述的VLP,其中所述融合蛋白包含与第二多肽融合的第一多肽,所述第一多肽包含来自存在于HCMV中的gB蛋白之N端胞外结构域,所述第二多肽包含不天然地存在于HCMV中的跨膜结构域、胞质结构域或二者。

[0356] 61.项60所述的VLP,其中所述第二多肽天然地存在于水泡性口炎病毒(VSV)中。

[0357] 62.项60的VLP,其中所述融合蛋白包含SEQ ID NO:10。

[0358] 63.VLP,其包含:

[0359] 融合蛋白,其包含与存在于HCMV中的pp65蛋白之表位融合的存在于MLV中的gag蛋白之N端部分;以及

[0360] 多肽,其包含存在于HCMV中的gH蛋白之表位。

[0361] 64.项63所述的VLP,其中所述gag蛋白的N端部分不包含存在于MLV中的C端聚合酶序列。

- [0362] 65. 项63或项64所述的VLP,其中所述gag蛋白的N端部分包含全长蛋白的1至538位残基。
- [0363] 66. 项63至65中任一项所述的VLP,其中所述gH蛋白的表位存在于全长gH蛋白中。
- [0364] 67. 项63至65中任一项所述的VLP,其中所述gH蛋白的表位被结合全长gH中表位的抗体识别。
- [0365] 68. 项63至65中任一项所述的VLP,其中所述gH蛋白的表位被结合SEQ ID NO:13中表位的抗体识别。
- [0366] 69. 项63至65中任一项所述的VLP,其中包含gH表位的所述多肽具有包含全长gH蛋白中1至751位残基的氨基酸序列。
- [0367] 70. 项63至65中任一项所述的VLP,其中包含gH表位的所述多肽具有包含SEQ ID NO:13之至少15个连续残基的氨基酸序列。
- [0368] 71. 项63至65中任一项所述的VLP,其中包含gH表位的所述多肽具有包含SEQ ID NO:13的氨基酸序列。
- [0369] 72. 项63至65中任一项所述的VLP,其中包含gH表位的所述多肽包含融合蛋白。
- [0370] 73. 项72所述的VLP,其中所述融合蛋白包含与第二多肽融合的第一多肽,所述第一多肽包含来自存在于HCMV中的gB蛋白之N端胞外结构域,所述第二多肽包含不天然地存在于HCMV蛋白中的跨膜结构域、胞质结构域或二者。
- [0371] 74. 项73所述的VLP,其中所述第二多肽为gB蛋白。
- [0372] 75. 项73所述的VLP,其中所述在第二多肽天然地存在于水泡性口炎病毒(VSV)中。
- [0373] 76. VLP,其包含:
- [0374] 融合蛋白,其包含融合在人巨细胞病毒(HCMV)中所存在pp65蛋白之表位上游的鼠白血病病毒(MLV)中所存在gag蛋白之N端部分;
- [0375] 多肽,其包含存在于HCMV中的糖蛋白B(gB)蛋白之表位;以及
- [0376] 多肽,其包含存在于HCMV中的糖蛋白H(gH)蛋白之表位。
- [0377] 77. 项76所述的VLP,其中所述gag蛋白的N端部分不包含存在于鼠白血病病毒(MLV)中的C端聚合酶序列。
- [0378] 78. 项76或项77所述的VLP,其中所述gag蛋白的N端部分包含全长蛋白的1至538位残基。
- [0379] 79. 项76至78中任一项所述的VLP,其中所述gB蛋白的表位存在于全长gB蛋白中。
- [0380] 80. 项76至78中任一项所述的VLP,其中所述gB蛋白的表位被结合全长gB中表位的抗体识别。
- [0381] 81. 项76至78中任一项所述的VLP,其中所述gB蛋白的表位被结合SEQ ID NO:7中表位的抗体识别。
- [0382] 82. 项76至78中任一项所述的VLP,其中包含gB表位的所述多肽具有包含全长gB蛋白中1至751位残基的氨基酸序列。
- [0383] 83. 项76至78中任一项所述的VLP,其中包含gB表位的所述多肽具有包含SEQ ID NO:7之至少15个连续残基的氨基酸序列。
- [0384] 84. 项76至78中任一项所述的VLP,其中包含gB表位的所述多肽具有包含SEQ ID NO:7的氨基酸序列。

- [0385] 85. 项76至78中任一项所述的VLP,其中包含gB表位的所述多肽包含融合蛋白。
- [0386] 86. 项85所述的VLP,其中所述融合蛋白包含与第二多肽融合的第一多肽,所述第一多肽包含来自存在于HCMV中的gB蛋白之N端胞外结构域,所述第二多肽包含不天然地存在于HCMV蛋白中的跨膜结构域、胞质结构域或二者。
- [0387] 87. 项86所述的VLP,其中所述第二多肽天然地存在于水泡性口炎病毒(VSV)中。
- [0388] 88. 项86所述的VLP,其中所述融合蛋白包含SEQ ID NO:10。
- [0389] 89. 项76至88中任一项所述的VLP,其中所述gH蛋白的表位存在于全长gH蛋白中。
- [0390] 90. 项76至88中任一项所述的VLP,其中所述gH蛋白的表位被结合全长gH中表位的抗体识别。
- [0391] 91. 项76至88中任一项所述的VLP,其中所述gH蛋白的表位被结合SEQ ID NO:13中表位的抗体识别。
- [0392] 92. 项76至88中任一项所述的VLP,其中包含gH表位的所述多肽具有包含全长gH蛋白中1至717位残基的氨基酸序列。
- [0393] 93. 项76至88中任一项所述的VLP,其中包含gH表位的所述多肽具有包含SEQ ID NO:13之至少15个连续残基的氨基酸序列。
- [0394] 94. 项76至88中任一项所述的VLP,其中包含gH表位的所述多肽具有包含SEQ ID NO:13的氨基酸序列。
- [0395] 95. 项76至88中任一项所述的VLP,其中包含gH表位的所述多肽包含融合蛋白。
- [0396] 96. 项95所述的VLP,其中所述融合蛋白包含与第二多肽融合的第一多肽,所述第一多肽包含来自存在于HCMV中的gH蛋白之N端胞外结构域,所述第二多肽包含不天然地存在于HCMV蛋白中的跨膜结构域、胞质结构域或二者。
- [0397] 97. 项96所述VLP,其中所述第二多肽为gB蛋白。
- [0398] 98. 项96所述VLP,其中所述第二多肽天然地存在于水泡性口炎病毒(VSV)中。
- [0399] 99. 项1至15、26至32和34至98中任一项所述的VLP,其中所述VLP的特征在于其直径范围的下限由20、30、40、50、60、70、80、90或100nm界定并且上限由300、290、280、270、260、250、240、230、220、210、200、190、180或170nm界定。
- [0400] 100. 根据项1至15、26至32和34至98中任一项所述VLP的群体,其中所述群体中VLP平均直径范围的下限由20、30、40、50、60、70、80、90或100nm界定并且上限由300、290、280、270、260、250、240、230、220、210、200、190、180或170nm界定。
- [0401] 101. 根据项100所述的VLP之群体,其中所述群体中VLP的多分散指数小于0.5。
- [0402] 102. 项1至15、26至32和34至98中任一项所述的VLP,其中所述VLP的特征在于所述糖蛋白的表位暴露在VLP的表面上并且所述融合蛋白位于VLP的被膜/内腔中。
- [0403] 103. 免疫原性组合物,其包含项1至15、26至32和34至98中任一项所述的VLP。
- [0404] 104. 项103所述的免疫原性组合物,其中所述组合物的特征在于当施用至对象时其诱导体液免疫应答和细胞免疫应答二者。
- [0405] 105. 项104所述的免疫原性组合物,其中所述体液免疫应答在对象中持续至少约1个月、至少约2个月、至少约3个月、至少约4个月、至少约5个月、至少约6个月、至少约7个月、至少约8个月、至少约9个月、至少约10个月、至少约11个月或至少12个月。
- [0406] 106. 项104所述的免疫原性组合物,其中所述细胞免疫应答在对象中持续至少约1

个月、至少约2个月、至少约3个月、至少约4个月、至少约5个月、至少约6个月、至少约7个月、至少约8个月、至少约9个月、至少约10个月、至少约11个月或至少12个月。

[0407] 107. 免疫原性组合物,其包含项50所述之VLP与项63所述之重组VLP的混合物。

[0408] 108. 免疫原性组合物,其包含项50所述之VLP与项75所述之VLP的混合物。

[0409] 109. 载体,其包含:

[0410] 编码融合蛋白的多核苷酸,所述融合蛋白包含融合在人巨细胞病毒 (HCMV) 中所存在pp65蛋白之表位上游的鼠白血病病毒 (MLV) 中所存在gag蛋白之N端部分;

[0411] 其中所述编码融合蛋白的多核苷酸在单个启动子的控制之下;并且

[0412] 其中所述载体包含在编码融合蛋白载体之多核苷酸下游的多聚腺苷酸化信号。

[0413] 110. 项109所述的载体,其中所述gag蛋白的N端部分不包含存在于鼠白血病病毒 (MLV) 中的C端聚合酶序列。

[0414] 111. 项109或项110所述的载体,其中所述gag蛋白的N端部分包含全长蛋白的1至538位残基。

[0415] 112. 项109至111中任一项所述的载体,其包含SEQ ID NO:5、SEQ ID NO:6或SEQ ID NO:22的多核苷酸序列。

[0416] 113. 项109至111中任一项所述的载体,其中所述载体如图1中所描绘。

[0417] 114. 载体,其包含:

[0418] 编码多肽的多核苷酸,所述多肽包含存在于HCMV中的糖蛋白B(gB) 蛋白之表位;

[0419] 其中所述多核苷酸在单个启动子的控制之下;并且

[0420] 其中所述载体包含所述多核苷酸上游的信号序列和所述多核苷酸下游的多聚腺苷酸化信号。

[0421] 115. 项114所述的载体,其中包含所述gB蛋白之表位的所述多肽包含全长gB蛋白的1至751位残基。

[0422] 116. 项114或项115所述的载体,其包含SEQ ID NO:8或SEQ ID NO:9的多核苷酸序列。

[0423] 117. 项114至116中任一项所述的载体,其中所述载体如图1中所描绘。

[0424] 118. 载体,其包含:

[0425] 编码多肽的多核苷酸,所述多肽包含存在于HCMV中的糖蛋白H(gH) 蛋白之表位;

[0426] 其中所述多核苷酸在单个启动子的控制下;并且

[0427] 其中所述载体包含所述多核苷酸上游的信号序列和所述多核苷酸下游的多聚腺苷酸化信号。

[0428] 119. 项118所述的载体,其中包含所述gH蛋白之表位的所述多肽不包含来自存在于HCMV中的gH蛋白之跨膜结构域或胞质结构域。

[0429] 120. 项118或项119所述的载体,其中包含所述gH蛋白之表位的所述多肽包含融合蛋白。

[0430] 121. 项120所述的载体,其中所述融合蛋白包含融合在第二多肽上游的来自HCMV的gH蛋白质之N端胞外结构域,所述第二多肽包含不天然地存在于HCMV蛋白中的跨膜结构域、胞质结构域或二者。

[0431] 122. 项121所述的载体,其中所述第二多肽为gB蛋白。

- [0432] 123. 项121所述的载体,其中所述第二多肽天然地存在于水泡性口炎病毒(VSV)中。
- [0433] 124. 项118至123中任一项所述的载体,其中包含gH表位的所述多肽包含全长gH蛋白的1至717位残基。
- [0434] 125. 项118至124中任一项所述的载体,其包含SEQ ID NO:14中1至2148位核苷酸或SEQ ID NO:15中1至2148位核苷酸的多核苷酸序列。
- [0435] 126. 项118至125中任一项所述的载体,其中载体如图1所描绘。
- [0436] 127. 细胞,其包含项109至126中任一项所述的一种或更多种载体。
- [0437] 128. 项127所述的细胞,其中所述细胞用一种或更多种所述载体瞬时转染。
- [0438] 129. 项127所述的细胞,其中所述细胞用一种或更多种所述载体稳定转染。
- [0439] 130. 用于生产VLP的方法,所述方法包括:
- [0440] 用项109至113中任一项所述的载体和项114至117中任一项所述的载体共转染宿主细胞;以及
- [0441] 在合适的培养基中于允许表达由所述载体编码之蛋白的条件下培养所述宿主细胞。
- [0442] 131. 用于生产VLP的方法,所述方法包括:
- [0443] 用项109至113中任一项所述的载体和项118至126中任一项所述的载体共转染宿主细胞;以及
- [0444] 在合适的培养基中于允许表达由所述载体编码之蛋白的条件下培养所述宿主细胞。
- [0445] 132. 用于生产VLP的方法,所述方法包括:
- [0446] 用项109至113中任一项所述的载体、项114至117中任一项所述的载体以及项118至126中任一项所述的载体共转染宿主细胞;以及
- [0447] 在合适的培养基中于允许表达由所述载体编码之蛋白的条件下培养所述宿主细胞。
- [0448] 133. 项130至132中任一项所述的方法,其还包括从所述培养基和/或所述宿主细胞回收VLP的步骤。
- [0449] 134. 药物组合物,其包含项1至15、26至32和34至98中任一项所述的VLP和可药用载体。
- [0450] 135. 项134所述的药物组合物,其还包含佐剂。
- [0451] 136. 项135所述的药物组合物,其中所述佐剂选自细胞因子、凝胶型佐剂、微生物佐剂、油乳剂和基于乳化剂的佐剂、颗粒佐剂、合成佐剂、聚合物佐剂和/或其组合。
- [0452] 137. 用于在对象中减少HCMV感染之症状的频率、严重程度或延迟所述症状发作的方法,所述方法包括施用项134至136中任一项所述的药物组合物。
- [0453] 138. 项137所述的方法,其中所述对象处于HCMV感染的风险中。
- [0454] 139. 项138所述的方法,其中所述对象为免疫抑制的对象。
- [0455] 140. 项139所述的方法,其中所述免疫抑制的对象选自:HIV感染的对象、AIDS患者、移植受者、儿童对象和妊娠对象。
- [0456] 141. 项137至140中任一项所述的方法,其中所述对象已经暴露于HCMV感染。

[0457] 142. 项137至141中任一项所述的方法,其中所述对象为人。

[0458] 143. 项137所述的方法,其中与施用所述药物组合物前所述对象中HCMV感染之症状的频率、严重程度或发作相比,所述对象中HCMV感染之症状的频率、严重程度或发作减少约50%或更多。

[0459] 144. 项143所述的方法,其中与施用所述药物组合物前所述对象中HCMV感染之症状的频率、严重程度或发作相比,所述对象中HCMV感染之症状的频率、严重程度或发作减少约60%或更多、约70%或更多、或约75%或更多。

[0460] 145. 项137至144中任一项所述的方法,其中所述药物组合物以初始剂量和至少一次加强剂量施用。

[0461] 146. 项145所述的方法,其中所述药物组合物以初始剂量和两次加强剂量施用。

[0462] 147. 项145所述的方法,其中所述药物组合物以初始剂量和三次加强剂量施用。

[0463] 148. 项145所述的方法,其中所述药物组合物以初始剂量和四次加强剂量施用。

[0464] 149. 项145所述的方法,其中所述药物组合物以初始剂量和在所述初始剂量后约1个月、约2个月、约3个月、约4个月、约5个月或约6个月的至少一次加强剂量施用。

[0465] 150. 项149所述的方法,其中所述药物组合物在所述初始剂量后约6个月、约7个月、约8个月、约9个月、约10个月、约11个月或约1年以第二加强剂量施用。

[0466] 151. 项137至150中任一项所述的方法,其中所述药物组合每1年、每2年、每3年、每4年、每5年、每6年、每7年、每8年、每9年或每10年以加强剂量施用。

## 序列表

<110> 变异生物技术公司

<120> 用于治疗巨细胞病毒的组合物和方法

<130> PAT 7372W-90

<140> PCT/IB2012/002854

<141> 2012-11-09

<150> US 61/558,800

<151> 2011-11-11

<160> 22

<170> PatentIn version 3.5

<210> 1

<211> 538

<212> PRT

<213> 鼠白血病病毒

[0001]

<220>

<221> MISC\_特征

<222> (1)..(538)

<223> MMLV Gag 氨基酸序列

<400> 1

Met Gly Gln Thr Val Thr Thr Pro Leu Ser Leu Thr Leu Gly His Trp  
1 5 10 15

Lys Asp Val Glu Arg Ile Ala His Asn Gln Ser Val Asp Val Lys Lys  
20 25 30

Arg Arg Trp Val Thr Phe Cys Ser Ala Glu Trp Pro Thr Phe Asn Val  
35 40 45

Gly Trp Pro Arg Asp Gly Thr Phe Asn Arg Asp Leu Ile Thr Gln Val

50	55	60
Lys Ile Lys Val Phe Ser Pro Gly Pro His Gly His Pro Asp Gln Val		
65	70	75 80
Pro Tyr Ile Val Thr Trp Glu Ala Leu Ala Phe Asp Pro Pro Pro Trp		
	85	90 95
Val Lys Pro Phe Val His Pro Lys Pro Pro Pro Leu Pro Pro Ser		
	100	105 110
Ala Pro Ser Leu Pro Leu Glu Pro Pro Arg Ser Thr Pro Pro Arg Ser		
	115	120 125
Ser Leu Tyr Pro Ala Leu Thr Pro Ser Leu Gly Ala Lys Pro Lys Pro		
	130	135 140
[0002]		
Gln Val Leu Ser Asp Ser Gly Gly Pro Leu Ile Asp Leu Leu Thr Glu		
145	150	155 160
Asp Pro Pro Pro Tyr Arg Asp Pro Arg Pro Pro Pro Ser Asp Arg Asp		
	165	170 175
Gly Asn Gly Gly Glu Ala Thr Pro Ala Gly Glu Ala Pro Asp Pro Ser		
	180	185 190
Pro Met Ala Ser Arg Leu Arg Gly Arg Arg Glu Pro Pro Val Ala Asp		
	195	200 205
Ser Thr Thr Ser Gln Ala Phe Pro Leu Arg Ala Gly Gly Asn Gly Gln		
	210	215 220
Leu Gln Tyr Trp Pro Phe Ser Ser Ser Asp Leu Tyr Asn Trp Lys Asn		
225	230	235 240



Asn Asn Pro Ser Phe Ser Glu Asp Pro Gly Lys Leu Thr Ala Leu Ile  
245 250 255

Glu Ser Val Leu Ile Thr His Gln Pro Thr Trp Asp Asp Cys Gln Gln  
260 265 270

Leu Leu Gly Thr Leu Leu Thr Gly Glu Glu Lys Gln Arg Val Leu Leu  
275 280 285

Glu Ala Arg Lys Ala Val Arg Gly Asp Asp Gly Arg Pro Thr Gln Leu  
290 295 300

Pro Asn Glu Val Asp Ala Ala Phe Pro Leu Glu Arg Pro Asp Trp Asp  
305 310 315 320

[0003]

Tyr Thr Thr Gln Ala Gly Arg Asn His Leu Val His Tyr Arg Gln Leu  
325 330 335

Leu Leu Ala Gly Leu Gln Asn Ala Gly Arg Ser Pro Thr Asn Leu Ala  
340 345 350

Lys Val Lys Gly Ile Thr Gln Gly Pro Asn Glu Ser Pro Ser Ala Phe  
355 360 365

Leu Glu Arg Leu Lys Glu Ala Tyr Arg Arg Tyr Thr Pro Tyr Asp Pro  
370 375 380

Glu Asp Pro Gly Gln Glu Thr Asn Val Ser Met Ser Phe Ile Trp Gln  
385 390 395 400

Ser Ala Pro Asp Ile Gly Arg Lys Leu Glu Arg Leu Glu Asp Leu Lys  
405 410 415

Asn Lys Thr Leu Gly Asp Leu Val Arg Glu Ala Glu Lys Ile Phe Asn  
420 425 430

Lys Arg Glu Thr Pro Glu Glu Arg Glu Glu Arg Ile Arg Arg Glu Thr  
435 440 445

Glu Glu Lys Glu Glu Arg Arg Arg Thr Glu Asp Glu Gln Lys Glu Lys  
450 455 460

Glu Arg Asp Arg Arg Arg His Arg Glu Met Ser Lys Leu Leu Ala Thr  
465 470 475 480

Val Val Ser Gly Gln Lys Gln Asp Arg Gln Gly Gly Glu Arg Arg Arg  
485 490 495

Ser Gln Leu Asp Arg Asp Gln Cys Ala Tyr Cys Lys Glu Lys Gly His  
500 505 510

[0004]

Trp Ala Lys Asp Cys Pro Lys Lys Pro Arg Gly Pro Arg Gly Pro Arg  
515 520 525

Pro Gln Thr Ser Leu Leu Thr Leu Asp Asp  
530 535

&lt;210&gt; 2

&lt;211&gt; 1614

&lt;212&gt; DNA

&lt;213&gt; 鼠白血病病毒

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (1).. (1614)

&lt;223&gt; MMLV Gag 核苷酸序列

&lt;400&gt; 2

	atgggccaga ctgttaccac tcccttaagt ttgaccttag gtcactggaa agatgtcgag	60
	cggatcgcctc acaaccagtc ggtagatgtc aagaagagac gttgggttac cttctgctct	120
	gcagaatggc caacctttaa cgtcggatgg ccgcgagacg gcacctttaa ccgagaccic	180
	atcaccagg ttaagatcaa ggtcttttca cctggccccg atggacaccc agaccaggtc	240
	ccctacatcg tgacctggga agccttggt tttgaccccc ctccctgggt caagcccttt	300
	gtacacccta agcctccgcc tcctcttctt ccctccgcc cgtctctccc ccttgaacct	360
	cctcgttcga ccccgctcg atcctccctt tatccagccc tcactccttc tctaggcgcc	420
	aaacctaaac ctcaagttct ttctgacagt ggggggccgc tcatcgacct acttacagaa	480
	gacccccgc cttataggga cccaagacca ccccttccg acaggacgg aatggtgga	540
	gaagcgacc ctgcgggaga ggcaccggac cctcccca tgcatctcg cctacgtggg	600
	agacgggagc cccctgtggc cgactccact acctcgcagg catteccct ccgcgcagga	660
[0005]	ggaaacggac agcttcaata ctggccgttc tcctcttctg acctttacaa ctggaaaaat	720
	aataaccctt ctttttctga agatccagg aaactgacag ctctgatcga gtctgttctc	780
	atcaccatc agccacctg ggacgactgt cagcagctgt tggggactct gctgaccgga	840
	gaagaaaaac aacgggtgct cttagaggct agaaaggcgg tgcggggcga tgatgggcgc	900
	cccatcaac tgcccaatga agtcgatgcc gcttttccc tcgagcgccc agactgggat	960
	tacaccacc aggaggttag gaaccaccta gtccactatc gccagttgct cctagcgggt	1020
	ctccaaaacg cgggcagaag cccaccaat ttggccaagg taaaaggaat aacacaagg	1080
	cccaatgagt ctccctcggc cttcctagag agacttaagg aagcctatcg caggtaact	1140
	ccttatgacc ctgaggacc agggcaagaa actaatgtgt ctatgtcttt catttgccag	1200
	tctgccccag acattgggag aaagttagag aggttagaag atttaaaaa caagacgtt	1260
	ggagatttgg ttagagaggc agaaaagatc tttaataaac gagaaacccc ggaagaaaga	1320
	gaggaacgta tcaggagaga aacagaggaa aaagaagaac gccgtaggac agaggatgag	1380

	cagaaagaga aagaaagaga tcgtaggaga catagagaga tgagcaagct attggccact	1440
	gtcgttagtg gacagaaaca ggatagacag ggaggagaac gaaggaggtc ccaactcgat	1500
	cgcgaccagt gtgcctactg caaagaaaag gggcactggg ctaaagattg tcccaagaaa	1560
	ccacgaggac ctccggggacc aagaccccag acctccctcc tgaccctaga tgac	1614
	<210> 3	
	<211> 1614	
	<212> DNA	
	<213> 鼠白血病病毒	
	<220>	
	<221> misc_特征	
	<222> (1)..(1614)	
	<223> 密码子优化的 MMLV Gag 核苷酸序列	
	<400> 3	
[0006]	atgggacaga cagtcactac acccctgagc ctgacactgg gacattggaa agacgtggag	60
	aggattgcac ataaccagag cgtggacgtg aagaaacgga gatgggtcac cttttgctcc	120
	gccgagtggc caacattcaa tgtgggatgg ccccgagatg gcaccttcaa ccgggacctg	180
	atcactcagg tgaagatcaa ggtcttctct ccaggacccc acggccatcc agatcaggtg	240
	ccctacatcg tcacctggga ggctctggca tttagacccc ctccatgggt gaagcctttc	300
	gtccacccaa aaccacctcc accactgcct ccatctgccc ctagtctgcc actggaaccc	360
	ctcgggtcaa cccacccag aagctccctg tatcccgcac tgacacctag cctggggggc	420
	aagcctaaac cacaggtgct gtctgatagt ggccggcctc tgatcgatct gctgaccgag	480
	gacctccac cataccgga cccacgacct ccaccaagcg accgggacgg aaacggagga	540
	gaggctacac ccgcaggcga agccccgat cctagtccaa tggcatcaag gctgcgcggg	600
	aggcgcgaac ctccagtggc cgactcaacc acaagccagg catttcact gagggccggg	660
	ggaaatggac agctccagta ttggcccttc tctagttcag atctgtacaa ctggaagaac	720

aataacccta gcttcagcga ggaccacggc aaactgaccg ccctgatcga atccgtgctg	780
attaccacac agcccacatg ggacgattgt cagcagctcc tgggcaccct gctgaccgga	840
gaggaaaagc agagagtgtc gctggaggct aggaaagcag tccgcgggga cgatggaagg	900
ccaacacagc tcccaatga ggtggatgcc gcttccctc tggaacggcc agattgggac	960
tatactaccc aggtggacg caaccacctg gtgcattacc ggcagctcct gctggctgga	1020
ctgcagaatg cagggcgag cccactaac ctggccaagg tgaaaggaat caccagggc	1080
cccaatgagt ccccttctgc attcctggag cggtgaagg aagcctaccg acggtatact	1140
ccctacgac ctgaggaccc aggcaggaa accaacgtga gtatgagctt catctggcag	1200
tccgtcctg acattggccg aaaactggag cggtggaag atctgaaga caagaccctg	1260
ggcgacctgg tgcgggagc agaaaagac ttcaacaaaa gggagactcc agaggaacgg	1320
gaggaaagaa ttagaaggga aacagaggaa aaggaggaa gccgacggac tgaggatgaa	1380
[0007] cagaaggaga aagaaagaga ccggcggcgg caccgggaga tgictaagct gctggccacc	1440
gtggtcagt gccagaaaca ggtcgcagc ggaggagagc gacggagaag ccagctcgat	1500
cgggaccagt gcgcctattg taaggaaaaa gggcattggg ctaaggactg cccaagaaa	1560
cccagaggcc cagcgggccc ccgacctcag acttccctgc tgacctgga cgat	1614

<210> 4

<211> 1099

<212> PRT

<213> 鼠白血病病毒

<220>

<221> MISC\_特征

<222> (1)..(538)

<223> MMLV Gag 氨基酸序列

<220>

<221> MISC\_特征

<222> (539).. (1099)

<223> CMV pp65 氨基酸序列

<400> 4

Met Gly Gln Thr Val Thr Thr Pro Leu Ser Leu Thr Leu Gly His Trp  
1 5 10 15

Lys Asp Val Glu Arg Ile Ala His Asn Gln Ser Val Asp Val Lys Lys  
20 25 30

Arg Arg Trp Val Thr Phe Cys Ser Ala Glu Trp Pro Thr Phe Asn Val  
35 40 45

Gly Trp Pro Arg Asp Gly Thr Phe Asn Arg Asp Leu Ile Thr Gln Val  
50 55 60

[0008] Lys Ile Lys Val Phe Ser Pro Gly Pro His Gly His Pro Asp Gln Val  
65 70 75 80

Pro Tyr Ile Val Thr Trp Glu Ala Leu Ala Phe Asp Pro Pro Pro Trp  
85 90 95

Val Lys Pro Phe Val His Pro Lys Pro Pro Pro Pro Leu Pro Pro Ser  
100 105 110

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

Ser Leu Tyr Pro Ala Leu Thr Pro Ser Leu Gly Ala Lys Pro Lys Pro  
130 135 140

Gln Val Leu Ser Asp Ser Gly Gly Pro Leu Ile Asp Leu Leu Thr Glu  
145 150 155 160

Asp Pro Pro Pro Tyr Arg Asp Pro Arg Pro Pro Pro Ser Asp Arg Asp  
165 170 175

Gly Asn Gly Gly Glu Ala Thr Pro Ala Gly Glu Ala Pro Asp Pro Ser  
180 185 190

Pro Met Ala Ser Arg Leu Arg Gly Arg Arg Glu Pro Pro Val Ala Asp  
195 200 205

Ser Thr Thr Ser Gln Ala Phe Pro Leu Arg Ala Gly Gly Asn Gly Gln  
210 215 220

Leu Gln Tyr Trp Pro Phe Ser Ser Ser Asp Leu Tyr Asn Trp Lys Asn  
225 230 235 240

Asn Asn Pro Ser Phe Ser Glu Asp Pro Gly Lys Leu Thr Ala Leu Ile  
245 250 255

[0009]

Glu Ser Val Leu Ile Thr His Gln Pro Thr Trp Asp Asp Cys Gln Gln  
260 265 270

Leu Leu Gly Thr Leu Leu Thr Gly Glu Glu Lys Gln Arg Val Leu Leu  
275 280 285

Glu Ala Arg Lys Ala Val Arg Gly Asp Asp Gly Arg Pro Thr Gln Leu  
290 295 300

Pro Asn Glu Val Asp Ala Ala Phe Pro Leu Glu Arg Pro Asp Trp Asp  
305 310 315 320

Tyr Thr Thr Gln Ala Gly Arg Asn His Leu Val His Tyr Arg Gln Leu  
325 330 335

Leu Leu Ala Gly Leu Gln Asn Ala Gly Arg Ser Pro Thr Asn Leu Ala

340	345	350
Lys Val Lys Gly Ile Thr Gln Gly Pro Asn Glu Ser Pro Ser Ala Phe		
355	360	365
Leu Glu Arg Leu Lys Glu Ala Tyr Arg Arg Tyr Thr Pro Tyr Asp Pro		
370	375	380
Glu Asp Pro Gly Gln Glu Thr Asn Val Ser Met Ser Phe Ile Trp Gln		
385	390	395 400
Ser Ala Pro Asp Ile Gly Arg Lys Leu Glu Arg Leu Glu Asp Leu Lys		
405	410	415
Asn Lys Thr Leu Gly Asp Leu Val Arg Glu Ala Glu Lys Ile Phe Asn		
420	425	430

[0010]

Lys Arg Glu Thr Pro Glu Glu Arg Glu Glu Arg Ile Arg Arg Glu Thr		
435	440	445
Glu Glu Lys Glu Glu Arg Arg Arg Thr Glu Asp Glu Gln Lys Glu Lys		
450	455	460
Glu Arg Asp Arg Arg Arg His Arg Glu Met Ser Lys Leu Leu Ala Thr		
465	470	475 480
Val Val Ser Gly Gln Lys Gln Asp Arg Gln Gly Gly Glu Arg Arg Arg		
485	490	495
Ser Gln Leu Asp Arg Asp Gln Cys Ala Tyr Cys Lys Glu Lys Gly His		
500	505	510
Trp Ala Lys Asp Cys Pro Lys Lys Pro Arg Gly Pro Arg Gly Pro Arg		
515	520	525



	Pro	Gln	Thr	Ser	Leu	Leu	Thr	Leu	Asp	Asp	Cys	Glu	Ser	Arg	Gly	Arg	
	530						535					540					
	Arg	Cys	Pro	Glu	Met	Ile	Ser	Val	Leu	Gly	Pro	Ile	Ser	Gly	His	Val	
	545					550				555						560	
	Leu	Lys	Ala	Val	Phe	Ser	Arg	Gly	Asp	Thr	Pro	Val	Leu	Pro	His	Glu	
				565					570						575		
	Thr	Arg	Leu	Leu	Gln	Thr	Gly	Ile	His	Val	Arg	Val	Ser	Gln	Pro	Ser	
			580					585							590		
	Leu	Ile	Leu	Val	Ser	Gln	Tyr	Thr	Pro	Asp	Ser	Thr	Pro	Cys	His	Arg	
		595						600					605				
[0011]	Gly	Asp	Asn	Gln	Leu	Gln	Val	Gln	His	Thr	Tyr	Phe	Thr	Gly	Ser	Glu	
	610						615						620				
	Val	Glu	Asn	Val	Ser	Val	Asn	Val	His	Asn	Pro	Thr	Gly	Arg	Ser	Ile	
	625					630					635					640	
	Cys	Pro	Ser	Gln	Glu	Pro	Met	Ser	Ile	Tyr	Val	Tyr	Ala	Leu	Pro	Leu	
				645					650						655		
	Lys	Met	Leu	Asn	Ile	Pro	Ser	Ile	Asn	Val	His	His	Tyr	Pro	Ser	Ala	
			660						665						670		
	Ala	Glu	Arg	Lys	His	Arg	His	Leu	Pro	Val	Ala	Asp	Ala	Val	Ile	His	
		675						680						685			
	Ala	Ser	Gly	Lys	Gln	Met	Trp	Gln	Ala	Arg	Leu	Thr	Val	Ser	Gly	Leu	
		690						695					700				

Ala Trp Thr Arg Gln Gln Asn Gln Trp Lys Glu Pro Asp Val Tyr Tyr  
705 710 715 720

Thr Ser Ala Phe Val Phe Pro Thr Lys Asp Val Ala Leu Arg His Val  
725 730 735

Val Cys Ala His Glu Leu Val Cys Ser Met Glu Asn Thr Arg Ala Thr  
740 745 750

Lys Met Gln Val Ile Gly Asp Gln Tyr Val Lys Val Tyr Leu Glu Ser  
755 760 765

Phe Cys Glu Asp Val Pro Ser Gly Lys Leu Phe Met His Val Thr Leu  
770 775 780

[0012]

Gly Ser Asp Val Glu Glu Asp Leu Thr Met Thr Arg Asn Pro Gln Pro  
785 790 795 800

Phe Met Arg Pro His Glu Arg Asn Gly Phe Thr Val Leu Cys Pro Lys  
805 810 815

Asn Met Ile Ile Lys Pro Gly Lys Ile Ser His Ile Met Leu Asp Val  
820 825 830

Ala Phe Thr Ser His Glu His Phe Gly Leu Leu Cys Pro Lys Ser Ile  
835 840 845

Pro Gly Leu Ser Ile Ser Gly Asn Leu Leu Met Asn Gly Gln Gln Ile  
850 855 860

Phe Leu Glu Val Gln Ala Ile Arg Glu Thr Val Glu Leu Arg Gln Tyr  
865 870 875 880

Asp Pro Val Ala Ala Leu Phe Phe Phe Asp Ile Asp Leu Leu Leu Gln  
885 890 895

Arg Gly Pro Gln Tyr Ser Glu His Pro Thr Phe Thr Ser Gln Tyr Arg  
900 905 910

Ile Gln Gly Lys Leu Glu Tyr Arg His Thr Trp Asp Arg His Asp Glu  
915 920 925

Gly Ala Ala Gln Gly Asp Asp Asp Val Trp Thr Ser Gly Ser Asp Ser  
930 935 940

Asp Glu Glu Leu Val Thr Thr Glu Arg Lys Thr Pro Arg Val Thr Gly  
945 950 955 960

Gly Gly Ala Met Ala Gly Ala Ser Thr Ser Ala Gly Arg Lys Arg Lys  
965 970 975

[0013]

Ser Ala Ser Ser Ala Thr Ala Cys Thr Ala Gly Val Met Thr Arg Gly  
980 985 990

Arg Leu Lys Ala Glu Ser Thr Val Ala Pro Glu Glu Asp Thr Asp Glu  
995 1000 1005

Asp Ser Asp Asn Glu Ile His Asn Pro Ala Val Phe Thr Trp Pro  
1010 1015 1020

Pro Trp Gln Ala Gly Ile Leu Ala Arg Asn Leu Val Pro Met Val  
1025 1030 1035

Ala Thr Val Gln Gly Gln Asn Leu Lys Tyr Gln Glu Phe Phe Trp  
1040 1045 1050

Asp Ala Asn Asp Ile Tyr Arg Ile Phe Ala Glu Leu Glu Gly Val

1055

1060

1065

Trp Gln Pro Ala Ala Gln Pro Lys Arg Arg Arg His Arg Gln Asp  
 1070 1075 1080

Ala Leu Pro Gly Pro Cys Ile Ala Ser Thr Pro Lys Lys His Arg  
 1085 1090 1095

Gly

<210> 5  
 <211> 3300  
 <212> DNA  
 <213> 鼠白血病病毒

[0014]

<220>  
 <221> misc\_特征  
 <222> (1)..(1614)  
 <223> MMLV Gag 核苷酸序列

<220>  
 <221> misc\_特征  
 <222> (1615)..(3300)  
 <223> CMV pp65 核苷酸序列

<400> 5  
 atgggccaga ctgttaccac tcccttaagt ttgaccttag gtcactggaa agatgtcgag 60  
 cggatcgtc acaaccagtc ggtagatgtc aagaagagac gttgggttac cttctgtct 120  
 gcagaatggc caacctttaa cgtcggatgg ccgcgagacg gcacctttaa ccgagacctc 180  
 atcaccagg ttaagatcaa ggtcttttca cctggcccgc atggacaccc agaccaggtc 240  
 ccctacatcg tgacctggga agccttgget ttgaccccc ctcctgggt caagcccttt 300  
 gtacacccta agcctccgcc tcctcttctt ccatccgcc cgtctctccc ccttgaacct 360  
 cctcgttcga ccccgccctg atcctccctt tatccagccc tcactecttc tctaggcgcc 420

	aaacctaaac ctcaagttct ttctgacagt ggggggcccgc tcatcgacct acttacagaa	480
	gacccccccgc cttataggga cccaagacca ccccttccg acagggacgg aaatggtgga	540
	gaagcgaccc ctgcgggaga ggcaccggac cctccccaa tggcatctcg cctacgtggg	600
	agacgggagc cccctgtggc cgactccact acctcgagg catccccct ccgcgcagga	660
	ggaaacggac agctticaata ctggccgttc tctcttctg acctttacaa ctggaaaaat	720
	aataaccctt ctttttctga agatccaggt aaactgacag ctctgatcga gctgtttctc	780
	atcacccatc agcccacctg ggacgactgt cagcagctgt tggggactct gctgaccgga	840
	gaagaaaaac aacgggtgct cttagaggct agaaaggcgg tgcggggcga tgatgggcgc	900
	cccactcaac tgcccaatga agtcgatgcc gcttttccc tcgagcgccc agactgggat	960
	tacaccaccc aggcaggtag gaaccaceta gtccactatc gccagttgct cctagcgggt	1020
	ctccaaaacg cgggcagaag cccaccaat ttggccaagg taaaaggaa aacacaagg	1080
[0015]	cccaatgagt ctcctcggc ctccctagag agacttaagg aagcctatcg caggtacact	1140
	ccttatgacc ctgaggaccc agggcaagaa actaatgtgt ctatgtcttt catittgcag	1200
	tctgccccag acattgggag aaagttagag aggttagaag atttaaaaa caagacgett	1260
	ggagatttgg ttagagaggc agaaaagatc tttataaac gagaaacccc ggaagaaaga	1320
	gaggaacgta tcaggagaga aacagaggaa aaagaagaac gccgtaggac agaggatgag	1380
	cagaaagaga aagaaagaga tcgtaggaga catagagaga tgagcaagct attggccact	1440
	gtcgttagtg gacagaaaca ggatagacag ggaggagaac gaaggaggtc ccaactcgat	1500
	cgcgaccagt gtgcctactg caaagaaaag gggcactggg ctaaagattg tccaagaaa	1560
	ccacgaggac ctcggggacc aagaccccag acctccctcc tgaccctaga tgactgtgag	1620
	tgcgcgggtc gccgttgtcc cgaaatgata tccgtactgg gtccatttc ggggcacgtg	1680
	ctgaaagccg tgtttagtcg cggcgacacg ccggtgtgc cgcacgagac gcgactcctg	1740

	cagacgggta tccacgtgcg cgtgagccag ccctcgtga tcctgggtgc gcagtacacg	1800
	cccgactcga cgccatgcca ccgcggcgac aatcagctgc aggtgcagca cacgtacttt	1860
	acgggcagcg aggtggagaa cgtgtcggtc aacgtgcaca accccacggg cgggagcacc	1920
	tgccccagcc aagagcccat gtcgacttat gtgtacgcgc tgccgctcaa gatgtgaac	1980
	atccccagca tcaacgtgca ccactacccg tcggcgcccg agcgcaaaaca ccgacacctg	2040
	cccgtagctg acgctgtgat tcacgcgtcg ggcaagcaga tgtggcaggc gcgtctcacg	2100
	gtctcgggac tggcctggac gcgtcagcag aaccagtgga aagagcccg cgtctactac	2160
	acgtcagcgt tcgtgtttcc caccaaggac gtggcactgc ggcacgtggt gtgcgcgcac	2220
	gagctgggtt gtcctatgga gaacacgcgc gcaaccaaga tgcaggtgat aggtgaccag	2280
	tacgtcaagg tgtacctgga gtccttctgc gaggacgtgc cctccggcaa gctctttatg	2340
	cacgtcacgc tgggctctga cgtggaagag gacctgacga tgaccgcaa ccgcaaccc	2400
[0016]	ttcatgcgcc cccacgagcg caacggcttt acggtgttgt gtcccaaaaa tatgataatc	2460
	aaaccgggca agatctcgca catcatgctg gatgtggctt ttacctcaca cgagcatttt	2520
	gggctgctgt gtcccaagag catccgggc ctgagcatct caggtaacct attgatgaac	2580
	gggcagcaga tcttcctgga ggtgcaagcg atacgcgaga ccgtggaact gcgtcagtac	2640
	gatcccgtgg ctgcgctctt ctttttcgat atcgacttgc tgctgcagcg cgggcctcag	2700
	tacagcgaac accccacctt caccagccag tatcgcatcc agggcaagct tgagtaccga	2760
	cacacctggg accggcacga cgagggtgcc gccacggcg acgacgacgt ctggaccagc	2820
	ggatcggact ccgacgagga actcgtaacc accgagcgca agacgccccg cgttaccggc	2880
	ggcggcgcca tggcgggcgc ctccacttcc gcgggcgca aacgcaaac agcatcctcg	2940
	gcgacggcgt gcacggcggg cgttatgaca cgcggccgcc ttaaggccga gtccaccgtc	3000
	gcgcccgaag aggacaccga cgaggattcc gacaacgaaa tccacaatcc ggccgtgttc	3060
	acctggccgc cctggcaggc cggcalcctg gcccgcaacc tggtgcccat ggtggctacg	3120

gttcagggtc agaattctgaa gtaccaggag ttctttctggg acgccaacga catctaccgc	3180
atcttcgceg aattggaagg cgtatggcag cccgtgcgc aacccaaacg tcgccgccac	3240
cggcaagacg ccttgcccgg gccatgcacg gcctcgacgc ccaaaaagca ccgaggttag	3300

&lt;210&gt; 6

&lt;211&gt; 3300

&lt;212&gt; DNA

&lt;213&gt; 鼠白血病病毒

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (1)..(1614)

&lt;223&gt; 密码子优化的 MMLV Gag 核苷酸序列

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (1615)..(3300)

[0017] &lt;223&gt; 密码子优化的 CMV pp65 核苷酸序列

&lt;400&gt; 6

atgggacaga cagtcactac acccctgagc ctgacactgg gacattggaa agacgtggag	60
aggattgcac ataaccagag cgtggacgtg aagaaacgga gatgggtcac cttttgetcc	120
gccgagtggc caacattcaa tgtgggatgg ccccgagatg gcacattcaa ccgggacctg	180
atcactcagg tgaagatcaa ggtcttctct ccaggacccc acggccatcc agatcaggtg	240
ccctacatcg tcacctggga ggctctggca tttagacccc ctccatgggt gaagcctttc	300
gtccacccaa aaccacctcc accactgcct ccatctgcc ctagctgcc actggaaccc	360
cctcgggtcaa ccccaccag aagctccctg tatccgcac tgacacctag cctgggggcc	420
aagcctaaac cacaggtgct gtctgatagt ggccggccctc tgatcgatct gctgaccgag	480
gacctccac cataccgcga cccacgacct ccaccaagcg accgggacgg aaacggagga	540
gaggctacac ccgcaggcga agccccgat cctagtccaa tggcatcaag gctgcgcggg	600

	aggcgcgaac ctccagtggc cgactcaacc acaagccagg catttcact gagggccggg	660
	ggaaatggac agctccagta ttggcccttc tctagttcag atctgtacaa ctggaagaac	720
	aataacceta gcttcagcga ggacccaggc aaactgaccg ccttgatcga atccgtgctg	780
	attaccacac agcccacatg ggacgattgt cagcagctcc tgggcaccct gctgaccgga	840
	gaggaaaagc agagagtgtc gctggaggct aggaaagcag tccgcgggga cgatggaagg	900
	ccaacacagc tccccaatga ggtggatgcc gctttccctc tggaacggcc agattgggac	960
	tatactaccc aggctggacg caaccacctg gtgcattacc ggcagctcct gctggctgga	1020
	ctgcagaatg cagggcgcag cccactaac ctggccaagg tgaaaggaat caccaggggc	1080
	cccaatgagt ccccttctgc attcctggag cggctgaagg aagcctaccg acggtatact	1140
	ccctacgacg ctgaggaccc aggccaggaa accaacgtga gtatgagctt catctggcag	1200
	tccgctcctg acattggccg aaaactggag cggctggaag atctgaagaa caagaccctg	1260
[0018]	ggcgacctgg tgcgggaggc agaaaagatc ttcaacaaaa gggagactcc agaggaacgg	1320
	gaggaaagaa ttagaaggga aacagaggaa aaggaggaac gccgacggac tgaggatgaa	1380
	cagaaggaga aagaaagaga ccggcggcgg caccgggaga tgtctaagct gctggccacc	1440
	gtggtcagtg gccagaaaca ggatcgacag ggaggagagc gacggagaag ccagctcgat	1500
	cgggaccagt gcgcctattg taaggaaaaa gggcattggg ctaaggactg cccaagaaa	1560
	cccagaggcc caccggggcc ccgacctcag acttccttgc tgacctgga cgattgcgag	1620
	agccggggcc ggccgtgccc agaaatgac tctgtgttgg ggcccattag tggacatgtg	1680
	ctgaaggccg tctttctcag gggagacacc cccgtgtctc ctcacgagac tcgactgctg	1740
	cagaccggca tccatgtgcg ggtctcccag cctctcttga ttctgggtgc acagtataca	1800
	ccagatagca ctccctgcc aagaggagac aatcagctcc aggtgcagca tacctacttt	1860
	acaggctccg aggtcgaaaa cgtgtctgtc aatgtgcaca accctaccgg caggagcatc	1920
	tglcctagcc aggagccaat gagcatctac gtgtacgccc tgcctctgaa gatgtgaa	1980



	atcccatcaa ttaacgtcca ccattaccct agcgcagccg aacggaagca cagacatctg	2040
	ccagtggccg acgctgtcat ccatgccagc ggcaaacaga tgtggcaggc aagactgacc	2100
	gtgtccgggc tggcctggac aaggcagcag aatcagtgga aggagcccg cgtgtactat	2160
	accagcgctt tcgtgttccc taccaaagac gtggccctga gacatgtggt gtgcgcacat	2220
	gagctgggtgt gcagcatgga aaacactagg gccaccaaga tgcaggtcac cggcgatcag	2280
	tatgtcaaag tgtacctgga gagtttttgc gaagacgtgc catcagggaa gctgttcattg	2340
	catgtgaccc tgggcagcga tgtcgaggaa gacctgacca tgacaagaaa tccacagccc	2400
	tttatgagac cccacgagag gaatgggttc actgtgtctg gcccgaagaa catgatcatt	2460
	aagcctggaa aaatcagtca tattatgctg gatgtggcct ttacatcaca cgagcatttc	2520
	ggactgctgt gccccaaatc catccctgga ctgagcattt ccggcaatct gctgatgaac	2580
	ggccagcaga tcttcctgga agtgcaggcc atccgggaga ccgtcgaact gcgacagtat	2640
[0019]	gacccagtgg ctgcactgtt ctttttcgac atcgacctgc tgcctgcagc aggaccacag	2700
	tacagcgagc accctacttt tacctcccag tatcggatic aggggaagct ggagtacagg	2760
	cacacctggg atcgccatga cgaaggagcc gctcaggggg acgatgacgt gtggacatct	2820
	ggcagtgatt cagacgagga actggtgaca actgagcgaa aaaccccccg ggtgacagga	2880
	ggagggggcaa lggcaggggc cagcaccagc gcagggcgga agcgaaaaag cgccagcagc	2940
	gccacagcat gtaccgccgg cgtgatgact agaggaaggc tgaaggccga gtctacagtc	3000
	gtccccgagg aagatactga cgaggatagt gacaatgaaa tccacaaccc cgccgtgttc	3060
	acctggccac cttggcagc agggattctg gctcgcaacc tggcctccat ggtggcaacc	3120
	gtccagggac agaattctgaa gtatcaggag tttttctggg atgctaacga catctaccgg	3180
	atttttgcag agctggaagg cgtgtggcag ccagcagccc agcccaaacg acggagacat	3240
	cgacaggacg ctcgccagg acctgtatc gccagcacac caaagaagca caggggctaa	3300

<210> 7  
<211> 907  
<212> PRT  
<213> 人巨细胞病毒

<220>  
<221> MISC\_特征  
<222> (1).. (907)  
<223> HCMV gB 氨基酸序列

<220>  
<221> MISC\_特征  
<222> (751).. (907)  
<223> TM 和 CD

<400> 7

Met Glu Ser Arg Ile Trp Cys Leu Val Val Cys Val Asn Leu Cys Ile  
1 5 10 15

[0020]

Val Cys Leu Gly Ala Ala Val Ser Ser Ser Thr Arg Gly Thr Ser  
20 25 30

Ala Thr His Ser His His Ser Ser His Thr Thr Ser Ala Ala His Ser  
35 40 45

Arg Ser Gly Ser Val Ser Gln Arg Val Thr Ser Ser Gln Thr Val Ser  
50 55 60

His Gly Val Asn Glu Thr Ile Tyr Asn Thr Thr Leu Lys Tyr Gly Asp  
65 70 75 80

Val Val Gly Val Asn Thr Thr Lys Tyr Pro Tyr Arg Val Cys Ser Met  
85 90 95

Ala Gln Gly Thr Asp Leu Ile Arg Phe Glu Arg Asn Ile Val Cys Thr  
100 105 110

Ser Met Lys Pro Ile Asn Glu Asp Leu Asp Glu Gly Ile Met Val Val  
 115 120 125

Tyr Lys Arg Asn Ile Val Ala His Thr Phe Lys Val Arg Val Tyr Gln  
 130 135 140

Lys Val Leu Thr Phe Arg Arg Ser Tyr Ala Tyr Ile His Thr Thr Tyr  
 145 150 155 160

Leu Leu Gly Ser Asn Thr Glu Tyr Val Ala Pro Pro Met Trp Glu Ile  
 165 170 175

His His Ile Asn Ser His Ser Gln Cys Tyr Ser Ser Tyr Ser Arg Val  
 180 185 190

[0021]

Ile Ala Gly Thr Val Phe Val Ala Tyr His Arg Asp Ser Tyr Glu Asn  
 195 200 205

Lys Thr Met Gln Leu Met Pro Asp Asp Tyr Ser Asn Thr His Ser Thr  
 210 215 220

Arg Tyr Val Thr Val Lys Asp Gln Trp His Ser Arg Gly Ser Thr Trp  
 225 230 235 240

Leu Tyr Arg Glu Thr Cys Asn Leu Asn Cys Met Val Thr Ile Thr Thr  
 245 250 255

Ala Arg Ser Lys Tyr Pro Tyr His Phe Phe Ala Thr Ser Thr Gly Asp  
 260 265 270

Val Val Asp Ile Ser Pro Phe Tyr Asn Gly Thr Asn Arg Asn Ala Ser  
 275 280 285

Tyr Phe Gly Glu Asn Ala Asp Lys Phe Phe Ile Phe Pro Asn Tyr Thr  
 290 295 300

Ile Val Ser Asp Phe Gly Arg Pro Asn Ser Ala Leu Glu Thr His Arg  
 305 310 315 320

Leu Val Ala Phe Leu Glu Arg Ala Asp Ser Val Ile Ser Trp Asp Ile  
 325 330 335

Gln Asp Glu Lys Asn Val Thr Cys Gln Leu Thr Phe Trp Glu Ala Ser  
 340 345 350

Glu Arg Thr Ile Arg Ser Glu Ala Glu Asp Ser Tyr His Phe Ser Ser  
 355 360 365

[0022]

Ala Lys Met Thr Ala Thr Phe Leu Ser Lys Lys Gln Glu Val Asn Met  
 370 375 380

Ser Asp Ser Ala Leu Asp Cys Val Arg Asp Glu Ala Ile Asn Lys Leu  
 385 390 395 400

Gln Gln Ile Phe Asn Thr Ser Tyr Asn Gln Thr Tyr Glu Lys Tyr Gly  
 405 410 415

Asn Val Ser Val Phe Glu Thr Thr Gly Gly Leu Val Val Phe Trp Gln  
 420 425 430

Gly Ile Lys Gln Lys Ser Leu Val Glu Leu Glu Arg Leu Ala Asn Arg  
 435 440 445

Ser Ser Leu Asn Leu Thr His Asn Arg Thr Lys Arg Ser Thr Asp Gly  
 450 455 460

Asn Asn Ala Thr His Leu Ser Asn Met Glu Ser Val His Asn Leu Val  
465 470 475 480

Tyr Ala Gln Leu Gln Phe Thr Tyr Asp Thr Leu Arg Gly Tyr Ile Asn  
485 490 495

Arg Ala Leu Ala Gln Ile Ala Glu Ala Trp Cys Val Asp Gln Arg Arg  
500 505 510

Thr Leu Glu Val Phe Lys Glu Leu Ser Lys Ile Asn Pro Ser Ala Ile  
515 520 525

Leu Ser Ala Ile Tyr Asn Lys Pro Ile Ala Ala Arg Phe Met Gly Asp  
530 535 540

Val Leu Gly Leu Ala Ser Cys Val Thr Ile Asn Gln Thr Ser Val Lys  
545 550 555 560

[0023]

Val Leu Arg Asp Met Asn Val Lys Glu Ser Pro Gly Arg Cys Tyr Ser  
565 570 575

Arg Pro Val Val Ile Phe Asn Phe Ala Asn Ser Ser Tyr Val Gln Tyr  
580 585 590

Gly Gln Leu Gly Glu Asp Asn Glu Ile Leu Leu Gly Asn His Arg Thr  
595 600 605

Glu Glu Cys Gln Leu Pro Ser Leu Lys Ile Phe Ile Ala Gly Asn Ser  
610 615 620

Ala Tyr Glu Tyr Val Asp Tyr Leu Phe Lys Arg Met Ile Asp Leu Ser  
625 630 635 640

Ser Ile Ser Thr Val Asp Ser Met Ile Ala Leu Asp Ile Asp Pro Leu

645

650

655

Glu Asn Thr Asp Phe Arg Val Leu Glu Leu Tyr Ser Gln Lys Glu Leu  
660 665 670

Arg Ser Ile Asn Val Phe Asp Leu Glu Glu Ile Met Arg Glu Phe Asn  
675 680 685

Ser Tyr Lys Gln Arg Val Lys Tyr Val Glu Asp Lys Val Val Asp Pro  
690 695 700

Leu Pro Pro Tyr Leu Lys Gly Leu Asp Asp Leu Met Ser Gly Leu Gly  
705 710 715 720

Ala Ala Gly Lys Ala Val Gly Val Ala Ile Gly Ala Val Gly Gly Ala  
725 730 735

[0024]

Val Ala Ser Val Val Glu Gly Val Ala Thr Phe Leu Lys Asn Pro Phe  
740 745 750

Gly Ala Phe Thr Ile Ile Leu Val Ala Ile Ala Val Val Ile Ile Ile  
755 760 765

Tyr Leu Ile Tyr Thr Arg Gln Arg Arg Leu Cys Met Gln Pro Leu Gln  
770 775 780

Asn Leu Phe Pro Tyr Leu Val Ser Ala Asp Gly Thr Thr Val Thr Ser  
785 790 795 800

Gly Asn Thr Lys Asp Thr Ser Leu Gln Ala Pro Pro Ser Tyr Glu Glu  
805 810 815

Ser Val Tyr Asn Ser Gly Arg Lys Gly Pro Gly Pro Pro Ser Ser Asp  
820 825 830

Ala Ser Thr Ala Ala Pro Pro Tyr Thr Asn Glu Gln Ala Tyr Gln Met  
835 840 845

Leu Leu Ala Leu Val Arg Leu Asp Ala Glu Gln Arg Ala Gln Gln Asn  
850 855 860

Gly Thr Asp Ser Leu Asp Gly Gln Thr Gly Thr Gln Asp Lys Gly Gln  
865 870 875 880

Lys Pro Asn Leu Leu Asp Arg Leu Arg His Arg Lys Asn Gly Tyr Arg  
885 890 895

His Leu Lys Asp Ser Asp Glu Glu Glu Asn Val  
900 905

[0025]

<210> 8  
<211> 2724  
<212> DNA  
<213> 人巨细胞病毒

<220>  
<221> misc\_特征  
<222> (1).. (2724)  
<223> HCMV gB 核苷酸序列

<220>  
<221> misc\_特征  
<222> (2253).. (2724)  
<223> TM 和 CD

<400> 8  
atggaatcca ggatctgggtg cctggtagtc tgcgttaact tgtgtatcgt ctgtctgggt 60  
gctgcgggttt cctcatcttc tactcgtgga acttctgcta ctcacagtc ccatctctct 120  
catacgacgt ctgctgctca ttctcgatcc gggttcagtct ctcaacgcgt aacttcttcc 180

	caaacgggtca gccatggtgt taacgagacc atctacaaca ctacctcaa gtacggagat	240
	gtggtggggg tcaacaccac caagtacccc tatcgctgt gttctatggc acagggtacg	300
	gatcittatc gctttgaacg laatatcgtc tgcacctcga tgaagcccat caatgaagac	360
	ctggacgagg gcacatggt ggtctacaaa cgcaacatcg tcgcgcacac cttaaggtta	420
	cgagtctacc agaaggtttt gacgtttcgt cgtagctacg cttacatcca caccattat	480
	ctgctgggca gcaacacgga atacgtggcg cctcctatgt gggagattca tcatatcaac	540
	agtcacagtc agtgctacag ttcctacagc cgcgttatag caggcacggt tttcgtggct	600
	tatcataggg acagctatga aaacaaaacc atgcaattaa tgcccacga ttattccaac	660
	accacagta cccgttacgt gacggtaag gatcaatggc acagccgagg cagcacctgg	720
	ctctatcgtg agacctgtaa tctgaattgt atggtgacca tcaactactgc gcgtccaag	780
	tatccctate attttttcgc aacttcacg ggtgatgtgg ttgacatttc tcctttctac	840
[0026]	aacggaacta atcgcaatgc cagctatitt ggagaaaacg ccgacaagtt tttcattttt	900
	ccgaactaca ctatcgtctc cgactttgga agaccgaatt ctgcgttaga gaccacagg	960
	ttggtggctt ttcttgaacg tgcggactca gtgatctcct gggatatata ggacgagaag	1020
	aatgttactt gtcaactcac tttctgggaa gcctcggaac gcaccattcg ttccgaagcc	1080
	gaggactcgt atcaactttc ttctgcaaaa atgaccgcca ctttcttate taagaagcaa	1140
	gaggtgaaca tgtccgactc tgcgtggac tgtgtacgtg atgaggccat aaataagtta	1200
	cagcagattt tcaatacttc atacaatcaa acatatgaaa aatatggaaa cgtgtccgtc	1260
	tttgaacca ctggtggttt ggtggtgttc tggcaaggta tcaagcaaaa atctctggtg	1320
	gaactcgaac gtttggccaa ccgtccagt ctgaatctta ctcataatag aacaaaaaga	1380
	agtacagatg gcaacaatgc aactcattta tccaacatgg agtcggtgca caatctggtc	1440
	taccccagc tgcagttcac ctatgacacg ttgcgcggtt acatcaaccg ggcgctggcg	1500
	caaatcgcag aagcciggtg tglggatcaa cggcgcaccc tagaggcttt caaggaactt	1560



	agcaagatca acccgtcagc tattctctcg gccatctaca acaaaccgat tgccgcgcgt	1620
	ttcatgggtg atgtcctggg tctggccagc tgcgtgacca ttaaccaaac cagcgtcaag	1680
	gtgtcgcgtg atatgaatgt gaaggaatcg ccaggacgct gctactcacg accagtggtc	1740
	atctttaatt tcgccaacag ctctgtacgtg cagtacggtc aactgggcga ggataacgaa	1800
	atctgtttgg gcaaccaccg cactgaggaa tgtcagcttc ccagcctcaa gatcttcac	1860
	gccggcaact cggcctacga gtacgtggac tacctcttca aacgcattgat tgacctcagc	1920
	agcatctcca ccgtcgacag catgatcgcc ctagacatcg acccgctgga aaacaccgac	1980
	ttcagggtac tggaacttta ctgcagaaa gaattgcgtt ccatcaacgt ttttgatctc	2040
	gaggagatca tgcgcgagtt caattcgtat aagcagcggg taaagtacgt ggaggacaag	2100
	gtagtgcacc cgtgcgcc ctacctcaag ggtctggacg acctcatgag cggcctgggc	2160
	gccgcgggaa aggccgttgg cgtagccatt ggggccgtgg gtggcgcgtt ggcctccgtg	2220
[0027]	gtcgaaggcg ttgccacctt cctcaaaaac cccttcggag ccttcacat catcctcgtg	2280
	gccatagccg tcgtcattat catttatttg atciatactc gacagcggcg tctctgcatg	2340
	cagccgtgc agaacctctt tccctatctg gtgtccgccg acgggaccac cgtgacgtcg	2400
	ggcaacacca aagacacgtc gttacaggct ccgccttctt acgaggaaag tgtttataat	2460
	tctggtcgca aaggaccggg accaccgtcg tcigtatcat ccacggcggc tccgccttac	2520
	accaacgagc aggcctacca gatgcctctg gccctggtcc gtctggacgc agagcagcga	2580
	gcgcagcaga acggtacaga ttctttggac ggacagactg gcacgcagga caaggacag	2640
	aagcccaacc tgctagaccg actgcgacac cgcaaaaacg gctaccgaca cttgaaagac	2700
	tccgacgaag aagagaacgt ctga	2724

&lt;210&gt; 9

&lt;211&gt; 2724

&lt;212&gt; DNA

## &lt;213&gt; 人巨细胞病毒

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (1).. (2724)

&lt;223&gt; 密码子优化的 HCMV gB 核苷酸序列

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (2253).. (2724)

&lt;223&gt; TM 和 CD

&lt;400&gt; 9

	atggagtcaa ggatttgggtg cctggtcgtg tgcgtcaatc tgtgcatcgt ctgtctgggg	60
	gctgccgtgt catcaagttc tacaagaggc accagcgcca cccactcaca ccatagctcc	120
	cataccacat ccgccgctca ctcccggtct ggcagcgtga gccagagagt cacatctagt	180
	cagaccgtga gccacggggt caacgagacc atctacaata ctaccctgaa gtatggcgac	240
[0028]	gtggtcgggg tgaacacaac taaataccca tatagggtct gcagtatggc ccagggcact	300
	gatctgattia gattcgaaag gaacatcgtg tgcaccagca tgaagcccat taatgaggac	360
	ctggatgaag ggatcatggt ggtctacaaa cgcaatattg tggcccatat cttcaagggtg	420
	cgagtctatc agaaagtgtc gacatttcgg agatcttacg catatatcca caccacatac	480
	ctgctgggga gtaacaccga gtatgtggct cccctatgt gggaaattca ccatatcaat	540
	agccattccc agtctactc aagctacagc agagtgatcg ctggaacagt gttegtcgca	600
	taccacagag acitcttatga gaacaagact atgcagctca tgcgccagca ttacagcaat	660
	acacattcca ctagatatgt gacagtcaaa gatcagtggt actcaagggg cagcacctgg	720
	ctgtaccgag agacatgcaa cctgaattgt atggtgacta tcactaccgc tagatccaag	780
	taccctatc acttctttgc aattccacc ggggacgtgg tcgatatttc tctttctac	840
	aacggcacaa accggaatgc atcttatttt ggggagaacg ccgacaagtt ctttattttc	900
	ccaaattaca ccatcgtgic tgatlltggc agaccaaca gtgccctgga gacacatcga	960

	ctggtggcat tcttggaaacg ggccgactcc gtcattttctt gggacatcca ggatgagaag	1020
	aatgtgacct gccagctcac cttctgggag gccagcgaac gcaccatccg atccgaggct	1080
	gaagattctt accacttctc ctctgccaaa atgacagcta cttttctgag caagaaacag	1140
	gaggigaaca tgtctgacag tgctctggat tgcgtgcggg acgaagcaat taataagctg	1200
	cagcagatct tcaacacatc atacaaccag acttacgaga agtacggaaa cgtgagcgtc	1260
	ttcgaaacaa ctggcgggct ggtggtcttt tggcagggca tcaagcagaa atccctggtg	1320
	gagctggaaa ggctggccaa tcgcagtcca ctgaacctga ctcataatcg gaccaagaga	1380
	tctacagacg gaaacaatgc cacacatctg tctaactgg agagtgtgca caatctggtc	1440
	tacgctcagc tccagtttac ctacgacaca ctgagaggct atattaacag ggcaactggc	1500
	cagatcgctg aagcatggtg cgtggatcag aggcgcaccc tggaggtctt caaggaactg	1560
	tcctaaatca acccttcagc aattctgagc gccatctaca ataagccaat tgcagccagg	1620
[0029]	tttatgggag acgtgctggg cctggccagt tgcgtcacta tcaaccagac ctgagtgaag	1680
	gtccigcgcg atatgaatgt gaaagagagt cccggcagat gctattcacg gccigtggtc	1740
	atcttcaact ttgctaatag ctctacgtg cagtatggac agctcggcga ggacaacgaa	1800
	attctgctgg ggaatcacag gaccgaggaa tgtcagctcc ctagectgaa gattttcac	1860
	gctggaaact ccgcatacga gtaigtggat tacctgttca agcggatgat tgacctgtct	1920
	agtatctcca ctgtggattc tatgattgcc ctggacatcg atccactgga aaataccgac	1980
	ttcagggtgc tggagctgta tagccagaag gaactgcgt ccatcaacgt gtctgatctg	2040
	gaggaaatta tgagagagtt taatagctac aagcagaggg tgaaatatgt cgaagataag	2100
	gtggtcgacc ccctgccacc ctacctgaaa ggccctggacg atctgatgag cgggctggga	2160
	gctgcaggga aggcagtggt agtcgtatc ggccgagtgg gaggagccgt ggccagcgtg	2220
	gtcaggggag tggcaacatt cctgaaaaac cccttcgggg ccttcacat cattctggtg	2280

```

gcaatcgccg tggatcatcat tatctacctg atctacacaa ggcagcgccg gctgtgcatg      2340
cagcctctgc agaacctggt tccataacctg gtgagcgccg acgggaccac agtcacctca      2400
ggaaataacta aggataacctc tctgcaggcc cccccaagtt acgaggaatc agtgtataac      2460
agcggcagaa aaggaccagg accaccttca agcgacgcca gcactgccgc tccaccctac      2520
accaatgagc aggccatca gatgctgctg gctctggtgc gcctggatgc cgaacagcga      2580
gctcagcaga acgggaccga ctccctggat ggacagaccg gaacacagga caaggacag      2640
aaacctaatc tgctggatcg gctgcggcac agaaaaaacg ggtataggca cctgaaggac      2700
tccgacgaag aagaaaatgt ctaa                                             2724

```

<210> 10  
 <211> 795  
 <212> PRT  
 <213> 人巨细胞病毒

[0030]

<220>  
 <221> MISC\_特征  
 <222> (1)..(795)  
 <223> HCMV gB-G 氨基酸序列

<220>  
 <221> MISC\_特征  
 <222> (751)..(795)  
 <223> TM 和 CTD

<400> 10

```

Met Glu Ser Arg Ile Trp Cys Leu Val Val Cys Val Asn Leu Cys Ile
1             5             10             15

```

```

Val Cys Leu Gly Ala Ala Val Ser Ser Ser Thr Arg Gly Thr Ser
                20             25             30

```

```

Ala Thr His Ser His His Ser Ser His Thr Thr Ser Ala Ala His Ser
35             40             45

```

	Arg	Ser	Gly	Ser	Val	Ser	Gln	Arg	Val	Thr	Ser	Ser	Gln	Thr	Val	Ser
	50						55					60				
	His	Gly	Val	Asn	Glu	Thr	Ile	Tyr	Asn	Thr	Thr	Leu	Lys	Tyr	Gly	Asp
	65				70					75					80	
	Val	Val	Gly	Val	Asn	Thr	Thr	Lys	Tyr	Pro	Tyr	Arg	Val	Cys	Ser	Met
				85					90					95		
	Ala	Gln	Gly	Thr	Asp	Leu	Ile	Arg	Phe	Glu	Arg	Asn	Ile	Val	Cys	Thr
				100					105					110		
	Ser	Met	Lys	Pro	Ile	Asn	Glu	Asp	Leu	Asp	Glu	Gly	Ile	Met	Val	Val
		115						120					125			
[0031]	Tyr	Lys	Arg	Asn	Ile	Val	Ala	His	Thr	Phe	Lys	Val	Arg	Val	Tyr	Gln
		130						135					140			
	Lys	Val	Leu	Thr	Phe	Arg	Arg	Ser	Tyr	Ala	Tyr	Ile	His	Thr	Thr	Tyr
	145					150					155				160	
	Leu	Leu	Gly	Ser	Asn	Thr	Glu	Tyr	Val	Ala	Pro	Pro	Met	Trp	Glu	Ile
				165					170					175		
	His	His	Ile	Asn	Ser	His	Ser	Gln	Cys	Tyr	Ser	Ser	Tyr	Ser	Arg	Val
				180					185					190		
	Ile	Ala	Gly	Thr	Val	Phe	Val	Ala	Tyr	His	Arg	Asp	Ser	Tyr	Glu	Asn
		195						200					205			
	Lys	Thr	Met	Gln	Leu	Met	Pro	Asp	Asp	Tyr	Ser	Asn	Thr	His	Ser	Thr
		210						215					220			

Arg Tyr Val Thr Val Lys Asp Gln Trp His Ser Arg Gly Ser Thr Trp  
225 230 235 240

Leu Tyr Arg Glu Thr Cys Asn Leu Asn Cys Met Val Thr Ile Thr Thr  
245 250 255

Ala Arg Ser Lys Tyr Pro Tyr His Phe Phe Ala Thr Ser Thr Gly Asp  
260 265 270

Val Val Asp Ile Ser Pro Phe Tyr Asn Gly Thr Asn Arg Asn Ala Ser  
275 280 285

Tyr Phe Gly Glu Asn Ala Asp Lys Phe Phe Ile Phe Pro Asn Tyr Thr  
290 295 300

[0032]

Ile Val Ser Asp Phe Gly Arg Pro Asn Ser Ala Leu Glu Thr His Arg  
305 310 315 320

Leu Val Ala Phe Leu Glu Arg Ala Asp Ser Val Ile Ser Trp Asp Ile  
325 330 335

Gln Asp Glu Lys Asn Val Thr Cys Gln Leu Thr Phe Trp Glu Ala Ser  
340 345 350

Glu Arg Thr Ile Arg Ser Glu Ala Glu Asp Ser Tyr His Phe Ser Ser  
355 360 365

Ala Lys Met Thr Ala Thr Phe Leu Ser Lys Lys Gln Glu Val Asn Met  
370 375 380

Ser Asp Ser Ala Leu Asp Cys Val Arg Asp Glu Ala Ile Asn Lys Leu  
385 390 395 400

Gln Gln Ile Phe Asn Thr Ser Tyr Asn Gln Thr Tyr Glu Lys Tyr Gly  
405 410 415

Asn Val Ser Val Phe Glu Thr Thr Gly Gly Leu Val Val Phe Trp Gln  
420 425 430

Gly Ile Lys Gln Lys Ser Leu Val Glu Leu Glu Arg Leu Ala Asn Arg  
435 440 445

Ser Ser Leu Asn Leu Thr His Asn Arg Thr Lys Arg Ser Thr Asp Gly  
450 455 460

Asn Asn Ala Thr His Leu Ser Asn Met Glu Ser Val His Asn Leu Val  
465 470 475 480

Tyr Ala Gln Leu Gln Phe Thr Tyr Asp Thr Leu Arg Gly Tyr Ile Asn  
485 490 495

[0033]

Arg Ala Leu Ala Gln Ile Ala Glu Ala Trp Cys Val Asp Gln Arg Arg  
500 505 510

Thr Leu Glu Val Phe Lys Glu Leu Ser Lys Ile Asn Pro Ser Ala Ile  
515 520 525

Leu Ser Ala Ile Tyr Asn Lys Pro Ile Ala Ala Arg Phe Met Gly Asp  
530 535 540

Val Leu Gly Leu Ala Ser Cys Val Thr Ile Asn Gln Thr Ser Val Lys  
545 550 555 560

Val Leu Arg Asp Met Asn Val Lys Glu Ser Pro Gly Arg Cys Tyr Ser  
565 570 575

Arg Pro Val Val Ile Phe Asn Phe Ala Asn Ser Ser Tyr Val Gln Tyr

580	585	590
Gly Gln Leu Gly Glu Asp Asn Glu Ile Leu Leu Gly Asn His Arg Thr		
595	600	605
Glu Glu Cys Gln Leu Pro Ser Leu Lys Ile Phe Ile Ala Gly Asn Ser		
610	615	620
Ala Tyr Glu Tyr Val Asp Tyr Leu Phe Lys Arg Met Ile Asp Leu Ser		
625	630	635 640
Ser Ile Ser Thr Val Asp Ser Met Ile Ala Leu Asp Ile Asp Pro Leu		
645	650	655
Glu Asn Thr Asp Phe Arg Val Leu Glu Leu Tyr Ser Gln Lys Glu Leu		
660	665	670
[0034]		
Arg Ser Ile Asn Val Phe Asp Leu Glu Glu Ile Met Arg Glu Phe Asn		
675	680	685
Ser Tyr Lys Gln Arg Val Lys Tyr Val Glu Asp Lys Val Val Asp Pro		
690	695	700
Leu Pro Pro Tyr Leu Lys Gly Leu Asp Asp Leu Met Ser Gly Leu Gly		
705	710	715 720
Ala Ala Gly Lys Ala Val Gly Val Ala Ile Gly Ala Val Gly Gly Ala		
725	730	735
Val Ala Ser Val Val Glu Gly Val Ala Thr Phe Leu Lys Asn Pro Phe		
740	745	750
Phe Phe Ile Ile Gly Leu Ile Ile Gly Leu Phe Leu Val Leu Arg Val		
755	760	765



Gly Ile His Leu Cys Ile Lys Leu Lys His Thr Lys Lys Arg Gln Ile  
 770 775 780

Tyr Thr Asp Ile Glu Met Asn Arg Leu Gly Lys  
 785 790 795

<210> 11

<211> 2388

<212> DNA

<213> 人巨细胞病毒

<220>

<221> misc\_特征

<222> (1).. (2388)

<223> HCMV gB - G 核苷酸序列

<220>

<221> misc\_特征

<222> (2253).. (2388)

<223> TM 和 CTD

[0035]

<400> 11

atggaatcca ggatctgggtg cctggtagtc tgcgttaact tgtgtatcgt ctgtctgggt 60

gctgcgggttt cctcatcttc tactcgtgga acttctgcta ctcacagtca ccattcctct 120

catacgacgt ctgctgctca ttctcgatcc gggtcagtct ctcaacgcgt aacttcttcc 180

caaacgggtca gccatgggtgt taacgagacc atctacaaca ctaccctcaa gtacggagat 240

gtggtggggg tcaacaccac caagtacccc tatcgctgtgt gttctatggc acagggtacg 300

gatcttattc gctttgaacg taatatcgtc tgcacctega tgaagcccat caatgaagac 360

ctggacgagg gcatcatggt ggtctacaaa cgcaacatcg tcgcgcacac cttaaggtta 420

cgagtcctacc agaaggtttt gacgtttcgt cgtagctacg cttacatcca caccacttat 480

ctgctgggca gcaacacgga atacgtggcg cctcciatgt gggagattca tcatatcaac 540

	agtcacagtc agtgcctacag ttcctacagc cgcgttatag caggcacggt ttctgtggct	600
	tatcataggg acagctatga aaacaaaacc atgcaattaa tgcccgaaga ttattccaac	660
	accacacagta cccgttacgt gacggtaag gatcaatggc acagccgagg cagcacctgg	720
	ctctatcgtg agacctgtaa tctgaattgt atggtagaca tcaactactgc gcgtccaag	780
	tatccctatc attttttgcg aacttcacg ggtgatgtgg ttgacatttc tcctttctac	840
	aacggaacta atcgcaatgc cagctatatt ggagaaaacg ccgacaagtt ttcatatttt	900
	ccgaactaca ctatcgtctc cgactttgga agaccgaatt ctgcgttaga gaccacagg	960
	ttggtagctt ttcttgaacg tgcggactca gtgatctcct gggatataca ggacgagaag	1020
	aatgttactt gtcaactcac tttctgggaa gcctcggaac gcaccattcg ttccgaagcc	1080
	gaggactcgt atcaactttc ttctgcaaaa atgaccgcca ctttcttata taagaagcaa	1140
	gaggtgaaca tgtccgactc tgcgtggac tgtgtacgtg atgaggccat aaataagtta	1200
[0036]	cagcagattt tcaatacttc atacaatcaa acatatgaaa aatatggaaa cgtgtccgtc	1260
	tttgaacca ctgggtggtt ggtgggttgc tggcaaggta tcaagcaaaa atctctggtg	1320
	gaactcgaac gtttggccaa ccgtccagt ctgaatctta ctcataatag aacaaaaga	1380
	agtacagatg gcaacaatgc aactcattta tccaacatgg agtcgggtgca caatctggtc	1440
	tacgcccage tgcagttcac ctatgacacg ttgcgagggt acatcaaccg ggcgtggcg	1500
	caaategcag aagcctgggtg tgtggatcaa cggcgacccc tagaggtctt caaggaactt	1560
	agcaagatca acccgtcage tattctctcg gccatctaca acaaaccgat tgcgcgcgt	1620
	ttcatgggtg atgtcctggg tctggccage tgcgtgacca ttaacaaac cagcgtcaag	1680
	gtgctgcgtg atatgaatgt gaaggaatcg ccaggacgt gctactcacg accagtggtc	1740
	atctttaatt tcgccaacag ctctgacgtg cagtacggtc aactgggcga ggataacgaa	1800
	atctgttgg gcaaccaccg cactgaggaa tgtcagcttc ccagcctcaa gatcttcac	1860
	gccggcaact cggcctacga gtacgtggac tacctcttca aacgcatgat tgacctcagc	1920

```

agcatctcca ccgtcgacag catgatcgcc ctagacatcg acccgctgga aaacaccgac      1980

ttcagggtac tggaacttta ctgcagaaa gaattgcgtt ccatcaacgt tttgatctc      2040

gaggagatca tgcgcgagtt caaticgtat aagcagcggg taaagtacgt ggaggacaag      2100

gtagtcgacc cgctgccgcc ctaccicaag ggtctggacg acctcatgag cggcctgggc      2160

gccgcgggaa aggccgttgg cgtagccatt ggggcgctgg gtggcgcggt ggcctccgtg      2220

gtcgaaggcg ttgccacctt cctcaaaaac cctttttctt ttatcatagg gttaatcatt      2280

ggactattct tggttctcgg agttggtatc catctttgca ttaaattaaa gcacaccaag      2340

aaaagacaga ttatacaga catagagatg aaccgacttg gaaagtaa      2388

```

<210> 12

<211> 2388

<212> DNA

<213> 人巨细胞病毒

[0037]

<220>

<221> misc\_特征

<222> (1).. (2388)

<223> 密码子优化的 HCMV gB - G 核苷酸序列

<220>

<221> misc\_特征

<222> (2253).. (2388)

<223> TM 和 CTD

<400> 12

```

atggagtcaa ggatttgglg tctggtcgtc tgcgtcaacc tgtgcattgt ctgcctggga      60

gccgcgctct catcatcatc tacccgaggc acatccgcca ctcaacttca ccatagctcc      120

cataccacat ccgccgtca ctcaagaagc gggtcctgtg ctcaagggtg cacatctagt      180

cagaccgtga gccatggagt caacgagaca atctacaata ctaccctgaa gtatggagac      240

gtggtcggcg tgaacacaac taaatacccc tatagggtct gctctatggc ccaggggaca      300

```

	gatctgatcc gatttgaacg gaacatcgtg tgcactagca tgaagcctat caatgaggac	360
	ctggatgaag gaattatggt ggtctacaaa cgaaatatcg tggcccatatc ttttaaggtg	420
	agagtctatc agaaagtgtg gacctccgg agaagctacg ctlatattca caccacatac	480
	ctgctggggg ccaacaccga giatgtggca cccctatgt gggaaatcca ccatattaat	540
	agtcattcac agtctactc aagctacagc agagtgatcg ctggaaccgt gttcgtcgca	600
	taccacagag acagttatga gaacaagaca atgcagctca tgcccagcga ttacagtaat	660
	accattcaa caagatatgt gaccgtcaaa gatcagtggc actctcggc cagtacctgg	720
	ctgtaccgag agacatgcaa cctgaattgt atggtgacaa ttactaccgc cagaagcaag	780
	tacccttate acttctttgc tactcaaca ggggacgtgg tcgacatcag ccccttctac	840
	aacggaacaa accggaatgc ctctatttc ggcgagaacg ctgacaaatt ctttatcttc	900
	cccaactaca ctatcgtgag cgatttcggc agacctaca gtgccctgga gacctcgg	960
[0038]	ctggtggcat ttctggaaag agccgacagc gtgatctcct gggacattca ggaatgagaag	1020
	aatgtgacct gccagctcac cttctgggag gccagcgaaa gaacctcag gtccgaggca	1080
	gaagattctt accacttttc ctctgcaaaa atgactgcca ccttctgtc caagaacag	1140
	gaggtgaaca tgagcgactc cgcactggat tgcgtcggg acgaagccat caataagctg	1200
	cagcagatct tcaacacatc ttacaaccag acttacgaga agtacggcaa cgtgagtgtc	1260
	tttgaacaa ctggcgggct ggtggtcttc tggcagggga tcaagcagaa atctctggtg	1320
	gagctggaac ggctggccaa tagaagtcca ctgaacctga ctcataatcg caccaagcga	1380
	tccacagacg gaaacaatgc aactcatctg agcaacatgg agtccgtgca caatctggtc	1440
	tacgcccagc tccagttcac ttacgacacc ctgcgaggct atatcaaccg ggccttggt	1500
	cagattgcag aagcctggtg cgtggatcag aggcgcaccc tggaggtctt taaggaaactg	1560
	agcaaaatta acccatctgc tatcctgagt gcaatctaca ataagccat cgcagccagg	1620
	ttcatggggg acgtgctggg actggcctcc tgcgtcacta tcaaccagac ctctglgaag	1680

gtcctgcgcg atatgaatgt gaaagagagt cctggcaggt gttattcacg cccagtggtc	1740
atcttcaact tcgctaatag ctccctacgtg cagtatggcc agctcgggga ggacaacgaa	1800
atcctgctgg gaaatcacag gaccgaggaa tgtcagctcc caagtctgaa gatctttatt	1860
gccggcaact cagcttacga gtatgtggat tacctgttca aacgcatgat cgacctgtct	1920
agtatttcaa cagtggatag catgatgcc ctggacattg atcccctgga aaatactgac	1980
ttcagggtgc tggagctgta tagccagaag gaactgcgct ccattaacgt gtttgatctg	2040
gaggaaatca tgaggaggtt caaticctac aagcagcgcg tgaaatatgt cgaagataag	2100
gtggtcgacc ctctgccacc ctacctgaaa ggccctggacg atctgatgag cgggctggga	2160
gctgcaggca aggcagtgga agtcgccatc ggagctgtgg gaggcgctgt cgcacccgtg	2220
gtcgaggagag tggctacctt tctgaagaac ccattctttt tcatcatcgg cctgatcatt	2280
gggctgttcc tgggtctgag agtcggcatc cacctgtgca ttaagctgaa gcacaccaag	2340
[0039] aagaggcaga tctacaccga tattgaaatg aacagactgg gcaagtga	2388

<210> 13

<211> 742

<212> PRT

<213> 人巨细胞病毒

<220>

<221> MISC\_特征

<222> (1).. (742)

<223> HCMV gH 氨基酸序列

<220>

<221> MISC\_特征

<222> (717).. (742)

<223> TM 和 CTD

<400> 13

Met Arg Pro Gly Leu Pro Ser Tyr Leu Ile Val Leu Ala Val Cys Leu

1	5	10	15
Leu Ser His	Leu Leu Ser Ser Arg Tyr	Gly Ala Glu Ala Ile Ser Glu	
20	25	30	
Pro Leu Asp Lys Ala Phe His	Leu Leu Leu Asn Thr Tyr Gly Arg Pro		
35	40	45	
Ile Arg Phe Leu Arg Glu Asn Thr Thr	Gln Cys Thr Tyr Asn Ser Ser		
50	55	60	
Leu Arg Asn Ser Thr Val Val Arg Glu Asn Ala Ile Ser Phe Asn Phe			
65	70	75	80
Phe Gln Ser Tyr Asn Gln Tyr Tyr Val Phe His Met Pro Arg Cys Leu			
85	90	95	
[0040]			
Phe Ala Gly Pro Leu Ala Glu Gln Phe Leu Asn Gln Val Asp Leu Thr			
100	105	110	
Glu Thr Leu Glu Arg Tyr Gln Gln Arg Leu Asn Thr Tyr Ala Leu Val			
115	120	125	
Ser Lys Asp Leu Ala Ser Tyr Arg Ser Phe Ser Gln Gln Leu Lys Ala			
130	135	140	
Gln Asp Ser Leu Gly Glu Gln Pro Thr Thr Val Pro Pro Pro Ile Asp			
145	150	155	160
Leu Ser Ile Pro His Val Trp Met Pro Pro Gln Thr Thr Pro His Gly			
165	170	175	
Trp Thr Glu Ser His Thr Thr Ser Gly Leu His Arg Pro His Phe Asn			
180	185	190	

	Gln Thr Cys Ile Leu Phe Asp Gly His Asp Leu Leu Phe Ser Thr Val	
	195	200 205
	Thr Pro Cys Leu His Gln Gly Phe Tyr Leu Ile Asp Glu Leu Arg Tyr	
	210	215 220
	Val Lys Ile Thr Leu Thr Glu Asp Phe Phe Val Val Thr Val Ser Ile	
	225	230 235 240
	Asp Asp Asp Thr Pro Met Leu Leu Ile Phe Gly His Leu Pro Arg Val	
		245 250 255
	Leu Phe Lys Ala Pro Tyr Gln Arg Asp Asn Phe Ile Leu Arg Gln Thr	
	260	265 270
[0041]	Glu Lys His Glu Leu Leu Val Leu Val Lys Lys Asp Gln Leu Asn Arg	
	275	280 285
	His Ser Tyr Leu Lys Asp Pro Asp Phe Leu Asp Ala Ala Leu Asp Phe	
	290	295 300
	Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg Tyr	
	305	310 315 320
	Ala Val Asp Val Leu Lys Ser Gly Arg Cys Gln Met Leu Asp Arg Arg	
		325 330 335
	Thr Val Glu Met Ala Phe Ala Tyr Ala Leu Ala Leu Phe Ala Ala Ala	
	340	345 350
	Arg Gln Glu Glu Ala Gly Ala Gln Val Ser Val Pro Arg Ala Leu Asp	
	355	360 365

Arg Gln Ala Ala Leu Leu Gln Ile Gln Glu Phe Met Ile Thr Cys Leu  
370 375 380

Ser Gln Thr Pro Pro Arg Thr Thr Leu Leu Leu Tyr Pro Thr Ala Val  
385 390 395 400

Asp Leu Ala Lys Arg Ala Leu Trp Thr Pro Asn Gln Ile Thr Asp Ile  
405 410 415

Thr Ser Leu Val Arg Leu Val Tyr Ile Leu Ser Lys Gln Asn Gln Gln  
420 425 430

His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala Leu  
435 440 445

[0042]

Lys Leu His Lys Thr His Leu Ala Ser Phe Leu Ser Ala Phe Ala Arg  
450 455 460

Gln Glu Leu Tyr Leu Met Gly Ser Leu Val His Ser Met Leu Val His  
465 470 475 480

Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys Ser  
485 490 495

Leu Ala Glu Leu Ser His Phe Thr Gln Leu Leu Ala His Pro His His  
500 505 510

Glu Tyr Leu Ser Asp Leu Tyr Thr Pro Cys Ser Ser Ser Gly Arg Arg  
515 520 525

Asp His Ser Leu Glu Arg Leu Thr Arg Leu Phe Pro Asp Ala Thr Val  
530 535 540



Pro Thr Thr Val Pro Ala Ala Leu Ser Ile Leu Ser Thr Met Gln Pro  
545 550 555 560

Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly Glu  
565 570 575

Ser Phe Ser Ala Leu Thr Val Ser Glu His Val Ser Tyr Val Val Thr  
580 585 590

Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr Val  
595 600 605

Val Gly Gln Ser Leu Ile Ile Thr Gln Thr Asp Ser Gln Thr Lys Cys  
610 615 620

Glu Leu Thr Arg Asn Met His Thr Thr His Ser Ile Thr Ala Ala Leu  
625 630 635 640

[0043]

Asn Ile Ser Leu Glu Asn Cys Ala Phe Cys Gln Ser Ala Leu Leu Glu  
645 650 655

Tyr Asp Asp Thr Gln Gly Val Ile Asn Ile Met Tyr Met His Asp Ser  
660 665 670

Asp Asp Val Leu Phe Ala Leu Asp Pro Tyr Asn Glu Val Val Val Ser  
675 680 685

Ser Pro Arg Thr His Tyr Leu Met Leu Leu Lys Asn Gly Thr Val Leu  
690 695 700

Glu Val Thr Asp Val Val Val Asp Ala Thr Asp Ser Arg Leu Leu Met  
705 710 715 720

Met Ser Val Tyr Ala Leu Ser Ala Ile Ile Gly Ile Tyr Leu Leu Tyr

725

730

735

Arg Met Leu Lys Thr Cys  
740

&lt;210&gt; 14

&lt;211&gt; 2229

&lt;212&gt; DNA

&lt;213&gt; 人巨细胞病毒

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (1)..(2229)

&lt;223&gt; HCMV gH 核苷酸序列

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (2151)..(2229)

&lt;223&gt; TM 和 CTD

[0044]

&lt;400&gt; 14

atggggccag gcctcccctc ctacctcacc gtccctgcgc tctgtctcct cagecaccta 60

ctttcgtcac gatatggcgc agaagccata tccgaaccgc tggacaaagc gtttcaccta 120

ctgctcaaca cctacgggag acccatccgc ttctgtcgtg aaaacaccac ccagtgtacc 180

tacaatagca gcctccgtaa cagcacggtc gtcagggaaa acgccatcag ttcaacttt 240

ttccaaagct ataataata ctatgtatc cataigcctc gatgtctttt tgcgggtcct 300

ctggcggagc agtttctgaa ccaggtagat ctgaccgaaa ccttggaag ataccaacag 360

agacttaaca cttacgcgct ggtatccaaa gacctggcca gctaccgac tttttcgcag 420

cagctaaagg cacaggacag cctaggtgaa cagcccacca ctgtgccacc acccatgac 480

ctgtcaatac ctacgcttg gatgccaccg caaacaccac cacacggctg gacagaatca 540

cataccacct caggactaca ccgaccacac tttaaccaga cctgtatcct ctttgatgga 600

cagcatctac tattcagcac cgtcacacct tgtttgcacc aaggctttta cctcatcgac 660

[0045]	gaactacgtt acgttaaaat aacactgacc gaggacttct tcgtagttac ggtgtccata	720
	gacgacgaca cacccatgct gcttatcttc ggccatcttc cagcggtact ctttaaagcg	780
	ccctatcaac gcgacaactt tatactacga caaactgaaa aacacgagct cctggtgcta	840
	gttaagaaag atcaactgaa ccgtcactct tatcicaaag acccggaatt tcttgacgcc	900
	gcacttgact tcaactacct ggacctcagc gcactactac gtaacagctt tcaccgttac	960
	gccgtggatg tactcaaaag cggtcgatgt cagatgctgg accgccgcac ggtagaaatg	1020
	gccttcgcct acgcattagc actgttcgca gcagcccgac aagaagaggc cggcgcccaa	1080
	gtctccgtcc caggggcct agaccgccag gccgcactct tacaatata agaatttatg	1140
	atcacctgcc tctcacaac accaccagc accacgttgc tgctgtatcc cagggccgtg	1200
	gacctggcca aacgagccct ttggacaccg aatcagatca ccgacatcac cagcctcgta	1260
	cgcttggtct acatactctc taaacagaat cagcaacatc tcatecccca gtgggcacta	1320
	cgacagatcg ccgactttgc cctaaaacta caaaaacgc acctggcctc tttctttca	1380
	gccttcgcgc gteaagaact ctacctatg ggcagcctcg tccactccat gctagtacat	1440
	acgacggaga gacgcgaaat cttcatcgta gaaacgggcc tctgttcatt agccgageta	1500
	tcacacttta cgcagttgct agctcatccg caccacgaat acctcagcga cctgtacaca	1560
	ccctgttcca gtagcgggag acgcgatcac tcgctcgaa cgcctcacag tctcttcccc	1620
	gatgccaccg tccccactac cgttcccgcc gccctctcca tcttatctac catgcaacca	1680
	agcacgctag aaaccttccc cgacctgttt tgtctgccgc tcggcgaatc cttctccgcg	1740
	ctgaccgtct ccgaacacgt cagttatgtc gtaacaaacc agtacctgat caaaggtatc	1800
	tcttacctg tctccaccac cgtcgttagc cagagcctca tcatcacca gacggacagt	1860
	caaactaaat gcgaactgac gcgcaacatg cataccacac acagcatcac agcggcgctc	1920
	aacatttccc tagaaaactg cgccttttgc caaagcgcgc tactagaata cgacgacacg	1980

caaggcgtca tcaacatcat gtacatgcac gactcggacg acgtcctttt cggcctggat	2040
cctacaacg aagtgggtgt ctcatctccg cgaactcaact acctcatgct ttgaaaaac	2100
ggtacgggcc tagaagtaac tgacgtcgtc gggacgcta ccgacagtcg tctcctcatg	2160
atgtccgtct acgcgctatc ggccatcacc ggcatctatc tgcctaccg catgctcaag	2220
acatgctga	2229

<210> 15  
 <211> 2229  
 <212> DNA  
 <213> 人巨细胞病毒

<220>  
 <221> misc\_特征  
 <222> (1)..(2229)  
 <223> 密码子优化的 HCMV gH 核苷酸序列

[0046] <220>  
 <221> misc\_特征  
 <222> (2151)..(2229)  
 <223> TM 和 CTD

<400> 15	
atgagacctg gactgccttc ttatctgatt gtgtggccg tctgcctgct gtcacatctg	60
ctgagttcac gctatggggc tgaggctatc tccgagccac tggacaaggc ttttcacctg	120
ctgtgaaca cctacgggag acccattagg ttcttgcgcg agaataccac acagtgcaca	180
tataacagct ccttgcggaa cagcactgtg gtccgcgaaa acgcatctc ttttaatttc	240
tttcagagtt acaaccagta ctacgtgttc catatgccac gctgtctgtt tgcaggaccc	300
ctggccgagc agttcctgaa ccaggtggac ctgaccgaga cactggaaag ataccagcag	360
aggctgaata cctatgccct ggtgagtaag gatctggctt catatcggtc tttcagtcag	420
cagctcaagg cccaggactc actgggcgag cagcctacta ccgtgcccc tccaatcgat	480
ctgagcattc cacacgtctg gatgccccct cagacaactc cccacggctg gaccgaaagc	540

	cataccacat ccgggctgca cagaccccat ttcaaccaga catgcatcct gtttgatggg	600
	cacgacctgc tgttcagcac tgtgaccct tgtctgcatc agggattcta cctgatcgat	660
	gagctgagat atgtgaaaat tacactgact gaagacttct ttgtggtcac cgtgagcatc	720
	gacgatgaca caccaatgct gctgattttt ggacacctgc ccgggtgct gttcaaggcc	780
	ccctaccage gagacaactt tattctgcgg cagaccgaga aacacgaact gctggtgctg	840
	gtcaagaaag atcagctcaa caggcatagc tatctgaagg accccgactt tctggatgcc	900
	gctctggact tcaactacct ggacctgta gcactgctgc ggaatagctt ccacagatat	960
	gccgtggatg tcttgaaac cggaagatgc cagatgctgg accggagaac cgtggagatg	1020
	gcatttgctt acgctctggc actgttcgca gccgctaggc aggaggaagc aggcgctcag	1080
	gtgtccgtcc ctgcgcact ggatcgacag gcagccctgc tgcagatcca ggagttcatg	1140
	attacctgtc tgtctcagac accaccaga actacctgc tgcgtaccc cactgccgtg	1200
[0047]	gacctggcta agagggcact gtggaccct aaccagatca ctgatattac ctctctggtg	1260
	cgcttggtct atatcctgag taaacagaat cagcagcacc tgatccaca gtgggccctg	1320
	cgacagattg ccgacttcgc tctgaagctg cacaaaacc atctggcttc ctctctgtct	1380
	gcatttgccc gccaggagct gtacctgatg ggtctctctg tgcacagtat gctggtccat	1440
	acaactgaga ggcgcgaaat ctttattgtg gagacagggc tgtgcagcct ggctgaactg	1500
	tcccattca ctcagctcct ggcccatcct caccatgagt acctgtccga tctgtatacc	1560
	ccatgittta gttcaggcgc acgggaccac tctctggaac gactgactcg gctgtttcct	1620
	gatgcaaccg tgcctaccac cgtgcccgcc gccctgagta tctgtcaac aatgcagcca	1680
	agcacactgg agactttccc cgacctgttt tgctgcctc tgggggagtc attcagcgcc	1740
	ctgacctgtt cagaacatgt cagctacgtg gtcacaaacc agtatctgat caagggaatt	1800
	tcttaccceg tgtctactac cgtggctggc cagagtctga tcattacca gacagattca	1860

```

cagactaaat gtgagctgac ccggaatatg cacacaactc atagcatcac cgccgctctg      1920

aacatttccc tggagaattg cgctttttgt cagagtgcac tgctggaata cgatgacaca      1980

cagggcgtga tcaacattat gtatatgcac gatagegatg acgtgctglt cgctctggac      2040

ccctacaacg aggtggctgt gagctcccct cgcactcatt atctgatgct gctgaagaat      2100

ggaacagtgc tggaagtcac tgatgtcgtg gtcgatgcca cagactcccg gctgctgatg      2160

atgtctgtgt acgcactgtc cgccatcatc ggcatctatc tgctgtatcg aatgctgaaa      2220

acctgttga                                     2229

```

<210> 16  
 <211> 761  
 <212> PRT  
 <213> 人巨细胞病毒

[0048] <220>  
 <221> MISC\_特征  
 <222> (1)..(761)  
 <223> HCMV gH - G 氨基酸序列

<220>  
 <221> MISC\_特征  
 <222> (717)..(761)  
 <223> TM 和 CTD

<400> 16

Met Arg Pro Gly Leu Pro Ser Tyr Leu Ile Val Leu Ala Val Cys Leu  
 1                    5                    10                    15

Leu Ser His Leu Leu Ser Ser Arg Tyr Gly Ala Glu Ala Ile Ser Glu  
                   20                    25                    30

Pro Leu Asp Lys Ala Phe His Leu Leu Leu Asn Thr Tyr Gly Arg Pro  
                   35                    40                    45

Ile Arg Phe Leu Arg Glu Asn Thr Thr Gln Cys Thr Tyr Asn Ser Ser  
50 55 60

Leu Arg Asn Ser Thr Val Val Arg Glu Asn Ala Ile Ser Phe Asn Phe  
65 70 75 80

Phe Gln Ser Tyr Asn Gln Tyr Tyr Val Phe His Met Pro Arg Cys Leu  
85 90 95

Phe Ala Gly Pro Leu Ala Glu Gln Phe Leu Asn Gln Val Asp Leu Thr  
100 105 110

Glu Thr Leu Glu Arg Tyr Gln Gln Arg Leu Asn Thr Tyr Ala Leu Val  
115 120 125

Ser Lys Asp Leu Ala Ser Tyr Arg Ser Phe Ser Gln Gln Leu Lys Ala  
130 135 140

[0049]

Gln Asp Ser Leu Gly Glu Gln Pro Thr Thr Val Pro Pro Pro Ile Asp  
145 150 155 160

Leu Ser Ile Pro His Val Trp Met Pro Pro Gln Thr Thr Pro His Gly  
165 170 175

Trp Thr Glu Ser His Thr Thr Ser Gly Leu His Arg Pro His Phe Asn  
180 185 190

Gln Thr Cys Ile Leu Phe Asp Gly His Asp Leu Leu Phe Ser Thr Val  
195 200 205

Thr Pro Cys Leu His Gln Gly Phe Tyr Leu Ile Asp Glu Leu Arg Tyr  
210 215 220

Val Lys Ile Thr Leu Thr Glu Asp Phe Phe Val Val Thr Val Ser Ile

225	230	235	240
Asp Asp Asp Thr Pro Met Leu Leu Ile Phe Gly His Leu Pro Arg Val			
245	250	255	
Leu Phe Lys Ala Pro Tyr Gln Arg Asp Asn Phe Ile Leu Arg Gln Thr			
260	265	270	
Glu Lys His Glu Leu Leu Val Leu Val Lys Lys Asp Gln Leu Asn Arg			
275	280	285	
His Ser Tyr Leu Lys Asp Pro Asp Phe Leu Asp Ala Ala Leu Asp Phe			
290	295	300	
Asn Tyr Leu Asp Leu Ser Ala Leu Leu Arg Asn Ser Phe His Arg Tyr			
305	310	315	320
[0050]			
Ala Val Asp Val Leu Lys Ser Gly Arg Cys Gln Met Leu Asp Arg Arg			
325	330	335	
Thr Val Glu Met Ala Phe Ala Tyr Ala Leu Ala Leu Phe Ala Ala Ala			
340	345	350	
Arg Gln Glu Glu Ala Gly Ala Gln Val Ser Val Pro Arg Ala Leu Asp			
355	360	365	
Arg Gln Ala Ala Leu Leu Gln Ile Gln Glu Phe Met Ile Thr Cys Leu			
370	375	380	
Ser Gln Thr Pro Pro Arg Thr Thr Leu Leu Leu Tyr Pro Thr Ala Val			
385	390	395	400
Asp Leu Ala Lys Arg Ala Leu Trp Thr Pro Asn Gln Ile Thr Asp Ile			
405	410	415	



Thr Ser Leu Val Arg Leu Val Tyr Ile Leu Ser Lys Gln Asn Gln Gln  
420 425 430

His Leu Ile Pro Gln Trp Ala Leu Arg Gln Ile Ala Asp Phe Ala Leu  
435 440 445

Lys Leu His Lys Thr His Leu Ala Ser Phe Leu Ser Ala Phe Ala Arg  
450 455 460

Gln Glu Leu Tyr Leu Met Gly Ser Leu Val His Ser Met Leu Val His  
465 470 475 480

Thr Thr Glu Arg Arg Glu Ile Phe Ile Val Glu Thr Gly Leu Cys Ser  
485 490 495

[0051]

Leu Ala Glu Leu Ser His Phe Thr Gln Leu Leu Ala His Pro His His  
500 505 510

Glu Tyr Leu Ser Asp Leu Tyr Thr Pro Cys Ser Ser Ser Gly Arg Arg  
515 520 525

Asp His Ser Leu Glu Arg Leu Thr Arg Leu Phe Pro Asp Ala Thr Val  
530 535 540

Pro Thr Thr Val Pro Ala Ala Leu Ser Ile Leu Ser Thr Met Gln Pro  
545 550 555 560

Ser Thr Leu Glu Thr Phe Pro Asp Leu Phe Cys Leu Pro Leu Gly Glu  
565 570 575

Ser Phe Ser Ala Leu Thr Val Ser Glu His Val Ser Tyr Val Val Thr  
580 585 590

Asn Gln Tyr Leu Ile Lys Gly Ile Ser Tyr Pro Val Ser Thr Thr Val  
595 600 605

Val Gly Gln Ser Leu Ile Ile Thr Gln Thr Asp Ser Gln Thr Lys Cys  
610 615 620

Glu Leu Thr Arg Asn Met His Thr Thr His Ser Ile Thr Ala Ala Leu  
625 630 635 640

Asn Ile Ser Leu Glu Asn Cys Ala Phe Cys Gln Ser Ala Leu Leu Glu  
645 650 655

Tyr Asp Asp Thr Gln Gly Val Ile Asn Ile Met Tyr Met His Asp Ser  
660 665 670

[0052]

Asp Asp Val Leu Phe Ala Leu Asp Pro Tyr Asn Glu Val Val Val Ser  
675 680 685

Ser Pro Arg Thr His Tyr Leu Met Leu Leu Lys Asn Gly Thr Val Leu  
690 695 700

Glu Val Thr Asp Val Val Val Asp Ala Thr Asp Ser Arg Phe Phe Phe  
705 710 715 720

Ile Ile Gly Leu Ile Ile Gly Leu Phe Leu Val Leu Arg Val Gly Ile  
725 730 735

His Leu Cys Ile Lys Leu Lys His Thr Lys Lys Arg Gln Ile Tyr Thr  
740 745 750

Asp Ile Glu Met Asn Arg Leu Gly Lys  
755 760

<210> 17  
 <211> 2286  
 <212> DNA  
 <213> 人巨细胞病毒

<220>  
 <221> misc\_特征  
 <222> (1).. (2286)  
 <223> HCMV gH - G 核苷酸序列

<220>  
 <221> misc\_特征  
 <222> (2151).. (2286)  
 <223> TM 和 CTD

<400> 17

[0053]

```

atgcgccag gcctccctc ctacctcgc gtctcgccg tctgtctct cagccaccta      60
ctttcgtcac gatatggcgc agaagccata tccgaaccgc tggacaaagc gtttcaccta      120
ctgtcaaca cctacgggag acccatccgc ttctgcgtg aaaacaccac ccagtgtacc      180
tacaatagca gcctccgtaa cagcacggtc gtcagggaaa acgccatcag tttcaacttt      240
ttccaaagct ataatcaata ctatgtattc catatgcctc gatgtctttt tgcgggtcct      300
ctggcgggag agtttctgaa ccaggtagat ctgaccgaaa ccttggaag ataccaacag      360
agacttaaca cttacgcgct ggtatccaaa gacctggcca gctaccgac tttttcgcag      420
cagctaaagg cacaggacag cctagglgaa cagcccacca ctgtgccacc acctatgac      480
ctgtcaatac ctcacgtttg gatgccaccg caaaccactc cacacggctg gacagaatca      540
cataccacct caggactaca ccgaccacac tttaaccaga cctgtatcct ctttgatgga      600
cacgatctac tattcagcac cgtcacacct tgtttgcacc aaggctttta cctcatcgac      660
gaactacgtt acgttaaaat aacactgacc gaggacttct tcgtagttac ggtgtccata      720
gacgacgaca caccatgct gcttatcttc ggccatcttc cacgcgtact ctttaaagcg      780
ccctatcaac gcgacaactt tatactacga caaactgaaa aacacgagct cctggtgcta      840

```

	gttaagaaag atcaactgaa cgcactct tatctcaaag acccgactt tcttgacgc	900
	gcacttgact tcaactacct ggacctcagc gcactactac gtaacagctt tcaccgttac	960
	gccgtggatg tactcaaaag cggtcgatgt cagatgctgg accgccgcac ggtagaaatg	1020
	gccttcgcct acgcattagc actgttcgca gcagcccgac aagaagagge cggegcccaa	1080
	gtctccgtcc caccggccct agaccgccag gccgcactct tacaaatata agaatttatg	1140
	atcacctgcc tctcaciaac accaccacgc accacgttgc tgctgtatcc caccgccgtg	1200
	gacctggcca aacgagccct ttggacaccg aatcagatca ccgacatcac cagcctcgta	1260
	cgcttggctt acatactctc taaacagaat cagcaacatc tcatcccca gtgggcacta	1320
	cgacagatcg ccgactttgc cctaaaacta caaaaacgc acctggcctc ttttcttca	1380
	gccttcgcgc gtcaagaact ctacctatg ggcagcctcg tccactccat gctagtacat	1440
	acgacggaga gacgcgaaat cttcatcgta gaaacgggcc tctgttcatt agccgagcta	1500
[0054]	tcacacttta cgcagttgct agtcatccg caccacgaat acctcagcga cctgtacaca	1560
	ccctgttcca gtacggggcg accgatcac tcgtcgaac gcctcacacg tctcttccc	1620
	gatgccaccg tcccactac cgttccgcc gccctctcca tctatctac catgcaacca	1680
	agcacgctag aaaccttccc cgacctgtt tgtctgccgc tcggcgaatc cttctccgcg	1740
	ctgaccgtct ccgaacacgt cagttatgtc gtaacaaacc agtacctgat caaaggatc	1800
	tcctacctg tctccaccac cgtcgtaggc cagagcctca tcatcccca gacggacagt	1860
	caactaaat ggaacigac gcgcaacatg cataccacac acagatcac agcggcgctc	1920
	aacatttccc tagaaaactg cgccttttgc caaagcggc tactagaata cgacgacag	1980
	caaggcgta tcaacatcat gtacatgcac gactcggacg acgtcctttt cgcctggat	2040
	ccctacaacg aagtgggtgt ctcctctccg cgaactcact acctcatgct ttgaaaaac	2100
	ggtacggctc tagaagtaac tgacgtcgtc gtggacgcta ccgacagtcg tttttcttt	2160
	atcatagggt taatcatlgg actatcttg gtctccgag ttggatatcca tctttgcatt	2220

	aaattaaagc acaccaagaa aagacagatt tatacagaca tagagatgaa ccgacttgga	2280
	aagtaa	2286
	<210> 18	
	<211> 2286	
	<212> DNA	
	<213> 人巨细胞病毒	
	<220>	
	<221> misc_特征	
	<222> (2151).. (2286)	
	<223> TM 和 CTD	
	<220>	
	<221> misc_特征	
	<222> (2151).. (2286)	
	<223> 密码子优化的 HCMV gH - G 核苷酸序列	
[0055]	<400> 18	
	atgcgaccgc gactgccaag ctacctgatt gtccctggctg tctgtctgct gtcacacctg	60
	ctgagttcaa gatatggggc cgaagccatc agcgagccac tggacaaggc tttccacctg	120
	ctgtcgaaca cctacggcag acccattagg ttctctgcgc agaataccac acagtgcaca	180
	tataacagct cctgaggaa tagcactgtg gtccgcgaaa acgcatctc ttcaatttc	240
	tttcagagtt acaaccagta ctacgtgtc catatgccac gctgtctgtt cgcaggaccc	300
	ctggccgagc agttttctgaa ccaggtggac ctgaccgaga cactggaaag ataccagcag	360
	aggetgaata cctatgccct ggtgagtaag gatctggctt catatcggtc tticagtcag	420
	cagctcaagg cccaggactc tctgggagag cagcctacta ccgtgcccc tccaatcgat	480
	ctgagtattc cacacgtctg gatgccccct cagacaactc cccacggatg gaccgaaagc	540
	cataccacat ccggcctgca cagaccccac ttcaaccaga catgcatcct gttcgatggc	600
	cacgacctgc tgttttccac tgtgaccct tgtctgcac aggggttcta cctgatcgat	660

	gagctgagat atgtgaagat tacactgact gaagacttct ttgtggtcac cgtgtctatc	720
	gacgatgaca caccaatgct gctgatttcc ggacacctgc cccgggtgct gttcaaggcc	780
	ccctaccagc gagacaactt catcctgcgg cagaccgaga aacacgaact gctgggtgctg	840
	gtcaagaaag atcagctcaa ccggcattcc tatctgaagg accccgactt cctggatgcc	900
	gctctggact ttaactacct ggacctgtca gcactgtgc ggaatagctt tcacagatat	960
	gccgtggatg tcttgaatc tgggcgctgc cagatgtgg accggagaac cgtggagatg	1020
	gcattcgctt acgctctggc actgtttgca gccgctcggc aggaggaagc aggagctcag	1080
	gtgagtgtcc ctgcgcact ggatcgacag gcagccctgc tgcagatcca ggagttcatg	1140
	attacctgtc tgagccagac accaccaga actacctgc tgcgtaccc cactgccgtg	1200
	gacctggcta agagggcact gtggacctt aaccagatca ctgatattac cagcctgggtg	1260
	agactggtct atatctgtc caaacagaat cagcagcacc tgatcccaca gtgggccctg	1320
[0056]	aggcagattg ccgactttgc tctgaagctg cacaaaacc atctggcttc ctttctgtct	1380
	gcattcgcca gacaggagct gtacctgatg ggatctctgg tgcacagtat gctgggccat	1440
	acaactgaga ggccgcaaat cttcattgtg gagacaggcc tgtgcagcct ggctgaactg	1500
	tcccacttta ctacgtcct ggcccatcct caccatgagt acctgtcaga tctgtatacc	1560
	ccatgttcta gttcaggacg acgggaccac agcctggaac gactgactcg gctgttccct	1620
	gatgcaaccg tgcctaccac cgtgcccgcc gccctgagta tctgtcaac aatgcagcca	1680
	agcacactgg agacttttcc cgacctgttc tgcttgctc tgggcgagag ctccagcgcc	1740
	ctgacctga gcgaacatgt cagctacgtg gtcacaaacc agtatctgat caaggggatt	1800
	tcctaccccg tgtctactac cgtggctgga cagtcctga tcattacca gacagattct	1860
	cagactaat gtgagctgac cagaaatatg cacacaactc atagtatcac cgccgctctg	1920
	aacatttcac tggagaattg cgctttctgt cagtcgcac tgctggaata cgatgacaca	1980
	cagggcgiga tcaacattat gtatgacac gattcigat acgtgctgtt tgctcggac	2040

	ccctacaacg aggtggctgt gagctcccc agaactcatt atctgatgt gctgaagaat	2100
	ggcacagtgc tggaaagtcac tgatgtcgtg gtcatgccca cagactcccg cttcttttc	2160
	atcattggcc tgatcattgg gctgttcctg gtgctgcgag tcggcatcca cctgtgcac	2220
	aagctgaagc atacaaagaa gagacagatc tacaccgata ttgaaatgaa caggctgggc	2280
	aaatga	2286
	<210> 19	
	<211> 4895	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> Propol II 表达质粒	
	<220>	
[0057]	<221> misc_特征	
	<222> (1)..(4895)	
	<223> Propol II 表达质粒核苷酸序列	
	<400> 19	
	ctagagagct tggccattg catacgttg atccatatca taatatgtac atttatattg	60
	gctcatgtcc aacattaccg ccatgttgac attgattatt gactagtat taatagtaat	120
	caattacggg gtcattagt catagcccat atattggagt ccgcgttaca taacttacgg	180
	taaatggccc gcctggctga ccgccaacg accccgccc attgacgta ataatacgt	240
	atgttcccat agtaacgcca atagggaatt tccattgacg tcaatgggtg gattatttac	300
	ggtaaaactgc ccacttgga gtacatcaag tgtatcatat gccaaagtac cccctattg	360
	acgtcaatga cggtaaatgg ccgcctggc attatgcca gtacatgacc ttatgggact	420
	ttcctacttg gcagtacac tacgtattag tcacgctat taccatggig atgcggtttt	480
	ggcagtacat caatgggcgt ggatagcgt ttgactcacg gggatttcca agtctccacc	540

	ccattgacgt caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc	600
	gtaacaactc cgccccattg acgcaaatgg gcggtaggcg tgtacggtgg gaggtctata	660
	taagcagagc tcgttttagtg aaccgicaga tcgcctggag acgccatcca cgctgttttg	720
	acctccatag aagacaccgg gaccgatcca gcctccggtc gaccgatcct gagaacttca	780
	gggtgagitt ggggaccett gattgttctt tctttttcgc tattgtaaaa ttcattgtat	840
	atggaggggg caaagttttc aggggtgtgt ttagaatggg aagatgtccc ttgtatcacc	900
	atggaccctc atgataatit tgtttctttc actttctact ctgttgacaa ccattgtctc	960
	ctcttatttt cttttcattt tcttgtaact ttttcgttaa actttagctt gcatttgtaa	1020
	cgaattttta aattcacttt tgtttatttg tcagattgta agtactttct ctaatcactt	1080
	ttttttcaag gcaatcaggg tatattatat tgtacttcag cacagtttta gagaacaatt	1140
	gttataatta aatgataagg tagaatatit ctgcatataa attctggctg gcgtggaaat	1200
[0058]	attcttattg gtagaaacaa ctacatcctg gtcacatcc tgcctttctc tttatggtta	1260
	caatgatata cactgtttga gatgaggata aaatactctg agtccaaacc ggccccctct	1320
	gctaaccatg ttcattgcctt cttctttttc ctacagctcc tgggcaacgt gctggttatt	1380
	gtgctgtctc atcattttgg caaagaattc ctcgagcgta cgcctagggg atccagcgct	1440
	attlaaatgc tagcatgcat gttaacctg caggggtacc gcggccgcaa gcttagatcc	1500
	gtcgaggaat tcactcctca ggtgcaggct gcctatcaga aggtgggtggc tgggtgtggc	1560
	aatgccttgg ctacaaaata ccactgagat ctttttccct ctgccaaaaa ttatggggac	1620
	atcatgaagc cccctgagca tctgacttct ggctaataaa ggaaatttat tttcattgca	1680
	atagtgtgtt ggaatttttt gtgtctctca ctcggaagga catatgggag ggcaaatcat	1740
	ttaaaacatc agaatgagta tttggtttag agtttgcaa catatgccca tatgctggct	1800
	gccatgaaca aaggttggct ataaagaggt catcagtata tgaacagcc cctgtctgtc	1860
	cattcciat tccatagaaa agccttgact tgaggttaga ttttttttat attttgtttt	1920



	gtgttatttt ttcttttaac atccctaaaa ttctcttac atgttttact agccagattt	1980
	ttctctctct cctgactact cccagtcata gctgtccctc ttctcttatg gagatccctc	2040
	gacggatcgg ccgaattcg taatcatgtc atagctgttt cctgtgtgaa attgttatcc	2100
	gctcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta	2160
	atgagtgagc taactcacat taattgcgtt gcgtcactg cccgctttcc agtcgggaaa	2220
	cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat	2280
	tgggcgctct tccgcttctt cgtcactga ctcgtcgcgc tcggtcgttc ggctgcggcg	2340
	agcggtatca gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc	2400
	aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggcgcggtt	2460
	gctggcgttt ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgtcaag	2520
	tcagaggtgg cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc	2580
[0059]	cctcgtcgcg tctctgttc cgacctgcc gcttaccgga tacctgtccg cttttctccc	2640
	ttcgggaagc gtggcgcttt ctcatagctc acgtgttagg tatctcagtt cgggtgtaggt	2700
	cgttcgtcc aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt	2760
	atccggtaac tategtcttg agtccaacct ggtaagacac gacttatcgc cactggcagc	2820
	agccactggt aacaggatta gcagagcgag gtaigttagc ggtgclacag agttcttgaa	2880
	gtggtggcct aactacggt acactagaag aacagtatit ggtatctgcg ctctgtgaa	2940
	gccagttacc ttcggaaaaa gagtggtag ctcttgatcc ggcaaaaaa ccaccgtgg	3000
	tagcggtggt tttttgttt gcaagcagca gattacgcgc agaaaaaag gatctcaaga	3060
	agatcctttg atcttttcta cggggtctga cgtcagtgg aacgaaaact cacgttaagg	3120
	gattttggtc atgagattat caaaaaggat ctccacctag atccttttaa attaaaaatg	3180
	aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt	3240

	aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact	3300
	ccccgtcgtg tagataacta cgatacggga gggettacca tctggcccca gtgctgcaat	3360
	gataccgga gacccaogct caccggctcc agatttatca gcaataaacc agccagccgg	3420
	aagggccgag cgcagaagtg gtcttgcaac tttatccgcc tccatccagt ctattaattg	3480
	ttgcgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat	3540
	tgctacaggc atcgtgggtg cagctcgtc gtttggtatg gcttcattca gctccggttc	3600
	ccaacgatca aggcgagtta catgatcccc catgtttgtc aaaaaagcgg gttagctcct	3660
	tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg	3720
	cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct gtgactgggtg	3780
	agtactcaac caagtcattc tgagaatagt gtatggcgcg accgagttgc tcttgccgg	3840
	cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa	3900
[0060]	aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt	3960
	aaccactcg tgcaccaaac tgatcttcag catcttttac ttccaccagc gtttctgggt	4020
	gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aaggcgaca cggaaatgtt	4080
	gaatactcat actcttcctt ttccaatatt attgaagcat ttatcagggt tattgtctca	4140
	tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat	4200
	ttccccgaaa agtgccacct aaattgtaag cgtaaatatt ttgttaaaat tcgcgttaaa	4260
	tttttgtaa atcagctcat tttttaacca ataggccgaa atcggcaaaa tcccttataa	4320
	atcaaaagaa tagaccgaga tagggttag tgttggtcca gtttgaaca agagtccact	4380
	attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc	4440
	actacgtgaa ccatcacct aatcaagttt tttggggtcg aggtgccgta aagcactaaa	4500
	tcggaacctt aaaggagcc cccgatttag agcttgacgg ggaaagccgg cgaacgtggc	4560
	gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa gtgtagcgg	4620

```

cacgctgcgc gtaaccacca caccgcgcgc gcttaatgcg ccgctacagg gcgcgtccca 4680

ttcgccattc aggctgcgca actgttggga agggcgatcg gtgcgggcct cttcgctatt 4740

acgccagctg gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa cgccagggtt 4800

ttcccagtca cgacgttgta aaacgacggc cagtgcgcgc gcgtaatacg actcactata 4860

gggcgaattg gagctccacc gcggtggcgg ccgct 4895

```

<210> 20  
 <211> 2622  
 <212> DNA  
 <213> 人巨细胞病毒

<220>  
 <221> sig\_肽  
 <222> (1)..(69)  
 <223> gH 肽信号

[0061]

<220>  
 <221> misc\_特征  
 <222> (1)..(2622)  
 <223> HCMV gH - HCMV gB TM/CTD 核苷酸序列

<220>  
 <221> misc\_特征  
 <222> (2151)..(2622)  
 <223> gB TM-CTD

```

<400> 20
atgcggccag gcctcccctc ctacctcacc gtccctgccg tctgtctcct cagccaccta 60

ctttcgtcac gatatggcgc agaagccata tccgaaccgc tggacaaagc gtttcaccta 120

ctgctcaaca cctacgggag acccatccgc ttctgcgtg aaaacaccac ccagtgtacc 180

tacaatagca gcctccgtaa cagcacggtc gtcagggaaa acgccaicag tticaacttt 240

ttccaaagct ataatcaata ctatgtatc catatgcctc gatgtctttt tgcgggtcct 300

ctggcgggagc agtttctgaa ccaggtagat ctgaccgaaa ccttggaaag ataccaacag 360

```

	agacttaaca cttacgcgt ggtatccaaa gacctggcca gctaccgate tttttcgag	420
	cagctaaagg cacaggacag cctaggtgaa cagcccacca ctgtgccacc acccattgac	480
	ctgtcaatac ctcacgtttg gatgccaccg caaaccactc cacacggctg gacagaaatca	540
	cataccacct caggactaca ccgaccacac ttaaccaga cctgtatcct ctttgatgga	600
	cacgatctac tattcagcac cgtcacacct tgtttgcacc aaggtttta cctcatcgac	660
	gaactacgtt acgttaaaat aacactgacc gaggacttct tcgtagttac ggtgtccata	720
	gacgacgaca caccatgct gcttatcttc ggccatcttc cacgcgtact cttaaagcg	780
	ccctatcaac gcgacaactt tatactacga caaactgaaa aacacgagct cctgggtcta	840
	gttaagaaag atcaactgaa ccgtcactct tatctcaaag acccgactt tcttgacgc	900
	gcacttgact tcaactacct ggacctcagc gcactactac gtaacagctt tcaccgttac	960
	gccgtggatg tactcaaaag cggtcgatgt cagatgctgg accgccgcac ggtagaaatg	1020
[0062]	gccttcgcct acgcattagc actgttcgca gcagcccgac aagaagaggc cggcgcccaa	1080
	gtctccgtcc caggggccct agaccgccag gccgcactct tacaataca agaatttatg	1140
	atcacctgcc tctcacaac accaccacgc accacgttgc tgetgtatcc cagggccgtg	1200
	gacctggcca aacgagccct ttggacaccg aatcagatca ccgacatcac cagcctcgta	1260
	cgcctggtct acatactctc taaacagaat cagcaacatc tcatcccca gtgggcaata	1320
	cgacagatcg ccgactttgc cctaaaacta cacaaaacgc acctggcctc tttttttca	1380
	gccttcgcgc gtcaagaact ctacctatg ggcagcctcg tccactccat getagtacat	1440
	acgacggaga gacgcgaaat cttcatcgta gaaacgggcc tctgttcatt agccgageta	1500
	tcacacttta cgcagttgct agctcatccg caccacgaat acctcagcga cctgtacaca	1560
	ccctgttcca gtacggggcg acgcgatcac tgetcgaac gcctcacag tctcttcccc	1620
	gatgccaccg tccccactac cgttcccgcc gccctctcca tcctatctac catgcaacca	1680

	agcacgctag aaaccttccc cgacctgttt tgtctgccgc tcggcgaatc cttctccgcg	1740
	ctgaccgtct ccgaacacgt cagttatgtc gtaacaaacc agtacctgat caaaggatc	1800
	tcctaccctg tcctccaccac cgtcgtaggc cagagcctca tcctaccca gacggacagt	1860
	caactaaat gcgaactgac gcgcaacatg cataccacac acagcatcac agcggcgctc	1920
	aacatttccc tagaaaactg cgccttttgc caaagcgccc tactagaata cgacgacacg	1980
	caaggcgta tcaacatcat gtacatgcac gactcggacg acgtcctttt cgccttgat	2040
	ccctacaacg aagtgggtgt ctcctctccg cgaactcact acctcatgct ttgaaaaac	2100
	ggtacgggcc tagaagtaac tgacgtcgtc gtggacgcta ccgacagtcg ttccggagcc	2160
	ttcaccatca tcctcgtggc catagccgtc gtcattatca tttatttgat ctatactcga	2220
	cagcggcgctc tctgcatgca gccgtgcag aacctcttcc cctatctggt gtccgcgcac	2280
	gggaccaccg tgacgtcggg caacaccaa gacacgtcgt tacaggctcc gccttcctac	2340
[0063]	gaggaaagtg tttataattc tggctcgaaa ggaccgggac caccgtcgtc tgatgcatcc	2400
	acggcggctc cgccttacac caacgagcag gcttaccaga tgcttctggc cctgggccgt	2460
	ctggacgcag agcagcgagc gcagcagaac ggtacagatt ctttgacgg acagactggc	2520
	acgcaggaca agggacagaa gcccaacctg ctagaccgac tgcgacaccg caaaaacggc	2580
	taccgacact tgaaagactc cgacgaagan gagaacgtct ga	2622

<210> 21

<211> 1614

<212> DNA

<213> 鼠白血病病毒

<220>

<221> misc\_特征

<222> (1)..(1614)

<223> 密码子优化的 MMLV Gag 核苷酸序列

<400> 21

	atgggacaga ccgtcacaac acccctgagc ctgaccctgg gacattggaa agacgtggag	60
	aggatcgcac ataaccagag cgtggacgtg aagaaacgga gatgggtcac attctgcagt	120
	gctgagtggc caacttttaa tgtgggatgg ccccgagacg gcacittcaa cagggatctg	180
	atcaccagg tgaagatcaa ggtctttage ccaggacctc acggacatcc agaccagggtg	240
	ccttatatcg tcacctggga ggcactggcc ttcgatcccc ctccatgggt gaagccattt	300
	gtccacccaa aaccacctcc accactgcct ccaagtgcct cttcaactgcc actggaacca	360
	ccccggagca caccaccccg cagctccctg tatcctgctc tgactccatc tctgggcgca	420
	aagccaaaac cacagggtgt gagcgactcc ggaggaccac tgattgacct gctgacagag	480
	gacccccac cataccgaga tcctcggcct ccaccaagcg accgcgatgg aaatggagga	540
	gaggctactc ctgccggcga agccctgac ccattctcaa tggctagtag gctgcgcggc	600
	aggcgcgagc ctccagtggc agatagcacc acatcccagg ccttccctct gagggctggg	660
[0064]	ggaaatgggc agctccagta ttggccattt tctagttcag acctgtacaa ctggaagaac	720
	aataaccctt ctttcagiga ggaccccggc aaactgaccg ccttgatcga atccgtgctg	780
	attaccatc agcccatatg ggacgattgt cagcagctcc tgggcacctt gctgaccgga	840
	gaggaaaagc agcgcgtgct gctggaggct cgcaaagcag tccgagggga cgatggacgg	900
	cccacacagc tcctaataa ggtggacgcc gcttttccac tggaaagacc cgactgggat	960
	tatactaccc aggcagggag aaaccacctg gtccattaca ggcagctcct gctggcaggc	1020
	ctgcagaatg ccgggagatc cccaccaac ctggccaagg tgaaaggcat cacacagggg	1080
	cctaagtagt caccaagcgc ctttctggag aggetgaagg aagcttaccg acggtatacc	1140
	ccatacgacc ctgaggacce cggacaggaa acaaacgtct ccatgtcttt catctggcag	1200
	tctgccccag acattgggcg gaagctggag agactggaag acctgaagaa caagaccctg	1260
	ggcgacctgg tgcgggaggc tgaaaagatc ttcaacaaac gggagacccc cgaggaaaga	1320
	gaggaaagga tiagaaggga aactgaggaa aaggaggaac gccgacggac cgaggacgaa	1380

cagaaggaga aagaacgaga tcggcggcgg caccgggaga tgtcaaagct gctggccacc	1440
gtggtcagcg gacagaaaca ggacagacag ggaggagagc gacggagaag ccagctcgac	1500
agggatcagt gcgcatactg taaggaaaaa ggccattggg ccaaggattg ccccaaaaag	1560
ccaagaggac caagaggacc aagaccacag acatcactgc tgaccctgga cgac	1614

&lt;210&gt; 22

&lt;211&gt; 3300

&lt;212&gt; DNA

&lt;213&gt; 鼠白血病病毒

&lt;220&gt;

&lt;221&gt; misc\_特征

&lt;222&gt; (1)..(1614)

&lt;223&gt; 密码子优化的 MMLV Gag 核苷酸序列

&lt;220&gt;

[0065]

&lt;221&gt; misc\_特征

&lt;222&gt; (1615)..(3300)

&lt;223&gt; 密码子优化的 CMV pp65 核苷酸序列

&lt;400&gt; 22

atgggacaga ccgtcacaa acccctgagc ctgacctgg gacattggaa agacgtggag	60
aggatcgcac ataaccagag cgtggacgtg aagaaacgga gatgggtcac attctgcagt	120
gctgagtggc caacttttaa tgtgggatgg ccccgagacg gcactttcaa cagggatctg	180
atcaccaggg tgaagatcaa ggtcttttagc ccaggacctc acggacatcc agaccaggtg	240
ccttatatcg tcacctggga ggcactggcc ttgatcccc ctccatgggt gaagccattt	300
gtccacccaa aaccacctcc accactgcct ccaagtcccc cttcactgcc actggaacca	360
ccccggagca caccaccccg cagctccctg tatectgctc tgactccatc totgggcgca	420
aagccaaaac cacaggtgct gagcgactcc ggaggaccac tgattgacct gctgacagag	480
gacccccac cataccgaga tcctcggcct ccaccaagcg accgcgatgg aaatggagga	540

	gaggctactc ctgccggcga agcccttgac ccattccaa tggctagtag gctgcgcggc	600
	aggcgcgagc ctccagtggc agatagcacc acatcccagg ccttccctct gagggctggg	660
	ggaaatgggc agctccagta ttggccattt tctagttcag acctgtacaa ctggaagaac	720
	aataacccct ctctcagtag ggaccccggc aaactgaccg ccctgatcga atccgtgctg	780
	attaccatc agccacatg ggacgattgt cagcagctcc tgggcaccct gctgaccgga	840
	gaggaanaag agcgcgtgct gctggaggct cgcaaacgag tccgagggga cgatggacgg	900
	cccacacagc tccctaatga ggtggacgcc gcttttccac tggaaagacc cgactgggat	960
	tatactacc aggcagggag aaaccacctg gtccattaca ggacagctct gctggcaggc	1020
	ctgcagaatg ccgggagatc cccaccaac ctggccaagg tgaaaggcat cacacagggg	1080
	cctaattgagt caccaagcgc ctttctggag aggtgaagg aagcttaccg acggtatacc	1140
	ccatacgacc ctgaggacc cggacaggaa acaaacgtct ccatgtcttt catctggcag	1200
[0066]	tctgcccag acattgggag gaagctggag agactggaag acctgaagaa caagaccctg	1260
	ggcgacctgg tgcgggaggc tgaaaagatc ttcaacaac gggagacccc cgaggaaaga	1320
	gaggaanaag ttagaaggga aactgaggaa aaggaggaa gccgacggac cgaggacgaa	1380
	cagaaggaga aagaacgaga tcggcggcgg caccgggaga tgtcaaagct gctggccacc	1440
	gtggtcagcg gacagaaaca ggacagacag ggaggagagc gacggagaag ccagctcgac	1500
	agggatcagt gcgcatactg taaggaaaaa ggccattggg ccaaggattg ccccaaaaag	1560
	ccaagaggac caagaggacc aagaccacag acatcaactg tgaccctgga cgactgcgag	1620
	agccggggcc ggccgtgccc agaaatgac tctgtgctgg ggccattag tggacatgtg	1680
	ctgaaggccg tcttctccag gggagacacc ccgtgctgc ctcacgagac tcgactgctg	1740
	cagaccggca tccatgtgag ggtctcccag ccctctctga ttctgggtgc acagtataca	1800
	ccagatagca ctccctgcca cagaggagac aatcagctcc aggtgcagca tacctacttt	1860
	acaggctccg aggtcgaaaa cgtgtctgtc aatgtgcaca accctaccgg caggagcatc	1920



	tgtcctagcc aggagccaat gagcatctac gtgtacgccc tgcccttgaa gatgctgaat	1980
	atcccatcaa ttaacgtcca ccattaccct agcgagccg aacggaagca cagacatctg	2040
	ccagtggcgg acgtgtcat ccattgccagc ggcaaacaga tgtggcaggc aagactgacc	2100
	gtgtccgggc tggcctggac aaggcagcag aatcagtggc aggagcccg cgtgtactat	2160
	accagcgctt tcgtgttccc taccaaagac gtggccctga gacatgtggt gtgcgcacat	2220
	gagctggtgt gcagcatgga aaacactagg gccaccaaga tgcaggtcat cggcgatcag	2280
	tatgtcaaag tgiacctgga gattttttgc gaagacgtgc catcaggga gctgttcatg	2340
	catgtgacct tgggcagcga tgcgaggaa gacctgacca tgacaagaaa tccacagccc	2400
	tttatgagac cccacgagag gaatgggttc actgtgctgt gcccgaagaa catgatcatt	2460
	aagcctggaa aatcagtca tattatgctg gatgtggcct ttacatcaca cgagcatttc	2520
	ggactgctgt gcccgaatc catccctgga ctgagcattt ccggcaatct gctgatgaac	2580
[0067]	ggccagcaga tcttctgga agtgcaggcc atccgggaga ccgtcgaact gcgacagtat	2640
	gaccagtggt ctgcactgtt ctttttcgac atcgacctgc tgcctgcagc aggaccacag	2700
	tacagcgagc accctacttt tacctcccag tatcggtatc aggggaagct ggagtacagg	2760
	cacacctggg atcgccatga cgaaggagcc gtcaggggg acgatgacgt gtggacatct	2820
	ggcagtgatt cagacgagga actgggtgaca actgagcgaa aaaccccccg ggtgacagga	2880
	ggaggggcaa tggcaggggc cagcaccagc gcaggggcga agcgaaaaag cgccagcagc	2940
	gccacagcat gtaccgccgg cgtgatgact agaggaaggc tgaaggccga gtctacagtc	3000
	gtccccgagg aagatactga cgaggatagt gacaatgaaa tccacaacct cgcctgttc	3060
	acctggccac cttggcagc agggtattctg gctcgcaacc tggccccat ggtggcaacc	3120
	gtccagggac agaatttgaa gtatcaggag tttttctggg atgctaacga catctaccgg	3180
	atttttgcag agctggaagg cgtgtggcag ccagcagccc agcccaaacg acggagacat	3240
[0068]	cgacaggacg ctctgccagg acctgtatc gccagcacac caaagaagca caggggctaa	3300

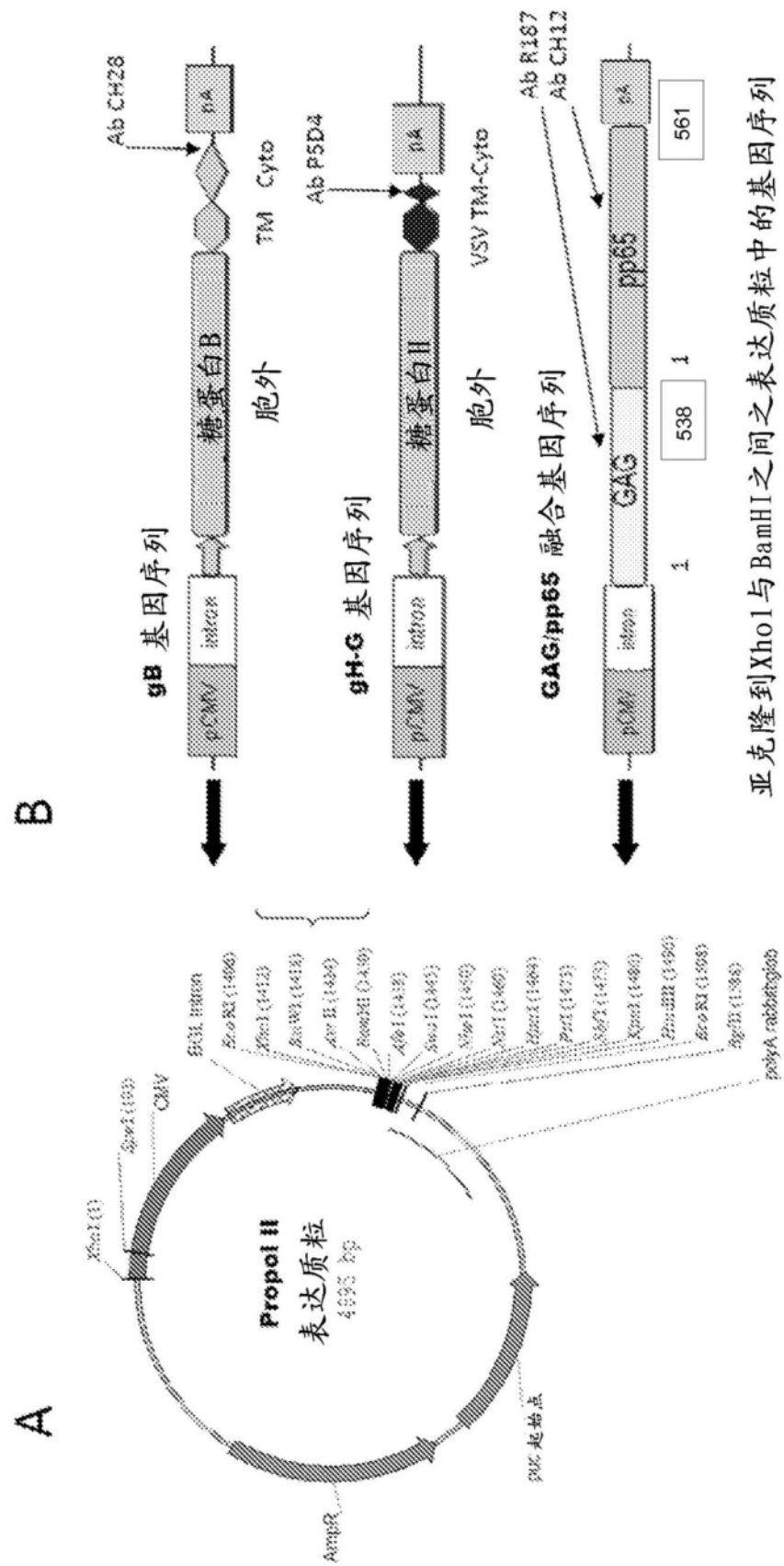


图1

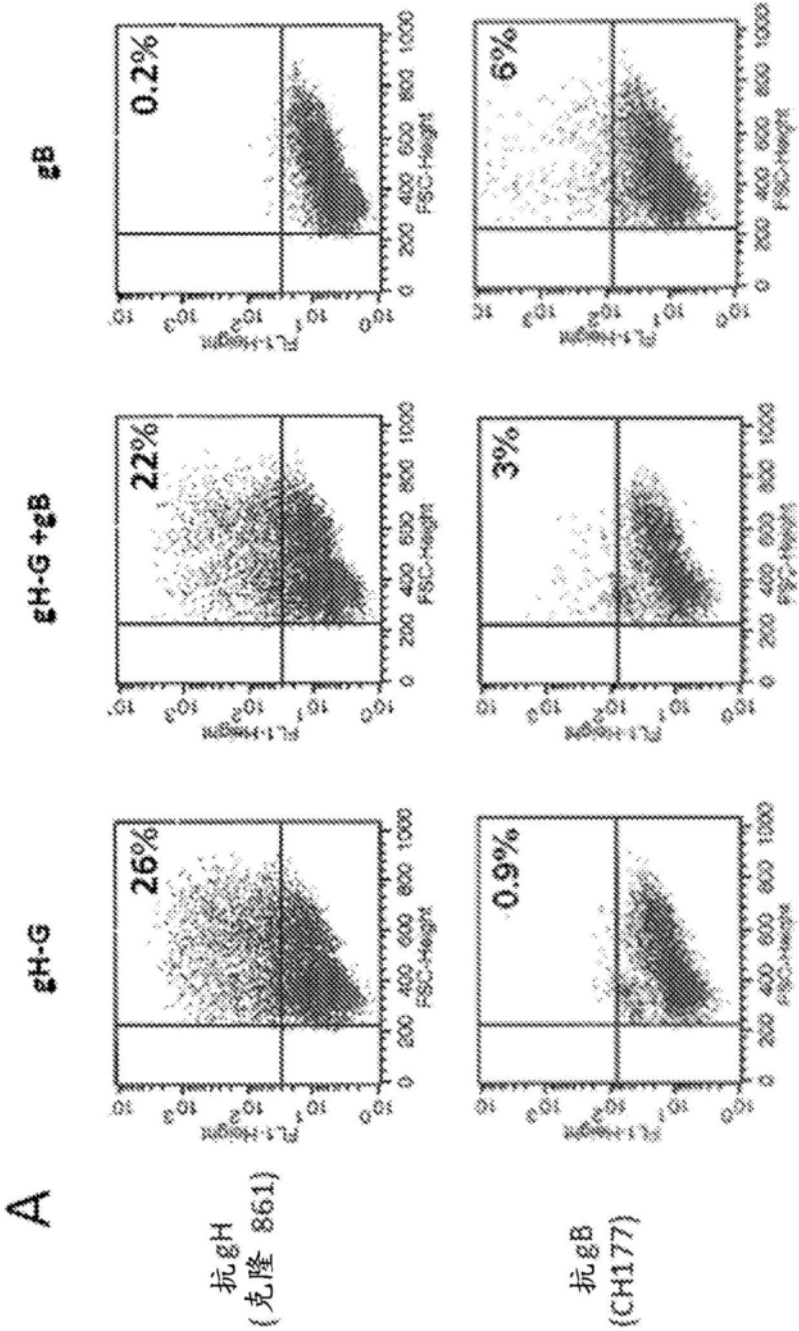


图2

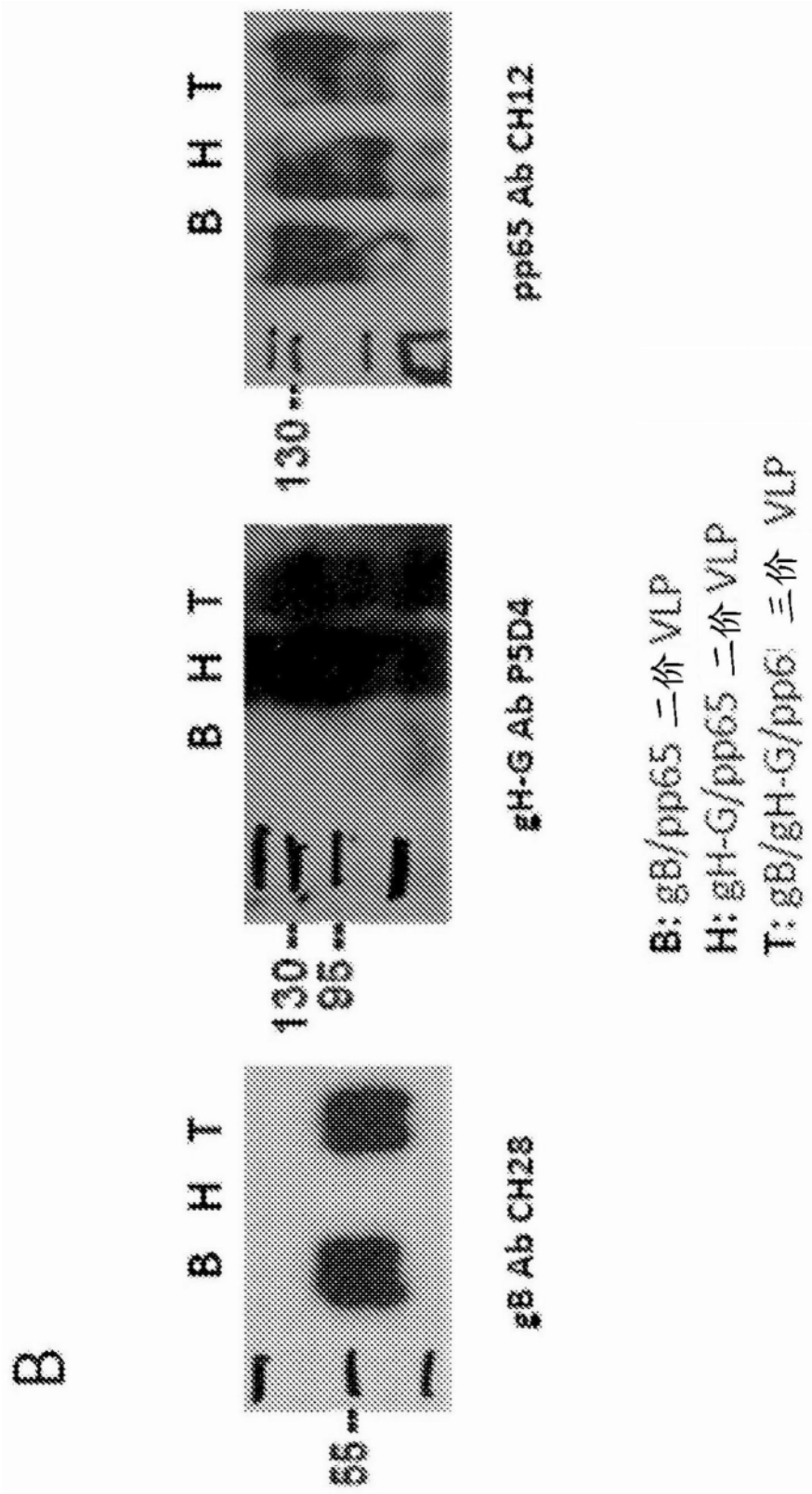


图2(续)

A

Z-平均 (4.000)	853.4	直径 (mm)	% 数目	宽度 (mm)
Peak 1	0.000	峰值 1:	100.0	37.82
Peak 2	0.000	峰值 2:	0.0	0.000
Peak 3	0.000	峰值 3:	0.0	0.000

结果质量: 好

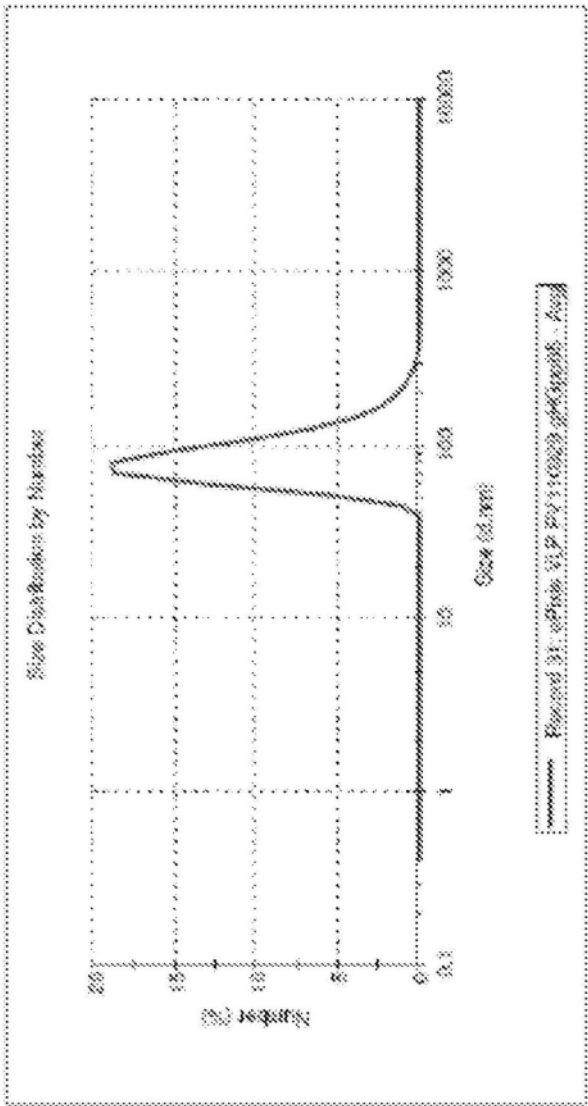


图3

B

Z-平均 (nm): 102.4		直径 (nm)	% 数目	宽度 (nm)
Pw: 0.214		峰值 1: 0.000	0.0	45.21
Mn: 0.000		峰值 2: 0.000	0.0	0.000
		峰值 3: 0.000	0.0	0.000

结果质量: 好

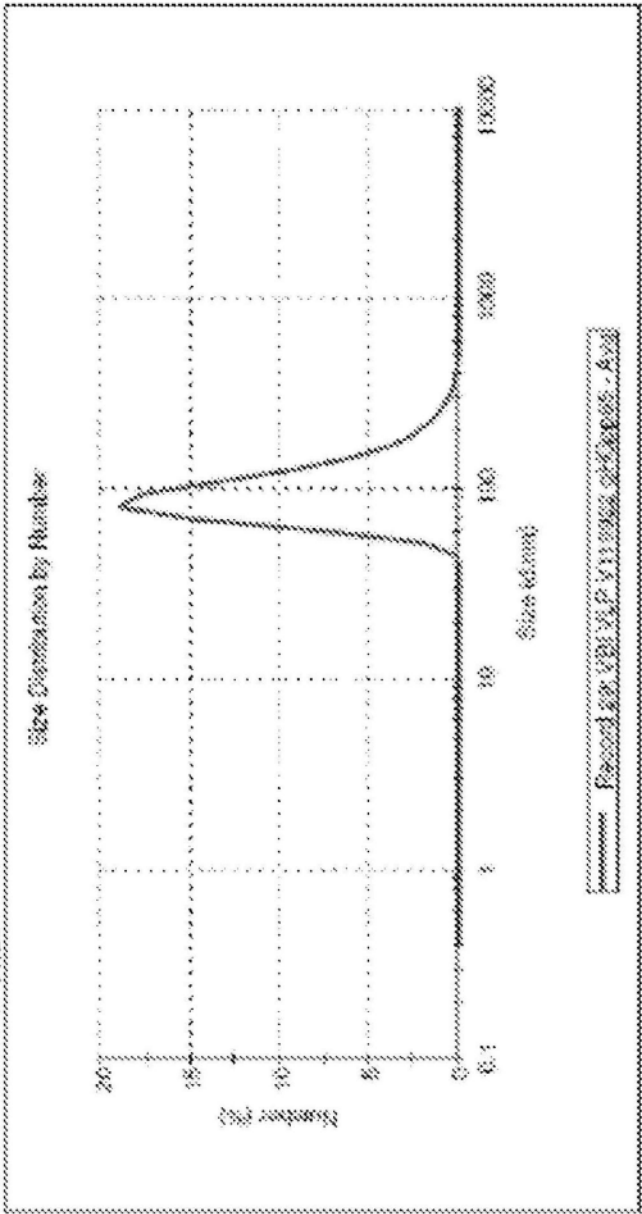


图3 (续)

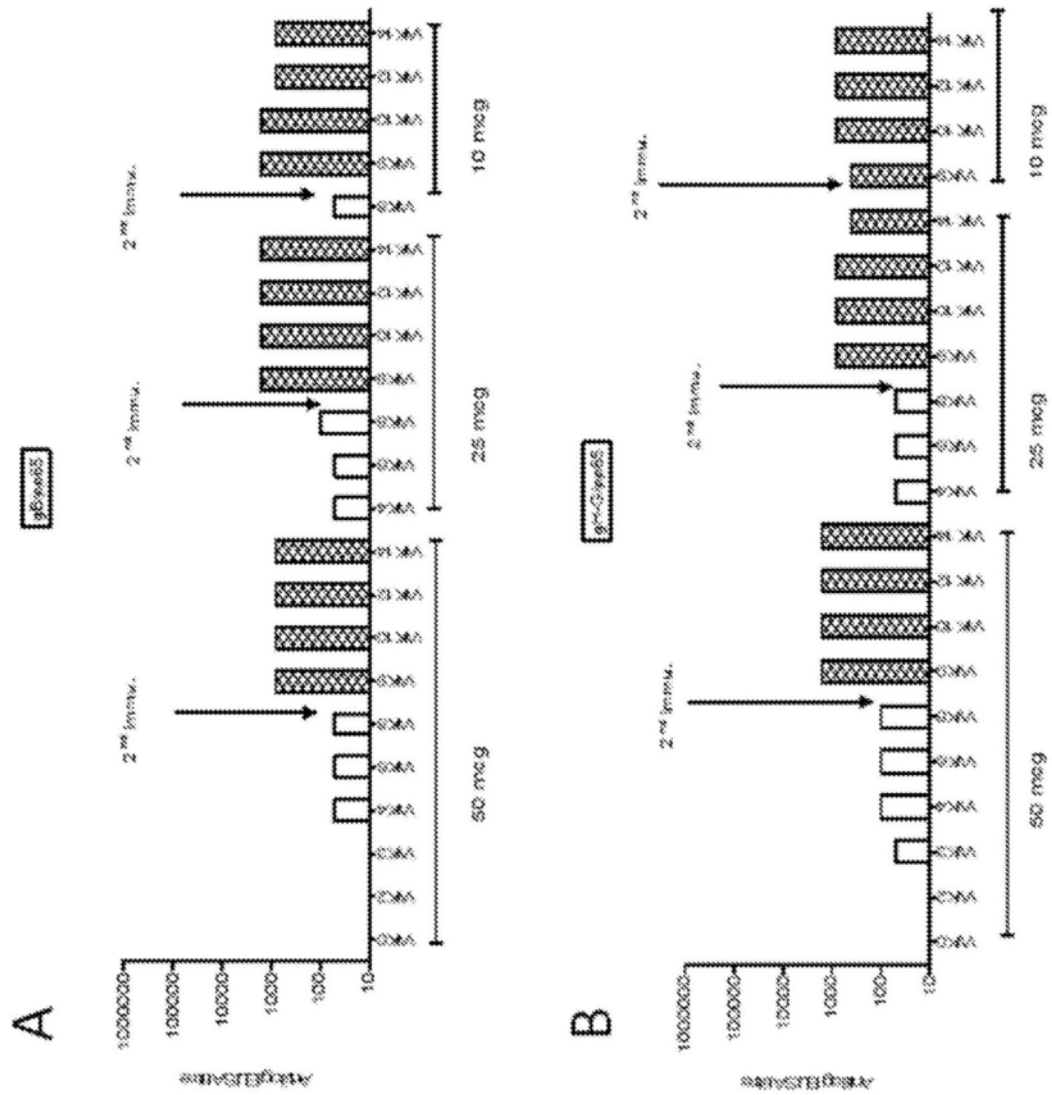


图4

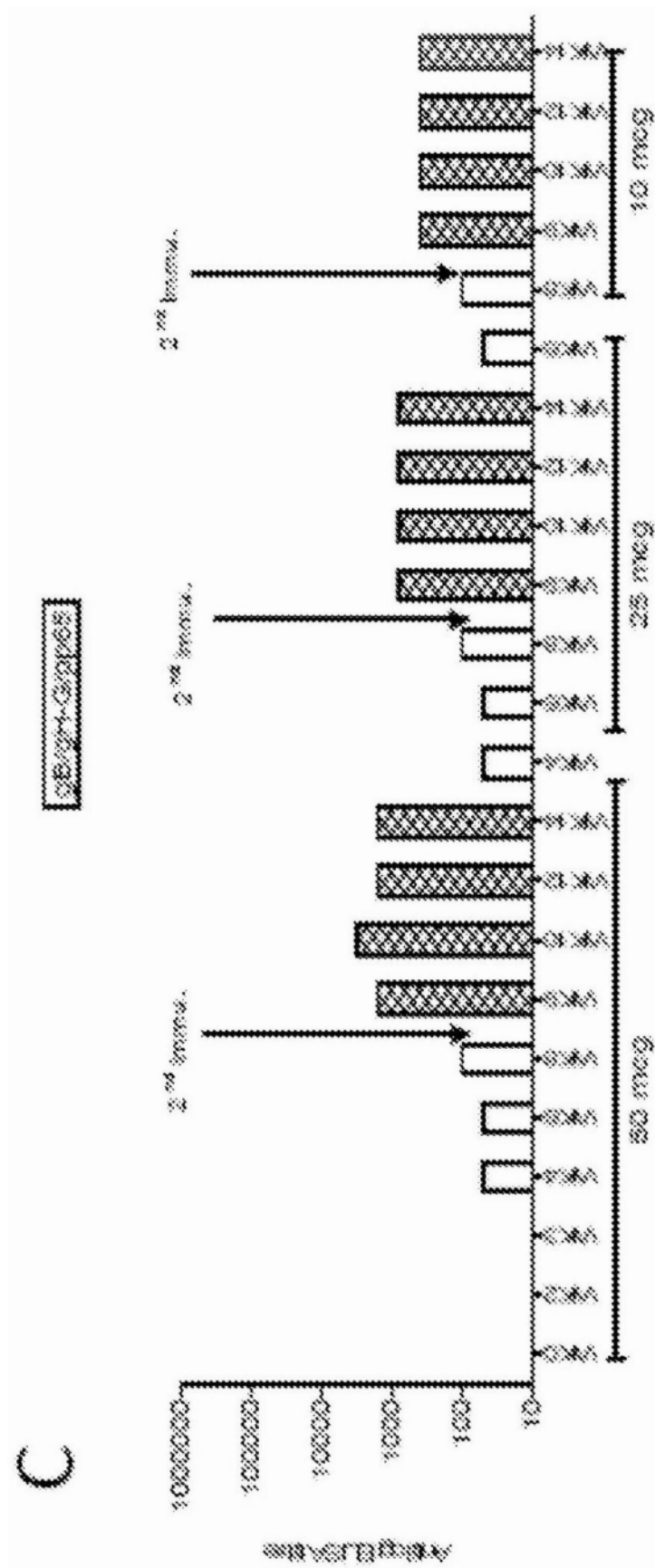


图4(续)



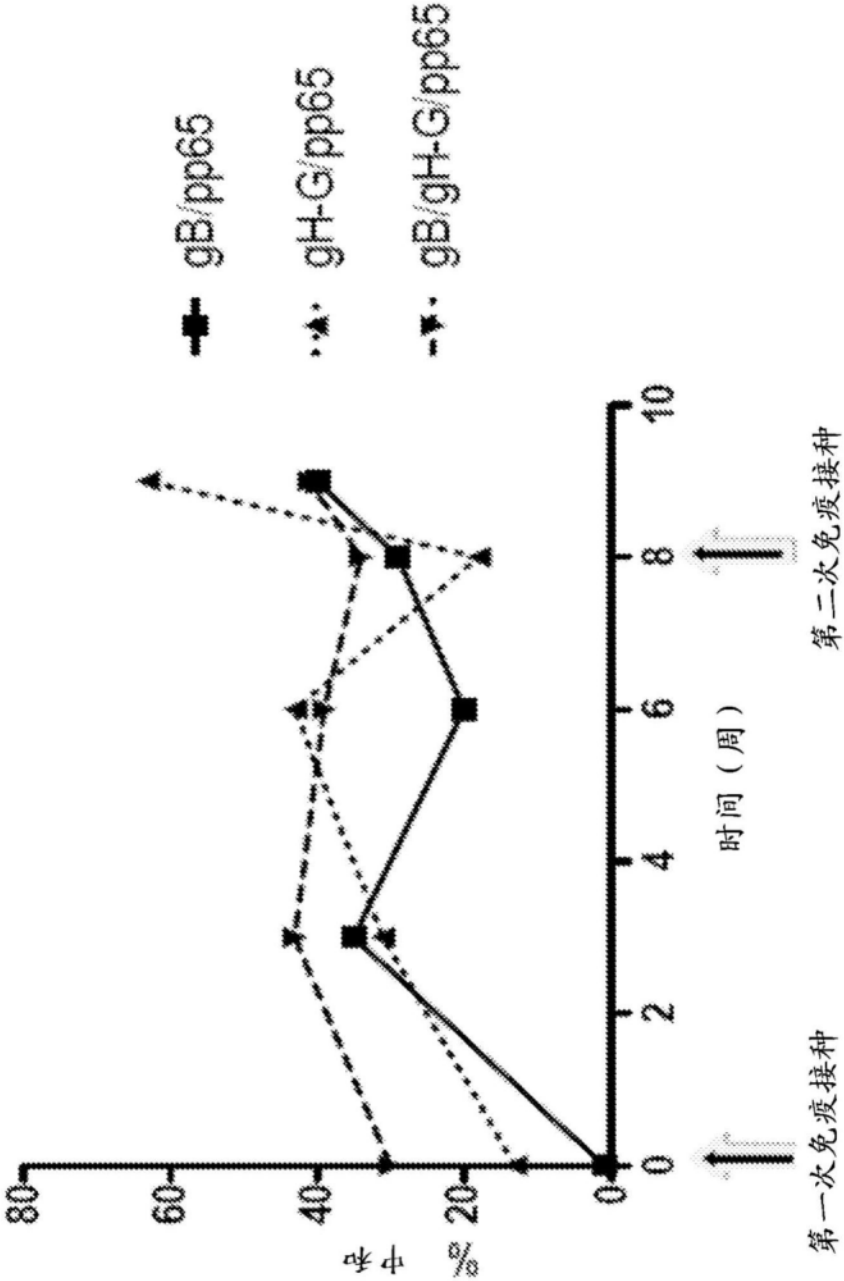


图5

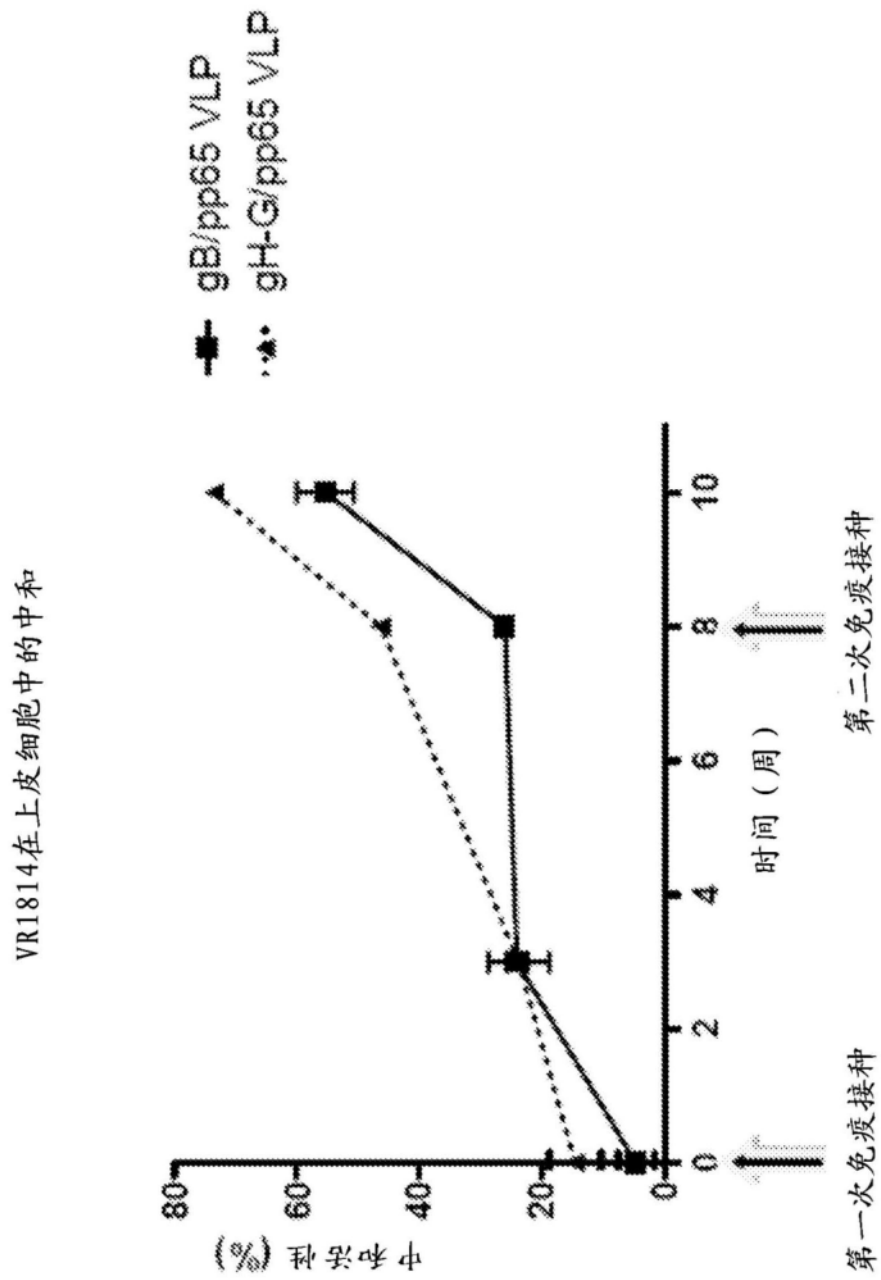


图6

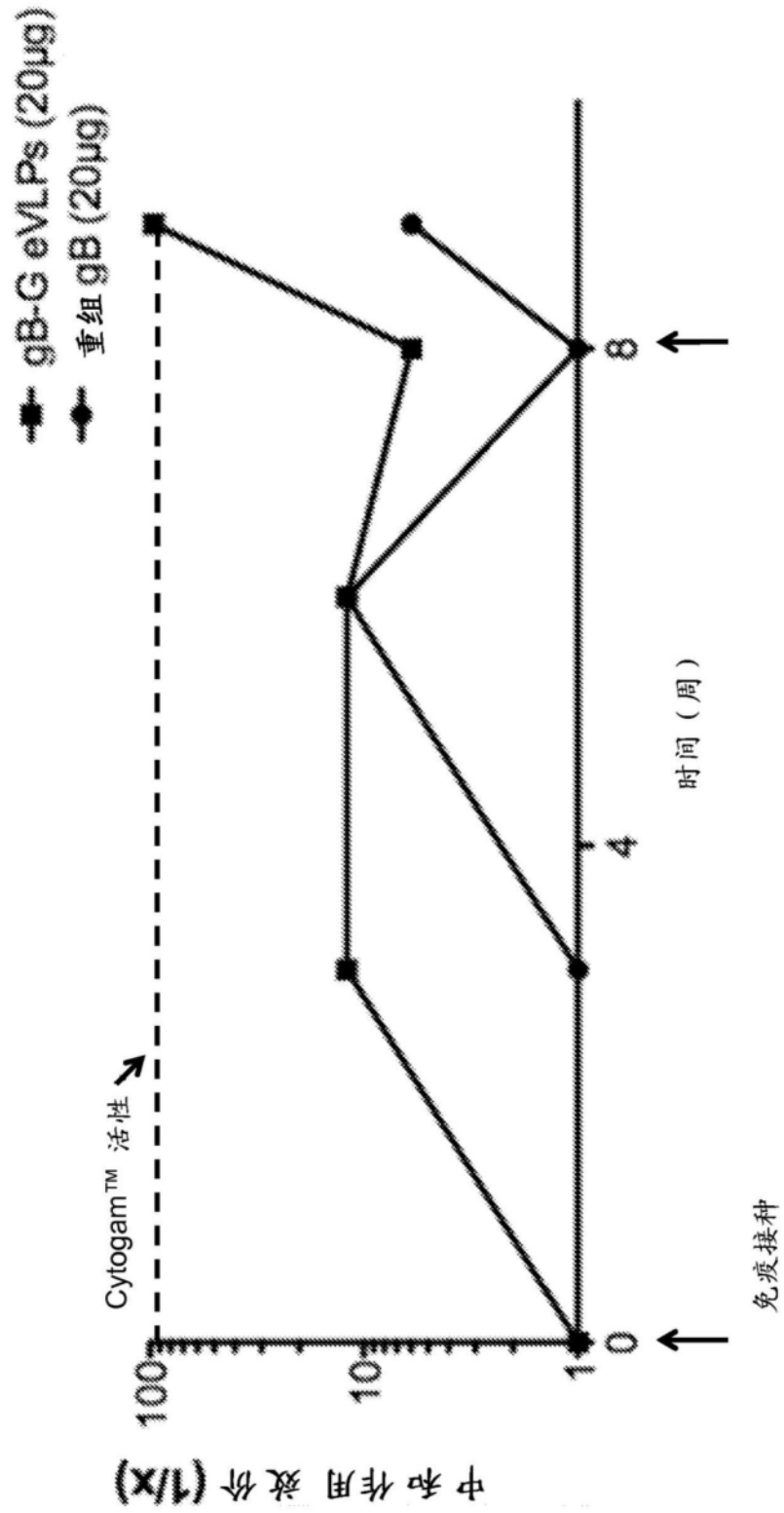


图7

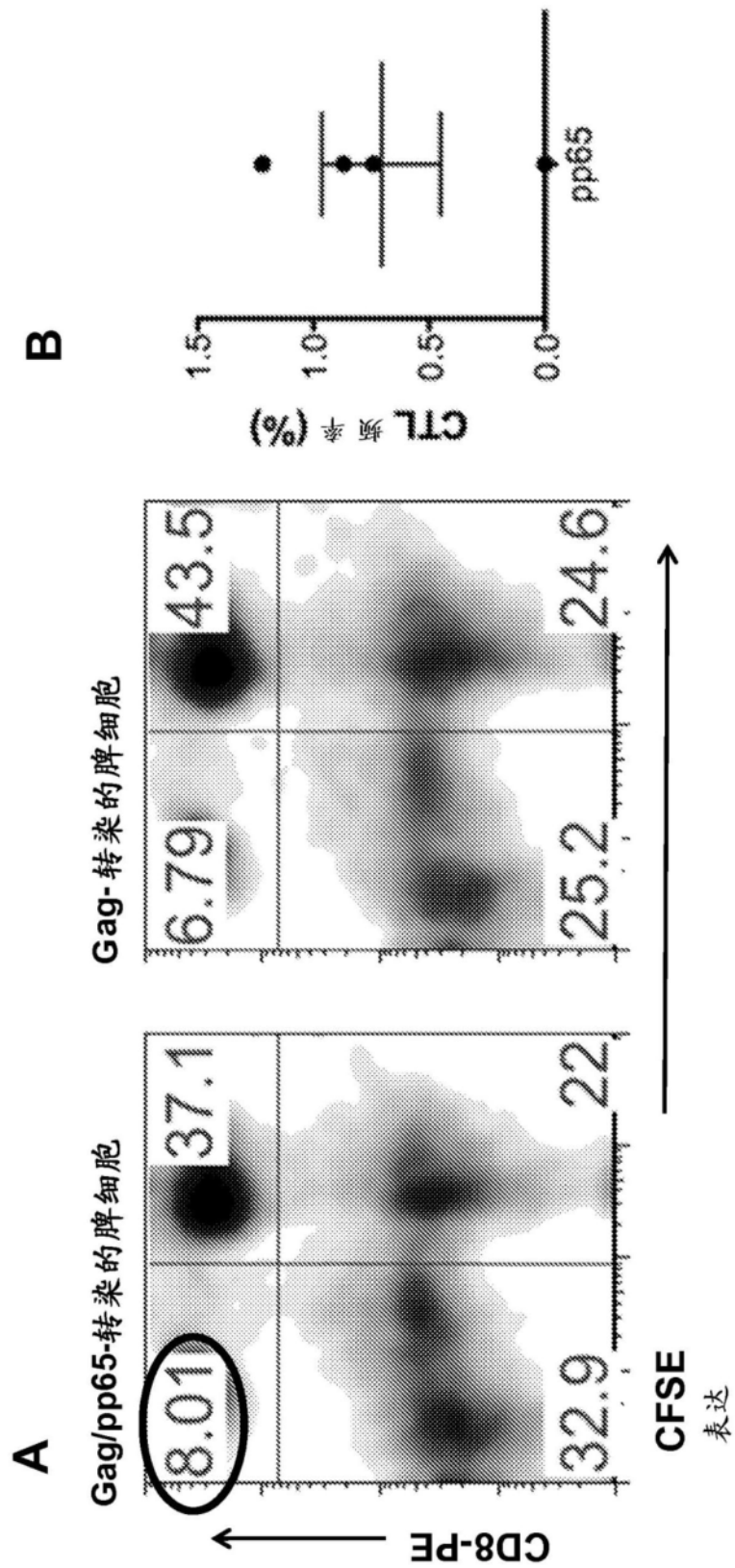


图8

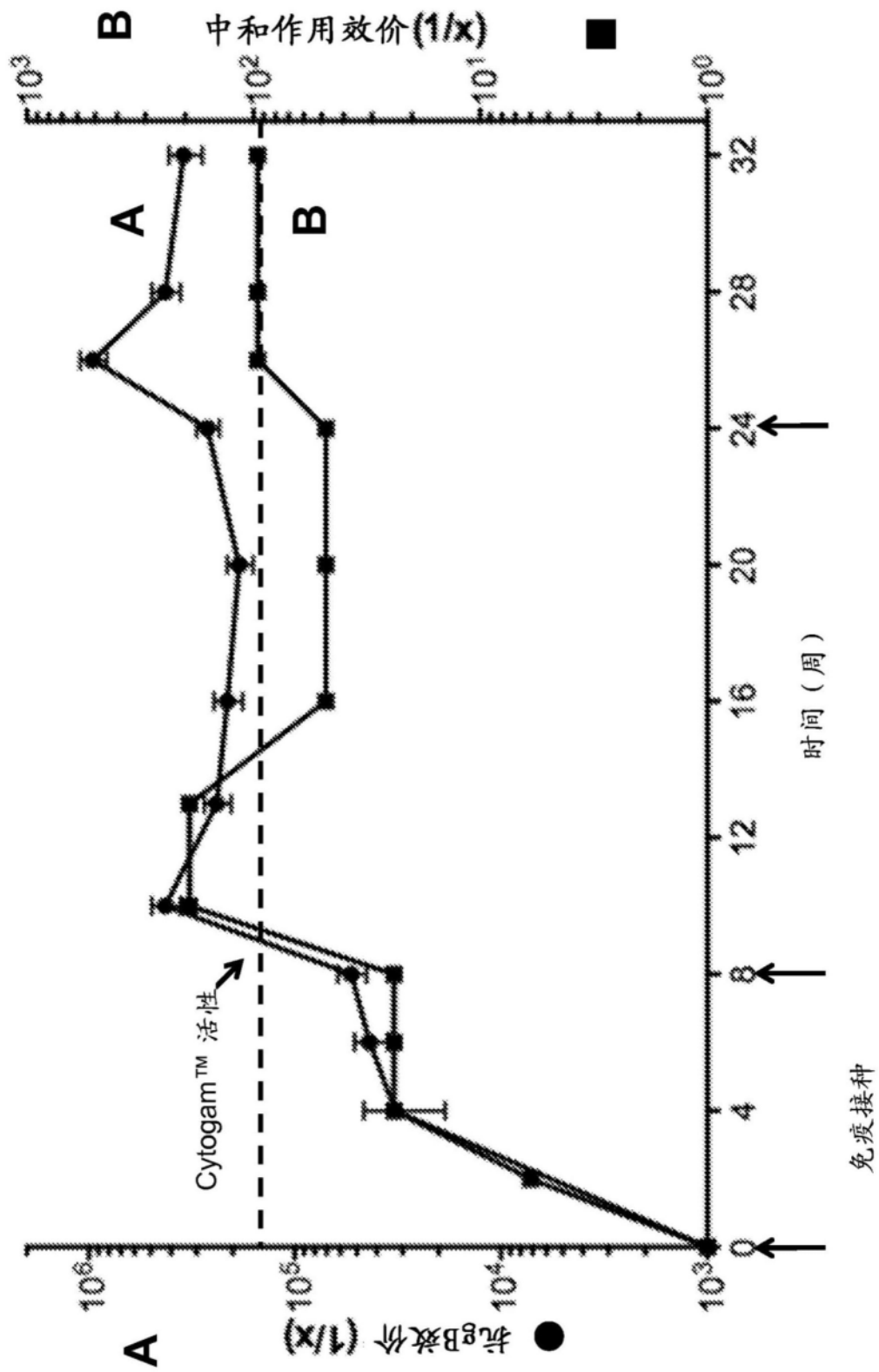


图9

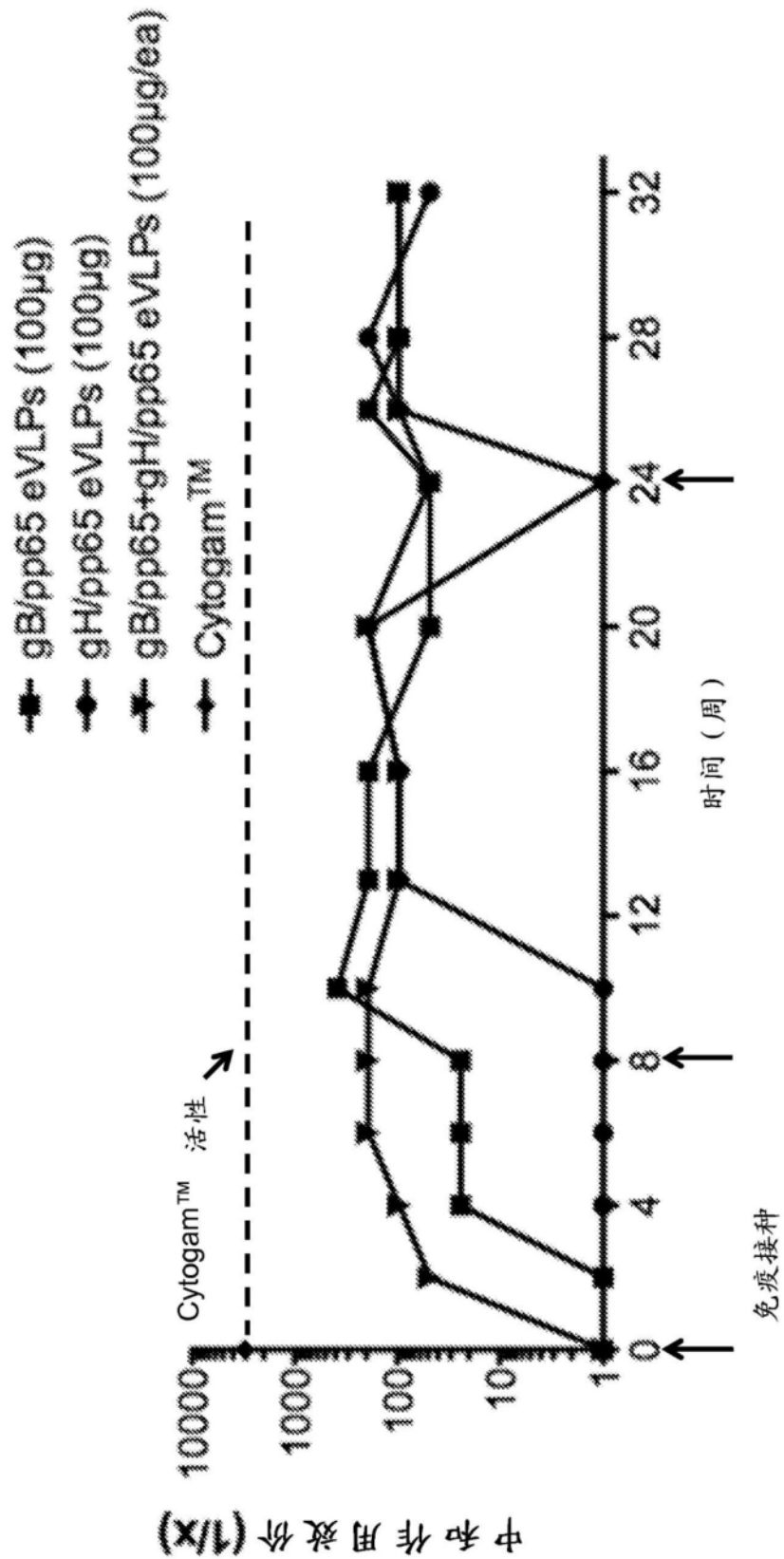
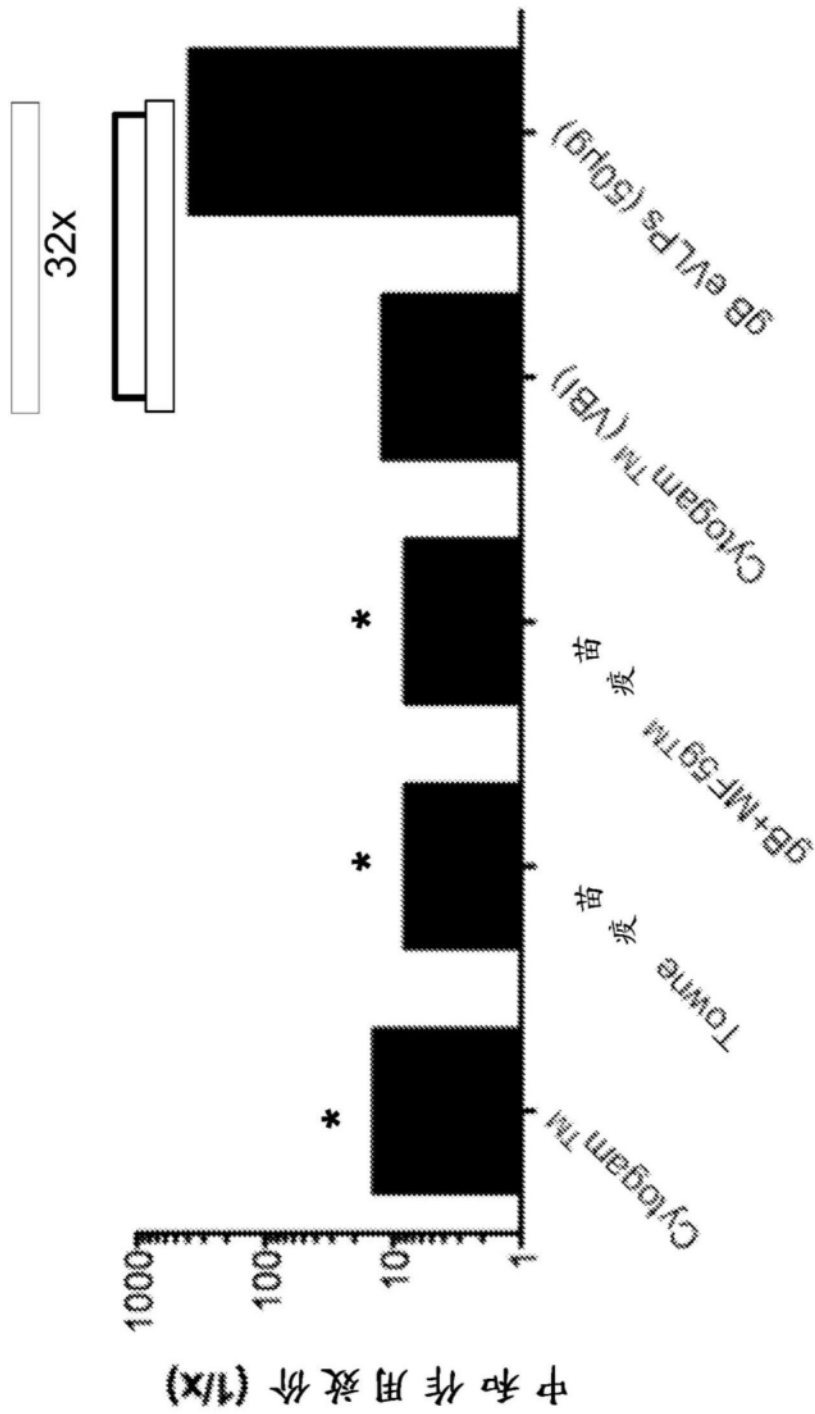


图10

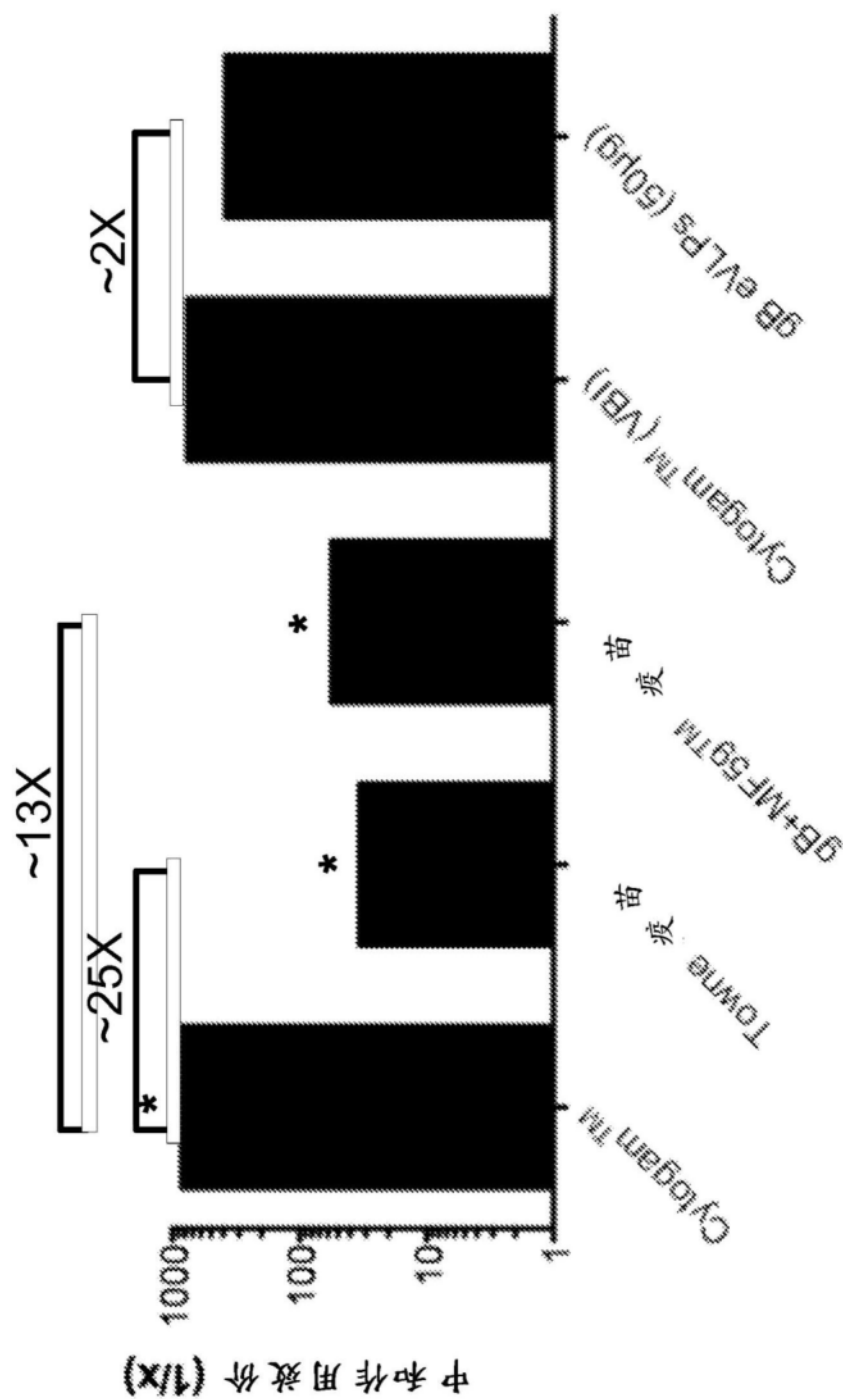




\* 数据来源: Cytogam™、活的减毒Towne疫苗和佐剂的gB亚单位疫苗的效力数据, 改编自Cui X等, 2008 Vaccine 26: 5760-5766

图12





\* 数据来源: Cytogam™、活的减毒Towne疫苗和佐剂的gB亚单位疫苗的效力数据, 改编自Cui X等, 2008 Vaccine 26: 5760-5766

图13

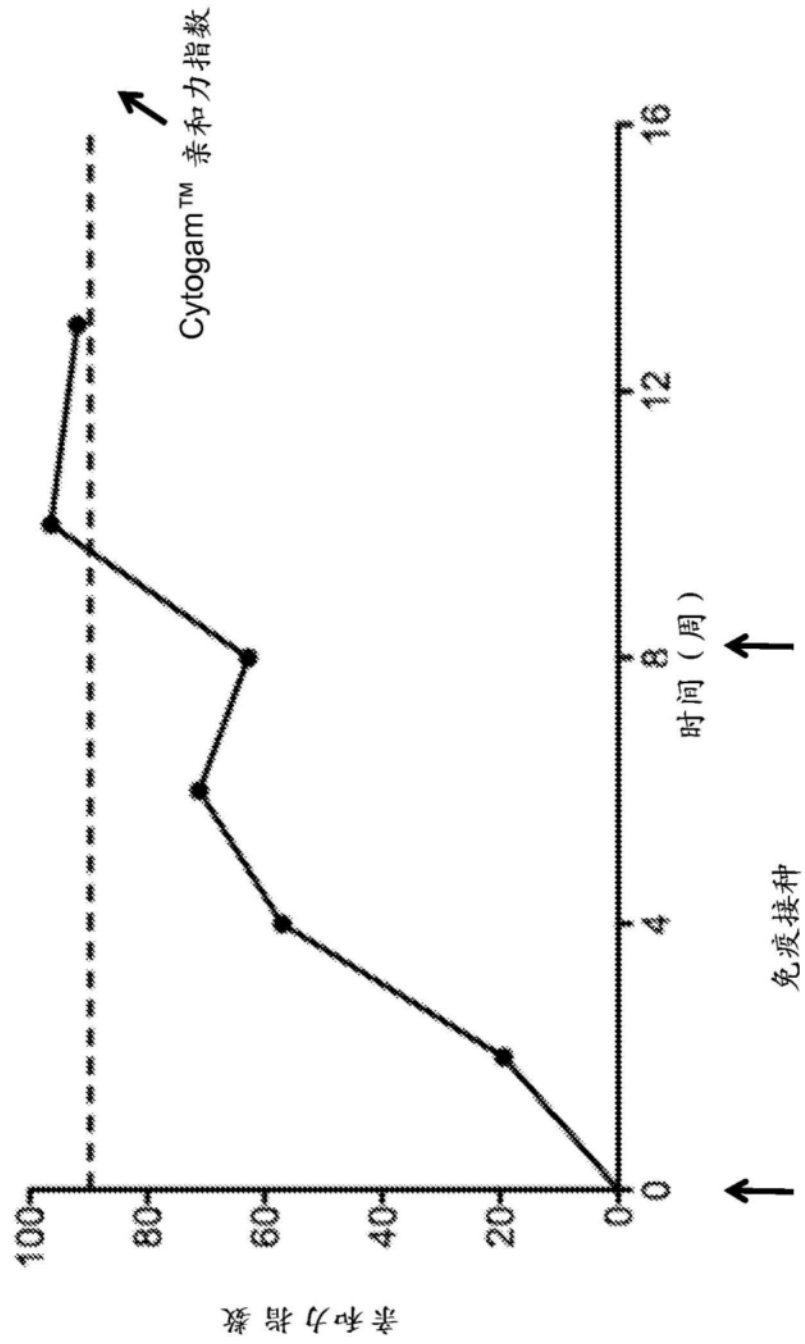


图14