



# [12] 发明专利申请公开说明书

[21]申请号 94193881.6

[51]Int.Cl<sup>6</sup>

C07H 17/00

[43]公开日 1996年10月16日

[22]申请日 94.8.19

[30]优先权

[32]93.8.20 [33]US[31]08 / 110522

[32]94.6.23 [33]US[31]08 / 264537

[86]国际申请 PCT / US94 / 09303 94.8.19

[87]国际公布 WO95 / 06055 英 95.3.2

[85]进入国家阶段日期 96.4.22

[71]申请人 史密丝克莱恩比彻姆公司

地址 美国宾夕法尼亚州

[72]发明人 A·G·迪莱拉

C·M·迪布克

[74]专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 姜建成

C12N 15 / 00 C12N 5 / 00

C12Q 1 / 70 A61K 38 / 00

A61K 35 / 14

权利要求书 4 页 说明书 47 页 附图页数 3 页

[54]发明名称 单纯性疱疹病毒-2UL26基因,衣壳蛋白,免疫测定和蛋白酶抑制剂

[57]摘要

公开了基本上纯的 HSV-2UL26 基因产物和其片段,包括成熟的 HSV-2 蛋白酶和其活性片段。公开了在实质上纯的 HSV-2UL26.5 基因产物和其片段,包括成熟的 HSV-2 衣壳蛋白和功能片段。公开了包括 HSV-2UL26 基因的全部或一部分和 / 或 HSV-2UL26.5 基因的分离到的核酸分子。公开了表达载体和包含这些核酸分子的宿主细胞。公开了鉴定抑制 HSV-2 蛋白酶活性的化合物的方法和鉴定抑制 HSV-2 病毒粒子装配的化合物的方法。公开了合成的 HSV-2 底物。公开了选择性地与 HSV-2 蛋白酶加工过的底物相结合但不与未加工的底物结合的抗体,或与未加工过的底物结合但不与加工过的底物结合的抗体。公开了区别 HSV-1 DNA 或蛋白与 HSV-2 DNA 或蛋白的方法和试剂盒以及在这些方法和试剂盒中有用的试剂。

(BJ)第 1456 号

# 权利要求书

---

1. 一个基本上纯的由HSV-2 UL 26基因编码的蛋白及其功能片段。

2. 权利要求1中基本上纯的蛋白，其中所说的蛋白选自由HSV-2蛋白酶前体蛋白，成熟的HSV-2蛋白酶以及所说成熟的HSV-2蛋白酶的功能片段所组成的组中。

3. 权利要求1中基本上纯的蛋白，其中所说的蛋白是成熟的HSV-2蛋白酶。

4. 一个基本上纯的由HSV-2 UL 26.5基因所编码的蛋白或其片段。

5. 权利要求4中基本上纯的蛋白，其中所说的蛋白选自由HSV-2衣壳前体蛋白，成熟的HSV-2衣壳蛋白和它的功能片段组成的组中。

6. 权利要求1中基本上纯的蛋白，其中所说的蛋白是成熟的HSV-2衣壳蛋白。

7. 分离到的包含一个HSV-2 UL 26基因或它的功能片段的核酸分子。

8. 权利要求7的分离到的核酸分子，包含一个SEQ ID NO: 1或它的一个功能片段的核苷酸序列。

9. 权利要求7的分离到的核酸分子，包含一个编码成熟HSV-2蛋白酶的核苷酸序列。

10. 权利要求7的分离到的核酸分子，包含一个HSV-2 UL 26.5基因或它的一个功能片段。

1 1 . 权利要求 1 0 的分离到的核酸分子，包含一个编码成熟 HSV-2 衣壳蛋白的核苷酸序列。

1 2 . 权利要求 1 0 的分离到的核酸分子，包含 HSV-2 26.5 启动子。

1 3 . 一个包含 HSV-2 UL 26 基因或它的功能片段的表达载体。

1 4 . 权利要求 1 3 的表达载体，其中所说 UL 26 基因在 SEQ ID NO: 1 中公开。

1 5 . 权利要求 1 3 的表达载体，其中所说 UL 26 基因的所说片段选自由编码成熟 HSV-2 蛋白酶的一个核苷酸序列，一个编码成熟 HSV-2 衣壳蛋白的核苷酸序列，一个编码 HSV-2 UL 26.5 基因的核苷酸序列，一个编码成熟 HSV-2 衣壳蛋白的核苷酸序列和 HSV-2 UL 26.5 启动子组成的组中。

1 6 . 已用权利要求 1 3 的表达载体转化的宿主细胞，所说的宿主细胞能表达所说的 UL 26 基因或它的功能片段。

1 7 . 一种鉴定抑制 HSV-2 蛋白酶活性的化合物的方法包括下列步骤：

a ) 在一种试验化合物存在下，用一个 HSV-2 蛋白酶底物与 HSV-2 蛋白酶或他的功能片段相接触；

b ) 检测所说底物的蛋白水解裂解的水平，以及

c ) 在 ( b ) 水平与在没有试验化合物存在下，用一个 HSV-2 蛋白酶底物与 HSV-2 蛋白酶或其功能片段接触时发生的蛋白水解活性的水平相比较。

1 8 . 一个鉴定抑制 HSV-2 病毒粒子装配的化合物的方法，包含

a ) 在一种受试化合物存在下，把二个或二个以上存在于一个受试化合物中的至少包含 HSV-2 衣壳蛋白部分的蛋白相接触，

b) 检测衣壳-衣壳联合的水平, 以及

c) 以所说的衣壳-衣壳联合的水平与在没有受试化合物存在下, 当二个或二个以上至少包括HSV-2衣壳蛋白的部分蛋白相接触时发生的衣壳-衣壳联合的水平相比较。

19. 一种合成的HSV-2蛋白酶底物, 它具有通式 $R_1$ -SEQ ID NO: 3- $R_2$  或 $R_1$ -SEQ ID NO: 4- $R_2$ 。

20. 权利要求19中合成的HSV-2蛋白酶底物, 选自SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 6; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 9; SEQ ID NO: 10; SEQ ID NO: 11; SEQ ID NO: 12; SEQ ID NO: 13; SEQ ID NO: 14和SEQ ID NO: 15。

21. 一种选择性地与一个未经加工的HSV-2蛋白酶相结合的抗体, 其中所说的抗体是不能与加工过的HSV-2底物相结合的。

22. 一种区别HSV-1 DNA和HSV-2 DNA的方法, 包括下列步骤:

a) 应用扩增HSV-1 DNA而非HSV-2 DNA的引物扩增样品中的DNA, 和/或应用扩增HSV-2 DNA而非HSV-1 DNA的引物扩增样品中的DNA,

b) 检测扩增DNA的存在。

23. 一组PCR引物, 包含只能用于扩增HSV-1 DNA而非HSV-2 DNA的核苷酸序列或只能用于扩增HSV-2 DNA而非HSV-1 DNA的核苷酸序列。

24. 一种区别HSV-1 DNA和HSV-2 DNA的试剂盒, 包括一个含有权利要求23的一组PCR引物的容器和一个含有一个DNA大小标志分子的容器。

25. 一种区别HSV-1蛋白和HSV-2蛋白的方法, 包括下列步骤:

a ) 应用能选择地与HSV-1蛋白结合但不能与HSV-2蛋白结合的抗体进行免疫测定, 和 / 或应用能选择地与HSV-2蛋白结合但不能与HSV-1蛋白结合的抗体进行免疫测定; 以及

b ) 检测结合抗体的存在。

2 6 . 一种能选择地与HSV-2蛋白结合但不能与HSV-1蛋白结合的抗体, 和一种能选择地与HSV-1蛋白结合而不能与HSV-2蛋白结合的抗体。

2 7 . 一种区别HSV-1蛋白和HSV-2蛋白的试剂盒, 包括一个含有权利要求 2 6 的一个抗体的容器和 / 或一个含有能选择地结合HSV-1蛋白而不能结合HSV-2蛋白的抗体和 / 或一个含有能选择地结合HSV-2蛋白而不能结合HSV-1蛋白的抗体的容器。

2 8 . 用权利要求 1 7 的方法鉴定的HSV-2蛋白酶抑制剂化合物。

2 9 . 用权利要求 1 8 的方法鉴定的抑制HSV-2病毒粒子装配的化合物。

# 说明书

---

单纯性疱疹病毒-2 UL 26基因，衣壳蛋白，免疫测定和蛋白酶抑制剂

## 有关申请的相互参考

这个申请是1993年8月20日提交的共同未决的美国专利申请系列号08/110,522的一个部分继续申请，该申请整个内容在此引入作为参考。

## 发明领域

本发明涉及HSV-2（单纯性疱疹病毒-2）UL26和HSV-2 UL26.5的基因；涉及基本上纯的HSV-2 UL 26和HSV-2 UL 26.5的基因各产物；涉及生产和利用HSV-2 UL 26和HSV-2 UL 26.5 DNA序列和基因产物的组成和方法。

## 发明背景

疱疹病毒由大约二十面体包封的内含一条双链基因组的病毒粒子组成。当前，六种人的疱疹病毒已被分离到，并已知它们从亚临床感染到致死病害状态的多种免疫受损状态有关。有一种人的疱疹病毒，单纯性疱疹病毒类型2，叫作HSV-2，常是通过性接触被传染的并引起生殖器疱疹。第二次生殖器疱疹复发率范围为每年1到6次之间。据估计，在美国有1千万到6千万人发生生殖器HSV-2感染。当今，还没有可以用来预防HSV-2感染的疫苗。

关于HSV-2的基因组组成所知甚少。然而，HSV-2引出了一个重

要的公共健康问题。多数个人继续被病毒感染而且还没有完全满意的抗病毒药剂或疫苗。需要有一种鉴定抗HSV-2剂的方法以及在这样方法中使用的试剂。需要有一种鉴定化合物的方法，该种化合物能调节HSV-2蛋白的活性和影响病毒在一个感染的细胞中复制和产生多重感染的病毒粒子的能力。还需要有区别HSV-2感染与其他疱疹病毒感染的方法和检别试剂盒。

### 发明概要

本发明涉及基本上纯的HSV-2 UL-26 基因产物和其片段，他们包括HSV-2蛋白酶前体蛋白，成熟的HSV-2蛋白酶及其活性片段，HSV 衣壳前体蛋白和成熟的HSV-2衣壳蛋白。

本发明涉及基本上纯的包括HSV-2衣壳前体蛋白和成熟的HSV-2衣壳蛋白的HSV-2 UL 26.5基因产物及其片段。

本发明涉及分离的包含HSV-2 UL-26 基因或它的部份的核酸分子，该基团或它的部份的核酸分子包括编码成熟的HSV-2蛋白酶和其活性片段的核酸分子以及编码前体或成熟的HSV-2衣壳蛋白，调节区域，例如启动子区域或其功能片段的核酸分子。

本发明涉及包含HSV-2 UL-26 基因或其一部分的表达载体，其中包括编码成熟HSV-2蛋白酶和它的活性片段的核苷酸序列，以及编码前体或成熟的HSV-2衣壳蛋白或它的功能片段的核苷酸序列。

本发明涉及含有包括HSV-2 UL-26 基因或他的一部分的表达载体的宿主细胞，其中包括编码成熟的HSV-2蛋白酶和他的活性片段的核苷酸序列和编码前体或成熟的HSV-2衣壳蛋白或他的功能片段的核苷酸序列。

本发明涉及包括HSV-2 UL 26.5基因或他的一部分的分离到的核酸分子，该UL 26.5 基因或他的一部分包含编码成熟的HSV-2衣壳蛋

白、调节区（例如启动子区或他的片段）的分离到的核酸，和编码前体或成熟的HSV-2衣壳蛋白或他的功能片段的核苷酸序列。

本发明涉及包括HSV-2 UL 26.5基因或他的一部分的表达载体，该UL 26.5基因或他的一部分包含编码成熟的HSV-2衣壳蛋白或他的片段的核苷酸序列和编码前体或成熟的HSV-2衣壳蛋白或他的功能片段的核苷酸序列。

本发明涉及含有包括HSV-2 UL 26.5基因或他的一部分的表达载体的宿主细胞，该UL 26.5基因或它的一部分包括编码成熟的HSV-2衣壳蛋白或他的片段的核苷酸序列和编码前体或成熟的HSV-2衣壳蛋白或他的功能片段的核苷酸序列。

本发明涉及鉴定抑制HSV-2蛋白酶活性的化合物的方法，包括在一种被试化合物存在下，用一个HSV-2蛋白酶的底物与HSV-2蛋白酶或他的活性片段接触以检测底物被蛋白酶裂解的水平，并与不存在被试化合物下酶解的水平相比较。

本发明涉及鉴定抑制HSV-2病毒粒子装配的化合物的方法，通过在一种被试化合物存在下与HSV-2衣壳蛋白的接触，检测衣壳-衣壳联合的水平，并以此水平与在没有被试化合物下发生的水平相比较。

本发明涉及用化学合成手段产生HSV-2蛋白酶底物，或重组生产断定是全部UL 26 基因产物或部分片段的HSV-2蛋白酶底物。

本发明涉及抗体，他选择性地结合HSV-2蛋白酶处理过的底物但不是没有处理过的底物，或选择性地与不经处理过的底物结合但不是经过处理的底物。

本发明涉及区别HSV-1 DNA和HSV-2 DNA间的方法，包括扩增HSV-1 DNA但不是HSV-2 DNA的引物用PCR 扩增DNA 和/或用扩增HSV-2 DNA但不是HSV-1 DNA的引物用PCR 扩增DNA。

本发明涉及扩增HSV-1 DNA但不是HSV-2 DNA的PCR 引物和扩增HSV-2 DNA但不是HSV-1 DNA的PCR 引物。

本发明涉及用以区别HSV-1 DNA与HSV-2 DNA的试剂盒，包括一个内装有扩增HSV-1 DNA但不是HSV-2 DNA的PCR 引物和一个阳性对照以及确定HSV-1 DNA是否通过引物扩增的分子大小标志的容器，和/或一个内装有扩增HSV-2 DNA但不是HSV-1 DNA的PCR 引物和一个阳性对照以及用以确定HSV-2 DNA是否已通过引物扩增的分子大小标志的容器。

本发明涉及区别HSV-1蛋白和HSV-2蛋白的方法包括采用选择性地与HSV-1蛋白但不是HSV-2蛋白相结合的抗体的免疫测定法和/或利用选择性地与HSV-2蛋白但不是HSV-1蛋白相结合的抗体的一个免疫测定法。

本发明涉及选择性地与HSV-1蛋白但不是HSV-2蛋白相结合的抗体或选择性地与HSV-2蛋白但不是HSV-1蛋白相结合的抗体。

本发明涉及用以区别HSV-1蛋白和HSV-2蛋白的试剂盒。所说的试剂盒包括一个有分隔空间以接受一系列严格规定的容器的运载箱，它装有一个包含选择性地与HSV-1蛋白但不是HSV-2蛋白相结合的抗体和一套检测抗体是否被结合到HSV-1蛋白上工具的第一容器，和/或一个包含选择性地与HSV-2蛋白但不是HSV-1蛋白相结合的抗体和一套检测抗体是否被结合到HSV-2蛋白上的工具的第二容器。

本发明涉及HSV-2蛋白酶启动子和/或增强子元件和他们的应用。

本发明涉及HSV-2衣壳蛋白启动子和/或增强子元件和他们的应用。

#### 附图简述

图1说明HSV-2 UL 26基团。符号<>表示HSV-2 UL 26基团产

物的限界线。公认的末端密码子为在字下划线。

符号 [[]] 表示 HSV-2 UL 26 基因产物的限界线。

符号 [] 表示二个主要蛋白水解位点的限界线。

裂开键用 \* 表示。

符号 || 表示 HSV-2 UL 26.5 基因的启动子区域，一个公认的“TATA 盒”用在字下划线表示。

图 2 说明在 HSV-2 UL 26.5 启动子调节下，氯霉素乙酰转移酶的表达。

#### 发明详述

这里所用的术语 UL 26 基因指的是包含一个编码 HSV-2 蛋白酶和一种 HSV-2 衣壳蛋白的核苷酸序列的一个 DNA 分子。UL 26 基因被公开在 SEQ ID NO: 1 (序列鉴定 1 号) 中。UL 26 基因的编码区域由 SEQ ID NO: 1 的 534-2447 间核苷酸组成。UL 26 基因表达时编码一个 638 个氨基酸的活性蛋白酶前体 (在 SEQ ID NO: 1 和 SEQ ID NO: 2 中公开)。

如在这里所用，术语“有活性的蛋白酶前体”指的是不经加工的 UL 26 翻译的产物。有活性的蛋白酶前体是一个有活性的 HSV-2 蛋白酶。当产生时，有活性的蛋白酶前体位在 247 和 248 氨基酸之间的一个内部的蛋白酶裂开位点上自剪切。氨基酸 247 个氨基酸部分仍保留蛋白酶活性。

在这里所用的，术语“成熟的蛋白酶”指的是有活性的蛋白酶前体的通过自剪切而产生的氨基酸 247 个氨基酸的蛋白。成熟蛋白酶的氨基酸序列以 SEQ ID NO: 1 和 SEQ ID NO: 2 的 1-247 氨基酸公开。

如在这里所用的，术语“HSV-2 蛋白酶”意思就是，交替地表示为有活性的蛋白酶前体，成熟的蛋白酶或其活性的片段。

如在这里所用的，术语“UL 26.5”基因指的是一个包含编码HSV-2衣壳蛋白核苷酸序列的DNA分子。UL 26.5基因是UL 26基因内被分开转录的一个内部序列。UL 26.5基因在SEQ ID NO: 1中公开，包括从1461-2447个核苷酸的编码区域。当表达时，UL 26.5基因编码一个329个氨基酸的衣壳前体（在SEQ ID NO: 1和SEQ ID NO: 2中以310-638间氨基酸公开）。

如在这里所用的，术语“衣壳前体”指的是没有加工过的UL 26.5翻译产物。在希望有关本发明基因产物的功能不被任何特殊机制理论约束的同时，但根据有关HSV-1的部分文献，可相信的是衣壳前体产生之后由HSV-2蛋白酶于SEQ ID NO: 1和SEQ ID NO: 2的613与614氨基酸残基之间的一个内部蛋白酶裂解位点上被切开。304氨基酸部分是用于病毒装配和病毒DNA包被的衣壳蛋白。UL 26.5的C末端加工促使病毒DNA包被成为成熟衣壳。这个加工事件的抑制会导致不能将DNA包装成成熟衣壳。

如在这里所用，术语“成熟的衣壳蛋白”指的是由HSV-2蛋白酶通过对衣壳前体的裂解所产生的304氨基酸蛋白。成熟衣壳蛋白的氨基酸序列作为SEQ ID NO: 1和SEQ ID NO: 2的310-613氨基酸被公开。

如在这里所用，术语“HSV-2衣壳蛋白”意思就是衣壳前体和成熟衣壳蛋白（可相互变换）。

如在这里所用，术语“功能片段”，当用于修饰特殊的基因或基因产物时，意指一个比基因或基因产物全长部分短的而其实际上保留有与全长基因相关连的有关全长基因产物的所有生物功能。要决定一个特殊基因或基因产物的一个片段是否是一个功能片段，只不过用熟知的核酸分解和蛋白分解的技术产生一个片段和测定如此产生的片段的生物学功能。

本发明涉及基本上纯的HSV-2蛋白酶，涉及其组成生产方法及应用，涉及编码HSV-2蛋白酶的核酸分子以及生产和应用此编码HSV-2蛋白酶核酸分子的方法。本发明涉及基本上纯的HSV-2衣壳蛋白，涉及其成分和生产和应用HSV-2衣壳蛋白的方法，涉及编码HSV-2衣壳蛋白的核酸分子，涉及产生和应用编码HSV-2衣壳蛋白的核酸分子的方法。本发明涉及用作HSV-2蛋白酶裂解的底物，涉及抑制HSV-2蛋白酶活性的化合物的鉴定方法。涉及抑制HSV-2衣壳装配的化合物的鉴定方法，涉及区别含有HSV-1 DNA的样本和含有HSV-2 DNA样本的方法，涉及区别含有HSV-1蛋白样本和含有HSV-2蛋白样本的方法，以及涉及用以完成这样方法的试剂，包括寡核苷酸和抗体。

本发明的一些实施方案提供了鉴定抑制或在其他方面调节HSV-2蛋白酶活性化合物的方法。因此，本发明提供了鉴定用作抗HSV-2剂的化合物的方法，因为HSV-2蛋白酶的活性是病毒生活史中必不可少的。根据本发明，在存在有一个要测定的化合物时，HSV-2蛋白酶与一个HSV-2蛋白酶底物（底物）相接触以确定受测定化合物对蛋白分解活性有或没有影响。受试化合物对HSV-2蛋白酶的影响也可用存在有受试化合物时的蛋白分解活性与不存在有受试化合物时将被观察到的蛋白分解活性相比较而决定。

蛋白分解活性指的是HSV-2蛋白酶酶解加工底物成产物的能力，即分解一个单一的底物肽分子成两个或多个肽分子（蛋白分解产物）。在病毒生活史中，通过这样的蛋白裂解，蛋白酶前体被加工成成熟的蛋白酶，衣壳前体被加工成成熟的衣壳。这种变换对病毒粒子装配和病毒DNA包装是必须的。蛋白分解活性的水平可被在本领域中那些具有常规技术人员所熟知的多种方法进行确定。基本上提供了一种把没有加工的底物与蛋白分解产物区别开的方法。因此，底物加入以后，

HSV-2蛋白酶与受试化合物相接触，通过检测保留的没有加工过的底物量，没有加工过底物的消耗量，或蛋白分解产物的生成量就能观察到蛋白分解活性的水平。

本发明提供了基本上纯的HSV-2蛋白酶，它在鉴定调节HSV-2蛋白酶活性的化合物的分析中是有用的。本发明也提供了产生基本上纯的HSV-2蛋白酶的方法。HSV-2蛋白酶的氨基酸序列在SEQ ID NO: 1和SEQ ID NO: 2中公开。如上所述，638 氨基酸的活性蛋白酶前体也在SEQ ID NO: 1和SEQ ID NO: 2中公开。有活性的蛋白酶前体是一个有活性的HSV-2蛋白酶，它是通过在氨基酸残基247 和248 之间的一个蛋白酶内部裂解位点的自动裂解而进行加工以产生一个247 个氨基酸的蛋白称为成熟的蛋白酶。通过常规的肽的合成方法或者通过用在SEQ ID NO: 1中提供的信息，应用重组DNA 技术可以生产纯的有活性的蛋白酶前体，成熟蛋白酶和它的活性片段。应用标准的操作和容易得到的起始材料，本领域中具有一般技术者就能生产HSV-2蛋白酶。另一方面，用标准的操作和容易得到的起始材料，本领域中具有一般技术者就能确定一个片段和/或有蛋白酶活性的前体的衍生物或成熟的蛋白酶是否为一个有活性的片段。

确定一个蛋白或肽能不能裂解一个特殊底物的测定法在此公开。要确定一个HSV-2蛋白酶片段是否有蛋白分解活性，则本领域中具有一般技术者就能如在此所叙述的没有试验化合物条件下，用蛋白酶的片段或衍生物代替与SEQ ID NO: 2相同的蛋白酶完成蛋白酶活性测定。如果片段或衍生物裂解底物，则他是有活性的，即这片段或衍生物拥有蛋白分解活性。因此，本领域中具有一般技术者就能常规地确定蛋白酶的一个片段或衍生物是一个有活性的片段或一个有活性的衍生物。

本发明涉及编码HSV-2蛋白酶的核苷酸序列和涉及编码HSV-2衣壳蛋白的核苷酸序列。UL 26 基因包括编码HSV-2蛋白酶和HSV-2衣壳蛋白的一个前体形式的核苷酸序列在SEQ ID NO: 1中被公开。UL 26 基因包括一个编码HSV-2衣壳蛋白的核苷酸序列也在SEQ ID NO: 1中被公开。在本领域中具有一般技术者，应用标准技术和容易得到的起始材料，利用在这里公开包括SEQ ID NO: 1的信息可以得到或合成一个编码HSV-2蛋白酶的核酸分子，或一个编码HSV-2衣壳蛋白的核酸分子。而且，应用标准技术、容易得到的起始材料和在这里公开的包括SEQ ID NO: 1的信息，在本领域中一个具有一般技术者就能生产基本上纯的HSV-2蛋白酶，包括有活性的前体蛋白酶，成熟蛋白酶或有活性的HSV-2蛋白酶片段。同样地，应用标准的技术，容易得到的起始材料和在这里公开的包括SEQ ID NO: 1的信息，在本领域中一个具有一般技术者就能生产基本上纯的HSV-2衣壳蛋白包括衣壳前体，成熟的衣壳或能装配成功能片段的HSV-2衣壳片段。本领域中一个具有一般技术者，应用标准的技术和容易得到的起始材料，能利用在这里公开的包括在SEQ ID NO: 1的信息利用表达系统中给定宿主细胞以产生理想且白质产品的密码子获得或合成编码HSV-2蛋白酶或HSV-2衣壳蛋白的核酸分子。

在本领域中那些具有一般技术者利用不同的技术和适当的实验也能产生编码HSV-2蛋白酶或HSV-2衣壳蛋白的核酸分子。例如，应用多聚酶链反应（PCR）方法学。可设计引物用以产生编码HSV-2蛋白酶或HSV-2衣壳蛋白的多拷贝核苷酸序列。通过病毒DNA的扩增也可能常规地得到编码有活性的蛋白酶前体的完整的核苷酸序列。同样，通过病毒DNA的扩增也可以常规地获得编码成熟蛋白酶的核苷酸序列。同样地，通过病毒DNA的扩增，也可以常规地得到编码一个有活性的

HSV-2蛋白酶片段的核苷酸序列。在一个相同的情况中，通过病毒DNA的扩增也可以常规地得到编码衣壳前体，成熟衣壳或它的功能片段的完整核苷酸序列。换一种方法，利用限制性酶，编码HSV-2蛋白酶的DNA，包括有活性的蛋白酶前体，成熟蛋白酶，或其有活性片段或HSV-2衣壳蛋白包括衣壳前体，成熟衣壳或其功能片段也可从把病毒DNA克隆到载体中得到，用从公开的核苷酸序列设计的探针通过杂交后鉴定。此外，编码HSV-2蛋白酶或HSV-2衣壳蛋白的核酸分子也可以用在本领域中那些具有一般技术者所熟知的技术进行合成。为HSV-2蛋白酶或HSV-2衣壳蛋白编码的密码子可用于优化选作HSV-2蛋白酶或HSV-2衣壳蛋白主组生产的宿主细胞的蛋白产品。HSV-2基因组有高富含量的G+C核苷酸。这一点对编码HSV-2蛋白酶的UL26基因来说是特别正确的。这样高含量G+C的特征在E. coli (大肠杆菌)中过度表达基因中由于密码子使用和移码突变机会的增加提出了一个问题。在努力改进UL 26在E. coli中表达中，变更UL 26基因和其片段以提供在E. coli中表达仍保留蛋白酶的可靠的氨基酸序列的优选密码子。有关优选密码子使用的参考文献是：Wada等，(1992)“从基因银行遗传序列资料中制作成的密码子使用表”，Nucleic Acid Research, Vol. 20 Supplement, 2111-2118页，在这里以参考文献编入。密码子的最适使用已为大家所熟知，并根据本发明能应用于设计核酸分子，使其在一个选择过的宿主中能被表达到一个效率增高的水平。

在本领域中一个具有一般技术者，应用大家熟知的技术能把这样的DNA分子插入到像商业上可以得到的熟知作表达系统用的载体那样的载体中。例如，商业上可以得到的象psE 420 (Invitrogen, San Diego, CA)或pET-16 (b) (Novagen, Modison N.I.)可以用为在

*E. coli*中生产HSV-2蛋白酶。商业上可以得到的质粒pYES2(Invitrogen, San Diego, CA)例如,在酵母菌的*S. cerevisiae*(啤酒酵母菌)也可用作生产。商业上可以得到的MAXBAC™完全的杆菌状病毒表达系统(Invitrogen, San Diego, CA)例如在昆虫细胞中也可用以生产。商业上可以得到的质粒pcDNA1(Invitrogen, San Diego, CA),例如,在哺乳类细胞中如中国仓鼠卵巢细胞中也可用作生产。在本领域中一个有一般技术的人员,应用常规的技术和容易得到的起始材料,能够利用这些商业上的表达载体和系统或其他的以生产HSV-2蛋白酶或HSV-2衣壳蛋白。(请见,例如, Sambrook等,分子克隆,一个实验室手册,第二版,冷泉港出版社(1989),此书以参考文献在这里编入)因此。在原核的和真核的二个系统中能够制备所期望的蛋白,产生一个蛋白加工形式的系列。

适合期望宿主的表达系统的构建细节,本领域中的那些人是知道的。简短的说,为重组体蛋白的生产,编码多肽的DNA要合适地连接到所选择好的表达载体上。DNA可操作连接到在选择宿主中DNA表达所必需的所有调节元件。本领域中一个有一般技术人员,用熟知的技术能够制备出重组产生多肽的表达载体。

包含编码HSV-2蛋白酶或HSV-2衣壳蛋白的DNA的表达载体被用于去转化或转染合适宿主,然后把该宿主培养和保持在外源DNA表达发生的条件下。从培养物,或通过裂解细胞,或从适当的及为本领域中那些人知道的培养基中回收本发明产生的蛋白。本领域中一个有一般技术的人员,应用熟知的技术,能分离应用这样表达系统所产生的蛋白。

根据本发明的一个实施例,蛋白也可以被生产和纯化如下。包括编码HSV-2蛋白酶或HSV-2衣壳蛋白的核苷酸序列的一个DNA分子可

制成包括在蛋白的一个末端部分编码多个组氨酸残基的核苷酸序列。这个DNA分子被组合到一个表达载体，并把该载体引入适合的宿主细胞中。DNA被表达，所产生的蛋白包含末端的组氨酸残基，他在这里被称作组氨酸尾标或His-tag。收集这种细胞放入pH 8.5的磷酸缓冲液生理盐水中并保持的冰上。然后用超声波把细胞裂开。经过超声波振裂的细胞物质在30,000 xg离心。然后把上清液通过一个2微米滤膜过滤。把过滤过的上清液用一种金属螯合树脂（例如，氨基三乙酸镍树脂是一个对这样目的有用的多种树脂之一）在室温共孵2小时，经过这2小时后，通过离心把树脂从没有结合的物质中分离开来。然后把树脂装柱并用50mM咪唑液清洗以除去没有特异结合的蛋白。然后用150mM咪唑缓冲液把末端标记组氨酸蛋白酶从Ni柱上洗脱下来。把来自柱的洗脱液，用Pharmacia Superdex 75大小的柱在磷酸盐缓冲生理盐水中通过柱层析进一步纯化。

DNA分子可设计为在尾标组氨酸与可靠的HSV-2蛋白酶间包含一个特异裂解位点，使能从表达的蛋白上除去尾标记组氨酸。尾标记的组氨酸的去除也可被完成如下：当尾标记组氨酸位在HSV-2蛋白酶氨基末端时，可将（天冬氨酸）<sub>4</sub>赖氨酸序列置于尾标记组氨酸后面及HSV-2序列之前。肠激酶在（天冬氨酸）<sub>4</sub>赖氨酸序列之后特异地切割因而产生可信HSV-2蛋白酶。

除了用重组体技术生产这些蛋白以外，也可以应用自动的肽合成仪生产HSV-2蛋白酶或HSV-2衣壳蛋白。本领域中具有一般技术的那些人，对这样的技术是熟知的。

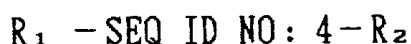
本发明提供了基本上纯的用于HSV-2蛋白酶裂解活性的底物，包括合成的底物。一个HSV-2蛋白酶底物是一个肽，被HSV-2蛋白酶引起的蛋白水解作用裂解成至少二个分开的肽。在一些实施例中，裂解

的和不裂解的底物间大小的差别能被用于检测蛋白酶是活性的或没有活性的。在一些实施例中，本发明的底物是被标记的因此他们也能被测知。在一些实施例中，底物是被固定在固相上。在本发明的一些实施例中，或是底物或是一个蛋白分解产物有生物学上的或化学的活性但在其他部分活性则不存在，这一点能用于把这一个与其他区别开来。生物活性的范例包括酶活性和与特异抗体的结合能力。

在已被鉴定有天然裂解位点的UL 26 中含有两个氨基酸序列。第一个是LQAS (SEQ ID NO: 3) 在那里HSV-2蛋白酶在A和S之间把肽裂解。第二个是VNAS (SEQ ID NO: 4)，在那里HSV-2蛋白酶在A和S之间把肽裂开。天然的或合成的底物能被产生含有这二个裂解位点的任一个。因此，按照本发明的底物有通式



或通式



在式中  $R_1$  和  $R_2$  分别地表示为氢或一个或多个氨基酸。在有些实施例中，底物是UL 26 基因的产物，它含有两个蛋白酶裂解位点；一个包含SEQ ID NO: 3和另一个包含SEQ ID NO: 4。在有些实施例中，底物是UL 26.5基因产物，它含有一个包含SEQ ID NO: 4 的蛋白酶裂解位点。在有些实施例中， $R_1$  优选1 - 20个氨基酸，更好是1 - 10个最好为3、4、5、6、7、8或9个氨基酸。在有些实施例中， $R_2$  可以是1 - 20个氨基酸，更好是1 - 10个，最好是3、4、5、6、7、8或9个氨基酸。本领域中一个具有一般技术人员按照上述的通式能容易设计底物。下列各肽已被设计作为底物。

1. 包含内部裂解位点SEQ ID NO: 3 (LQA\*S) 的各肽:

AHTYLQA*SEKFK	SEQ ID NO:5
AGIAGHTYLQA*SEKFK	SEQ ID NO:6
GIAGHTYLQA*SEKFK	SEQ ID NO:7
IAGHTYLQA*SEKFK	SEQ ID NO:8
GHTYLQA*SEKFK	SEQ ID NO:9
HTYLQA*SEKFKM	SEQ ID NO:10
HTYLQA*SEKFKMW	SEQ ID NO:11
HTYLQA*SEKFKMWG	SEQ ID NO:12
HTYLQA*SEKFKMWGA	SEQ ID NO:13
HTYLQA*SEKFKMWGAE	SEQ ID NO:14

2. 包含末端裂解位点SEQ ID NO: 4 (VNA\*S) 的各肽

ALVNA\*SSAAHVDVD SEQ ID NO: 15

星号(\*)表示裂开键, HSV-2蛋白酶就在这里使链裂开。

底物也可从UL 26 的蛋白水解的裂解中或UL 26.5蛋白产物中得到。他们可通过UL 26 基因或UL 26.5基因的表达重组产生或含有裂解位点的它的片段; 或用本领域中熟知的标准肽合成步骤, 例如 Merrifield合成法, 依靠有机化学合成手段合成底物。

本领域中一个具有一般技术人员, 应用HSV-2蛋白酶和底物就能容易地设计测定方法去鉴定调节HSV-2蛋白酶活性的化合物。如在这里所用的术语“测定试验”指的是对包括HSV-2蛋白酶, 底物和受试化合物的一个混合物的测定; 术语“对照试验”指的是对包括HSV-2蛋白酶和底物但没有受试化合物的一个混合液的测定。要决定一个受试化合物是否能调节HSV-2蛋白酶活性, 在一个测定试验中HSV-2蛋白酶活性的水平可以与在一个对照试验中HSV-2蛋白酶活性的水平相

比较。

在本发明的一些实施例中，当存在有一个受试化合物与HSV-2蛋白酶相接触时，裂解的与不裂解的底物之间大小的差别是用于确定底物是否被裂解。在有些实施例中，采用HPLC（高压液相层析）完成。含有蛋白酶的样本与底物，例如HTYLQASEKFKMWGAE（SEQ ID NO：14），在磷酸盐缓冲盐水中在37℃保温4小时，保温之后用三氟乙酸终止反应。然后把反应液经过一个HPLC柱，用肽的裂解产物表明活性。

本发明的有些实施例中，当一个受试化合物存在，并与HSV-2蛋白酶接触时用免疫测定法检测底物是否被裂解。在一些实施例中，提供的抗体特异地与未经裂解的底物结合但不与蛋白酶裂解的产物结合。这样的抗体在这里被称为“底物-特异抗体”。在一些实施例中，提供的抗体特异地与HSV-2蛋白酶裂解产物相结合，但不与未经裂解底物结合。这样的抗体在这里被称作为“产物-特异抗体”。或一个产物，或一个底物反应但不与二者（即底物-特异抗体和产物-特异抗体的群体）都反应的抗体在这里被称为“非交叉反应抗体”。在有些实施例中，抗体被固定在一个固相上。在一些实施例中，抗体被标记。

例如，一个含有HSV-2蛋白酶，底物和受试验化合物的混合液被保持在合适的条件下并经一个充分的时间任其发生蛋白水解作用，除非受试化合物影响其作用。能把混合液加到一个内部表面贴附有无交叉反应抗体的容器中。如果无交叉反应抗体是底物特异的抗体，则保留在混合液中的任何未经裂解的底物将与抗体结合。如果底物标记的，容器可以被淋洗，而存在的标记总量可以被测知。因此HSV-2蛋白酶活性的水平就被确定。如果非交叉反应抗体是产物-特异抗体，在混合液中任何HSV-2蛋白酶产物将与抗体结合。如果底物是被标记在作

为产物释放出来的一部分上，容器可以被淋洗，存在的标记量就可被测定。因此HSV-2蛋白酶活性水平就被确定。

ICP35抗体（目录号：B-118-100；Rivers Park, 9108 Gulford Rd. Columbia Maryland）也可以用于测定裂解的底物。这样的抗体是产物特异的，它只与被HSV-2蛋白酶水解加工过的衣壳蛋白相结合。

换一种方法，代替应用标记的底物，把范例化的免疫测定可以被修改作夹心面包式测定，在方法中对被束缚住抗原复合物特异的抗体就能被测定。这样的抗体在这里被称作“复合-特异抗体”。容器再次被淋洗，并以充分的时间让复合-特异抗体与存在的任何复合物的结合。测定复合-特异抗体的水平就表示出HSV-2蛋白酶活性水平。

在一些免疫测定法实施例中，不标记的底物在反应混合液中应用。在把反应混合液加到包含一个非交叉反应抗体的容器之后并保持充分的时间利于非交叉反应抗体与底物或与产物相结合，又把标记的底物或标记的产物，分别地加入并将使他与还没有被混合液中的底物或产物结合的非交叉反应抗体相结合。测定标记底物或标记产物的总量就能表明蛋白水解裂解的水平。

在一些实施例中，底物是被标记的，当底物改换成蛋白水解产物时标记被释放出来。测知标记的释放，就能表示HSV-2蛋白酶活性，这可用一类熟知的手段完成。在一些实施例中，被标记的底物在固相上固定。由于HSV-2蛋白酶对其裂解，连接在部分底物上的标记被释放出来，这部分底物为未连接标记的产物。比较存在于反应液中前和后的标记水平就知道多少标记被释放，因而也知道HSV-2蛋白酶活性的水平。换一种方式，测定脱离固相的标记的总量，也表示HSV-2蛋白酶活性的水平。

在另外一个实施例中，检测HSV-2蛋白酶活性的方法包括荧光析出测定法，其中，底物在裂开键邻近含有荧光标记。在这样的位置上，在不裂解底物中标记不能被测定到。然而，当底物被HSV-2蛋白酶在裂解位点上被裂解时，荧光基团成为暴露的，荧光便变成可被测定的。因此，在一个受试化合物存在时把底物与HSV-2蛋白酶接触之后，用度量可检测的荧光法可以量度蛋白水解活性的水平。

在另一个实施例中，测定HSV-2蛋白酶活性的方法包括闪烁近邻测定法，在此法中，放射标记的底物被结合到固体小珠上，当小珠紧密邻近放射标记时被激发，通过闪烁法就可检测的。当底物被裂解时，放射标记不再紧密邻近小珠，而小珠不被激发所以用闪烁法也不能检测了。因此当存在一个受试化合物时把结合的底物与HSV-2蛋白酶接触之后，能用闪烁法量度小珠的激发以测定蛋白水解活性的水平。

除了这些实施例之外，本领域中一个具有一般技术的人员能应用熟知的技术采用测定HSV-2蛋白酶裂解或其缺失的各种方法，去设计鉴定调节HSV-2蛋白酶活性的化合物的其他方法。

本发明涉及鉴定调节HSV-2蛋白酶活性的化合物的试剂盒。这样的试剂盒包含有分开的容器，容器中包括HSV-2蛋白酶，底物和任意选择的抗体，或其他用以检测HSV-2蛋白酶活性的试剂或用于区别不裂解的底物和产物的试剂。底物或抗体也可能是被固定在容器的内部表面。底物或抗体也可能是被标记。

本发明的一些实施例，也提供一个用多体化测定法以鉴定抑制或其他调节HSV-2衣壳蛋白装配的化合物的方法。本发明提供鉴别作为抗HSV-2药剂有用的化合物的方法，因为衣壳装配对病毒复制和感染是必需的。按照本发明，提供的嵌合基因包含一个编码HSV-2衣壳蛋白的HSV-2 UL 26.5基因（或其一部分）连接到编码酵母菌GAL4 DNA

结合蛋白的序列上；或编码HSV-2衣壳蛋白的HSV-2 UL 26.5基因（或其一部分）连接到编码酵母GAL4活化蛋白序列上。编码HSV-2衣壳蛋白的嵌合基因部分最好编码成熟衣壳，也可用编码衣壳前体蛋白的基因。把嵌合基因插入到啤酒酵母菌（*Saccharomyces Cerevisiae*）质粒上并将此质粒引入含有一个整合的GAL4-应答的lacZ 指示基因的啤酒酵母菌中，当在质粒上嵌合基因被表达时，就产生融合蛋白。包含HSV-2衣壳蛋白的融合蛋白各部分在选定的条件下，将彼此结合并使GAL4的DNA-结合区和活化区汇合。当这二个很邻近的GAL4区与GAL4-反应的lacZ指示基因相互作用时，指示基因被表达在合适的条件下一种可检测的兰色就被观察到。如果融合蛋白被阻止结合，这二个GAL4区彼此不是邻近存在，故而指示基因不被活化。因此，看不到兰色。

因此，这个酵母系统提供了一个测定HSV-2衣壳蛋白在病毒粒子装配中发生的相互作用的快速和特异的方法。在存在有阻断或抑制HSV-2衣壳蛋白相互作用的化合物时，由嵌合基因表达产生的融合蛋白中GAL4区将不结合因此将不活化酵母菌系统中的lacZ基因。因此，通过转化过的酵母菌lacZ基因不出现活化可以鉴定抑制HSV-2衣壳装配的化合物，因此化合物具有抗病毒的特性。

本发明的有些实施例中，提供了区别含有HSV-1 DNA的样品和含有HSV-2 DNA 的样品的方法，或区别含有HSV-1 蛋白样品和含有HSV-2蛋白样品的方法。因此，本发明提供了诊断某个体是否感染HSV-1和/或HSV-2的一个诊断方法。公开的是鉴定一个人是否感染HSV-1和/或HSV-2的方法，其中HSV-1感染能与HSV-2感染区别开来。

按照本发明的一些实施例，用PCR 技术区别含有HSV-1 DNA样本

和含有HSV-2 DNA样本。这样的方法提供了一个区别HSV-1和HSV-2感染的手段，并考虑作为对一个人感染HSV类型的诊断方法。设计特异的引物以利于HSV-1 DNA但不是HSV-2 DNA的扩增和/或HSV-2 DNA但不是HSV-1 DNA的扩增。因此，用这样的引物通过扩增技术以及取自个体的生物样品例如细胞，血清或组织样本，特别是在发疱区或其他可被见到有病毒散发表现区所取的样本，人们就能确定样品中的DNA是否来自HSV-1或HSV-2，取样个体是否感染HSV-1或HSV-2就被确定。

UL26基因的核苷酸序列包括编码HSV-2蛋白酶和HSV-2衣壳蛋白的核苷酸序列在SEQ ID NO: 1被公开。编码HSV-1蛋白酶和HSV-1衣壳蛋白的核苷酸序列在SEQ ID NO: 16中被公开。设计一套PCR引物放大HSV-2序列但不放大HSV-1序列。因此，放大DNA的检测表明HSV-2是存在的。同样，设计一套PCR引物只放大HSV-1序列不放大HSV-2序列。因此，放大DNA的检测就表明HSV-1是存在的。最好的是提供的两套引物用在公开的放大方案中并以来自同一样本的材料以便供给一个额外的对照。另外选择的对照包括含有的DNA序列能被放大的阳性对照和/或不能被引物放大的阴性对照。在放大方案完成之后，放大的DNA用在电泳凝胶上跑电泳的方法可以被检测。一个放大产物的期望长度的DNA分子可被提供作为大小标志。

本发明也涉及区别一个样品含有的DNA是否来自HSV-1或HSV-2的试剂盒。本发明的试剂盒在诊断一个人是否感染HSV-1和/或HSV-2上是有用的。试剂盒含有包含有放大HSV-1 DNA但不是HSV-2 DNA的引物的各容器或将放大HSV-2 DNA但不是HSV-1 DNA的各容器。试剂盒可以任选地含有二套引物放在分开的容器中，用同一样本的不同部分进行分别的放大程序。试剂盒也可任选地含有阳性和/或阴性

对照于分开的容器中。试剂盒可任选地含有能作为大小标志的DNA分子于一个分开的容器中。DNA分子用引物可被放大成一个期望长度的DNA分子。

按照发明一些实施例，所用的免疫测定法去区别含有HSV-1蛋白样品与含有HSV-2蛋白的样品。免疫测定法被用于区别HSV-1和HSV-2的感染从而也可诊断一个个人持有感染HSV的类型。这样的免疫测定法是根据HSV-1和HSV-2的UL 26基因产物的不同或者根据HSV-1和HSV-2的UL 26.5基因产物的差别。免疫测定法也可根据HSV-1和HSV-2在蛋白酶和/或衣壳蛋白间的不同。提供的特异抗体它有选择地与HSV-1抗原上表位相结合但不与HSV-2抗原上存在的结合，或它选择地与在HSV-2抗原上的表位结合但不与HSV-1抗原上存在的结合。例如，提供的特异抗体它选择地与HSV-2蛋白酶但不是HSV-2蛋白酶相结合，或它选择地与HSV-2蛋白酶但不与HSV-1蛋白酶相结合。同样地，提供的特异抗体它特异地与HSV-1衣壳但不是HSV-2衣壳相结合，或它选择地与HSV-2衣壳但不是HSV-1衣壳相结合。

因此，用特异的抗体，通过抗体结合分析法的实行以及取自个人的生物样品例如细胞，血清或组织样品，特别是在发疱区或其他可被观察到的有病毒散发表现区所取的样品，人们就能确定HSV-1特异抗体或HSV-2特异抗体是否与样品中蛋白相结合，因而取样个体是否感染HSV-1和/或HSV-2就被确定。HSV-2活性蛋白酶前体的氨基酸序列在SEQ ID NO: 1和SEQ ID NO: 2中跨越1-638氨基酸。HSV-2成熟蛋白酶的氨基酸序列跨越SEQ ID NO: 1和SEQ ID NO: 2的1-247个氨基酸。HSV-2衣壳前体的氨基酸序列在SEQ ID NO: 1和SEQ ID NO: 2中跨越310-638氨基酸。HSV-2成熟衣壳的氨基酸序列跨越

SEQ ID NO: 1和SEQ ID NO: 2的310-613氨基酸。HSV-1蛋白酶和衣壳的氨基酸序列公开在SEQ ID NO: 17 中。HSV-1活性蛋白酶前体的氨基酸序列在SEQ ID NO: 17 中跨越1-635个氨基酸。HSV-1 成熟蛋白酶的氨基酸序列跨越SEQ ID NO: 17的1-247 氨基酸。HSV-1 衣壳前体的氨基酸序列在SEQ ID NO: 17 中跨越307-635氨基酸。HSV-1成熟衣壳的氨基酸序列跨越SEQ ID NO: 17 的307-610氨基酸。

本领域中有一般技术的那些人用常规的方法和能广泛地可得到的起始材料可以生产特异地与HSV-2蛋白酶结合但不与HSV-1蛋白酶结合的抗体。同样，本领域中有一般技术的那些人，用常规的方法和可广泛地得到的起始材料可以生产对HSV-2衣壳但不是HSV-1衣壳特异地结合的抗体。这些HSV-2特异抗体的任一个都可在一个区别HSV-1与HSV-2的免疫测定法中检测HSV-2。同样，在本领域中那些具有一般技术者用常规的方法和可广泛地得到的起始材料，可以生产对HSV-1蛋白酶但不是HSV-2蛋白酶特异地结合的抗体。同样地，在本领域中那些有一般技术者用常规的方法和可广泛地得到起始材料，可以生产不是对HSV-2衣壳而是对HSV-1衣壳特异地结合的抗体。这些HSV-1特异的抗体被用于在一个区别HSV-1和HSV-2的免疫测定法中检测HSV-1。最好的是，用从同一样品材料进行这二个免疫测定法以便提供一个附加的对照。其他选择的对照包括阳性对照，它包含有时与在免疫测定中所用的抗体相结合的肽，和/或阴性对照，它包含有不与在免疫测定中所用的抗体结合的肽。抗体可被标记，也可应用一个特异地与HSV 特异抗体结合的抗体。本领域中一个有一般技术的人应用这里提供的信息能容易地生产包含所有需要的试剂的免疫测定试剂盒。

HSV-1蛋白酶抗体由Serotech 作为抗体45KD 生产并在商业上可

从生物产品科学公司 (Bioproducts for Science Inc.) 得到, 其目录号为 MCA 406 (P. O. Box 29176, Indianapolis, IN), 它能在免疫测定中应用以从 HSV-1 中区别 HSV-2。这种 Serotech 抗体对 HSV-1 前体或成熟衣壳蛋白但不是 HSV-2 前体或成熟衣壳蛋白相结合。因此, 用 Serotech 抗体的一个免疫测定法可以完成确定一个样品中是否含有 HSV-1 或 HSV-2, 因而确定从其本身取得样品那个人是否感染 HSV-1 或 HSV-2。

本发明也涉及诊断某个体是否感染 HSV-1 或 HSV-2 的试剂盒。本发明的试剂盒可以包括含有与 HSV-1 蛋白酶结合但不与 HSV-2 结合的抗体之一个容器和 / 或含有与 HSV-2 蛋白酶结合但不与 HSV-1 蛋白酶结合的抗体的一个容器。最好的是试剂盒包含抗体二种类型存放于分开的容器中。用于试剂盒中的抗体可被标记。试剂盒含有完成一种用抗体的免疫测定的其他所有试剂和材料。试剂盒也可任选地在分开的容器中含有阳性对照和 / 或阴性对照。试剂盒也可任选地含有检测抗体的手段包括, 例如, 一个第二抗体, 他特异地与抗-HSV 蛋白酶抗体相结合。本发明的试剂盒也可包括含有与 HSV-1 衣壳结合但不与 HSV-2 衣壳结合的抗体的一个容器和 / 或一个含有与 HSV-2 衣壳结合但不与 HSV-1 衣壳结合的抗体的一个容器。最好的是试剂盒在分开的容器中包含抗体的二种类型。用于试剂盒的抗体可被标记。试剂盒含有完成一种用抗体的免疫测定用的其他所有试剂和材料。试剂盒也可任选地在分开的容器中含有阳性和 / 或阴性对照。试剂盒也可任选地包含检测抗体的手段包括, 例如, 一个第二抗体, 他特异地与抗-HSV-1 衣壳抗体相结合。试剂盒也可包括 Serotech 抗体。

本发明的另一方面涉及 HSV-2 蛋白酶启动子和 / 或增强子和他们的用途。HSV-2 蛋白酶启动子可以被合成或被分离连接到编码除

HSV-2蛋白酶以外蛋白的编码序列上。因此，本发明涉及重组体DNA分子，它至少包括SEQ ID NO: 1 1-534 核苷酸间核苷酸序列的一部分，并操作连接到一个编码除HSV-2蛋白酶外一个蛋白的核苷酸序列上。本发明涉及包含DNA分子的细胞，该DNA分子至少包含有SEQ ID NO: 1的在1和534间核苷酸序列的一部分，可操作连接到编码除HSV-2蛋白酶以外的一个蛋白的核苷酸序列上。

发明的另一方面适用于入噬体克隆，该克隆存有HSV-2 UL 26基因（SEQ ID NO: 1）和基因的上游和下游序列。因此，连接的序列能被用为UL 26 启动子调节区和/或启动子增强区的筛选。

本发明另一方面涉及HSV-2衣壳蛋白启动子和它的用途。HSV-2衣壳蛋白启动子位在SEQ ID NO: 1的1461核苷酸上游。它可以被合成或被分离并连接到除HSV-2衣壳蛋白外其它编码蛋白的序列上。因此，本发明涉及重组体DNA分子，它至少包含SEQ ID NO: 1的1461核苷酸上游核苷酸序列的一部分，并可操作连接到除HSV-2衣壳蛋白外的编码蛋白的一个核苷酸序列上。本发明涉及细胞，它包含的DNA分子至少含有SEQ ID NO: 1的第1461核苷酸的上游核苷酸序列的一部分，并可操作连接到除HSV-2衣壳蛋白外编码一个蛋白的核苷酸序列上。核苷酸1191到1461（SEQ ID NO: 1），例如，被连接到氯霉素乙酰转移酶基因上并当转染进入VERO细胞时表现出具有重要的启动子活性。

## 实施例

### 实施例 1

蛋白水解活性对疱疹病毒的病毒粒子成熟必不可少已被表示为HSV-2的特性。HSV-2蛋白酶也被称为HSV-2 UL26，有约67,028Da（道尔顿）的分子量（表观）和pI（等电点）=6.94。HSV-2蛋白酶用分子的和生化的技术通过合理的设计和筛选在体外实验中鉴定该活

性的抑制剂并在感染的细胞中测定这些抑制剂的抗病毒活性。

HSV-2 UL 26基因被克隆成一个NcoI-EcoRI 片段(1938碱基对), 它含有启动密码子, 整个开放阅读框架, 停止密码子和3' -不翻译序列的22对碱基。用pOTS载体系统在大肠杆菌(E. coli) 中表达HSV-2 UL 26全部长度, 在载体中基因被插入在来自pOTS-207 载体的入噬菌体的强烈而紧密调节的P<sub>2</sub> 启动子的下游。当表达着的基因很可能是对细胞有毒时, 例如蛋白酶, 启动子的紧密调节是重要的。27KD蛋白酶区域相当于来自HSV-2 UL 26原始翻译产物的自动蛋白水解产物的一个产物, 而该原始翻译产物是用含有T<sub>7</sub> 启动子的紧密调节的表达载体pET-16(b) (Novagen, Madison W. I.) 在大肠杆菌中产生。

设计的每个构建物包括位在HSV-2 UL 26启动子密码子前面的6个组氨酸密码子和(天门冬酸)<sub>4</sub> 赖氨酸密码子, 因此表达的蛋白在N-末端含有一个可裂解的组氨酸尾利于在镍柱进行蛋白纯化。也可用其他的螯合柱, 组氨酸-尾标记蛋白通过咪唑的加入从柱上洗脱下来。另外, 用其他的方法例如改变pH它也能被洗脱下来。对纯化蛋白有用的柱和技术方案也可从在商业上可得到的来源例如Qiagen处获得。

对P<sub>2</sub> 启动子载体, 把重组体构建物引入到用于表达和处理/纯化研究的大肠杆菌AR 120(茶啉酮酸诱导株) 和大肠杆菌AR 58(热诱导株) 中。对T<sub>7</sub> 启动子载体, 把重组构建物引入到大肠杆菌BL21, 一株IPTG(异丙基-β-D 硫代半乳糖苷) 诱导菌株。在镍螯合柱上通过色谱法蛋白很容易被纯化。

p27 蛋白酶片段像其能力所表现那样是有活力的, 其活力能从包含UL 26.5编码区大部分的一个构建物上除去最后25个氨基酸。

## 实施例 2

p27 蛋白酶基因可被合成使含有高表达的大肠杆菌基因为特征的密码子，但仍保留有p27 蛋白酶的氨基酸序列。合成的基因在表达载体pET-16(b)上位在紧密调节的T<sub>7</sub>启动子的下游。随着IPTG (1mM) 诱导，27 K Da 蛋白区域就在大肠杆菌中高度表达。

## 实施例 3

把上述HSV-2 UL 26基因 (NcoI-EcoRI 片段) 和p27 蛋白酶克隆到昆虫细胞表达载体pVL 1392上。然后把重组体构建物引入到来自草地夜蛾的昆虫细胞中。然后为蛋白酶活性分析制备高滴度病毒母液，继之按比例加大蛋白生产。

## 实施例 4

对分享有最少量相同性的HSV-1 UL 26基因和HSV-2 UL 26基团的DNA 区设计寡核苷酸PCR 引物，以保证测定的特异性。通过计算机分析比较分别公开在SEQ ID NO: 1和SEQ ID NO: 2中的二个DNA 序列，就能容易地观察到这样一个区域。在二个同源染色体间少量相同的区域位在UL 26.5 区域内，即编码衣壳基因的一部分。根据在输入的计算机分析中所提供的在核苷酸序列比较中给出的核苷酸数，提供了下列所用引物的序列和所给引物的位置。如下面所示，它对设计产生不同大小的HSV-1和HSV-2产物一个系统以改进分析是有帮助的。

5' - PCR引物 (正义 - 链序列) :

SEQ ID NO:18	HSV-1: 5'-CCGGTGCCCAATCGTCCGT-3' (#864-882)
SEQ ID NO:19	HSV-2: 5'-GTCCGTGCGCGTCAAGTCG-3' (#1397-1416)

3' - PCR引物 (反义 - 链序列) :

SEQ ID NO:20      HSV-1: 5'-TTCCGGCTCCCCCACCTGA-3' (#1560-1542)  
SEQ ID NO:21      HSV-2: 5'-ATTCCGGATCCTGGAGGCGA-3' (#2470-2452)

用这些引物组期望PCR 产物的大小：

HSV-1：696 碱基对

HSV-2：1073碱基对。

在HSV-1分析中用SEQ ID NO：18 和SEQ ID NO：20 或在HSV-2分析中用SEQ ID NO：19和SEQ ID NO：21，对怀疑含有HSV-1或HSV-2 DNA 的样品使用不同的PCR 扩增方法。如果一个696 碱基对的DNA 片段在HSV-1分析中产生，就表明HSV-DNA在样品中存在。要检测696 碱基对片段的的存在，可把扩增产物通过一个电泳基质上被迁移。大约696 碱基对的DNA 的一个大小标志也在同一基质上同时跑电泳。如果在HSV-2分析中产生的是一个1083碱基对的DNA 片段，则就表明样品中有HSV-2 DNA的存在。为要检测一个1073碱基对片段的的存在，可把扩增产物通过一个电泳基质上的迁移。把约1073碱基对的DNA 的一个大小标志也同时在相同的基质上电泳。

提供一种试剂盒，它包括一个含有HSV-1分析中用的SEQ ID NO：18和SEQ ID NO：20 的容器。提供一种试剂盒，它包括一个含有在HSV-2分析中用的SEQ ID NO：19和SEQ ID NO：21的容器。提供一种试剂盒，它包括这两者，一个容器含有在HSV-1分析中的SEQ ID NO：18和SEQ ID NO：20，而另一个容器中含有在HSV-2 分析中的SEQ ID NO：19和SEQ ID NO：21。也可任选地提供大小标志DNA。在有些试剂盒中提供一个696 碱基对大小的标志。在有些试剂盒中提供一个1073碱基对大小的标志。在一些试剂盒中提供一个696 碱基对大小的标志和一个1073碱基对大小的标志。

## 实施例 5

把包含在HSV-2 UL 26基因中推定的HSV-2 UL 26.5启动子的区域克隆以测试启动子的活性。被分析的256碱基对区跨越SEQ ID NO: 1的\*1191至\*1147。用正义链跨越SEQ ID NO: 1的核苷酸\*1191到\*1209的引物(5'-AACATGAGCTGCGTGACC-3')和跨越SEQ ID NO:1的\*1447至\*1429 核苷酸的反义链引物(5'-AAAGAAGAAGAAGAAGAC-3')通过多聚酶链反应把DNA 片段进行克隆。通过把256 碱基对PCR 片段克隆在商业上可得到的质粒pCAT Basic (Promega)的氯霉素乙酰转移酶(CAT) 报告基因的上游测试启动子的活性。然后把产生的构建物引入到一个合适哺乳类细胞系如Vero细胞中,用分析CAT 活性水平试验启动子活性。该细胞系缺乏内源的CAT, 因此,在引入启动子构建物进入这样的细胞品系之后,CAT 活性的水平就是HSV-2 UL 26.5启动子活性的一个直接量度。

把Vero细胞在含苄他霉素(10 $\mu$ g/ml)的DMEM+10% FCS 中生长,把15微克HSV-2 UL 26.5/pCAT构建物,用标准的方案进行电穿孔引入到5百万Vero细胞中。经过电穿孔后48小时,收集细胞放在100 微升0.25M pH 8.0 Tris 缓冲液中。用反复冻-熔把细胞裂解,以15000 rpm 离心并把上清液转移到新的试管中。用Bio-Rad蛋白测定染料试剂盒(目录号500-0006)确定总蛋白浓度。把5 微升D-苏[二氯乙酰-1-<sup>14</sup>C]氯霉素(Amersham, 56 mCi/mmol)和5 微升n-丁醛辅酶A(5 mg/ml)加到一个细胞提取上清液的试样中使其最后体积为100 微升,用乙酸乙酯提取液接着用薄层色谱法以测定CAT 活性。

除上述构建物外,也可把256 碱基对HSV-2 UL 26.5片段克隆在含有一个SV-40增强子的pCAT Enhancen 载体中的CAT 报告基因上游。

这个构建物在Vero细胞中用上述相同的方法也可测试CAT 活性。

对照载体pCAT（含有SV40启动子和增强子）可以用为HSV-2 UL 26.5启动子能力的比较。

图 2 总结了四个实验的结果。柱形 1 是一个阴性对照并代表在没有启动子和增强子的转录控制因子时的CAT 的表达。柱形 2，一个阳性的对照，应用SV40的启动子和SV40的增强因子驱使CAT 基因表达。柱形 3 代表由UL 26.5 启动子单独驱使CAT 基团表达，而柱形 4 代表当UL 26.5 启动子与SV40增强子联合应用时CAT 基团表达。

已建立起一个基本的UL 26.5 表达水平（柱形 3）只要通过从核苷酸1191回到核苷酸 1 上游分离合适长度的片段，并把此片段引入到相对于启动子区起作用位置的基本表达构建物中并测定他们超过基本水平的增强CAT 表达能力，这样就能用在图 1 中基因序列的附加片段以鉴定UL 26.5 增强因子。

这里所述的启动子当操作连接到那里则对异源基因表达的调节是有用的。

序列目录

( 1 ) 一般信息

(i) 申请者: Dilella, Anthony G.  
Debouck, Christine

(ii) 发明的题目: 新的基因

(iii) 序列数: 21

(iv) 相应的通讯地址:

(A) 收信人: Smithkline Beecham 公司法人专利 - US  
UW 2220

(B) 街道: P. O. Box 1539

(C) 城市: King of Prussia

(D) 州: Pennsylvania

(E) 国家: USA

(F) ZIP: 19406 - 0939

(V) 计算机可读型

(A) 基质类型: 软磁盘

(B) 计算机: IBM PC 兼容的

(C) 操作系统: PC - DOS / MS - DOS

(D) 软件: PatentIn 释放 \*1.0 文件, \*1.25, mmd

(vi) 本申请资料

(A) 申请号:

(B) 归档日期:

(C) 分类:

(viii) 代理人 / 业务代理人信息

(A) 姓名: Jervis, Herbert H.

(B) 注册号：31,171

(C) 参考文献/摘要号：P50188

(ix) 电信交流信息

(A) 电话：215-270-5019

(B) 传真：215-270-5090

(2) SEQ ID NO: 1 信息

(i) 序列特征：

(A) 长度：2472

(B) 类型：核酸

(C) 链：双链

(D) 拓扑学：线性

(ii) 分子类型：基团组DNA

(ix) 特征

(A) 名称/按键：CDS

(B) 位置：534..2447

(ix) 特征

(A) 名称/按键：misc特征

(B) 位置：1461, 2442

(xi) 序列描述：SEQ ID NO: 1

GTCGACGAGG CGCGTGGTGG ATATGTCGTC GGGCGCCCGC CAGGCGGCGC TCGTGCGCCT	60
CACCGCGCTG GAGCTCATCA ACCGCACCCG CACAAACACC ACCCCTGTGG GGGAGATTAT	120
TAACGCCCAC GATGCCTTGG GGATACAATA CGAACAGGGC CTGGGGCTGC TCGCCCAGCA	180
GGCACGCATC GGCTTGGCGT CGAACGCCAA GCGATTGCC ACGTTCAACG TGGGCAGCGA	240
CTACGACCTG TTGTACTTTT TGTGTCTCGG GTTCATTCCC CAGTACCTGT CCGTGGCCTA	300
GGGAAGGGTG GGGGTGGTGG TGGTGGGGTG TTTTCTGTT GTTGTGTTT CTGGTCCGCC	360
TGGTCACAAA AGGCACGGCG CCCCAGAACG CGGGCTTTAG TCCCGGCCCG GACGTCGGCG	420

GACACACAAC AACGGCGGGC CCCGTGGGTG GGTAAGTTGG TTCGGGGGCA TCGCTGTATT	480
CCCTTGCCCG CTTCCACCCC CCCTTCCCCT TTGGTTTGTG TGTGCGGGTG CCC ATG	536
	Met 1
GCG TCG GCG GAA ATG CGC GAG CGG TTG GAG GCG CCT CTG CCC GAC CGG	584
Ala Ser Ala Glu Met Arg Glu Arg Leu Glu Ala Pro Leu Pro Asp Arg	
	5 10 15
GCG GTG CCC ATC TAC GTG GCC GGG TTT TTG GCC CTG TAC GAC AGC GGG	632
Ala Val Pro Ile Tyr Val Ala Gly Phe Leu Ala Leu Tyr Asp Ser Gly	
	20 25 30
GAC CCG GGC GAG CTG GCC CTG GAC CCA GAC ACG GTG CGT GCG GCC CTG	680
Asp Pro Gly Glu Leu Ala Leu Asp Pro Asp Thr Val Arg Ala Ala Leu	
	35 40 45
CCT CCG GAG AAC CCC CTG CCG ATC AAC GTA GAC CAC CGC GCT CGG TGC	728
Pro Pro Glu Asn Pro Leu Pro Ile Asn Val Asp His Arg Ala Arg Cys	
	50 55 60 65
GAG GTG GGC CGG GTG CTC GCC GTG GTC AAC GAC CCT CGG GGG CCG TTT	776
Glu Val Gly Arg Val Leu Ala Val Val Asn Asp Pro Arg Gly Pro Phe	
	70 75 80
TTT GTG GGG CTG ATC GCG TGC GTG CAG CTG GAG CGC GTC CTC GAG ACG	824
Phe Val Gly Leu Ile Ala Cys Val Gln Leu Glu Arg Val Leu Glu Thr	
	85 90 95
GCC GCC AGC GCC GCT ATT TTT GAG CGC CGC GGA CCC GCG CTC TCC CGG	872
Ala Ala Ser Ala Ala Ile Phe Glu Arg Arg Gly Pro Ala Leu Ser Arg	
	100 105 110
GAG GAG CGT CTG CTG TAC CTG ATC ACC AAC TAC CTG CCA TCG GTC TCG	920
Glu Glu Arg Leu Leu Tyr Leu Ile Thr Asn Tyr Leu Pro Ser Val Ser	
	115 120 125
CTG TCC ACA AAA CGC CGG GGG GAC GAG GTT CCG CCC GAC CGC ACC CTG	968
Leu Ser Thr Lys Arg Arg Gly Asp Glu Val Pro Pro Asp Arg Thr Leu	
	130 135 140 145
TTT GCG CAC GTG GCC CTG TGC GCC ATC GGG CGG CGC CTT GGA ACC ATC	1016
Phe Ala His Val Ala Leu Cys Ala Ile Gly Arg Arg Leu Gly Thr Ile	
	150 155 160
GTC ACC TAC GAC ACC AGC CTA GAC GCG GCC ATC GCT CCG TTT CGC CAC	1064
Val Thr Tyr Asp Thr Ser Leu Asp Ala Ala Ile Ala Pro Phe Arg His	
	165 170 175
CTG GAC CCG GCG ACG CGC GAG GGG GTG CGA CGC GAG GCC GCC GAG GCC	1112
Leu Asp Pro Ala Thr Arg Glu Gly Val Arg Arg Glu Ala Ala Glu Ala	
	180 185 190
GAG CTC GCG CTG GCC GGG CGC ACC TGG GCC CCC GGC GTG GAG GCG CTC	1160
Glu Leu Ala Leu Ala Gly Arg Thr Trp Ala Pro Gly Val Glu Ala Leu	
	195 200 205

ACA Thr 210	CAC His	ACG Thr	CTG Leu	CTC Leu	TCC Ser 215	ACC Thr	GCC Ala	GTC Val	AAC Asn	AAC Asn 220	ATG Met	ATG Met	CTG Leu	CGT Arg	GAC Asp 225	1208
CGC Arg	TGG Trp	AGC Ser	CTC Leu	GTG Val 230	GCC Ala	GAG Glu	CGG Arg	CGG Arg	CGG Arg	CAG Gln 235	GCC Ala	GGG Gly	ATC Ile	GCC Ala 240	GGA Gly	1256
CAC His	ACG Thr	TAC Tyr	CTT Leu 245	CAG Gln	GCG Ala	AGC Ser	GAA Glu	AAA Lys 250	TTT Phe	AAA Lys	ATA Ile	TGG Trp	GGG Gly 255	GCG Ala	GAG Glu	1304
TCT Ser	GCC Ala	CCT Pro 260	GCG Ala	CCG Pro	GAG Glu	CGT Arg	GGG Gly 265	TAT Tyr	AAA Lys	ACC Thr	GGC Gly	GCC Ala 270	CCG Pro	GGT Gly	GCC Ala	1352
ATG Met	GAC Asp 275	ACA Thr	TCC Ser	CCC Pro	GCC Ala	GCG Ala	AGC Ser	GTT Val	CCC Pro	GCG Ala	CCG Pro	CAG Gln	GTC Val	GCC Ala	GTC Val	1400
CGT Arg 290	GCG Ala	CGT Arg	CAA Gln	GTC Val 295	GCG Ala	TCG Ser	TCG Ser	TCG Ser	TCT Ser	TCT Ser	TCT Ser	TCT Ser	TCT Ser	TTT Phe	CCG Pro 305	1448
GCA Ala	CCG Pro	GCC Ala	GAT Asp	ATG Met 310	AAC Asn	CCC Pro	GTT Val	TCG Ser	GCA Ala 315	TCG Ser	GGC Gly	GCC Ala	CCG Pro	GCC Ala	CCT Pro	1496
CCG Pro	CCG Pro	CCC Pro	GGC Gly 325	GAC Asp	GGG Gly	AGT Ser	TAT Tyr	TTG Leu 330	TGG Trp	ATC Ile	CCC Pro	GCC Ala	TCT Ser 335	CAT His	TAC Tyr	1544
AAT Asn	CAG Gln 340	CTC Leu	GTC Val	ACC Thr	GGG Gly	CAA Gln	TCC Ser 345	GCG Ala	CCC Pro	CGC Arg	CAC His	CCG Pro 350	CCG Pro	CTG Leu	ACC Thr	1592
GCG Ala	TGC Cys 355	GGC Gly	CTG Leu	CCG Pro	GCC Ala	GCG Ala	GGG Gly 360	ACG Thr	GTG Val	GCC Ala	TAC Tyr 365	GGA Gly	CAC His	CCC Pro	GGC Gly	1640
GCC Ala 370	GGC Gly	CCG Pro	TCC Ser	CCG Pro	CAC His 375	TAC Tyr	CCG Pro	CCT Pro	CCT Pro	CCC Pro 380	GCC Ala	CAC His	CCG Pro	TAC Tyr	CCG Pro 385	1688
GGT Gly	ATG Met	CTG Leu	TTC Phe	GCG Ala 390	GGC Gly	CCC Pro	AGT Ser	CCC Pro	CTG Leu 395	GAG Glu	GCC Ala	CAG Gln	ATC Ile	GCC Ala 400	GCG Ala	1736
CTG Leu	GTG Val	GGG Gly	GCC Ala 405	ATC Ile	GCC Ala	GCC Ala	GAC Asp	CGC Arg 410	CAG Gln	GCG Ala	GGT Gly	GGG Gly	CTT Leu 415	CCG Pro	GCG Ala	1784
GCC Ala	GCC Ala	GGA Gly 420	GAC Asp	CAC His	GGG Gly	ATC Ile	CGG Arg 425	GGG Gly	TCG Ser	GCG Ala	AAG Lys	CGC Arg 430	CGC Arg	CGA Arg	CAC His	1832

GAG Glu 435	GTG Val	GAG Glu	CAG Gln	CCG Pro	GAG Glu	TAC Tyr	GAC Asp	TGC Cys	GGC Gly	CGT Arg	GAG Asp	GAG Glu	CCG Pro	GAC Asp	CGG Arg	1880
GAC Asp 450	TTC Phe	CCG Pro	TAT Tyr	TAC Tyr	CCG Pro	GGC Gly	GAG Glu	GCC Ala	CGC Arg	CCC Pro	GAG Glu	CCG Pro	CGC Arg	CCG Pro	GTC Val	1928
GAC Asp 470	TCC Ser	CGG Arg	CGC Arg	GCC Ala	GCG Ala	CGC Arg	CAG Gln	GCT Ala	TCC Ser	GGG Gly	CCC Pro	CAC His	GAA Glu	ACC Thr	ATC Ile	1976
ACG Thr	GCG Ala	CTG Leu	GTG Val	GGG Gly	GCG Ala	GTG Val	ACG Thr	TCC Ser	CTG Leu	CAG Gln	CAG Gln	GAA Glu	CTG Leu	GCG Ala	CAC His	2024
ATG Met	CGC Arg	GCG Ala	CGT Arg	ACC Thr	CAC His	GCC Ala	CCC Pro	TAC Tyr	GGG Gly	CCG Pro	TAT Tyr	CCG Pro	CCG Pro	GTG Val	GGG Gly	2072
CCC Pro	TAC Tyr	CAC His	CAC His	CCC Pro	CAC His	GCA Ala	GAC Asp	ACG Thr	GAG Glu	ACC Thr	CCC Pro	GCC Ala	CAA Gln	CCA Pro	CCC Pro	2120
CGC Arg 530	TAC Tyr	CCC Pro	GCC Ala	GAG Glu	GCC Ala	GTC Val	TAT Tyr	CTG Leu	CCG Pro	CCG Pro	CCG Pro	CAC His	ATC Ile	GCC Ala	CCC Pro	2168
CCG Pro	GGG Gly	CCT Pro	CCT Pro	CTA Leu	TCC Ser	GGG Gly	GCG Ala	GTC Val	CCC Pro	CCA Pro	CCC Pro	TCG Ser	TAT Tyr	CCC Pro	CCA Pro	2216
GTT Val	GCG Ala	GTT Val	ACC Thr	CCC Pro	GGT Gly	CCC Pro	GCT Ala	CCC Pro	CCG Pro	CTA Leu	CAT His	CAG Gln	CCC Pro	TCC Ser	CCC Pro	2264
GCA Ala	CAC His	GCC Ala	CAC His	CCC Pro	CCT Pro	CCG Pro	CCG Pro	CCG Pro	CCG Pro	GGA Gly	CCC Pro	ACG Thr	CCT Pro	CCC Pro	CCC Pro	2312
GCC Ala	GCG Ala	AGC Ser	TTA Leu	CCC Pro	CAA Gln	CCC Pro	GAG Glu	GCG Ala	CCC Pro	GGC Gly	GCG Ala	GAG Glu	GCC Ala	GGC Gly	GCC Ala	2360
TTA Leu 610	GTT Val	AAC Asn	GCC Ala	AGC Ser	AGC Ser	GCG Ala	GCC Ala	CAC His	GTG Val	AAC Asn	GTG Val	GAC Asp	ACG Thr	GCC Ala	CGG Arg	2408
GCC Ala	GCC Ala	GAT Asp	CTG Leu	TTT Phe	GTG Val	TCA Ser	CAG Gln	ATG Met	ATG Met	GGG Gly	TCC Ser	CGC Arg	TAACTCGCCT			2457
CCAGGATCCG AATTC															2472	

( 2 ) SEQ ID NO : 2信息

(i) 序列特征 :

(A) 长度 : 638 氨基酸

(B) 类型：氨基酸

(C) 拓扑学：线性

(ii) 分子类型：蛋白质

(xi) 序列描述：SEQ ID NO: 2

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

Val Arg Ala Arg Gln Val Ala Ser Ser Ser Ser Ser Ser Ser Ser Phe  
 290 295 300  
 Pro Ala Pro Ala Asp Met Asn Pro Val Ser Ala Ser Gly Ala Pro Ala  
 305 310 315 320  
 Pro Pro Pro Pro Gly Asp Gly Ser Tyr Leu Trp Ile Pro Ala Ser His  
 325 330 335  
 Tyr Asn Gln Leu Val Thr Gly Gln Ser Ala Pro Arg His Pro Pro Leu  
 340 345 350  
 Thr Ala Cys Gly Leu Pro Ala Ala Gly Thr Val Ala Tyr Gly His Pro  
 355 360 365  
 Gly Ala Gly Pro Ser Pro His Tyr Pro Pro Pro Pro Ala His Pro Tyr  
 370 375 380  
 Pro Gly Met Leu Phe Ala Gly Pro Ser Pro Leu Glu Ala Gln Ile Ala  
 385 390 395 400  
 Ala Leu Val Gly Ala Ile Ala Ala Asp Arg Gln Ala Gly Gly Leu Pro  
 405 410 415  
 Ala Ala Ala Gly Asp His Gly Ile Arg Gly Ser Ala Lys Arg Arg Arg  
 420 425 430  
 His Glu Val Glu Gln Pro Glu Tyr Asp Cys Gly Arg Asp Glu Pro Asp  
 435 440 445  
 Arg Asp Phe Pro Tyr Tyr Pro Gly Glu Ala Arg Pro Glu Pro Arg Pro  
 450 455 460  
 Val Asp Ser Arg Arg Ala Ala Arg Gln Ala Ser Gly Pro His Glu Thr  
 465 470 475 480  
 Ile Thr Ala Leu Val Gly Ala Val Thr Ser Leu Gln Gln Glu Leu Ala  
 485 490 495  
 His Met Arg Ala Arg Thr His Ala Pro Tyr Gly Pro Tyr Pro Pro Val  
 500 505 510  
 Gly Pro Tyr His His Pro His Ala Asp Thr Glu Thr Pro Ala Gln Pro  
 515 520 525  
 Pro Arg Tyr Pro Ala Glu Ala Val Tyr Leu Pro Pro Pro His Ile Ala  
 530 535 540  
 Pro Pro Gly Pro Pro Leu Ser Gly Ala Val Pro Pro Pro Ser Tyr Pro  
 545 550 555 560  
 Pro Val Ala Val Thr Pro Gly Pro Ala Pro Pro Leu His Gln Pro Ser  
 565 570 575  
 Pro Ala His Ala His Pro Pro Pro Pro Pro Gly Pro Thr Pro Pro  
 580 585 590  
 Pro Ala Ala Ser Leu Pro Gln Pro Glu Ala Pro Gly Ala Glu Ala Gly  
 595 600 605  
 Ala Leu Val Asn Ala Ser Ser Ala Ala His Val Asn Val Asp Thr Ala  
 610 615 620

Arg Ala Ala Asp Leu Phe Val Ser Gln Met Met Gly Ser Arg  
625 630 635

( 2 ) SEQ ID NO: 3信息

(i) 序列特征:

(A) 长度: 4 氨基酸

(B) 类型: 氨基酸

(C) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 3

Leu Gln Ala Ser

1

( 2 ) SEQ ID NO: 4信息

(i) 序列特征:

(A) 长度: 4 氨基酸

(B) 类型: 氨基酸

(C) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 4

Val Asn Ala Ser

1

( 2 ) SEQ ID NO: 5信息

(i) 序列特征:

(A) 长度: 1 2 氨基酸

(B) 类型: 氨基酸

(C) 拓扑学: 线性

(ii)分子类型：肽

(xi)序列描述：SEQ ID NO: 5

Ala His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys  
1 5 10

( 2 ) SEQ ID NO: 6信息

(i) 序列特征：

(A) 长度：16 氨基酸

(B) 类型：氨基酸

(C) 拓扑学：线性

(ii)分子类型：肽

(xi)序列描述：SEQ ID NO: 6

Ala Gly Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys  
1 5 10 15

( 2 ) SEQ ID NO: 7信息

(i) 序列特征：

(A) 长度：15 氨基酸

(B) 类型：氨基酸

(C) 拓扑学：线性

(ii)分子类型：肽

(xi)序列描述：SEQ ID NO: 7

Gly Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys  
1 5 10 15

( 2 ) SEQ ID NO: 8信息

(i) 序列特征：

(A) 长度：14 氨基酸

(B) 类型：氨基酸

(C) 拓朴学：线性

(ii)分子类型：肽

(xi)序列描述：SEQ ID NO: 8

Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys  
1 5 10

( 2 ) SEQ ID NO: 9信息

(i) 序列特征：

(A) 长度：12氨基酸

(B) 类型：氨基酸

(C) 拓朴学：线性

(ii)分子类型：肽

(xi)序列描述：SEQ ID NO: 9

Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys  
1 5 10

( 2 ) SEQ ID NO: 10信息

(i) 序列特征：

(A) 长度：12氨基酸

(B) 类型：氨基酸

(C) 拓朴学：线性

(ii)分子类型：肽

(xi)序列描述：SEQ ID NO: 10

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met  
1 5 10

( 2 ) SEQ ID NO: 11信息

(i) 序列特征：

(A) 长度：13 氨基酸

(B) 类型：氨基酸

(C) 拓朴学：线性

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO：11

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp  
1 5 10

( 2 ) SEQ ID NO：12 信息

(i) 序列特征：

(A) 长度：14 氨基酸

(B) 类型：氨基酸

(C) 拓朴学：线性

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO：12

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly  
1 5 10

( 2 ) SEQ ID NO：13 信息

(i) 序列特征：

(A) 长度：15 氨基酸

(B) 类型：氨基酸

(C) 拓朴学：线性

(ii) 分子类型：肽

(xi) 序列描述：SEQ ID NO：13

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala  
1 5 10 15

( 2 ) SEQ ID NO: 14信息

(i) 序列特征:

(A) 长度: 16 氨基酸

(B) 类型: 氨基酸

(C) 拓朴学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 14

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala Glu  
1 5 10 15

( 2 ) SEQ ID NO: 15信息

(i) 序列特征:

(A) 长度: 14 氨基酸

(B) 类型: 氨基酸

(C) 拓朴学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 15

Ala Leu Val Asn Ala Ser Ser Ala Ala His Val Asp Val Asp  
1 5 10

( 2 ) SEQ ID NO: 16信息

(i) 序列特征:

(A) 长度: 1908 碱基对

(B) 类型: 核酸

(C) 链: 双链

(D) 拓朴学: 线性

(ii) 分子类型: cDNA

(ix) 特性

(A) 名称 / 按键 : CDS

(B) 位置 : 1 . . . . 1908

(ix) 特性 :

(A) 名称 / 按键 : misc特性

(B) 位置 : 919 . . . . 1908

(xi) 序列特征 : SEQ ID NO : 16

ATG	GCA	GCC	GAT	GCC	CCG	GGA	GAC	CGG	ATG	GAG	GAG	CCC	CTG	CCC	GAC	48
Met	Ala	Ala	Asp	Ala	Pro	Gly	Asp	Arg	Met	Glu	Glu	Pro	Leu	Pro	Asp	
1				5					10					15		
AGG	GCC	GTG	CCC	ATT	TAC	GTG	GCT	GGG	TTT	TTG	GCC	CTG	TAT	GAC	AGC	96
Arg	Ala	Val	Pro	Ile	Tyr	Val	Ala	Gly	Phe	Leu	Ala	Leu	Tyr	Asp	Ser	
			20					25					30			
GGG	GAC	TCG	GGC	GAG	TTG	GCA	TTG	GAT	CCG	GAT	ACG	GTG	CGG	GCG	GCC	144
Gly	Asp	Ser	Gly	Glu	Leu	Ala	Leu	Asp	Pro	Asp	Thr	Val	Arg	Ala	Ala	
		35					40					45				
CTG	CCT	CCG	GAT	AAC	CCA	CTC	CCG	ATT	AAC	GTG	GAC	CAC	CGC	GCT	GGC	192
Leu	Pro	Pro	Asp	Asn	Pro	Leu	Pro	Ile	Asn	Val	Asp	His	Arg	Ala	Gly	
	50					55					60					
TGC	GAG	GTG	GGG	CGG	GTG	CTG	GCC	GTG	GTC	GAC	GAC	CCC	CGC	GGG	CCG	240
Cys	Glu	Val	Gly	Arg	Val	Leu	Ala	Val	Val	Asp	Asp	Pro	Arg	Gly	Pro	
65					70					75					80	
TTT	TTT	GTG	GGG	CTG	ATC	GCC	TGC	GTG	CAG	CTG	GAG	CGC	GTC	CTC	GAG	288
Phe	Phe	Val	Gly	Leu	Ile	Ala	Cys	Val	Gln	Leu	Glu	Arg	Val	Leu	Glu	
				85					90					95		
ACG	GCC	GCC	AGC	GCT	GCG	ATT	TTC	GAG	CGC	CGC	GGG	CCG	CCG	CTC	TCC	336
Thr	Ala	Ala	Ser	Ala	Ala	Ile	Phe	Glu	Arg	Arg	Gly	Pro	Pro	Leu	Ser	
			100					105					110			
CGG	GAG	GAG	CGC	CTG	TTG	TAC	CTG	ATC	ACC	AAC	TAC	CTG	CCC	TCG	GTC	384
Arg	Glu	Glu	Arg	Leu	Leu	Tyr	Leu	Ile	Thr	Asn	Tyr	Leu	Pro	Ser	Val	
		115					120					125				
TCC	CTG	GCC	ACA	AAA	CGC	CTG	GGG	GGC	GAG	GCG	CAC	CCC	GAT	CGC	ACG	432
Ser	Leu	Ala	Thr	Lys	Arg	Leu	Gly	Gly	Glu	Ala	His	Pro	Asp	Arg	Thr	
	130					135					140					
CTG	TTC	GCG	CAC	GTC	GCG	CTG	TGC	GCG	ATC	GGG	CGG	CGC	CTC	GGC	ACT	480
Leu	Phe	Ala	His	Val	Ala	Leu	Cys	Ala	Ile	Gly	Arg	Arg	Leu	Gly	Thr	
145					150					155				160		
ATC	GTC	ACC	TAC	GAC	ACC	GGT	CTC	GAC	GCC	GCC	ATC	GCG	CCC	TTT	CGC	528
Ile	Val	Thr	Tyr	Asp	Thr	Gly	Leu	Asp	Ala	Ala	Ile	Ala	Pro	Phe	Arg	
				165					170					175		

CAC	CTG	TCG	CCG	GCG	TCT	CGC	GAG	GGG	GCG	CGG	CGA	CTG	GCC	GCC	GAG	576
His	Leu	Ser	Pro	Ala	Ser	Arg	Glu	Gly	Ala	Arg	Arg	Leu	Ala	Ala	Glu	
			180					185					190			
GCC	GAG	CTC	GCG	CTG	TCC	GGG	CGC	ACC	TGG	GCG	CCC	GGC	GTG	GAG	GCG	624
Ala	Glu	Leu	Ala	Leu	Ser	Gly	Arg	Thr	Trp	Ala	Pro	Gly	Val	Glu	Ala	
		195					200					205				
CTG	ACC	CAC	ACG	CTG	CTT	TCC	ACC	GCC	GTT	AAC	AAC	ATG	ATG	CTG	CGG	672
Leu	Thr	His	Thr	Leu	Leu	Ser	Thr	Ala	Val	Asn	Asn	Met	Met	Leu	Arg	
	210					215					220					
GAC	CGC	TGG	AGC	CTG	GTG	GCC	GAG	CGG	CGG	CGG	CAG	GCC	GGG	ATC	GCC	720
Asp	Arg	Trp	Ser	Leu	Val	Ala	Glu	Arg	Arg	Arg	Gln	Ala	Gly	Ile	Ala	
225					230					235					240	
GGA	CAC	ACC	TAC	CTC	CAG	GCG	AGC	GAA	AAA	TTC	AAA	ATG	TGG	GGG	GCG	768
Gly	His	Thr	Tyr	Leu	Gln	Ala	Ser	Glu	Lys	Phe	Lys	Met	Trp	Gly	Ala	
				245					250					255		
GAG	CCT	GTT	TCC	GCG	CCG	GCG	CGC	GGG	TAT	AAG	AAC	GGG	GCC	CCG	GAG	816
Glu	Pro	Val	Ser	Ala	Pro	Ala	Arg	Gly	Tyr	Lys	Asn	Gly	Ala	Pro	Glu	
			260					265					270			
TCC	ACG	GAC	ATA	CCG	CCC	GGC	TCG	ATC	GCT	GCC	GCG	CCG	CAG	GGT	GAC	864
Ser	Thr	Asp	Ile	Pro	Pro	Gly	Ser	Ile	Ala	Ala	Ala	Pro	Gln	Gly	Asp	
		275					280					285				
CGG	TGC	CCA	ATC	GTC	CGT	CAG	CGC	GGG	GTC	GCC	TTG	TCC	CCG	GTA	CTG	912
Arg	Cys	Pro	Ile	Val	Arg	Gln	Arg	Gly	Val	Ala	Leu	Ser	Pro	Val	Leu	
	290					295					300					
CCC	CCC	ATG	AAC	CCC	GTT	CCG	ACA	TCG	GGC	ACC	CCG	GCC	CCC	GCG	CCG	960
Pro	Pro	Met	Asn	Pro	Val	Pro	Thr	Ser	Gly	Thr	Pro	Ala	Pro	Ala	Pro	
305					310					315					320	
CCC	GGC	GAC	GGG	AGC	TAC	CTG	TGG	ATC	CCG	GCC	TCC	CAT	TAC	AAC	CAG	1008
Pro	Gly	Asp	Gly	Ser	Tyr	Leu	Trp	Ile	Pro	Ala	Ser	His	Tyr	Asn	Gln	
				325					330					335		
CTC	GTC	GCC	GGC	CAT	GCC	GCG	CCC	CAA	CCC	CAG	CCG	CAT	TCC	GCG	TTT	1056
Leu	Val	Ala	Gly	His	Ala	Ala	Pro	Gln	Pro	Gln	Pro	His	Ser	Ala	Phe	
			340					345					350			
GGT	TTC	CCG	GCT	GCG	GCG	GGG	TCC	GTG	GCC	TAT	GGG	CCT	CAC	GGT	GCG	1104
Gly	Phe	Pro	Ala	Ala	Ala	Gly	Ser	Val	Ala	Tyr	Gly	Pro	His	Gly	Ala	
		355					360					365				
GGT	CTT	TCC	CAG	CAT	TAC	CCT	CCC	CAC	GTC	GCC	CAT	CAG	TAT	CCC	GGG	1152
Gly	Leu	Ser	Gln	His	Tyr	Pro	Pro	His	Val	Ala	His	Gln	Tyr	Pro	Gly	
	370					375					380					
GTG	CTG	TTC	TCG	GGA	CCC	AGC	CCA	CTC	GAG	GCG	CAG	ATA	GCC	GCG	TTG	1200
Val	Leu	Phe	Ser	Gly	Pro	Ser	Pro	Leu	Glu	Ala	Gln	Ile	Ala	Ala	Leu	
385					390					395					400	
GTG	GGG	GCC	ATA	GCC	GCG	GAC	CGC	CAG	GCG	GGC	GGT	CAG	CCG	GCC	GCG	1248
Val	Gly	Ala	Ile	Ala	Ala	Asp	Arg	Gln	Ala	Gly	Gly	Gln	Pro	Ala	Ala	
				405					410					415		

GGA	GAC	CCT	GGG	GTC	CGG	GGG	TCG	GGA	AAG	CGT	CGC	CGG	TAC	GAG	GCG	1296
Gly	Asp	Pro	Gly	Val	Arg	Gly	Ser	Gly	Lys	Arg	Arg	Arg	Tyr	Glu	Ala	
			420					425					430			
GGG	CCG	TCG	GAG	TCC	TAC	TGC	GAC	CAG	GAC	GAA	CCG	GAC	GCG	GAC	TAC	1344
Gly	Pro	Ser	Glu	Ser	Tyr	Cys	Asp	Gln	Asp	Glu	Pro	Asp	Ala	Asp	Tyr	
			435				440					445				
CCG	TAC	TAC	CCC	GGG	GAG	GCT	CGA	GGC	GCG	CCG	CGC	GGG	GTC	GAC	TCC	1392
Pro	Tyr	Tyr	Pro	Gly	Glu	Ala	Arg	Gly	Ala	Pro	Arg	Gly	Val	Asp	Ser	
	450					455						460				
CGG	CGC	GCG	GCC	CGC	CAT	TCT	CCC	GGG	ACC	AAC	GAG	ACC	ATC	ACG	GCG	1440
Arg	Arg	Ala	Ala	Arg	His	Ser	Pro	Gly	Thr	Asn	Glu	Thr	Ile	Thr	Ala	
	465				470					475					480	
CTG	ATG	GGG	GCG	GTG	ACG	TCT	CTG	CAG	CAG	GAA	CTG	GCG	CAC	ATG	CGG	1488
Leu	Met	Gly	Ala	Val	Thr	Ser	Leu	Gln	Gln	Glu	Leu	Ala	His	Met	Arg	
				485					490					495		
GCT	CGG	ACC	AGC	GCC	CCC	TAT	GGA	ATG	TAC	ACG	CCG	GTG	GCG	CAC	TAT	1536
Ala	Arg	Thr	Ser	Ala	Pro	Tyr	Gly	Met	Tyr	Thr	Pro	Val	Ala	His	Tyr	
			500					505					510			
CGC	CCT	CAG	GTG	GGG	GAG	CCG	GAA	CCA	ACA	ACG	ACC	CAC	CCG	GCC	CTT	1584
Arg	Pro	Gln	Val	Gly	Glu	Pro	Glu	Pro	Thr	Thr	Thr	His	Pro	Ala	Leu	
		515					520					525				
TGT	CCC	CCG	GAG	GCC	GTG	TAT	CGC	CCC	CCA	CCA	CAC	AGC	GCC	CCC	TAC	1632
Cys	Pro	Pro	Glu	Ala	Val	Tyr	Arg	Pro	Pro	Pro	His	Ser	Ala	Pro	Tyr	
	530					535					540					
GGT	CCT	CCC	CAG	GGT	CCG	GCG	TCC	CAT	GCC	CCC	ACT	CCC	CCG	TAT	GCC	1680
Gly	Pro	Pro	Gln	Gly	Pro	Ala	Ser	His	Ala	Pro	Thr	Pro	Pro	Tyr	Ala	
	545				550					555					560	
CCA	GCT	GCC	TGC	CCG	CCA	GGC	CCG	CCA	CCG	CCC	CCA	TGT	CCT	TCC	ACC	1728
Pro	Ala	Ala	Cys	Pro	Pro	Gly	Pro	Pro	Pro	Pro	Pro	Cys	Pro	Ser	Thr	
				565					570					575		
CAG	ACG	CGC	GCC	CCT	CTA	CCG	ACG	GAG	CCC	GCG	TTC	CCC	CCC	GCC	GCC	1776
Gln	Thr	Arg	Ala	Pro	Leu	Pro	Thr	Glu	Pro	Ala	Phe	Pro	Pro	Ala	Ala	
			580					585					590			
ACC	GGA	TCC	CAA	CCG	GAG	GCA	TCC	AAC	GCG	GAG	GCC	GGG	GCC	CTT	GTC	1824
Thr	Gly	Ser	Gln	Pro	Glu	Ala	Ser	Asn	Ala	Glu	Ala	Gly	Ala	Leu	Val	
		595					600					605				
AAC	GCC	AGC	AGC	GCA	GCA	CAC	GTG	GAC	GTT	GAC	ACG	GCC	CGC	GCC	GCC	1872
Asn	Ala	Ser	Ser	Ala	Ala	His	Val	Asp	Val	Asp	Thr	Ala	Arg	Ala	Ala	
	610					615					620					
GAT	TTG	TTC	GTC	TCT	CAG	ATG	ATG	GGG	GCC	CGC	TGA					1908
Asp	Leu	Phe	Val	Ser	Gln	Met	Met	Gly	Ala	Arg						
	625				630					635						

( 2 ) SEQ ID NO: 17信息

(i) 序列特征:

(A) 长度: 635 氨基酸

(B) 类型: 氨基酸

(C) 拓朴学: 线性

(ii) 分子类型: 氨基酸

(xi) 序列描述: SEQ ID NO: 17

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

Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala  
 245 250 255  
 Glu Pro Val Ser Ala Pro Ala Arg Gly Tyr Lys Asn Gly Ala Pro Glu  
 260 265 270  
 Ser Thr Asp Ile Pro Pro Gly Ser Ile Ala Ala Ala Pro Gln Gly Asp  
 275 280 285  
 Arg Cys Pro Ile Val Arg Gln Arg Gly Val Ala Leu Ser Pro Val Leu  
 290 295 300  
 Pro Pro Met Asn Pro Val Pro Thr Ser Gly Thr Pro Ala Pro Ala Pro  
 305 310 315 320  
 Pro Gly Asp Gly Ser Tyr Leu Trp Ile Pro Ala Ser His Tyr Asn Gln  
 325 330 335  
 Leu Val Ala Gly His Ala Ala Pro Gln Pro Gln Pro His Ser Ala Phe  
 340 345 350  
 Gly Phe Pro Ala Ala Ala Gly Ser Val Ala Tyr Gly Pro His Gly Ala  
 355 360 365  
 Gly Leu Ser Gln His Tyr Pro Pro His Val Ala His Gln Tyr Pro Gly  
 370 375 380  
 Val Leu Phe Ser Gly Pro Ser Pro Leu Glu Ala Gln Ile Ala Ala Leu  
 385 390 395 400  
 Val Gly Ala Ile Ala Ala Asp Arg Gln Ala Gly Gly Gln Pro Ala Ala  
 405 410 415  
 Gly Asp Pro Gly Val Arg Gly Ser Gly Lys Arg Arg Arg Tyr Glu Ala  
 420 425 430  
 Gly Pro Ser Glu Ser Tyr Cys Asp Gln Asp Glu Pro Asp Ala Asp Tyr  
 435 440 445  
 Pro Tyr Tyr Pro Gly Glu Ala Arg Gly Ala Pro Arg Gly Val Asp Ser  
 450 455 460  
 Arg Arg Ala Ala Arg His Ser Pro Gly Thr Asn Glu Thr Ile Thr Ala  
 465 470 475 480  
 Leu Met Gly Ala Val Thr Ser Leu Gln Gln Glu Leu Ala His Met Arg  
 485 490 495  
 Ala Arg Thr Ser Ala Pro Tyr Gly Met Tyr Thr Pro Val Ala His Tyr  
 500 505 510  
 Arg Pro Gln Val Gly Glu Pro Glu Pro Thr Thr Thr His Pro Ala Leu  
 515 520 525  
 Cys Pro Pro Glu Ala Val Tyr Arg Pro Pro Pro His Ser Ala Pro Tyr  
 530 535 540  
 Gly Pro Pro Gln Gly Pro Ala Ser His Ala Pro Thr Pro Pro Tyr Ala  
 545 550 555 560  
 Pro Ala Ala Cys Pro Pro Gly Pro Pro Pro Pro Cys Pro Ser Thr  
 565 570 575  
 Gln Thr Arg Ala Pro Leu Pro Thr Glu Pro Ala Phe Pro Pro Ala Ala  
 580 585 590  
 Thr Gly Ser Gln Pro Glu Ala Ser Asn Ala Glu Ala Gly Ala Leu Val  
 595 600 605

Asn Ala Ser Ser Ala Ala His Val Asp Val Asp Thr Ala Arg Ala Ala  
610 615 620

Asp Leu Phe Val Ser Gln Met Met Gly Ala Arg  
625 630 635

( 2 ) SEQ ID NO: 18信息

(i) 序列特征:

(A) 长度: 19 碱基对

(B) 类型: 核酸

(C) 链: 单链

(D) 拓扑学: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 18

CCGGTGCCCA ATCGTCCGT

( 2 ) SEQ ID NO: 19信息

(i) 序列特征:

(A) 长度: 19 碱基对

(B) 类型: 核酸

(C) 链: 单链

(D) 拓扑学: 线性

(ii) 分子类型: cDNA

(xi) 序列描述: SEQ ID NO: 19

GTCCGTGCGC GTCAAGTGG

( 2 ) SEQ ID NO: 20信息

(i) 序列特征:

(A) 长度: 19 碱基对

(B) 类型: 核酸

(C) 链: 单链

(D) 拓扑学：线性

(ii) 分子类型：cDNA

(xi) 序列描述：SEQ ID NO：20

TTCCGGCTCC CCCACCTGA

( 2 ) SEQ ID NO：21 信息

(i) 序列特征：

(A) 长度：19 碱基对

(B) 类型：核酸

(C) 链：单链

(D) 拓扑学：线性

(ii) 分子类型：cDNA

(xi) 序列描述：SEQ ID NO：21

ATTCCGATCC TGGAGCGA





HSV-2 UL 26.5启动子活性

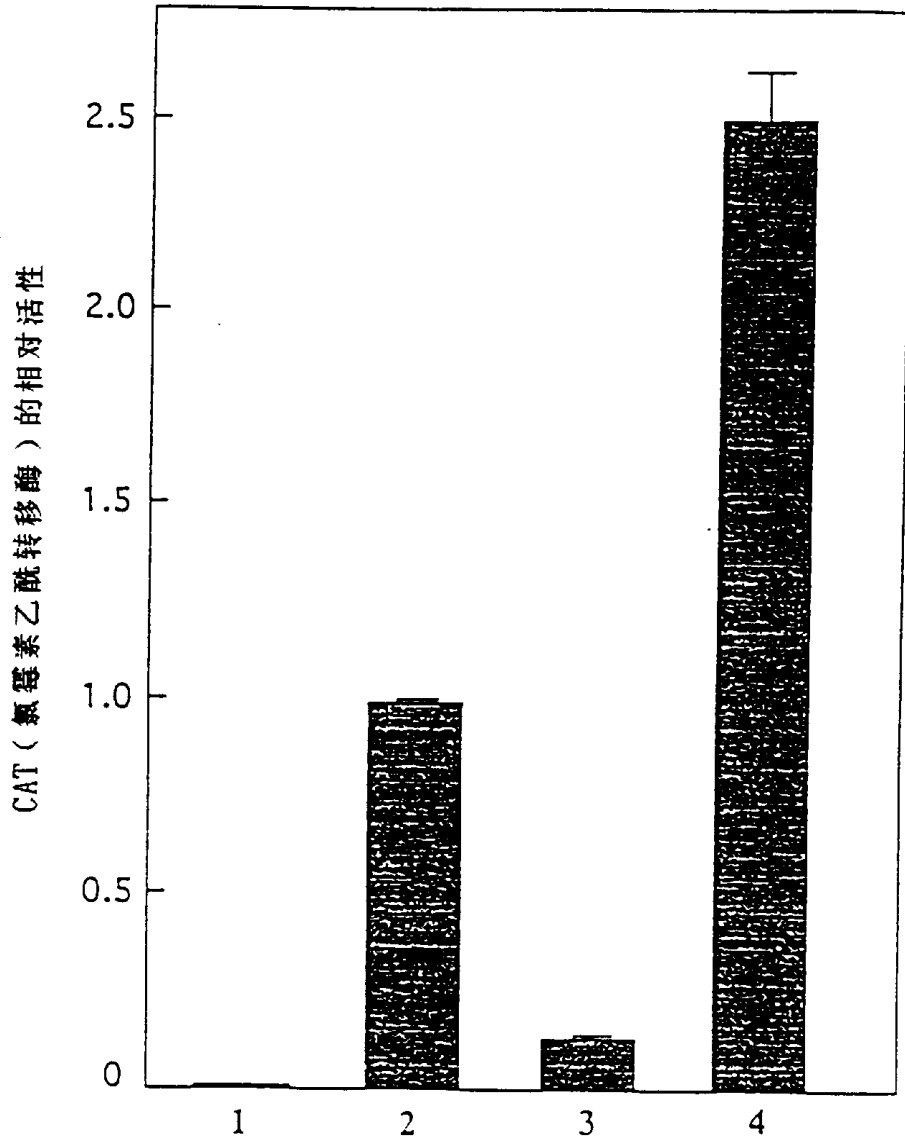


图 2