



(12) 发明专利申请

(10) 申请公布号 CN 104628867 A

(43) 申请公布日 2015. 05. 20

(21) 申请号 201510036266. 6

(22) 申请日 2015. 01. 23

(71) 申请人 同济大学苏州研究院

地址 215101 江苏省苏州市吴中区木渎镇金枫路 198 号

(72) 发明人 房健民 郭佳 蒋明 尹衍新  
于丽华 蒋韵

(74) 专利代理机构 上海泰能知识产权代理事务  
所 31233

代理人 黄志达

(51) Int. Cl.

C07K 19/00(2006. 01)

C12N 15/63(2006. 01)

C07K 16/28(2006. 01)

权利要求书1页 说明书5页

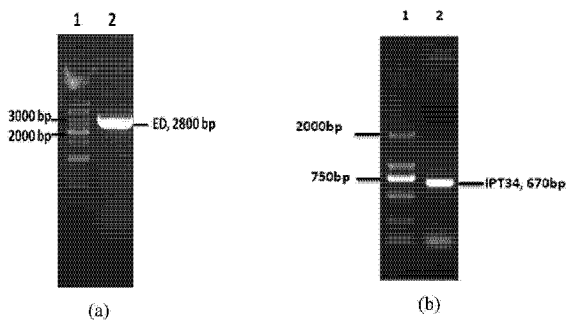
序列表17页 附图3页

(54) 发明名称

一种肝细胞生长因子受体关键结构域融合蛋白及其应用

(57) 摘要

本发明涉及一种肝细胞生长因子受体关键结构域融合蛋白及其应用,所述融合蛋白由表皮生长因子受体的强信号肽、肝细胞生长因子受体关键结构域及 His 标签构成。融合蛋白用于抗体制备。本发明通过慢病毒载体在 293T 细胞中长效、稳定的表达肝细胞生长因子受体关键结构域融合蛋白,获得融合蛋白可特异性的与肝细胞生长因子结合。在筛选阻断 C-Met/HGF 信号转导通路进而抑制肿瘤生长的中和抗体药物研发方面具有重大作用。



1. 一种肝细胞生长因子受体关键结构域融合蛋白,其特征在于:所述融合蛋白由表皮生长因子受体的强信号肽、肝细胞生长因子受体关键结构域及 His 标签构成;其中,肝细胞生长因子受体的关键结构域为胞外区 ED 或免疫球蛋白  $\gamma$ -plexin- 转录因子同源结构域 3-4IPT34;ED 的核苷酸序列如 SEQ ID NO:1 所示或是该序列的变体,ED 的氨基酸序列如 SEQ ID NO:2 所示或是该序列的变体;IPT34 的核苷酸序列如上述序列 SEQ ID NO:1 中 2152-2724 核苷酸序列所示或是该序列的变体,IPT34 的氨基酸序列如上述序列 SEQ ID NO:2 中 718-908 氨基酸序列所示或是该序列的变体。

2. 根据权利要求 1 所述的一种肝细胞生长因子受体关键结构域融合蛋白,其特征在于:所述表皮生长因子受体的强信号肽的核苷酸序列如 SEQ ID NO:3 所示,氨基酸序列如 SEQ ID NO:4 所示。

3. 根据权利要求 1 所述的一种肝细胞生长因子受体关键结构域融合蛋白,其特征在于:所述融合蛋白一共有 940 个氨基酸,如 SEQ ID NO:5 所示,其中 1-26 位氨基酸是表皮生长因子受体的强信号肽,27-934 位氨基酸是肝细胞生长因子受体 ED 蛋白,935-940 位氨基酸是 6 $\times$ His 标签。

4. 根据权利要求 1 所述的一种肝细胞生长因子受体关键结构域融合蛋白,其特征在于:所述融合蛋白一共有 223 个氨基酸,如 SEQ ID NO:6 所示,其中 1-26 位氨基酸是表皮生长因子受体的强信号肽,27-217 位氨基酸是肝细胞生长因子受体 IPT34 蛋白,218-223 位氨基酸是 6 $\times$ His 标签。

5. 编码有如权利要求 1-4 中任一所述的肝细胞生长因子受体关键结构域融合蛋白的 DNA 的质粒。

6. 编码有如权利要求 1-4 中任一所述的肝细胞生长因子受体关键结构域融合蛋白的表达,包括在原核和真核细胞的表达。

7. 根据权利要求 1-4 中任一所述的一种肝细胞生长因子受体关键结构域融合蛋白,其特征在于:所述融合蛋白与肝细胞生长因子特异性结合。

8. 一种如权利要求 1 所述的肝细胞生长因子受体关键结构域融合蛋白的应用,其特征在于:所述融合蛋白用于抗体制备。

## 一种肝细胞生长因子受体关键结构域融合蛋白及其应用

### 技术领域

[0001] 本发明属于融合蛋白领域,特别涉及一种肝细胞生长因子受体关键结构域融合蛋白及其应用。

### 背景技术

[0002] C-Met 是肝细胞生长因子 (Hepatocyte growth factor, HGF) 已知的惟一受体,也是一类具有自身磷酸化活性的跨膜受体,在多种恶性肿瘤的发生发展中具有重要作用。临床研究发现, C-Met 的过量表达可见于在我国发病率较高的肺癌、肝癌、大肠癌、乳腺癌中,并与肿瘤恶性化程度、转移及预后呈正相关。C-Met 主要包括三个功能不同的结构域:胞外区、跨膜区和胞内区。C-Met 胞外区 (Extracellular domain, ED) 包含一个信号素轴向蛋白样的结构域、一个富含 Cys 的 PSI 区域以及四个免疫球蛋白  $\gamma$ -plexin- 转录因子同源结构组成的 IPT1-4 结构域。胞内区则包含一个近膜区、一个酪氨酸激酶区和一个 C- 末端序列。ED 作为配体识别部位识别并结合 HGF,对后续发生的一系列生物调节效应具有重要作用。近年来研究还发现, C-Met 胞外区中的 IPT3-4 结构域也是 C-Met 与 HGF 结合的关键结构域。HGF 与 ED 结合使其构象发生改变,磷酸化激活胞内多条信号转导途径,促进肿瘤细胞的增殖、迁徙和血管生成,最终导致恶性肿瘤细胞侵入正常组织,穿透脏器组织层,最终扩散至全身。一旦肿瘤细胞中异常活化的 HGF/C-Met 信号通路被阻断,肿瘤细胞就会出现细胞增殖减缓、成瘤性降低、侵袭能力下降等一系列变化。阻断 HGF/C-Met 信号通路的酪氨酸激酶抑制剂 cabozantinib 已于 2012 年获 FDA 批准用于晚期甲状腺髓样癌的治疗。目前,还没有 FDA 批准的抗 C-Met 的中和抗体药物。

[0003] 中和性抗体因其特异性强、副作用小等特点,在多种肿瘤治疗中发挥重要作用。中和性抗体不仅能与相应抗原受体特异性结合,还能有效中和配体对受体的作用,具有发展成为治疗性抗体药物的潜力,在人类疾病治疗中发挥关键作用。如经 FDA 批准治疗乳腺癌的中和抗体类药物曲妥珠单抗 (trastuzumab) 能特异性与人表皮生长因子受体 2 结合,阻断配体介导的信号转导,抑制肿瘤生长。经 FDA 批准用于治疗结肠癌和肺癌的西妥昔单抗 (Cetuximab) 也属中和抗体类药物,可以竞争性的抑制 EGFR 及其配体的结合,进而抑制肿瘤细胞生长,诱导其凋亡。

[0004] 所谓 C-Met 的中和抗体,就是能与 C-Met 结合,占据 C-Met 与 HGF 结合的关键结构域,阻止两者结合,阻断 HGF/C-Met 信号转导,进而抑制肿瘤细胞增殖生长。而 C-Met 与 HGF 结合的关键结构域是其胞外区 ED,特别是胞外区的 IPT3-4 结构域。

### 发明内容

[0005] 本发明所要解决的技术问题是提供一种肝细胞生长因子受体关键结构域融合蛋白及其应用,该融合蛋白可特异性的与肝细胞生长因子结合,在筛选阻断 C-Met/HGF 信号转导通路进而抑制肿瘤生长的中和抗体药物研发方面具有重大作用。

[0006] 本发明的一种肝细胞生长因子受体关键结构域融合蛋白,所述融合蛋白由表皮生

长因子受体的强信号肽、肝细胞生长因子受体关键结构域及 His 标签构成 ;其中,肝细胞生长因子受体的关键结构域为胞外区 ED(Extracellular domain) 或免疫球蛋白  $\gamma$ -plexin- 转录因子同源结构域 3-4IPT34(Immunoglobulin plexin transcription factor homology Domains) ;ED 的核苷酸序列如 SEQ ID NO:1 所示或是该序列的变体,ED 的氨基酸序列如 SEQ ID NO:2 所示或是该序列的变体 ;IPT34 的核苷酸序列如上述序列 SEQ ID NO:1 中 2152-2724 核苷酸序列所示或是该序列的变体, IPT34 的氨基酸序列如上述序列 SEQ ID NO:2 中 718-908 氨基酸序列所示或是该序列的变体。

[0007] 所述表皮生长因子受体的强信号肽的核苷酸序列如 SEQ ID NO:3 所示,氨基酸序列如 SEQ ID NO:4 所示。

[0008] 所述融合蛋白一共有 940 个氨基酸,如 SEQ ID NO:5 所示,其中 1-26 位氨基酸是表皮生长因子受体的强信号肽,27-934 位氨基酸是肝细胞生长因子受体 ED 蛋白,935-940 位氨基酸是 6 $\times$ His 标签。

[0009] 所述融合蛋白一共有 223 个氨基酸,如 SEQ ID NO:6 所示,其中 1-26 位氨基酸是表皮生长因子受体的强信号肽,27-217 位氨基酸是肝细胞生长因子受体 IPT34 蛋白,218-223 位氨基酸是 6 $\times$ His 标签。

[0010] 编码有肝细胞生长因子受体关键结构域融合蛋白的 DNA 的质粒。

[0011] 编码有肝细胞生长因子受体关键结构域融合蛋白的表达,包括在原核和真核细胞的表达。

[0012] 所述融合蛋白与肝细胞生长因子特异性结合。

[0013] 本发明的一种肝细胞生长因子受体关键结构域融合蛋白的应用,所述融合蛋白用于抗体制备。

[0014] 本发明 C-Met 胞外区 (Extracellular domain, ED) 包含一个信号素轴向蛋白样的结构域、一个富含 Cys 的 PSI 区域以及四个免疫球蛋白  $\gamma$ -plexin- 转录因子同源结构组成的 IPT1-4 结构域。现有技术 201110215692.8 的 IPT-Fc 是由 IPT1-4 组成,而本发明构建的 IPT34 只包括 IPT3 和 IPT4。c-Met 只有与 HGF 结合才会被激活,发挥功效。目前,c-Met 胞外区与 HGF 结合的关键结合位点还不清楚, IPT1-4 被认为是关键区域,本发明构建的 IPT34 进一步缩小了搜索范围。

[0015] 与现有技术 201110215692.8 的融合蛋白相比,本发明更换了 c-MET 本身的信号肽,而用了表皮生长因子受体 (EGFR) 的强信号肽,此信号肽已经验证过多次,可实现融合蛋白的胞外表达,并且不改变融合蛋白的三维结构。

[0016] 有益效果

[0017] (1) 利用基因工程技术表达了 C-Met 的关键结构域 ED 和 IPT34 融合蛋白,该区域是 C-Met 与其配体 HGF 结合的关键区域,在其信号传导过程中发挥非常重要的作用。

[0018] (2) 为使融合蛋白在 293T 细胞中高效表达,选用慢病毒表达系统,因慢病毒载体能有效整合入宿主基因组,介导目的基因稳定、长期表达,为制备 C-Met 单克隆抗体及筛选阻断 HGF/C-Met 信号通路的中和性抗体奠定基础 ;构建方法简单,成本低,具有良好的应用前景。

附图说明

[0019] 图 1 为琼脂糖电泳鉴定人 C-Met 的 ED 或 IPT34 区基因 PCR 产物电泳图,其中 (a) 是 ED 区基因电泳结果 (泳道 1 为 500bp Ladder ;泳道 2 为目的基因条带 2700bp) ;(b) 是 IPT34 区基因电泳结果 (泳道 1 为 DL2000DNA Ladder ;泳道 2 为目的基因条带 670bp) ;

[0020] 图 2 为慢病毒载体双酶切后电泳图,其中 (a) 是 p RRL-CMV-ED 双酶切电泳图 (泳道 1 为 500-15000bp DNA Marker ;泳道 2 为 p RRL-CMV-ED) ;(b) 是 p RRL-CMV-IPT34 双酶切电泳图 (泳道 1 为 DL2000DNA Ladder ;泳道 2 为 p RRL-CMV-IPT34) ;

[0021] 图 3 为荧光显微镜 ( $\times 100$ ) 观察转染慢病毒质粒的 293T 细胞,其中 (a) 和 (b) 为显微镜下观察 p RRL-CMV-ED 组的绿色荧光蛋白表达 (a :普通光镜 ;b :荧光光镜) ;(c) 和 (d) 为显微镜下观察 p RRL-CMV-ED 组的绿色荧光蛋白表达 (c :普通光镜 ;d :荧光光镜) ;(e) 和 (f) 为显微镜下观察阳性对照组的绿色荧光蛋白表达 (e :普通光镜 ;f :荧光光镜) ;

[0022] 图 4 为实施例 3 中肝细胞生长因子受体融合蛋白 SDS-PAGE 考马斯亮蓝结果 ;其中 (a) 是 C-Met-ED 融合蛋白 ;(b) 是 C-Met-IPT34 融合蛋白 ;

[0023] 图 5 为实施例 4 中 293T 细胞表达的肝细胞生长因子受体融合蛋白的活性检测结果 ;其中 (a) 是 C-Met-ED 融合蛋白 ;(b) 是 C-Met-IPT34 融合蛋白。

### 具体实施方式

[0024] 下面结合具体实施例,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而并不用于限制本发明的范围。此外应理解,在阅读了本发明讲授的内容之后,本领域技术人员可以对本发明作各种改动或修改,这些等价形式同样落于本申请所附权利要求书所限定的范围。

[0025] 实施例 1

[0026] ED-His 融合蛋白慢病毒表达载体的构建

[0027] 以含有人 EGFR 全长序列的 pCR-Blunt-EGFR 为模板,根据 EGFR 基因信号肽序列设计并合成引物 P1 和 P2 (苏州金唯智公司合成),并在上游引物添加酶切位点 Age I,扩增得到 78bp 的信号肽。以含有人 C-Met 全长基因的 pCR-Blunt-C-Met 为模板,根据 GenBank 上登陆的人 C-Met 基因胞外区序列设计并合成 PCR 引物 P3 和 P4 (苏州金唯智公司合成),在下游引物上添加限制性酶切位点 Sal I 及 His 标签,扩增得到约 2800bp 的目的片段 ED。以这两个 PCR 产物为模板,以 P1 和 P4 为引物,进行第二次 PCR 扩增反应,通过重叠延伸 PCR 将两个 PCR 产物连接在一起。其中,引物序列如下表所示 :

[0028]

引物名称	序列
P1	5'GCCACCGGTCGCCACCATGGTCAGCTACTGGGACACCG3'
P2	5'TCCGGAAGTAGATCCTGTGA3'
P3	5'TCACAGGATCTAGTTCGAGAGTGTAAGAGGCACTAGC3'
P4	5'ATTGTCGACCTAGTGATGGTGATGGTGATGTGTGAAATTCTGATCTGGTTG3'

[0029] 上述 PCR 反应体系如下 :

[0030]

试剂名称	体积 ( $\mu$ l)
上游引物	2.5
下游引物	2.5
10 $\times$ buffer	10
d NTP(10mM)	1
模板 DNA	<100ng
NEB 高保真酶	0.5
dd H <sub>2</sub> O	加水至 50 $\mu$ l

[0031] PCR 反应条件:预变性 98 $^{\circ}$ C, 3min;变性 98 $^{\circ}$ C 15s,退火 56 $^{\circ}$ C 30s;延伸 72 $^{\circ}$ C, 1.5min,循环 30 次;延伸 72 $^{\circ}$ C 10min,反应结束。

[0032] PCR 反应结束后,将上述 PCR 产物通过核酸电泳鉴定,用割胶回收试剂盒回收特异性条带,获得纯化的人 C-Met 的 ED 区基因,通过 Eppendorf 核酸蛋白测定仪检测其浓度,-20 保存备用。

[0033] 将纯化后的 C-Met-ED PCR 产物通过 T4 连接酶平末端连入 pCR-Blunt 克隆载体,连接产物转化大肠杆菌感受态 DH5 $\alpha$ ,涂布于含卡那霉素的 LB 固体培养基上培养过夜,挑选单菌落于含有卡那霉素的 LB 液体培养基上培养过夜,通过菌液 PCR 及酶切初步鉴定阳性克隆,阳性克隆命名为 Blunt-C-Met-ED,将 Blunt-C-Met-ED 重组质粒进行测序鉴定。

[0034] 将测序正确的 Blunt-C-Met-ED 质粒和慢病毒表达载体 p RRL-CMV 分别用 Age I 和 Sal I 双酶切,经胶回收试剂盒纯化后,将酶切所得的目的片段和慢病毒表达载体连接,连接产物转化大肠杆菌感受态 DH5 $\alpha$ ,涂布于含氨苄青霉素的 LB 固体培养基上培养过夜,挑选单菌落于含有氨苄青霉素的 LB 液体培养基上培养过夜,通过酶切初步鉴定阳性克隆,阳性克隆命名为 p RRL-CMV-ED,将 pRRL-CMV-ED 慢病毒表达质粒进行测序鉴定,证实所构建的片段即为 C-Met-ED 的编码基因。

[0035] 实施例 2

[0036] IPT34-His 融合蛋白慢病毒表达载体的构建

[0037] 以含有人 EGFR 全长序列的 pCR-Blunt-EGFR 为模板,根据 EGFR 基因信号肽序列设计并合成引物 P1 和 P2(苏州金唯智公司合成)。以 pCR-Blunt-C-Met 为模板,根据人 C-Met 基因 IPT34 区的序列设计并合成引物 P5 和 P6(苏州金唯智公司合成)。分别进行第一次 PCR 扩增反应,获得两个 PCR 产物。以这两个 PCR 产物为模板,以 P1 和 P6 为引物,进行第二次 PCR 扩增反应,通过重叠延伸 PCR 将两个 PCR 产物连接在一起。扩增产物首先克隆到 pCR-Blunt 克隆载体中,之后经过酶切、连接等步骤,最终成功得到慢病毒表达载体 pRRL-CMV-IPT34,并通过测序验证。PCR 扩增的具体操作步骤及后续慢病毒载体构建步骤同实施例 1。引物序列如下:

[0038]

引物名称	序列
P5	5'TCACAGGATCTAGTTCCGGACCCATTGTCTATGAAATTCATCC3'
P6	5'ATTGTCGACCTAGTGATGGTGATGGTGATGTGTG3'

[0039] 实施例 3

[0040] 重组 C-Met-ED 及 C-Met-IPT34 融合蛋白表达及纯化

[0041] 接种处于对数生长期的 293T 细胞,待细胞生长至 50%~60% 融合时,利用磷酸钙法将 pRRL-CMV-C-Met-ED 与含有绿色荧光蛋白的对照载体 (p RRL-CMV-GFP) 共转染 293T 细胞以表达 C-Met-ED 蛋白。同时,为在 293T 细胞中表达 C-Met-IPT34 蛋白,将 pRRL-CMV-C-Met-IPT34 和 p RRL-CMV-GFP 共转染 293T 细胞。以 p RRL-CMV-GFP 转染 293T 细胞为阳性对照,以未转染的 293T 细胞作为空白对照。细胞培养 48h 后,如图 3 所示,通过荧光显微镜观察发现 C-Met-ED 组、C-Met-IPT34 组及阳性对照组细胞均有绿色荧光蛋白表达,未转染质粒的阴性对照组细胞则无荧光表达,说明慢病毒转染成功。

[0042] 分别收集上述 C-Met-ED 组和 C-Met-IPT34 组的 293T 细胞培养液,10000rpm 离心 20min 取上清,将其经 0.45  $\mu$ m 的滤膜过滤。将滤过的培养液与 PBS 缓冲液等体积混合,通过 Ni-NTAagarose 纯化柱来纯化蛋白。纯化后液体可用 BCA 蛋白浓度测定试剂盒测试蛋白含量,用纯化蛋白与肝细胞生长因子结合试验检测蛋白的生物学结合活性,采用 SDS-PAGE 电泳检测其纯度和分子量。

[0043] 实施例 4

[0044] 融合蛋白与肝细胞生长因子的结合

[0045] 用 CBS 将重组人 HGF 蛋白稀释至 2  $\mu$ g/ml,按照 100  $\mu$ L/孔 4 $^{\circ}$ C 过夜包被 96 孔酶标板。PBST 洗涤 3 次后,用 1% BSA 37 $^{\circ}$ C 封闭 2h。将不同体积实施例 1、2 制备的融合蛋白加入酶标板中与 HGF 混匀,37 $^{\circ}$ C 作用 2h,同时以 PBS 作为阴性对照,PBST 洗涤 3 次。以山羊抗人 C-Met 多克隆抗体作为一抗,37 $^{\circ}$ C 作用 1h,PBST 洗涤 3 次。HRP 标记兔抗山羊 IgG 抗体作为二抗进行检测,37 $^{\circ}$ C 作用 1h。PBST 洗涤后每孔加入 100  $\mu$ l TMB 显色液,37 $^{\circ}$ C 避光显色 10min,终止反应后酶标仪测定 450nm 处的吸光度值。检测结果如图 5,与阴性对照相比,试验组的吸光值均高于对照组,结果表明实施例 1、2 制备的融合蛋白可以有效的与 HGF 结合,且随着融合蛋白量的增多,与 HGF 结合率增大。证实 C-Met-ED 组、C-Met-IPT34 融合蛋白表达成功,且具有与人 HGF 特异结合的能力。

[0001]

## SEQUENCE LISTING

&lt;110&gt; 同济大学苏州研究院

&lt;120&gt; 一种肝细胞生长因子受体关键结构域融合蛋白及其应用

&lt;130&gt; 1

&lt;160&gt; 12

&lt;170&gt; PatentIn version 3.3

&lt;210&gt; 1

&lt;211&gt; 2724

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;400&gt; 1

```

gagtgtaaag aggcactagc aaagtccgag atgaatgtga atatgaagta tcagcttccc      60
aacttcaccg cggaaacacc catccagaat gtcattctac atgagcatca cattttcctt      120
ggtgccacta actacaffa tgttttaa at gaggaagacc ttcagaaggt tgctgagtac      180
aagactgggc ctgtgctgga acaccagat tgtttccat gtcaggactg cagcagcaaa      240
gccaatltat caggaggtgt ttgaaagat aacatcaaca tggtcttagt tgctgacacc      300
tactatgatg atcaactcat tagctgtggc agcgtcaaca gagggacctg ccagcgacat      360
gtctttccc acaatcatic tgctgacata cagtcggagg ttcactgcat attctccc      420
cagatagaag agcccagcca gttcctgac tgtgtggtga gcgccctggg agccaaagtc      480
ctttcatctg taaaggaccg gttcatcaac ttctttgtag gcaataccat aaattctct      540
tatttcccag atcatccatt gcattcgata tcagtgagaa ggctaaaagga aacgaaagat      600
ggttttatgt tttgacgga ccagtcctac attgatgttt tacctgagtt cagagattct      660
tacceatta agtatgtcca tgcctttgaa agcaacaatt ttattactt ctgacggtc      720
caaagggaaa ctctagatgc tcagaatttt cacacaagaa taatcaggtt etgttccata      780

```

[0002]



aactctggat tgcattccta catggaaatg cctctggagt gtattctcac agaaaagaga	840
aaaaagagat ccacaaagaa ggaagtgttt aatatacttc aggcctgcgta tgcagcaag	900
cctggggccc agcttgctag acaaatagga gccagcctga atgatgacat tctttcggg	960
gtgttcgcac aaagcaagcc agattctgcc gaaccaatgg atcgatctgc catgtgtgca	1020
tteccatca aatatgcaa cgacttctc aacaagatcg tcaacaaaaa caatgtgaga	1080
tgtctccage atttttacgg acccaatcat gagcaactgct ttaataggac acttctgaga	1140
aatcatcag gctgtgaagc gcgccgtgat gaatacgaag cagagtttac cacagctttg	1200
cagcgcgttg acttattcat gggcaattc agcgaagfcc tctaacatc talatccacc	1260
ttcattaaag gagacctcac catagctaact cttgggacat cagagggctc cttcatgcag	1320
gtgtgtgttt ctcatcagg accatcaacc cctcatgtga attttctcct ggactcctat	1380
ccagtgtctc cagaagtgat tgtggagcat acattaaacc aaaatggcta cacactggtt	1440
atcactggga agaagatcac gaagatccca ttgaatggct tgggctgcag acatttcag	1500
tctgcagtc aatgcctctc tgcctccacc ttgttcagt gtggctggtg ccacgacaaa	1560
tgtgtcggat cggaggaatg cctgagcggg acatggactc aacagatctg tctgcctgca	1620
atctacaagg ttctccaaa tagtgcacc cttgaaggag ggacaagget gaccatattg	1680
ggctgggact ttggatttcg gaggaataat aaattgatt taaagaaaac tagattctc	1740
cttgaaaatg agagctgcac cttgacttta agtgagagca ccatgaatac atfgaatgc	1800
acagtggtc ctgcatgaa taagcatctc aatatgtcca taattattc aatggccac	1860
gggacaacac aatacagtac attctctat gtggatctcg taataacaag tatttcgccc	1920
aaatacggtc ctatggtcgg tggcaattta ctactttaa ctggaaatta cctaaacagt	1980
gggaattcta gacacattc aattggctga aaaacatgta cttaaaaaag tgtgtcaaac	2040
agtattcttg aatgttatac cccagcccaa accatttcaa ctgagtttgc tgttaaattg	2100

[0003]

aaaattgact tagccaaccg agagacaagc atcttcagtt accgtgaaga tcccattgtc 2160  
 tatgaaattc atccaaccaa atcttttatt agtgggtggga gcacaataac aggtgttggg 2220  
 aaaaacctga attcagttag tgteccgaga atggtcataa atgtgcatga agcaggaagg 2280  
 aactttacag tggcatgtca acatcgctct aatteagaga taatctgttg taccactect 2340  
 fecctgcaac agctgaatct gcaactcccc ctgaaaacca aagcctttt catgttagat 2400  
 gggatccttt ccaaatactt tgatctcatt tatgtacata atcctgtgtt taagcctttt 2460  
 gaaaagccag tgatgatctc aatgggcaat gaaaatgtac tggaaattaa gggaaatgat 2520  
 atfgaccctg aagcagttaa aggtgaagtg ttaaagtg gaaataagag ctgtgagaat 2580  
 atacacttac attctgaagc cgttttatgc acggccccca atgacctgct gaaattgaac 2640  
 agcgagctaa atatagagtg gaagcaagca atttctcaa cegtccttgg aaaagtaata 2700  
 gttcaaccag atcagaattt caca 2724

<210> 2

<211> 908

<212> PRT

<213> 人工序列

<400> 2

Glu Cys Lys Glu Ala Leu Ala Lys Ser Glu Met Asn Val Asn Met Lys  
 1 5 10 15

Tyr Gln Leu Pro Asn Phe Thr Ala Glu Thr Pro Ile Gln Asn Val Ile  
 20 25 30

Leu His Glu His His Ile Phe Leu Gly Ala Thr Asn Tyr Ile Tyr Val  
 35 40 45

Leu Asn Glu Glu Asp Leu Gln Lys Val Ala Glu Tyr Lys Thr Gly Pro  
 50 55 60

[0004]

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

[0005]

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

[0006]

Ser Ile Ser Thr Phe Ile Lys Gly Asp Leu Thr Ile Ala Asn Leu Gly  
 420 425 430

Thr Ser Glu Gly Arg Phe Met Gln Val Val Val Ser Arg Ser Gly Pro  
 435 440 445

Ser Thr Pro His Val Asn Phe Leu Leu Asp Ser His Pro Val Ser Pro  
 450 455 460

Glu Val Ile Val Glu His Thr Leu Asn Gln Asn Gly Tyr Thr Leu Val  
 465 470 475 480

Ile Thr Gly Lys Lys Ile Thr Lys Ile Pro Leu Asn Gly Leu Gly Cys  
 485 490 495

Arg His Phe Gln Ser Cys Ser Gln Cys Leu Ser Ala Pro Pro Phe Val  
 500 505 510

Gln Cys Gly Trp Cys His Asp Lys Cys Val Arg Ser Glu Glu Cys Leu  
 515 520 525

Ser Gly Thr Trp Thr Gln Gln Ile Cys Leu Pro Ala Ile Tyr Lys Val  
 530 535 540

Phe Pro Asn Ser Ala Pro Leu Glu Gly Gly Thr Arg Leu Thr Ile Cys  
 545 550 555 560

Gly Trp Asp Phe Gly Phe Arg Arg Asn Asn Lys Phe Asp Leu Lys Lys  
 565 570 575

Thr Arg Val Leu Leu Gly Asn Glu Ser Cys Thr Leu Thr Leu Ser Glu  
 580 585 590

[0007]

Ser Thr Met Asn Thr Leu Lys Cys Thr Val Gly Pro Ala Met Asn Lys  
 595 600 605

His Phe Asn Met Ser Ile Ile Ile Ser Asn Gly His Gly Thr Thr Gln  
 610 615 620

Tyr Ser Thr Phe Ser Tyr Val Asp Pro Val Ile Thr Ser Ile Ser Pro  
 625 630 635 640

Lys Tyr Gly Pro Met Ala Gly Gly Thr Leu Leu Thr Leu Thr Gly Asn  
 645 650 655

Tyr Leu Asn Ser Gly Asn Ser Arg His Ile Ser Ile Gly Gly Lys Thr  
 660 665 670

Cys Thr Leu Lys Ser Val Ser Asn Ser Ile Leu Glu Cys Tyr Thr Pro  
 675 680 685

Ala Gln Thr Ile Ser Thr Glu Phe Ala Val Lys Leu Lys Ile Asp Leu  
 690 695 700

Ala Asn Arg Glu Thr Ser Ile Phe Ser Tyr Arg Glu Asp Pro Ile Val  
 705 710 715 720

Tyr Glu Ile His Pro Thr Lys Ser Phe Ile Ser Gly Gly Ser Thr Ile  
 725 730 735

Thr Gly Val Gly Lys Asn Leu Asn Ser Val Ser Val Pro Arg Met Val  
 740 745 750

Ile Asn Val His Glu Ala Gly Arg Asn Phe Thr Val Ala Cys Gln His  
 755 760 765

[0008]

Arg Ser Asn Ser Glu Ile Ile Cys Cys Thr Thr Pro Ser Leu Gln Gln  
 770 775 780

Leu Asn Leu Gln Leu Pro Leu Lys Thr Lys Ala Phe Phe Met Leu Asp  
 785 790 795 800

Gly Ile Leu Ser Lys Tyr Phe Asp Leu Ile Tyr Val His Asn Pro Val  
 805 810 815

Phe Lys Pro Phe Glu Lys Pro Val Met Ile Ser Met Gly Asn Glu Asn  
 820 825 830

Val Leu Glu Ile Lys Gly Asn Asp Ile Asp Pro Glu Ala Val Lys Gly  
 835 840 845

Glu Val Leu Lys Val Gly Asn Lys Ser Cys Glu Asn Ile His Leu His  
 850 855 860

Ser Glu Ala Val Leu Cys Thr Val Pro Asn Asp Leu Leu Lys Leu Asn  
 865 870 875 880

Ser Glu Leu Asn Ile Glu Trp Lys Gln Ala Ile Ser Ser Thr Val Leu  
 885 890 895

Gly Lys Val Ile Val Gln Pro Asp Gln Asn Phe Thr  
 900 905

<210> 3

<211> 78

<212> DNA

<213> 人工序列

<400> 3

[0009]

atggtcagct actgggacac cggggctctg ctgtgcgcgc tgcacagctg tctgcttctc 60

acaggatcta gttccgga 78

<210> 4

<211> 26

<212> PRT

<213> 人工序列

<400> 4

Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser  
1 5 10 15

Cys Leu Leu Leu Thr Gly Ser Ser Ser Gly  
20 25

<210> 5

<211> 940

<212> PRT

<213> 人工序列

<400> 5

Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser  
1 5 10 15

Cys Leu Leu Leu Thr Gly Ser Ser Ser Gly Glu Cys Lys Glu Ala Leu  
20 25 30

Ala Lys Ser Glu Met Asn Val Asn Met Lys Tyr Gln Leu Pro Asn Phe  
35 40 45

Thr Ala Glu Thr Pro Ile Gln Asn Val Ile Leu His Glu His His Ile  
50 55 60

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

[0010]



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

[0011]

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

[0012]

	420	425	430	
Met Gly Gln Phe Ser Glu Val Leu Leu Thr Ser Ile Ser Thr Phe Ile				
	435	440	445	
Lys Gly Asp Leu Thr Ile Ala Asn Leu Gly Thr Ser Glu Gly Arg Phe				
	450	455	460	
Met Gln Val Val Val Ser Arg Ser Gly Pro Ser Thr Pro His Val Asn				
465	470	475	480	
Phe Leu Leu Asp Ser His Pro Val Ser Pro Glu Val Ile Val Glu His				
	485	490	495	
Thr Leu Asn Gln Asn Gly Tyr Thr Leu Val Ile Thr Gly Lys Lys Ile				
	500	505	510	
Thr Lys Ile Pro Leu Asn Gly Leu Gly Cys Arg His Phe Gln Ser Cys				
	515	520	525	
Ser Gln Cys Leu Ser Ala Pro Pro Phe Val Gln Cys Gly Trp Cys His				
	530	535	540	
Asp Lys Cys Val Arg Ser Glu Glu Cys Leu Ser Gly Thr Trp Thr Gln				
545	550	555	560	
Gln Ile Cys Leu Pro Ala Ile Tyr Lys Val Phe Pro Asn Ser Ala Pro				
	565	570	575	
Leu Glu Gly Gly Thr Arg Leu Thr Ile Cys Gly Trp Asp Phe Gly Phe				
	580	585	590	
Arg Arg Asn Asn Lys Phe Asp Leu Lys Lys Thr Arg Val Leu Leu Gly				

[0013]

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

[0014]

770	775	780	
Gly Arg Asn Phe Thr Val Ala Cys Gln His Arg Ser Asn Ser Glu Ile			
785	790	795	800
Ile Cys Cys Thr Thr Pro Ser Leu Gln Gln Leu Asn Leu Gln Leu Pro			
	805	810	815
Leu Lys Thr Lys Ala Phe Phe Met Leu Asp Gly Ile Leu Ser Lys Tyr			
	820	825	830
Phe Asp Leu Ile Tyr Val His Asn Pro Val Phe Lys Pro Phe Glu Lys			
	835	840	845
Pro Val Met Ile Ser Met Gly Asn Glu Asn Val Leu Glu Ile Lys Gly			
850	855	860	
Asn Asp Ile Asp Pro Glu Ala Val Lys Gly Glu Val Leu Lys Val Gly			
865	870	875	880
Asn Lys Ser Cys Glu Asn Ile His Leu His Ser Glu Ala Val Leu Cys			
	885	890	895
Thr Val Pro Asn Asp Leu Leu Lys Leu Asn Ser Glu Leu Asn Ile Glu			
	900	905	910
Trp Lys Gln Ala Ile Ser Ser Thr Val Leu Gly Lys Val Ile Val Gln			
	915	920	925
Pro Asp Gln Asn Phe Thr His His His His His His			
930	935	940	

<210> 6

[0015]

&lt;211&gt; 223

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;400&gt; 6

Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser  
 1                                      5                                      10                                      15

Cys Leu Leu Leu Thr Gly Ser Ser Ser Gly Pro Ile Val Tyr Glu Ile  
                                         20                                      25                                      30

His Pro Thr Lys Ser Phe Ile Ser Gly Gly Ser Thr Ile Thr Gly Val  
                                         35                                      40                                      45

Gly Lys Asn Leu Asn Ser Val Ser Val Pro Arg Met Val Ile Asn Val  
                                         50                                      55                                      60

His Glu Ala Gly Arg Asn Phe Thr Val Ala Cys Gln His Arg Ser Asn  
 65                                      70                                      75                                      80

Ser Glu Ile Ile Cys Cys Thr Thr Pro Ser Leu Gln Gln Leu Asn Leu  
                                         85                                      90                                      95

Gln Leu Pro Leu Lys Thr Lys Ala Phe Phe Met Leu Asp Gly Ile Leu  
                                         100                                      105                                      110

Ser Lys Tyr Phe Asp Leu Ile Tyr Val His Asn Pro Val Phe Lys Pro  
                                         115                                      120                                      125

Phe Glu Lys Pro Val Met Ile Ser Met Gly Asn Glu Asn Val Leu Glu  
                                         130                                      135                                      140

Ile Lys Gly Asn Asp Ile Asp Pro Glu Ala Val Lys Gly Glu Val Leu  
 145                                      150                                      155                                      160

[0016]

Lys Val Gly Asn Lys Ser Cys Glu Asn Ile His Leu His Ser Glu Ala  
 165 170 175

Val Leu Cys Thr Val Pro Asn Asp Leu Leu Lys Leu Asn Ser Glu Leu  
 180 185 190

Asn Ile Glu Trp Lys Gln Ala Ile Ser Ser Thr Val Leu Gly Lys Val  
 195 200 205

Ile Val Gln Pro Asp Gln Asn Phe Thr His His His His His His  
 210 215 220

- <210> 7
- <211> 38
- <212> DNA
- <213> 人工序列

<400> 7  
 gccaccggtc gccaccatgg tcagctactg ggacaccg 38

- <210> 8
- <211> 20
- <212> DNA
- <213> 人工序列

<400> 8  
 tccggaacta gatcctgtga 20

- <210> 9
- <211> 40
- <212> DNA
- <213> 人工序列

<400> 9  
 tcacaggatc tagttccgga gagtgtaaag aggcactagc 40

[0017]

---

<210>	10	
<211>	51	
<212>	DNA	
<213>	人工序列	
<400>	10	
	attgtcgacc tagtgatggt gatggtgatg tgtgaaattc tgaictggtt g	51
<210>	11	
<211>	43	
<212>	DNA	
<213>	人工序列	
<400>	11	
	tcacaggate tagttccgga cccattgtct atgaaattca tcc	43
<210>	12	
<211>	34	
<212>	DNA	
<213>	人工序列	
<400>	12	
	attgtcgacc tagtgatggt gatggtgatg tgtg	34



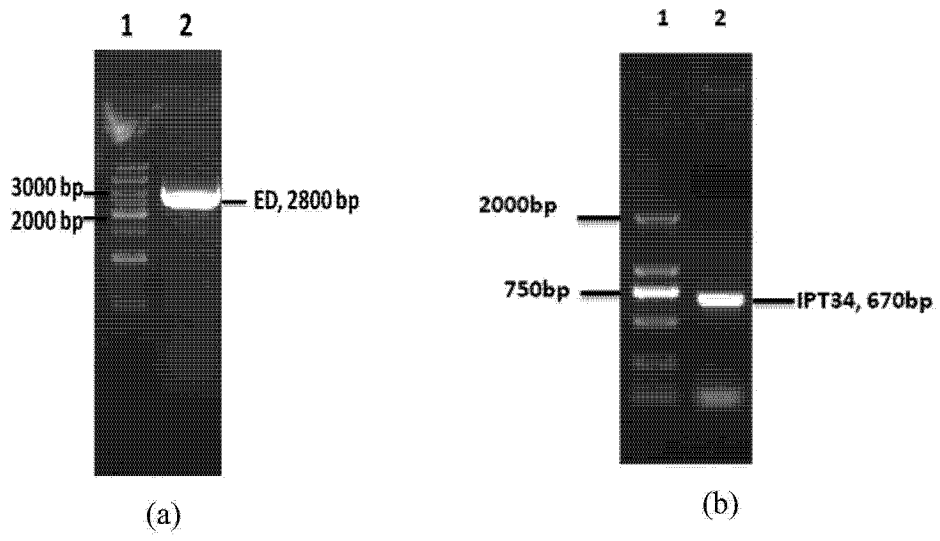


图 1

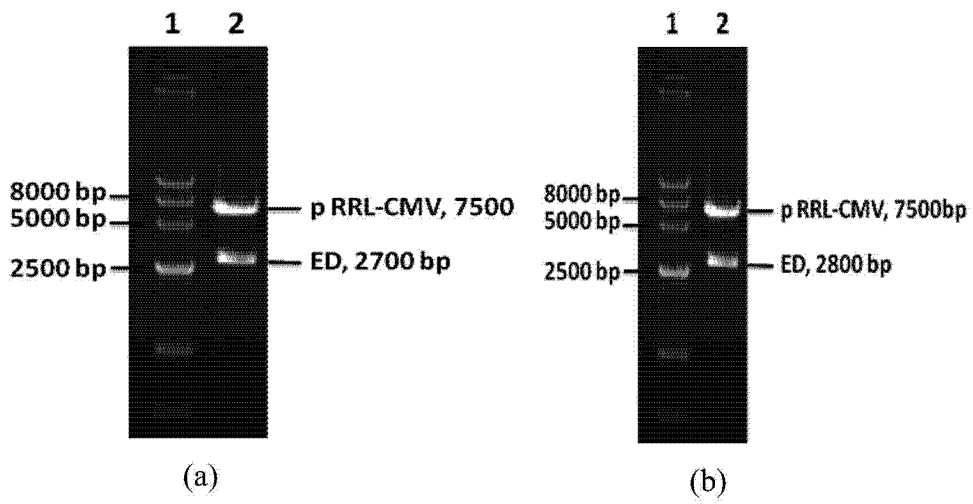


图 2

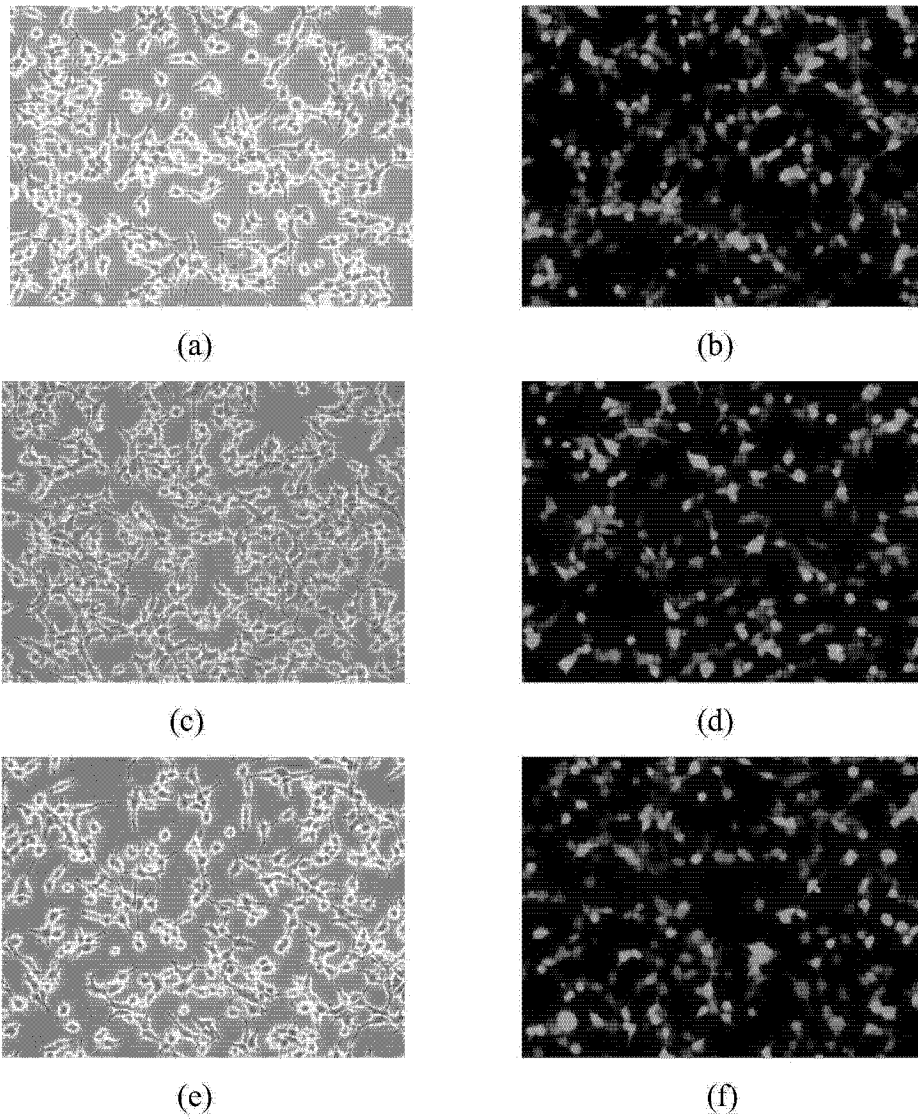


图 3

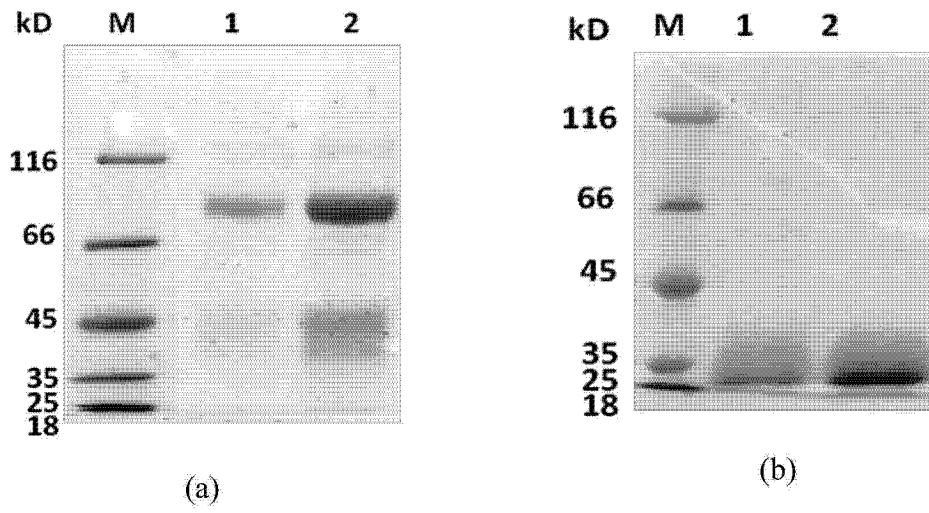


图 4

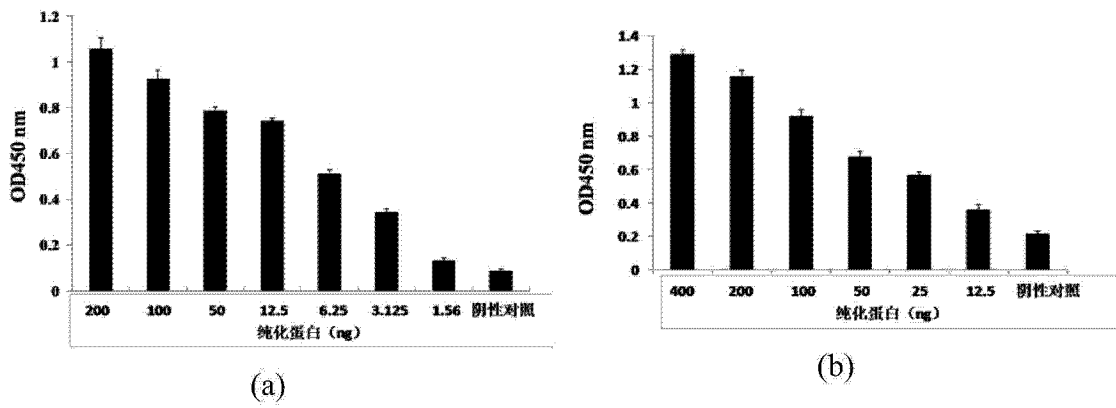


图 5