

[19] 中华人民共和国国家知识产权局

[ 51 ] Int. Cl<sup>7</sup>

A61K 38/00

A61K 39/00 C07K 1/00

C07K 14/00 C07K 17/00

G01N 33/53 G01N 33/567

G01N 33/574



# [12] 发明专利说明书

[21] ZL 专利号 97197769.0

[45] 授权公告日 2004 年 6 月 30 日

[11] 授权公告号 CN 1155402C

[22] 申请日 1997.7.18 [21] 申请号 97197769.0

[30] 优先权

[32] 1996.7.22 [33] US [31] 08/681,219

[86] 国际申请 PCT/US1997/012677 1997.7.18

[87] 国际公布 WO1998/005347 英 1998.2.12

[85] 进入国家阶段日期 1999.3.9

[71] 专利权人 纽约市哥伦比亚大学信托人

地址 美国纽约

[72] 发明人 塔卡-阿克·萨托 朱恩·柳-萨瓦

审查员 曲 燕

[74] 专利代理机构 中科专利商标代理有限责任公  
司

代理人 严 舫

权利要求书 4 页 说明书 50 页 附图 26 页

[54] 发明名称 一种抑制信号-传导蛋白和 GLGF  
(PDZ/DHR) 区域之间相互作用的  
化合物及其应用

[57] 摘要

本发明提供了一种组合物，它能够抑制信号传导蛋白和胞浆蛋白之间的特异结合。本发明也提供了一种鉴定一种能够抑制信号传导蛋白和胞浆蛋白之间的特异结合的化合物。本发明也提供了一种抑制癌细胞增殖的方法。本发明也提供了用一种化合物以可有效使细胞凋亡的量治疗癌症的方法。本发明也提供了一种抑制病毒感染的细胞繁殖的方法。本发明还提供了一种用一种化合物以可有效使细胞凋亡的量治疗病毒感染患者的方法。本发明也提供了药物组合物。

I S S N 1 0 0 8 - 4 2 7 4

1. 一种包含(S/T)-X-(V/I)-COOH 氨基酸序列的肽化合物, 在该序列中每个“-”代表一个肽键,每个括号包含可以相互替换的氨基酸, 以及括号内斜线分开可相互替换的氨基酸,“X”代表任何除 G 以外的氨基酸,该氨基酸是从一组含有 20 种天然存在的氨基酸中选择出来的, 其中该肽抑制一种信号-传导蛋白和一种含有(G/S/A/E)-L-G-(F/I/L)氨基酸序列的胞浆蛋白之间的特异结合。

2. 权利要求 1 的肽化合物, 其中这种肽含 DSENSNFRNEIQSLV 氨基酸序列。

3. 权利要求 1 的肽化合物, 其中这种肽含 RNEIQSLV 氨基酸序列。

4. 权利要求 1 的肽化合物, 其中这种肽含 NEIQSLV 氨基酸序列。

5. 权利要求 1 的肽化合物, 其中这种肽含 EIQSLV 氨基酸序列。

6. 权利要求 1 的肽化合物, 其中这种肽含 IQSLV 氨基酸序列。

7. 权利要求 1 的肽化合物, 其中这种肽含 QSLV 氨基酸序列。

8. 权利要求 1 的肽化合物, 其中这种肽含 SLV 氨基酸序列。

9. 权利要求 1 的肽化合物, 其中这种肽含 IPPDSEDGNEEQSLV 氨基酸序列。

10. 权利要求 1 的肽化合物, 其中这种肽含 DSEMYNFRSQLASVV 氨基酸序列。

11. 权利要求 1 的肽化合物, 其中这种肽含 SDSNMNMNELSEV 氨基酸序列。

12. 权利要求 1 的肽化合物, 其中这种肽含 QNFRITYIVSFV 氨基酸序列。

13. 权利要求 1 的肽化合物, 其中这种肽含 RETIESTV 氨基酸

序列。

14. 权利要求 1 的肽化合物, 其中这种肽含 RGFISLV 氨基酸序列。

15. 权利要求 1 的肽化合物, 其中这种肽含 TIQSVI 氨基酸序列。

16. 权利要求 1 的肽化合物, 其中这种肽含 ESLV 氨基酸序列。

17. 一种鉴定权利要求 1 的肽化合物的方法, 该方法包括:

(a) 在允许一种已知物和结合胞浆蛋白的信号-传导蛋白之间结合形成复合物的条件下, 所述已知物预先显示能 (i) 替换结合了胞浆蛋白的信号-传导蛋白和(ii)与胞浆蛋白形成信号-传导蛋白不再结合的复合物, 将已结合了信号-传导蛋白的胞浆蛋白与各种肽化合物接触; 和

(b) 检测步骤 (a) 中被替换的信号-传导蛋白或复合物, 其中, 任何被替换的信号-传导蛋白或者来自步骤 (a) 复合物的存在表示这种肽化合物能抑制信号-传导蛋白和胞浆蛋白之间的特异结合,

其中所述胞浆蛋白含有(G/S/A/E)-L-G-(F/I/L)氨基酸序列。

18. 权利要求 17 的方法, 其中, 该信号-传导蛋白和该胞浆蛋白之间的特异结合的抑制影响报告基因的转录活性。

19. 权利要求 17 的方法, 在步骤 (b) 中被替换的信号-传导蛋白或该复合物的检测是通过比较步骤 (a) 中与该肽化合物接触前后报告基因的转录活性来进行的, 其中该活性的改变表明该信号-传导蛋白和该胞浆蛋白之间的特异结合被抑制而该信号-传导蛋白被替换。

20. 权利要求 17 的方法, 该方法中胞浆蛋白是结合到固相支持物上。

21. 权利要求 17 的方法, 该方法中化合物是结合到固相支持物上。

22. 权利要求 17 的方法, 该方法中步骤(a)中的接触是在体外进行的。

23. 权利要求 17 的方法, 该方法中步骤(a)中的接触是在体内进

行的。

24. 权利要求 22 的方法，该方法中步骤(a)中的接触是在酵母细胞中进行的。

25. 权利要求 22 的方法，该方法中步骤(a)中的接触是在哺乳动物细胞中进行的。

26. 权利要求 17 的方法，其中，该信号-传导蛋白是一种细胞表面受体。

27. 权利要求 26 的方法，该方法中细胞表面受体是 CD4 受体。

28. 权利要求 26 的方法，该方法中细胞表面受体是 p75 受体。

29. 权利要求 26 的方法，该方法中细胞表面受体是血清素 2A 受体。

30. 权利要求 26 的方法，该方法中细胞表面受体是血清素 2B 受体。

31. 权利要求 1 的肽化合物在制备用于抑制癌细胞增殖的药物中的用途。

32. 权利要求 31 的用途，其中癌细胞来自胸腺、肝脏、肾脏、结肠、卵巢、乳腺、睾丸、脾脏、胃、前列腺、子宫、皮肤、头和颈。

33. 权利要求 31 的用途，其中癌细胞来自 T 细胞和 B 细胞。

34. 权利要求 1 的肽化合物在制备用于对患者治疗肿瘤的药物中的用途，其中制备该药物以向患者的癌细胞中以可有效使该癌细胞凋亡的量给药。

35. 权利要求 34 的用途，其中癌细胞来自胸腺、肝脏、肾脏、结肠、卵巢、乳腺、睾丸、脾脏、胃、前列腺、子宫、皮肤、头和颈。

36. 权利要求 34 的用途，其中癌细胞来自 T 细胞和 B 细胞。

37. 权利要求 1 的肽化合物在制备用于抑制病毒感染的细胞的增殖的药物中的用途。

38. 权利要求 37 的用途，其中的细胞被 B 型肝炎病毒、EB 病毒、流感病毒、乳头瘤病毒、腺病毒、人 T 细胞亲淋巴病毒 1 型或

HIV 感染。

39. 权利要求 1 的肽化合物在制备用于治疗病毒感染患者的药物中的用途，其中制备该药物以向患者的病毒感染的细胞中以可有效使该癌细胞凋亡的量给药。

40. 权利要求 39 的用途，其中的细胞被 B 型肝炎病毒、EB 病毒、流感病毒、乳头瘤病毒、腺病毒、人 T 细胞亲淋巴病毒 1 型或 HIV 感染。

41. 一种药物组合物，包含有效量的权利要求 1 的肽化合物和一种药用可接受的载体。

一种抑制信号-传导蛋白和GLGF (PDZ/DHR)  
区域之间相互作用的化合物及其应用

本发明是得到美国政府的支持,由美国卫生与人类事业局的国立卫生研究所 (NIH) 提供资助。基金号为: R01GM55147-01。因此,美国政府在此发明中拥有特定权利。

### 背景

在本申请全文中,各种出版物引证了其作者和年代,其全部引文是按字母顺序列于本申请专利说明书的结尾,紧靠在所列的序列和权利要求之前。这些文献的全部内容引入本申请,以便于更全面地叙述在本文所述的发明完成时普通技术人员已知的现有技术。

Fas (APO-1/CD95) 和它的配体被认为是凋亡(apoptosis)中重要的信号中介体 (Itoh, et al. 1991), Fas (Apo-1/CD95) 的结构组织形式提示它是肿瘤坏死因子受体超家族中的一个成员,该超家族还包括 p75神经生长因子受体 (NGFR) (Johnson, et al. 1986), T-细胞-活性标记 CD27 (Camerini, et al.. 1991), 何杰金-淋巴瘤-相关抗原 CD30 (Smith, et al. 1993), 人类B细胞抗原 CD40 (Stamenkovic, et al.. 1989), T细胞抗原 OX40 (Mallett, et al. 1990)。Fas以及它的配体的遗传突变均与小鼠的淋巴组织增殖失调和自身免疫失调有关 (Watanabe-Fukunaga, et al. 1992; Takahashi, et al. 1994)。而且,当T-细胞感染人类免疫缺陷病毒 (HIV) 时, Fas表达水平改变被认为是诱导凋亡 (Westendorp, et al. 1995)。

用酵母双杂交和生化方法已鉴定出几种Fas-相互作用的信号传导分子,如Fas-结合的磷酸酶-1 (FAP-1) (图1) (Sato, et al. 1995), FADD/MORT1/CAP-1/CAP-2 (Chinnaiyan, et al. 1995; Boldin, et

al. 1995;Kischkel, et al. 1995) 和 RIP ( Stanger, et al. 1995)。除了FAP-1以外, 所有分子都与Fas功能性细胞死亡区域有关, 当细胞转染了这些蛋白质时, 还与FADD/MORT1或RIP过量表达诱导凋亡有关。相反, FAP-1是唯一的与Fas的负调控区(C-端15氨基酸)有关(Ito, et al. 1993), 并且抑制Fas-诱导的凋亡的蛋白质。

FAP-1 (PTPN13) 有几种替换的剪接形式与PTP-BAS/hPTP1E/PTPL1一致 (Maekawa, et al. 1994; Banville, et al. 1994; Saras, et al. 1994)。并含有一个膜结合区, 它与在细胞骨架相关蛋白, ezrin (Gould et al. 1989), radixin (Funayama et al. 1991), moesin (Lankes, et al. 1991), II型神经纤维瘤基因产物 (NFII) (Rouleau, et al. 1993), 和4.1蛋白 (Conboy, et al. 1991), 以及在PTP酶PTPH1 (Yang, et al. 1991), PTP-MEG (Gu, et al. 1991), 和PTPD1 (Vogel, et al. 1993) 中发现的相似。FAP-1包括了令人感兴趣6个GLGF (PDZ/DHR) 重复, 被认为在蛋白区域中介导分子内和分子间的相互作用。FAP-1的第三个GLGF重复首先被确认为与Fas受体的C-端具有特异的相互作用区域 (Sato, et al. 1995)。提示了GLGF区在将这些蛋白定位于亚膜细胞骨架和/或调节生化活性上起着重要作用。GLGF重复最初是在鸟苷酸激酶和大鼠的后突触密度蛋白中 (PSD-95) (Cho, et al. 1992) 发现的, 与果蝇肿瘤抑制蛋白, Lethal-(1)-disc-large-1 [dlg-1] 同源, (Woods, et al. 1991; Kitamura, et al. 1994)。这些重复可以介导同源和异源二聚化, 此二聚化能强烈地影响PTP酶的活性、影响与Fas结合、和/或FAP-1同其他信号传导蛋白的相互作用。最近, 报道了这些蛋白的不同PDZ区域与离子通道和其他蛋白C-端的相互作用 (图1) (表1) (Kornau, et al. 1995; Kim, et al. 1995; Matsumine, et al. 1996)。

表1. 与PDZ区域相互作用的蛋白

蛋白	C-端序列	相关蛋白	参考文献
Fas (APO-1/CD95)	SLV	FAP-1	2
NMDA受体 NR2亚单位	SDV	PSD95	3
Shaker-型K <sup>+</sup> 通道	TDV	PSD95&DLG	4
APC	TEV	DLG	5

### 发明概述

本发明提供了一个肽化合物，它能够抑制一种信号-传导蛋白和一种含有 (G/S/A/E) -L-G- (F/I/L) (SEQ ID NO:1号) 氨基酸序列的胞浆蛋白间特异结合。而且，该胞浆蛋白可认包括 (K/R/Q) -X<sub>n</sub>- (G/S/A/E) -L-G- (F/I/L) 氨基酸序列 (SEQ ID NO:2号)，其中，“X”代表任何一种从一组包含20种天然存在的氨基酸中选择出来的氨基酸，“n”代表大于等于2小于等于4。在一个优选实施方案，该氨基酸序列是 SLGI (SEQ ID NO:3号)。此外，本发明提供了该信号-传导蛋白的羧基端含有 (S/T) -X- (V/I/L) (SEQ ID NO:4号) 氨基酸序列的一种肽化合物，其中，每个“-”代表一个肽键，每个括号包含了可相互替换的氨基酸，括号内每个斜线分开的氨基酸是可以相互替换的，“X”代表任何一种从一组包含20种天然存在的氨基酸中选择出来的氨基酸。

本发明也提供一种鉴定化合物的方法，该化合物能够抑制信号-传导蛋白和含有 (G/S/A/E) -L-G- (F/I/L) 氨基酸序列的胞浆蛋白间的特异结合。此外，本发明提供了一个鉴定化合物的方法，该化合物能够抑制羧基端含有 (S/T) -X- (V/L/I) 氨基酸序列的信号-传导蛋白和胞浆蛋白间的特异结合。

本发明也提供了一个抑制癌细胞增殖的方法，特别是，所述癌细胞



来自包括：结肠、肝脏、乳腺、卵巢、睾丸、肺、胃、脾、肾、前列腺、子宫、皮肤、头部、胸腺和颈部的器官，或这些细胞来自T-细胞或B-细胞。

本发明提供了一种对患者以可有效使所述细胞发生凋亡的量的组合物治疗癌症的方法，特别是，癌细胞来源于以下器官，包括：胸腺、结肠、肝脏、乳腺、卵巢、睾丸、肺、胃、脾、肾、前列腺、子宫、皮肤、头部、和颈部，或来自T-细胞或B-细胞。

本发明提供了一种抑制病毒感染细胞增殖的方法，特别是，病毒性感染的细胞是被B型肝炎病毒、EB病毒、流感病毒、乳头瘤病毒、腺病毒、人类T-细胞亲淋巴性病毒、I型或HIV感染的。

本发明也提供一种药用性组合物，该组合物包含能抑制信号-传导蛋白和胞浆蛋白间的特异结合的肽化合物。

本发明也提供一种药用性组合物，包含被鉴定为能抑制信号-传导蛋白和胞浆蛋白间的特异结合的肽化合物。

#### 附图简要说明

图1. Fas-结合的磷酸蛋白酶-1的图，表明有6个GLGF (PDZ/DHR) 区重复；与其他蛋白和含有GLGF (PDZ/DHR) 重复的蛋白的膜结合位点进行相似性比较。

图2A, 2B, 2C和2D. Fas C端与FAP-1结合所需的最小区域图谱。右侧的数字代表每个单独的克隆（图2C和2D）。

2A 用酵母双杂交系统筛选随机肽基因文库的策略。

2B 人类（SEQ ID NO:5号），大鼠（SEQ ID NO:6号），和小鼠（SEQ ID NO:7号）Fas C端15个氨基酸的排列。

2C 筛选半随机肽基因文库的结果。顶上一排表示人类和大鼠固定同源的氨基酸。破折号线表示未改变的氨基酸。

2D 筛选随机肽文库的结果（分别为：

SEQ ID NO:8号，	SEQ ID NO:9号，	SEQ ID NO:10号，
SEQ ID NO:11号，	SEQ ID NO:12号，	SEQ ID NO:13号，
SEQ ID NO:14号，	SEQ ID NO:15号，	SEQ ID NO:16号，

SEQ ID NO:17号, )

**3A, 3B 和 3C. Fas/FAP-1体外结合的抑制试验**

**3A** 使用Fas C端15个氨基酸对Fas/FAP-1结合的抑制试验。GST-Fas融合蛋白(191-355)用于体外结合试验(1, 3-10泳道)。GST-Fas融合蛋白(191-320)(2泳道)和1mM人类PAMP(肾上腺髓质素原的N-端20个氨基酸, 分子量2460.9)(3泳道)作为阴性对照。加入的C-端15个氨基酸浓度是1 $\mu$ M(4泳道), 3 $\mu$ M(5泳道), 10 $\mu$ M(6泳道), 30 $\mu$ M(7泳道), 100 $\mu$ M(8泳道), 300 $\mu$ M(9泳道), 1000 $\mu$ M(10泳道)。

**3B** 使用相应于Fas的C-末端15个氨基酸的肽进行的Fas/FAP-1结合抑制试验。在这个抑制试验中所有的合成肽都被乙酰化(分别为, SEQ ID NO:4号, SEQ ID NO:18号, SEQ ID NO:19号, SEQ ID NO:20号, SEQ ID NO:21号, SEQ ID NO:22号, SEQ ID NO:23号)。

**3C** 将三肽中的一种替换后, 对Fas/FAP-1结合的抑制效应。

图4A, 4B, 4C和4D.

**4A** 在酵母中, Fas C-端3个氨基酸与FAP-1的相互作用。

**4B** 在体外, Fas C-端3个氨基酸与FAP-1的相互作用。

**4C** 天然的Fas和GST-FAP-1的免疫沉淀反应。

**4D** Fas/FAP-1与Ac-SLV或Ac-SLY结合的抑制作用。

图5A, 5B, 5C, 5D, 5E和5F. 通过微注射将Ac-SLV导入DLD-1细胞系中。三角形表示是可被微注射Ac-SLV的细胞系并表现出可以鉴定的染色质凝集。另一方面, 当微注射Ac-SLY后, 这个视野内仅有一个细胞出现凋亡。

**5A** 相差显微镜显示在500ng/ml CH11存在下微注射Ac-SLY的细胞的范例。

**5B** 相差显微镜显示在500ng/ml CH11存在下微注射Ac-SLV的细胞的范例。

**5C** FITC染色显示在500ng/ml CH11存在下微注射Ac-SLY的细胞的范例。

**5D** FITC染色显示在500ng/ml CH11存在下微注射Ac-SLV的细胞的

范例。

5E Hoechst 33342 DNA荧光染色显示在500ng/ml CH11存在下微注射Ac-SLV的细胞的范例。

5F Hoechst 33342 DNA荧光染色显示在500ng/ml CH11存在下微注射Ac-SLY的细胞的范例。

图6 在微注射DLD-1的细胞中，细胞凋亡的定量测定。

图7A, 7B, 7C, 7D, 7E, 7F, 7G, 和7H.

7A 人类神经生长因子受体的氨基酸序列（序列I, D, 24号）。

7B 人类CD4受体的氨基酸序列（序列I, D, 25号）。

7C Fas-结合的磷酸蛋白酶-1和神经生长因子受体C-端（NGFR）（p75）的相互作用。

7D 人类结肠直肠突变癌蛋白的氨基酸序列（SEQ ID NO:26号）。

7E  $\alpha$ 型蛋白激酶C的氨基酸序列。

7F 5-羟色胺2A受体的氨基酸序列（SEQ ID NO:27号）。

7G 5-羟色胺2B受体的氨基酸序列（SEQ ID NO: 28号）。

7H 腺上皮结肠息肉瘤蛋白的氨基酸序列（I. D. 29号）。

图8 p75 NGFR（低-亲和性神经生长因子受体）的结构特征。

图9 Fas C-末端和p75 NGFR的比较。

图10  $^{35}\text{S}$ -标记的FAP-1与各种以GST融合蛋白形式表达的受体在体外的相互作用。该GST融合蛋白是固定在谷胱苷肽-琼脂糖凝胶珠上，与体外转录的， $^{35}\text{S}$ -标记的FAP-1蛋白一起孵育。洗过珠以后，通过SDS-PAGE和放射自显影分析保留的FAP-1蛋白。

图11A和11B.  $^{35}\text{S}$ -标记的FAP-1与GST- p75缺失突变在体外的相互作用。

11A表明含有p75和p75缺失突变的胞浆区的GST融合蛋白图解。右侧所描绘的是FAP-1与具有各种p75缺失突变的GST融合蛋白的结合，根据（11B）的数据。

11B 体外转录的 $^{35}\text{S}$ -标记FAP-1蛋白与固定在谷胱苷肽-琼脂糖凝胶珠上的各种GST融合蛋白的相互作用。洗过珠以后，用SDS-PAGE和放射自显影方法分析保留FAP-1蛋白。

图12 p75 NGFR的 LexA-C-端胞浆区和VP16-FAP-1之间的结合。酵母株是通过转化构建的,并检测克隆的生长。

+/-表示在His<sup>-</sup>板上克隆生长的情况

### 发明的详细说明

在本发明中,所用的氨基酸残基缩写如下: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp;和Y, Tyr.

为了便于了解以下的实验材料,常规的方法和/或专业术语,请参阅Sambrook, et al. 1989。

本发明提供了一种肽化合物,它能够抑制一种信号-传导蛋白和含有(G/S/A/E)-L-G-(F/I/L)氨基酸序列的胞浆蛋白间的特异结合,其中,每个“-”代表一个肽键,每个括号内包括可相互替换的氨基酸,括号内每个斜线分开的氨基酸是可以替换的。此外,该胞浆蛋白可以含(K/R/Q)-X<sub>n</sub>-(G/S/A/E)-L-G-(F/I/L)氨基酸序列,其中,“X”代表任何一种氨基酸,该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的,“n”代表大于等于2小于等于4。特别是,在一优选实施方案中,该胞浆蛋白含SLGI的氨基酸序列。

(K/R/Q)-X<sub>n</sub>-(G/S/A/E)-L-G-(F/I/L)氨基酸序列在现有技术中也被称作“GLGF (PDZ/DHR)氨基酸区域”。如本申请中所用的“GLGF (PDZ/DHR)氨基酸区域”意思是指(K/R/Q)-X<sub>n</sub>-(G/S/A/E)-L-G-(F/I/L)氨基酸序列。

在一优选实施方案中,该信号-传导蛋白在羧基端有(S/T)-X-(V/I/L)氨基酸序列,其中,每个“-”代表一个肽键,每个括号内包括可相互替换的氨基酸,括号里的每一个斜线分开的氨基酸是可以相互替换的。其中,“X”代表任何一种氨基酸,该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的。

本发明的组合物可以是,不仅限于,抗体,无机化合物,有机化合物,肽,模拟肽化合物,多肽或蛋白,共有一些或全部特性的片段或衍

生物，如，融合蛋白。该组合物可以是天然存在的并通过纯化得到，或非天然存在通过合成得到的。

特别是，该肽化合物可以是含有 (S/T) -X- (V/I/L) -COOH 序列的肽，其中，每个“-”代表一个肽键，每个括号内包括可相互替换的氨基酸，括号里的每一个斜线分开的氨基酸是可以替换的。“X”代表任何一种氨基酸，该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的。在一优选实施方案中，该肽包括下列序列之一：DSENSNFRNEIQSLV, RNEIQSLV, NEIQSLV, EIQLV, IQSLV, QSLV, SLV, IPPDSEdGNEEQSLV, DSEMYNFRSQLASVV, IDLASEFLFLSNSFL, PPTCSQANSGRISTL, SDSNMNMNELSEV, QNFRTYIVSFV, RETIESTV, RGFISLV, TIQSVI, ESLV。更优选的实施方案是一种有机化合物，它含有Ac-SLV-COOH序列，其中，Ac代表一个乙酰基，每个“-”代表一个肽键。

下文中提供了本发明的一个例子。乙酰化肽可以在改进的ChemTech ACT357仪器上，按以前发表的程序以类推的方法自动合成。每个循环用Wang树脂，所有的氨基酸用N<sup>α</sup>-Fmoc保护，然后是20%六氢吡啶/DMF，偶联是用DIC/HOBt，随后用HBTU/DIEA完成。最后的氨基酸偶联上后，在树脂上延伸的肽用Ac<sub>2</sub>O/DMF乙酰化。用HPLC纯化乙酰化的肽，通过FAB-MS和<sup>1</sup>H-NMR鉴定其特征。

而且，本领域普通技术人员通晓上述与非-乙酰基，例如，氨基，偶联的合成肽的构建方法。

本发明也提供了一种肽化合物，它能抑制羧基端带有 (S/T) -X- (V/I/L) 氨基酸序列的信号-传导蛋白和一种胞浆蛋白间特异结合，其中，每个“-”代表一个肽键，每个括号内包括可相互替换的氨基酸，括号内的每个斜线将可相互替代的氨基酸分开，“X”代表任何一种氨基酸，该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的。

本发明的组合物包括抗体，无机化合物，有机化合物，肽，模拟肽化合物，多肽或蛋白，共有一些或全部特性的片段或衍生物，例如，融合蛋白。

本发明也提供一种鉴定一种化合物的方法，该化合物能抑制一种信

号-传导蛋白和含有 (G/S/A/E)-L-G-(F/I/L) 氨基酸序列的胞浆蛋白间的特异结合, 其中, 每个“-”代表一个肽键, 每个括号内包括可相互替换的氨基酸, 括号里的每个斜线分开的氨基酸是可以替换的。该方法包括: (a) 在允许一种已知的化合物, 该化合物曾经显示能替换结合了胞浆蛋白的信号-传导蛋白, 和被结合胞浆蛋白形成复合物的条件下, 将已结合了信号-传导蛋白的胞浆蛋白与各种化合物接触; 和 (b) 检测步骤 (a) 中被替换的信号-传导蛋白或形成的复合物, 其中, 替换作用表示这种化合物能抑制信号-传导蛋白和胞浆蛋白之间的特异结合。

抑制信号-传导蛋白和胞浆蛋白之间的特异结合可以影响报告基因的转录活性。

此外, 在步骤 (b) 中, 通过比较在步骤 (a) 中连接化合物前后报告基因的转录活性来鉴定替换的胞浆蛋白或复合物。这种活性的改变表明信号-传导蛋白和胞浆蛋白间特异结合受抑制以及信号-传导蛋白被替换。

正如本文中所用的“报告基因的转录活性”表示当信号-传导蛋白和胞浆蛋白结合时, 报告基因的表达可在观察到的水平上改变。人们也可以通过检测基于信号-传导蛋白和胞浆蛋白之间结合的其他生物功能来鉴定这种化合物。报告基因的例子很多而且技术上是众所周知, 包括但不限于, 组氨酸抗性基因, 氨苄青霉素抗性基因,  $\beta$ -半乳糖苷酶基因。

另外, 胞浆蛋白是可以结合在固相支持物上。这种化合物也可以结合在固相支持物上并包括一种抗体, 一种无机化合物, 一种有机化合物, 一种肽, 一种模拟肽化合物, 一种多肽或蛋白。

下面提供一个方法的例子。人们可以用直接鉴定的方法去确认一种化合物, 该化合物能抑制信号-传导蛋白和胞浆蛋白之间的特异结合, 如该胞浆蛋白与结合了一种可检测标记的化合物之间的免疫-沉淀。此外, 可以使用非直接检测方法, 该方法检测基因表达水平的提高或减少。正如下面所讨论的, 人们能构建融合到LexA DNA结合区的合成肽。这些构建物被转化到L40-细胞株的带有报告基因的适当细胞系中。然后

可以通过检测报告基因的表达水平检测抑制是否发生。为了检测报告基因的表达水平，技术上熟练的人能应用各种熟知的方法，例如，用酵母、哺乳类或其他细胞中的双杂交系统。

另外，步骤(a)中的接触可以在体外、体内，具体地说是在适当的细胞中，如，酵母细胞或哺乳动物细胞中进行。哺乳动物细胞包括，不仅限于，小鼠成纤维细胞NIH3T3，CHO细胞，HeLa细胞，Ltk<sup>-</sup>细胞，Cos细胞等。

其他合适的细胞包括，但不仅限于，原核或真核细胞，如，细菌细胞（包括革兰氏阳性细胞），真菌细胞，昆虫细胞，或其他动物细胞。

此外，信号-传导蛋白可以是一个细胞表面受体，信号传导蛋白，或肿瘤抑制蛋白。特别是，该细胞表面蛋白是Fas受体，可以在下列细胞中表达，该细胞来源于下列器官，但不仅限于，胸腺，肝脏，肾脏，结肠，卵巢，乳腺，睾丸，脾，肺，胃，前列腺，子宫，皮肤，头部和颈部，或在包含T-细胞和B-细胞的细胞中表达。在一优选实施方案中，T-细胞是Jurkat T-细胞。

另外，细胞表面受体可以是CD4受体，p75受体，5-羟色胺2A受体，或5-羟色胺2B受体。

此外，该信号传导蛋白可以是 $\alpha$ -型蛋白激酶C。

另外，肿瘤抑制蛋白可以是腺上皮结肠息肉抑制蛋白或结肠直肠癌突变癌蛋白。

此外，该胞浆蛋白包含SLGI氨基酸序列，特别是，Fas-结合的磷酸酶-1。

本发明也提供一种鉴定一种化合物的方法，该化合物能抑制羧基端带有(S/T)-X-(V/I/L)氨基酸序列的信号-传导蛋白和一种胞浆蛋白之间的特异结合，其中，每个“-”代表一个肽键，每个括号内包括可相互替换的氨基酸，括号里的每个斜线分开的氨基酸是可以替换的。“X”代表任何一种氨基酸，该氨基酸是从一组包含20种天然存在的氨基酸中选出来的。该方法包括(a)在允许一种已知的化合物，该化合物曾经显示能替换结合了信号-传导蛋白的胞浆蛋白和被结合的信号-传导蛋白形成复合物的条件下，将已结合了胞浆蛋白的信号-传导蛋白与各

种化合物接触；和（b）鉴定步骤（a）中被替换的胞浆蛋白或形成的复合物。其中，替换表示这种化合物能抑制信号-传导蛋白和胞浆蛋白之间的特异结合。

抑制信号-传导蛋白和胞浆蛋白之间的特异结合可以影响报告基因的转录活性。此外，在步骤（b）中，通过比较在步骤（a）中接触化合物前后报告基因的转录活性来鉴定被替换的信号-传导蛋白或复合物，其中，活性的改变表明信号-传导蛋白和胞浆蛋白之间的特异结合受抑制和胞浆蛋白被替换。

另外，在步骤（b）中，通过比较在步骤（a）中接触化合物前后报告基因的转录活性来鉴定被替换的胞浆蛋白或复合物，其中，活性的改变表示信号-传导蛋白和胞浆蛋白间特异结合受到抑制和信号-传导蛋白被替换。

正如本文中所用的“报告基因的转录活性”表示当信号-传导蛋白和胞浆蛋白结合时，报告基因的表达可在观察到的水平上改变。人们也可以通过检测基于信号-传导蛋白和胞浆蛋白的结合的其他生物功能来鉴定这种化合物。报告基因的例子很多而且技术上是众所周知，包括，但不仅限于，组氨酸抗性基因，氨苄青霉素抗性基因， $\beta$ -半乳糖苷酶基因。

另外，胞浆蛋白是可以结合在固相支持物上，或这种化合物也可以结合在固相支持物上，包括一种抗体，一种无机化合物，一种有机化合物，一种肽，一种模拟肽化合物，一种多肽或蛋白。

下面提供一个方法的例子。人们可以用直接鉴定的方法去确认一种化合物，该化合物能抑制信号-传导蛋白和胞浆蛋白之间的特异结合，如该胞浆蛋白和结合了一种可检测标记的该化合物的免疫-沉淀反应。此外，人们也可用非直接鉴定方法，该方法检测基因表达水平的增加或减少。正如下面所讨论的，人们能构建融合到LexA DNA结合区的合成肽。这些构建物被转化到L40-株的带有报告基因的适当细胞系。然后，人们通过检测报告基因的表达水平来鉴定是否发生抑制。在本技术中应用各种不同众所周知的方法，例如，用酵母双杂交系统检测报告基因的表达水平。



另外，步骤(a)中的接触可以在体外或体内进行，具体地说是在酵母细胞或哺乳动物细胞中进行。哺乳动物细胞包括，但不限于，小鼠成纤维细胞NIH3T3，CHO细胞，HeLa细胞，Ltk<sup>-</sup>细胞，Cos细胞等。

其他合适的细胞包括，但不限于，原核或真核细胞，如，细菌细胞（包括革兰氏阳性细胞），真菌细胞，昆虫细胞，或其他动物细胞。

此外，信号-传导蛋白可以是一个细胞表面受体，信号传导蛋白，或肿瘤抑制蛋白。特别是，该细胞表面蛋白是Fas受体，可以在下列细胞中表达，该细胞来源于下列器官，但不限于，胸腺，肝脏，肾脏，结肠，卵巢，乳腺，睾丸，脾，肺，胃，前列腺，子宫，皮肤，头部和颈部，或在包含T-细胞和B-细胞的细胞中表达。在一优选实施方案中，T-细胞是Jurkat T-细胞。

另外，细胞表面受体可以是CD4受体，p75受体，5-羟色胺2A受体，或5-羟色胺2B受体。

此外，信号传导蛋白可以是 $\alpha$ -型蛋白激酶C。

另外，肿瘤抑制蛋白可以是腺上皮结肠息肉抑制蛋白或结肠直肠突变癌蛋白。

此外胞浆蛋白包含SLGI氨基酸序列，特别是，Fas-结合的磷酸酶-1。

本发明也提供一种这些细胞的抑制癌细胞增殖的方法，特别是所述癌细胞来源于，但不限于，胸腺，肝脏，肾脏，结肠，卵巢，乳腺，睾丸，脾脏，胃，前列腺，子宫，皮肤，头部和颈部，或这种癌细胞来源于包含T-细胞和B-细胞的细胞。

本发明提供一种包含用上述方法鉴定的化合物的抑制癌细胞增殖的方法，其中，癌细胞来源于以下器官，但不限于此，胸腺，肝脏，肾脏，结肠，卵巢，乳腺，睾丸，脾脏，胃，前列腺，子宫，皮肤，头部和颈部，或癌细胞来源于包含T-细胞和B-细胞的细胞。

本发明也提供一种治疗癌症的方法，包括向患者的癌细胞中导入以上描述的，能有效引起细胞凋亡的组合物。其中，癌细胞来源于下列器官，但不限于此，胸腺，肝脏，肾脏，结肠，卵巢，乳腺，睾丸，脾脏，胃，前列腺，子宫，皮肤，头部和颈部，或癌细胞来源于包含T-

细胞和B-细胞的细胞。

在本文中所用的，“凋亡”为细胞程序死亡。细胞程序死亡的机制和效应不同于细胞溶解。凋亡的一些可观察的效应是：DNA片段化和分解成称之为凋亡小体的小的膜结合片段。

鉴定组合物是否能有效地诱导细胞凋亡的方法在技术上是众所周知的。一种方法是用相差或荧光显微镜判断染色质形态学变化。

本发明也提供一种抑制病毒感染细胞增殖的方法，包含上面描述的组合物或上述方法鉴定出的化合物，其中，所述病毒感染的细胞包括B型肝炎病毒，EB病毒，流感病毒，乳头瘤病毒，腺病毒，人类T-细胞亲淋巴性病毒，1型或HIV感染的细胞。

本发明也提供了一种治疗病毒-感染患者的方法。包含向患者被病毒感染的细胞中导入以上描述的，能有效诱导细胞凋亡的组合物或以上描述的权利要求27中的方法鉴定的化合物。其中，病毒感染细胞包含B型肝炎病毒，EB病毒，流感病毒，乳头瘤病毒，腺病毒，人类T-细胞亲淋巴性病毒，1型或HIV感染的细胞。

鉴定组合物是否有效抑制细胞凋亡的方法在技术上是众所周知的。一个方法是用相差或荧光显微镜来判断染色质的形态学的变化。

本发明提供了一种药物性组合物，包含有效剂量和在一个合适药物性载体的上述的组合物。

本发明也提供一种药物性组合物，包含有效剂量的上述方法鉴定出的化合物和一种药学可接受的载体。

本发明进一步提供一种肽化合物，它能特异结合羧基端含有(S/T)-X-(V/L/I)氨基酸序列的信号-传导蛋白，其中，每个“-”代表一个肽键，每一个括号内包括了可以相互替换的氨基酸，括号里的每一个斜线分开的氨基酸是可以替换的，“X”代表任何一种氨基酸，该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的。这种肽化合物可以包括(G/S/A/E)-L-G-(F/I/L)氨基酸序列，其中，每个“-”代表一个肽键，每个括号内包括可相互替换的氨基酸，括号里的每一个斜线分开的氨基酸是可以替换的。在一优选实施方案中，该肽化合物包括(K/R/Q)-X<sub>n</sub>-(G/S/A/E)-L-G-(F/I/L)氨基酸序列。其中，“X”代表任何一

种氨基酸,该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的。“n”代表大于等于2小于等于4。在另一优选实施方案中,该组合物包括SLGI氨基酸序列。

本发明进一步提供了一种鉴定化合物的方法,该化合物能结合羧基端含(S/T)-X-(V/L/I)氨基酸序列的信号-传导蛋白,其中,每个“-”代表一个肽键,每个括号内包括可相互替换的氨基酸,括号里的每一个斜线分开的氨基酸是可以替换的,“X”代表任何一种氨基酸,该氨基酸是从一组包含20种天然存在的氨基酸中选择出来的。该方法包含(a),在允许与一种已知的化合物,该化合物曾显示能结合信号-传导蛋白并形成一复合物,结合的条件下,将信号-传导蛋白与多种化合物接触;和(b),检测步骤(a)中形成的复合物,以鉴别能与该信号-传导蛋白结合的化合物。特别是,被鉴定的化合物包括(G/S/A/E)-L-G-(F/I/L)氨基酸序列。在又一优选实施方案中,被鉴定的化合物包括SLGI氨基酸序列。

此外,以上描述的方法中,该信号-传导蛋白可以结合在一个固相支持物上。该化合物也可以结合在一个固相支持物上,包含抗体,无机化合物,有机化合物,肽,模拟肽化合物,多肽或蛋白。

另外,该信号-传导蛋白可以是一个细胞-表面受体或信号导体。特别是,该信号-传导蛋白可能是Fas受体,CD4受体,p75受体,5-羟色胺2A受体,5-羟色胺2B受体,或 $\alpha$ -型蛋白激酶-C。

本发明也提供了一种恢复细胞中凋亡的负调节的方法,包含以上描述的组合物或被以上方法鉴定的化合物。

在本文中,“恢复凋亡的负调节”是指使细胞能够阻止细胞程序死亡的进行。

例如,具有功能性Fas受体和Fas-结合的磷酸酶-1的细胞由于通过磷酸酶引起Fas的负调节,而不能进入程序细胞死亡或凋亡。然而,如果Fas-结合的磷酸酶1不能结合到Fas受体的羧基端((S/T)-X-(V/L/I)区),也就是,在(G/S/A/E)-L-G-(F/I/L)氨基酸序列中,至少一个发生突变或缺失,细胞进入凋亡。通过导入能结合Fas受体的羧基端的化合物,人们能模拟一个功能性磷酸酶的效应,这样就恢

复了凋亡的负调节。

本发明提供了一种预防细胞凋亡的方法，包含上述描述的组合物或被以上描述方法鉴定的化合物。

本发明提供了一种控制由相关细胞凋亡引起的疾病，包含以上所描述的组合物或由以上描述的方法鉴定的化合物。

本发明在以下的具体实验部分中有详细说明。这些是为了有助于了解本发明，但并不意味着，也不应当被解释为，限制本发明的权利要求。

## 第一部分实验

### 实验的详细操作

#### 材料和方法

##### 1. 筛选半随机和随机肽基因文库

为了在限制性的DNA序列中产生许多突变，根据文献(Hill等, 1987)所叙述的程序，用简并寡核苷酸引起突变的PCR方法，基于人与大鼠之间的同源性，设计两种回文序列，用于构建半随机的基因文库。所用的两个引物是

5'-CGGAATTCNNNNNNNNAACAGCNNNNNNNNAATGAANNCAAGTCTGNNNTGAGGATCCTCA-3'  
(Seq. I.D. No.:30) 和

5'-CGGAATTCGACTCAGAANNNNNNAACTTCAGANNNNNNATCNNNNNNNNGTCTGAGGATCCTCA-3'  
(Seq. I.D.:31)

简述如下，经过HPLC纯化的这两种引物(每种200pmol)在70℃退火，5分钟，于23℃冷却60分钟，然后在Klenow DNA聚合酶作用下(5u)，于23℃，60分钟，用4种核苷酸混合物补平，(每种核苷酸最终浓度为1mM)。加入1μl 0.5M EDTA终止反应。乙醇沉淀来纯化DNA。将得到的双链DNA用EcoRI和BamHI酶切。然后在非变性聚丙烯酰胺凝胶电泳中进一步纯化然后将双链寡核苷酸与pBTM116质粒在EcoRI- BamHI位点进行连接。将连接混合物电转移到E.Coli XL1-Blue MRF' (Stratagene) 菌中以构建质粒基因文库。按以前报道的方法进行大批量转化。质粒基因文库转化到L40 细胞中(MATa, *trp1*, *leu2*, *his3*, *ade2*, *LYS2*;

(*lexAop*)<sup>+</sup>-*HIS3, URA3* : ( *LexAop*)<sup>+</sup>-*LacZ*). 该细胞株携带有含有 FAP-1 cDNA的PVP16-31质粒(Sato等,1995). 将在组氨酸-缺陷的培养基中(His<sup>+</sup>)形成的克隆转移到含有40 μg/ml x-gal的平板上,以测定在平板滤膜实验所产生的蓝色反应产物(β-gal<sup>+</sup>). 对所筛选的His<sup>+</sup>和β-gal<sup>+</sup>克隆进行检测以用于进一步分析. 有回文结构的寡核苷酸5'-CGGAATTC-(NNH)<sub>4-15</sub>-TGAGGATCCTCA-3' (seq. I. D. No:32)用于构建随机的肽基因文库.

## 2. 肽的合成

肽的合成是在新型Chemtech ACT357自动合成仪按照以前发表的类似的操作程序自动合成(Schnorrenberg和Gerhardt,1989). 每次循环用Wang树脂(用量为0.2-0.3mmole),所有氨基酸用N<sup>α</sup>-Fmoc进行保护. 再用20%六氢吡啶/DMF处理去除保护,用DIC/HOBt然后用HBTU/DIEA进行偶联. 当最后一个氨基酸被偶联上之后,用Ac<sub>2</sub>O/DMF对树脂上延长的肽进行乙酰化. 最后用TFA处理共同去除所有的保护基团将该肽从树脂上分离下来. 乙酰化的肽经过HPLC纯化后,通过FAB-MS和<sup>1</sup>H-NMR鉴定其特征.

## 3. 利用Fas C端15个氨基酸对Fas/FAP-1结合进行抑制作用的检测

亚克隆到Bluescript pSK-II载体上的HFAP-10 cDNA的体外翻译是利用体外转录/翻译系统(Promega TNT lysate)和T7RNA聚合酶在<sup>35</sup>S-L-甲硫氨酸存在下,从内部的甲硫氨酸密码子开始翻译,得到的<sup>35</sup>S-标记蛋白与GST-Fas融合蛋白于4℃共同孵育16小时,该GST-Fas融合蛋白在含150mM NaCl, 50mM Tris (pH 8.0), 5 mM DTT, 2mM EDTA, 0.1%NP-40, 1mM PMSF, 50 μg/ml亮抑素, 1mM 联苯胺和7 μg/ml抑肽素的缓冲液中,固定在GST-Sepharose 4B亲和层析珠上(Pharmacia),用同样的缓冲液彻底洗4次,通过离心回收带有谷胱甘肽-Sepharose珠的结合蛋白,在煮沸的Laemmili缓冲液中洗脱下来,然后通过SDS-PAGE和荧光学的方法进行分析.

#### 4. Fas末端15个氨基酸的抑制作用的检测和利用各种三肽对Fas/FAP-1结合的抑制影响

体外翻译的 [<sup>35</sup>S] HFAP-1经过NAP-5柱的纯化后 (pharmacia), 与 3 μ M的GST-融合蛋白在4℃ 孵育16小时, 用结合缓冲液洗4次, 在b计数器测定参入的放射性活性。结合抑制的百分率按下式计算: 抑制百分数 %=[使用GST-Fas (19-335)和肽参入的放射活性-使用GST-Fas (191-320)和肽参入的放射活性]/[使用GST-Fas (191-335)没有肽时参入的放射活性-使用GST-Fas (191-320)没有肽时参入的放射活性]。 n=3 .

#### 5. 结合FAP-1的Fas C端3个氨基酸在体外和在酵母中的相互作用。

构建诱饵质粒pBTM116 (LexA)-SLV, -PLV, -SLY/, 和SLA并转化到带有pVP16-FAP-1或ras的L40菌株中, 从每种转化体中挑出6个单独的克隆, 在组氨酸-缺陷的培养基中分析其生长情况。用GST-Sepharose4B亲和层析珠 (Pharmacia) 纯化GST-Fas, -SLV和PLV。体外结合的方法如上所述。

#### 6. 天然的Fas与 GST-FAP-1的免疫沉淀反应和Fas/FAP-1与Ac-SLV结合的抑制作用

有FAP-和没有FAP-1的GST-融合蛋白与表达Fas的Jurkat T细胞的细胞提取物共同孵育。利用抗-Fas单克隆抗体进行Western杂交分析以检测结合的Fas (F22120, Transduction Lab)。AC-SLV和AC-SLY三肽用于Fas/FAP-1结合抑制作用的检测。

7. 显微注射Ac-SLV到DLD-1细胞株。将DLD-1人肠癌细胞在含10%胎牛血清的RPMI1640培养基中培养。微注射时, 在35mm的塑料培养皿上, 以 $1 \times 10^5$ 细胞/2ml的密度, 将细胞接种在CELLoate (Eppendorf)上, 生长一天。恰好在进行微注射之前, 加入500 ng/ml的Fas单克隆抗体CH11 (MBLinternational)。所有微注射实验均用自动微注射系统操作 (Eppendorf transjector 5246 micro-manipalator5171 和

Femtotips) (pantel, 等, 1995). 合成的三肽以100mM的浓度悬浮在0.1% (W/V) FITC-硫酸葡聚糖 (Sigma) /K-PBS中. 将样品微注射到DLD-1细胞浆内. 注射16-20小时后, 用PBS洗细胞, 然后用PBS配制的10  $\mu$ g/ml Hoechst33342染细胞. 37 $^{\circ}$ C解育30分钟后, 对这些细胞进行照相, 计数染色质浓缩的细胞作为凋亡的细胞

## 8 微注射的DLD-1细胞中凋亡的定量分析

每次实验, 微注射 25-100个细胞. 微注射后细胞的凋亡的检测是利用相差以及荧光显微镜对形态学改变进行评价 (Wang等, 1995) (McGahon, 等1995). 数据分析是统计2-3次单独测定的平均值 $\pm$ 标准误.

## 讨论

为了鉴定FAP-1结合所必须的Fas受体C端区域的最小肽, 我们利用一系列合成的肽和酵母双杂交系统肽的基因文库 (图24), 在体外检测Fas/FAP-1结合的抑制作用. 首先构建了15种氨基酸融合到LexA DNA结合区域的半随机基因文库 (基于人与大鼠之间Fas的同源性) (图2B和2C), 然后与起初作为FAP-1分离的pVP16-31 (Sato, 等, 1995) 共同转化到L40酵母菌中. 从 $5.0 \times 10^6$ 转化体中 (Johnson等1986) 筛选出200 His<sup>+</sup>克隆后, 挑出100个 $\beta$ -半乳糖苷酶阳性的克隆用于进一步分析. 编码C端15个氨基酸的文库质粒的序列分析揭示了所有的C端氨基酸或是缬氨酸, 亮氨酸或是异亮氨酸残基. 第二, 构建了4-15个氨基酸与LexA DNA结合区域相融合的随机基因文库, 然后按图2D的策略进行筛选. 令人惊奇的是所有C端的第三个氨基酸都是丝氨酸, 而且C端氨基酸分析结果与半随机cDNA文库筛选的结果是一致的. 在筛选这些基因文库过程中没有发现其他的有意义的氨基酸序列, 提示最后3个氨基酸 (ts-X-V/L/I) 基序对于与FAP-1的第三PDZ区域的结合非常重要, 并且在蛋白-蛋白之间相互作用以及调节由Fas诱导的凋亡过程中起至关重要的作用. 为了进一步确定是否这最后的3个氨基酸对于Fas/FAP-1的结合是必需的, 而且足以满足其结合, 我们

构建了LexA-SLV, -PLV, -PLY, -SLY和-SLA融合蛋白的质粒, 然后与PVP16-FAP-1共转化到酵母中。结果表明只有LexA-SLV与FAP-1结合, 而LexA-PLV, -PLY, -SLY和-SLA不能结合(图4A)。用各种GST-三肽融合体研究体外结合和在体外翻译的FAP-1所得的结果是一致的(图4B)。

除了酵母双杂交系统外, 也利用了Fas/FAP-1结合的体外抑制作用检测。首先, 检测C-端15个氨基酸的合成肽是否能在体外抑制Fas和FAP-1的结合(图3A)。体外翻译的FAP-1与GST-Fas的结合也显著地降低而且与合成的FAS的15个氨基酸的浓度有关。与这些结果相比, 在同样的生物化学条件下, 人的PAMP肽(Kitamuta, 等 1994)作为阴性对照对Fas/FAP-1的结合活性没有影响。第二, 检测截短的Fas C-末端合成肽对Fas/FAP-1体外结合的效应。正如图3B所示, 对于Fas/FAP-1的结合, 在C-端只有3个氨基酸(Ac-SLV)就足以得到与合成的4-15肽所起的抑制影响。进一步利用三肽的替换肽, 研究Fas/FAP-1的结合, 以决定对抑制作用所必须的重要的氨基酸(图3C)。这些结果揭示了C端的第三个氨基酸和C-末端氨基酸分别为丝氨酸或苏氨酸; 以及缬氨酸, 亮氨酸, 或异亮氨酸时具有最强的抑制效应。但是C端第二个氨基酸中对于它们的Fas/FAP-1结合抑制效应没有显著差异。这些结果与酵母双杂交系统所得结果(图2C, 2D)是一致的。因此, 结论是C端的三个氨基酸(SLV)对于Fas与FAP-1蛋白的第三个PDZ区域结合是决定性的。

为了进一步研究在更天然条件下PDZ区域与ts/T-X-V/L./I之间相互作用的重要意义, 我们检测了GST融合的FAP-1蛋白与Jurkat T细胞表达的Fas相互作用的能力。结果揭示了是Ac-SLV三肽, 而不是Ac-SLY三肽取消了FAP-1与从Jurkat T细胞中提取的Fas蛋白以剂量依赖的方式相结合的活性(图4C和4D)。这些结果揭示了C-端氨基酸tSLV对于FAP-1是最小的结合位点。而且丝氨酸和缬氨酸对于这种物理性的结合是至关重要的。

为了进一步验证这一推论, 即Fas C端的3个氨基酸与FAP-



1 第三个PDZ区域之间的物理结合对于体内FAP-1作为由Fas介导的信号传导的负调控因子的功能是必需的,应用微注射的实验,将合成的三肽注射到大肠癌细胞株DLD-1,该细胞株能表达Fas和FAP-1,并抵抗由Fas诱导的凋亡。这些实验包括直接将合成的三肽注射到单个细胞的胞浆,检测体内由Fas-诱导的凋亡的生理性反映。结果显示,将Ac-SLV微注射到DLD-1细胞时有Fas-单克隆抗体存在的情况下(CH11, 500 ng/ml),可以显著地诱导凋亡(图5A, 5E, 和图6),但是微注射Ac-SLY和PBS/K则不能诱导凋亡(图5B, 5F和图6)。这些结果有力地支持了这种推论,即FAP-1与FasC端物理性结合对于防止细胞被Fas-诱导引起凋亡是必需的。

总之,我们发现了Fas C端的SLV对于结合到FAP-1第三个PDZ区域是唯一必须的,并能满足其需要。第二,提出了新的tS/T-X-V/L/I的共用基序,而不是tS/T-X-V,被用于与PDZ区域的结合。因此很可能除了FAP-1与Fas物理性相互作用外,FAP-1在信号传导途径的调节作用中也起很重要的作用。第三,证明了在大肠癌细胞中直接注射三肽Ac-SLV可以定向地诱导由Fas-介导的凋亡。进一步的研究包括确定FAP-1的底物和结构功能的分析,在癌细胞中利用Fas/FAP-1的相互作用用于治疗以及提供证据更好地了解FAP-1对由Fas介导的信号传导的抑制作用。

## 第二部分实验

FAP-1原是鉴定为膜结合蛋白酪氨酸磷酸酶,它与Fas的C端相结合,有6个PDZ区域(也被称作DHR区域或GLGF重复)。最近研究表明PDZ区域是作为一种新的特异的蛋白-蛋白之间相互作用模件,而且看来在膜蛋白的装配,和在多蛋白的复合物中与信号分子连接的过程中起重要作用。

最近的综合研究结果发现FAP-1的第三个PDZ区域特异地识别一个序列基序t(S/T)-X-V,并与Fas的C-端3个氨基酸相互作用(图9)。为了研究FAP-1也与p75NGFR C端区域相互作用的可能性(图8),利用体外结合实验以及通过用酵母的双杂交系统将一系列的p75NGFR缺失突

变后进行分析。结果揭示p75NGFR C-端胞浆部分与FAP-1相互作用，该p75NGFR C-端胞浆部分在所有的种类中是高度保守的（图10）。而且，p75NGFR C-端的三个氨基酸SPV是必须的，且足以满足与FAP-1的第三个PDZ区域都相互作用（图11A, 11B）。由于FAP-1在胎儿脑的表达最高，所以这些结果暗示FAP-1与p75NGFR之间的相互作用在神经细胞中通过p75NGFR的途径进行信号传递的过程中以及在p75NGFR起始信号转变复合物的形成中起重要作用。

## 参考文献

1. Banville, D., et al. J. Biol. Chem. 269: 22320-22327 (1994).  
5
2. Boldin, M. P. et al. J. Biol. Chem. 270: 7795-7798 (1995).
3. Camerini, D., et al. J. Immunol. 147: 3165-3169 (1991).  
10
4. Chao, M.V. and B.L. Hempstead TINS 18: 321-326 (1995).
5. Chinnaiyan, A. M., et al. Cell 81: 505-512 (1995).  
15
6. Cho, K.-O., et al. Neuron 9: 929-942 (1992).
7. Conboy, J. G., et al. J. Biol. Chem. 266: 8273-8280 (1991).  
20
8. Doyle, D.A., et al. Cell 85: 1067-1076 (1996).
9. Funayama, N., et al. J. Cell Biol. 115: 1039-1048 (1991).  
25
10. Gould, K. L., et al. EMBO J. 8: 4133-4142 (1989).
11. Gu, M. X., et al. Proc. Natl. Acad. Sci. U.S.A. 88: 5867-5871 (1991).  
30
12. Hill, D. E., et al. Meth. Enzymol. 155, 558-568 (1987).
13. Ito, N., and Nagata, S. J. Biol. Chem. 268: 10932-10937 (1993).  
35

- 
14. Itoh, N. et al. Cell 66: 233-243 (1991).
15. Johnson, D. et al. Cell 47: 545-554 (1986).
- 5 16. Kim, E., et al. Nature 378: 85-88 (1995).
17. Kischkel, F. C. et al. EMBO J. 14: 5579-5588 (1995).
18. Kitamura, K. et al. FEBS Lett. 351: 35-37 (1994).
- 10 19. Kornau, H.-C., et al. Science 269:1737-1740 (1995).
20. Lankes, W. T., and Furthmayr, H. Proc. Natl. Acad. Sci. U.S.A. 88: 8297-8301 (1991).
- 15 21. Maekawa, K., et al. FEBS Letters 337: 200-206 (1994).
22. Mallett, S., et al. EMBO J. 9: 1063-1068 (1990).
- 20 23. Matsumine, A. et al. Science 272: 1020-1023 (1996).
24. McGahon, A. J. et al. Meth. Cell Biol. 46: 153-185 (1995).
- 25 25. Pantel, K. et al. J. Natl. Cancer Inst. 87: 1162-1168 (1995).
26. Rouleau, G. et al. Nature 363: 515-521 (1993).
- 30 27.
28. Sambrook, J., et al. (1989) Molecular Cloning: a laboratory manual. Second Edition. Cold Spring Harbor Laboratory Press.
- 35 29. Sato, T., et al. Science 268: 411-415 (1995).
30. Schnorrenberg, G. and Gerhardt H. Tetrahedron 45:

- 7759-7764 (1989).
31. Saras, J., et al. J. Biol. Chem. 269, 24082-24089 (1994).
- 5
32. Smith, C. A. et al. Cell 73: 1349-1360 (1993).
33. Stamenkovic, I., et al. EMBO J. 8: 1403-1410 (1989).
- 10
34. Stanger, B. Z., et al. Cell 81: 513-523 (1995).
35. Takahashi, T. et al. Cell 76: 969-976 (1994).
- 15
36. Vogel, W., et al. (1993). Science 259: 1611-1614 (1993).
37. Watanabe-Fukunaga, R., et al. Nature 356: 314-317 (1992).
- 20
38. Wang, X. W., et al. Cancer Res. 55: 6012-6016 (1995).
39. Westendorp, M. O. et al. Nature 375: 497-500 (1995).
- 25
40. Woods, D.F. and Bryant, P.J. Cell 66: 451-464 (1991).
- 30
41. Yang, Q., and Tonks, N. K. Proc. Natl. Acad. Sci. U.S.A. 88: 5949-5953 (1991).

## 序列表

- (1) 总信息:
- 5 (i) 申请人: 纽约市哥伦比亚大学信托人
- 10 (ii) 发明名称: 一种抑制信号-传导蛋白和  
GLGF (PDZ/DHR) 区域之间相  
互作用的化合物及其应用
- (iii) 序列数目: 33
- (iv) 通信地址:
- 15 (A) 收信人: Cooper & Dunham LLP  
(B) 街道: 1185 Avenue of the Americas  
(C) 城市: New York  
(D) 州: New York  
(E) 国家: U. S. A.  
20 (F) 邮编: 10036
- (v) 计算机可读形式:
- 25 (A) 介质类型: 软盘  
(B) 计算机: IBM PC 兼容  
(C) 操作系统: PC-DOS/MS-DOS  
(D) 软件: PatentIn Release #1. 0, Version #1. 30
- (vi) 目前申请状态:
- 30 (A) 申请号:  
(B) 申请日: 1997年 7月18日  
(C) 分类:
- (viii) 代理信息:
- 35 (A) 姓名: White, John P  
(B) 登记号: 28, 678  
(C) 案号: 0575/48962-A-PCT/JPW/JKM
- (ix) 通讯资料:
- 40 (A) 电话: (212) 278-0400  
(B) 传真: (212) 391-0525
- (2) SEQ ID NO: 1的信息:
- 45 (i) 序列特征:  
(A) 长度: 4 氨基酸  
(B) 类型: 氨基酸  
(C) 链型: 单链  
(D) 拓扑结构: 线性
- 50 (ii) 分子类型: 肽
- (iii) 假设: 否
- 55 (iv) 反义: 否
- (xi) 序列描述: SEQ ID NO: 1:
- 60 Gly/Ser/Ala/Glu Leu Gly Phe/Ile/Leu  
1
- (2) SEQ ID NO: 2的信息:
- 65 (i) 序列特征:  
(A) 长度: 6 氨基酸  
(B) 类型: 氨基酸

- (C) 链型: 单链  
(D) 拓扑结构: 线性
- 5 (ii) 分子类型: 肽  
(iii) 假设: 否  
(iv) 反义: 否
- 10 (xi) 序列描述: SEQ ID NO: 2:  
Lys/Arg/Gln Xaa(n) Gly/Ser/Ala/Glu Leu Gly Phe/Ile/Leu  
1 5
- 15 (2) SEQ ID NO: 3的信息:  
(i) 序列特征:  
20 (A) 长度: 4 氨基酸  
(B) 类型: 氨基酸  
(C) 链型: 单链  
(D) 拓扑结构: 线性  
(ii) 分子类型: 肽  
25 (iii) 假设: 否  
(iv) 反义: 否  
(xi) 序列描述: SEQ ID NO: 3:  
30 Ser Leu Gly Ile  
1
- 35 (2) SEQ ID NO: 4的信息:  
(i) 序列特征:  
40 (A) 长度: 6 氨基酸  
(B) 类型: 氨基酸  
(C) 链型: 单链  
(D) 拓扑结构: 线性  
(ii) 分子类型: 肽  
45 (iii) 假设: 否  
(iv) 反义: 否  
(xi) 序列描述: SEQ ID NO: 4:  
50 Ser/Thr Xaa Val/Ile/Leu  
1
- 55 (2) SEQ ID NO: 5的信息:  
(i) 序列特征:  
60 (A) 长度: 15 氨基酸  
(B) 类型: 氨基酸  
(C) 链型: 单链  
(D) 拓扑结构: 线性  
(ii) 分子类型: 肽  
65 (xi) 序列描述: SEQ ID NO: 5:

Asp Ser Glu Asn Ser Asn Phe Arg Asn Glu Ile Gln Ser Leu Val .  
 1 5 10 15

5 (2) SEQ ID NO: 6的信息:

- (i) 序列特征:  
 (A) 长度: 15 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 6:

Ser Ile Ser Asn Ser Arg Asn Glu Asn Glu Gly Gln Ser Leu Glu  
 1 5 10 15

20 (2) SEQ ID NO: 7的信息:

- (i) 序列特征:  
 (A) 长度: 15 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 7:

Ser Thr Pro Asp Thr Gly Asn Glu Asn Glu Gly Gln Cys Leu Glu  
 1 5 10 15

35 (2) SEQ ID NO: 8的信息:

- (i) 序列特征:  
 (A) 长度: 4 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 8:

Glu Ser Leu Val  
 1

50 (2) SEQ ID NO: 9的信息:

- (i) 序列特征:  
 (A) 长度: 6 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 9:

Thr Ile Gln Ser Val Ile  
 1 5



## (2) SEQ ID NO: 10的信息:

- (i) 序列特征:  
 (A) 长度: 8 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 10:

Arg Gly Phe Ile Ser Ser Leu Val  
 1 5

## (2) SEQ ID NO: 11的信息:

- (i) 序列特征:  
 (A) 长度: 8 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ I

Arg Glu Thr Ile Glu Ser Thr Val  
 1 5

## (2) SEQ ID NO: 12的信息:

- (i) 序列特征:  
 (A) 长度: 11 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 12:

Gln Asn Phe Arg Thr Tyr Ile Val Ser Phe Val  
 1 5 10

## (2) SEQ ID NO: 13的信息:

- (i) 序列特征:  
 (A) 长度: 13 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 13:

Ser Asp Ser Asn Met Asn Met Asn Glu Leu Ser Glu Val  
 1 5 10

## (2) SEQ ID NO: 14的信息:

- (i) 序列特征:

- (A) 长度: 15 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性
- 5 (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 14:
- 10 Pro Pro Thr Cys Ser Gln Ala Asn Ser Gly Arg Ile Ser Thr Leu  
 1 5 10 15
- (2) SEQ ID NO: 15的信息:
- 15 (i) 序列特征:  
 (A) 长度: 15 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性
- 20 (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 15:
- 25 Ile Asp Leu Ala Ser Glu Phe Leu Phe Leu Ser Asn Ser Phe Leu  
 1 5 10 15
- (2) SEQ ID NO: 16的信息:
- (i) 序列特征:  
 (A) 长度: 15 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性
- 35 (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 16:
- 40 Asp Ser Glu Met Tyr Asn Phe Arg Ser Gln Leu Ala Ser Val Val  
 1 5 10 15
- (2) SEQ ID NO: 17的信息:
- (i) 序列特征:  
 (A) 长度: 15 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性
- 50 (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 17:
- 55 Ile Pro Pro Asp Ser Glu Asp Gly Asn Glu Glu Gln Ser Leu Val  
 1 5 10 15
- (2) SEQ ID NO: 18的信息:
- (i) 序列特征:  
 (A) 长度: 4 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性
- 65

- (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 18:
- 5           Gln Ser Leu Val  
              1
- (2) SEQ ID NO: 19的信息:
- 10           (i) 序列特征:  
              (A) 长度: 5 氨基酸  
              (B) 类型: 氨基酸  
              (C) 链型: 单链  
              (D) 拓扑结构: 线性
- 15           (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 19:
- 20           Ile Gln Ser Leu Val  
              1                   5
- (2) SEQ ID NO: 20的信息:
- 25           (i) 序列特征:  
              (A) 长度: 6 氨基酸  
              (B) 类型: 氨基酸  
              (C) 链型: 单链  
              (D) 拓扑结构: 线性
- 30           (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 20:
- 35           Glu Ile Gln Ser Leu Val  
              1                   5
- (2) SEQ ID NO: 21的信息:
- 40           (i) 序列特征:  
              (A) 长度: 7 氨基酸  
              (B) 类型: 氨基酸  
              (C) 链型: 单链  
              (D) 拓扑结构: 线性
- 45           (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 21:
- 50           Asn Glu Ile Gln Ser Leu Val  
              1                   5
- (2) SEQ ID NO: 22的信息:
- 55           (i) 序列特征:  
              (A) 长度: 8 氨基酸  
              (B) 类型: 氨基酸  
              (C) 链型: 单链  
              (D) 拓扑结构: 线性
- 60           (ii) 分子类型: 肽
- (xi) 序列描述: SEQ ID NO: 22:
- 65

Arg Asn Glu Ile Gln Ser Leu Val  
1 5

5 (2) SEQ ID NO: 23的信息:

- (i) 序列特征:
  - (A) 长度: 15 氨基酸
  - (B) 类型: 氨基酸
  - (C) 链型: 单链
  - (D) 拓扑结构: 线性

10

(ii) 分子类型: 肽

15

(xi) 序列描述: SEQ ID NO: 23:

Asp Ser Glu Asn Ser Asn Phe Arg Asn Glu Ile Gln Ser Leu Val  
1 5 10 15

20

(2) SEQ ID NO: 24的信息:

- (i) 序列特征:
  - (A) 长度: 427 氨基酸
  - (B) 类型: 氨基酸
  - (C) 链型: 单链
  - (D) 拓扑结构: 线性

25

(ii) 分子类型: 肽

30

(xi) 序列描述: SEQ ID NO: 24:

Met Gly Ala Gly Ala Thr Gly Arg Ala Met Asp Gly Pro Arg Leu Leu  
1 5 10 15

35

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

40

Pro Thr Gly Leu Tyr Thr His Ser Gly Glu Cys Cys Lys Ala Cys Asn  
35 40 45

Leu Gly Glu Gly Val Ala Gln Pro Cys Gly Ala Asn Gln Thr Val Cys  
50 55 60

45

Glu Pro Cys Leu Asp Ser Val Thr Phe Ser Asp Val Val Ser Ala Thr  
65 70 75 80

50

Glu Pro Cys Lys Pro Cys Thr Glu Cys Val Gly Leu Gln Ser Met Ser  
85 90 95

Ala Pro Cys Val Glu Ala Asp Asp Ala Val Cys Arg Cys Ala Tyr Gly  
100 105 110

55

Tyr Tyr Gln Asp Glu Thr Thr Gly Arg Cys Glu Ala Cys Arg Val Cys  
115 120 125

Glu Ala Gly Ser Gly Leu Val Phe Ser Cys Gln Asp Lys Gln Asn Thr  
130 135 140

60

Val Cys Glu Glu Cys Pro Asp Gly Thr Tyr Ser Asp Glu Ala Asn His  
145 150 155 160

Val Asp Pro Cys Leu Pro Cys Thr Val Cys Glu Asp Thr Glu Arg Gln  
165 170 175

65

Leu Arg Glu Cys Thr Arg Trp Ala Asp Ala Glu Cys Glu Glu Ile Pro  
180 185 190

Gly Arg Trp Ile Thr Arg Ser Thr Pro Pro Glu Gly Ser Asp Ser Thr  
 195 200 205  
 5 Ala Pro Ser Thr Gln Glu Pro Glu Ala Pro Pro Glu Gln Asp Leu Ile  
 210 215 220  
 Ala Ser Thr Val Ala Gly Val Val Thr Thr Val Met Gly Ser Ser Gln  
 225 230 235 240  
 10 Pro Val Val Thr Arg Gly Thr Thr Asp Asn Leu Ile Pro Val Tyr Cys  
 245 250 255  
 Ser Ile Leu Ala Ala Val Val Val Gly Leu Val Ala Tyr Ile Ala Phe  
 260 265 270  
 15 Lys Arg Trp Asn Ser Cys Lys Gln Asn Lys Gly Gly Ala Asn Ser Arg  
 275 280 285  
 20 Pro Val Asn Gln Thr Pro Pro Pro Glu Gly Glu Lys Ile His Ser Asp  
 290 295 300  
 Ser Gly Ile Ser Val Asp Ser Gln Ser Leu His Asp Gln Gln Pro His  
 305 310 315 320  
 25 Thr Gln Thr Ala Ser Gly Gln Ala Leu Lys Gly Asp Gly Gly Leu Tyr  
 325 330 335  
 Ser Ser Leu Pro Pro Ala Lys Arg Glu Glu Val Glu Lys Leu Leu Asn  
 340 345 350  
 30 Gly Ser Ala Gly Asp Thr Trp Arg His Leu Ala Gly Glu Leu Gly Tyr  
 355 360 365  
 35 Gln Pro Glu His Ile Asp Ser Phe Thr His Glu Ala Cys Pro Val Arg  
 370 375 380  
 Ala Leu Leu Ala Ser Trp Ala Thr Gln Asp Ser Ala Thr Leu Asp Ala  
 385 390 395 400  
 40 Leu Leu Ala Ala Leu Arg Arg Ile Gln Arg Ala Asp Leu Val Glu Ser  
 405 410 415  
 Leu Cys Ser Glu Ser Thr Ala Thr Ser Pro Val  
 420 425  
 45

## (2) SEQ ID NO: 25的信息:

50 (i) 序列特征:  
 (A) 长度: 458 氨基酸  
 (B) 类型: 氨基酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性  
 55 (ii) 分子类型: 肽  
 (xi) 序列描述: SEQ ID NO: 25:  
 60 Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu  
 1 5 10 15  
 Ala Leu Leu Pro Ala Ala Thr Gln Gly Lys Lys Val Val Leu Gly Lys  
 20 25 30  
 65 Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser  
 35 40 45

	Ile	Gln	Phe	His	Trp	Lys	Asn	Ser	Asn	Gln	Ile	Lys	Ile	Leu	Gly	Asn
	50					55						60				
5	Gln	Gly	Ser	Phe	Leu	Thr	Lys	Gly	Pro	Ser	Lys	Leu	Asn	Asp	Arg	Ala
	65					70					75					80
	Asp	Ser	Arg	Arg	Ser	Leu	Trp	Asp	Gln	Gly	Asn	Phe	Pro	Leu	Ile	Ile
					85					90					95	
10	Lys	Asn	Leu	Lys	Ile	Glu	Asp	Ser	Asp	Thr	Tyr	Ile	Cys	Glu	Val	Glu
				100					105					110		
	Asp	Gln	Lys	Glu	Glu	Val	Gln	Leu	Leu	Val	Phe	Gly	Leu	Thr	Ala	Asn
			115					120					125			
15	Ser	Asp	Thr	His	Leu	Leu	Gln	Gly	Gln	Ser	Leu	Thr	Ile	Thr	Leu	Glu
	130						135					140				
	Ser	Pro	Pro	Gly	Ser	Ser	Pro	Ser	Val	Gln	Cys	Arg	Ser	Pro	Arg	Gly
20	145					150					155					160
	Lys	Asn	Ile	Gln	Gly	Gly	Lys	Thr	Leu	Ser	Val	Ser	Gln	Leu	Glu	Leu
					165					170					175	
25	Gln	Asp	Ser	Gly	Thr	Trp	Thr	Cys	Thr	Val	Leu	Gln	Asn	Gln	Lys	Lys
				180					185					190		
	Val	Glu	Phe	Lys	Ile	Asp	Ile	Val	Val	Leu	Ala	Phe	Gln	Lys	Ala	Ser
			195					200					205			
30	Ser	Ile	Val	Tyr	Lys	Lys	Glu	Gly	Glu	Gln	Val	Glu	Phe	Ser	Phe	Pro
		210					215					220				
	Leu	Ala	Phe	Thr	Val	Glu	Lys	Leu	Thr	Gly	Ser	Gly	Glu	Leu	Trp	Trp
35	225					230					235					240
	Gln	Ala	Glu	Arg	Ala	Ser	Ser	Ser	Lys	Ser	Trp	Ile	Thr	Phe	Asp	Leu
				245						250					255	
40	Lys	Asn	Lys	Glu	Val	Ser	Val	Lys	Arg	Val	Thr	Gln	Asp	Pro	Lys	Leu
				260					265					270		
	Gln	Met	Gly	Lys	Lys	Leu	Pro	Leu	His	Leu	Thr	Leu	Pro	Gln	Ala	Leu
			275					280					285			
45	Pro	Gln	Tyr	Ala	Gly	Ser	Gly	Asn	Leu	Thr	Leu	Ala	Leu	Glu	Ala	Lys
		290					295					300				
	Thr	Gly	Lys	Leu	His	Gln	Glu	Asn	Val	Leu	Val	Val	Met	Arg	Ala	Thr
50	305					310					315					320
	Gln	Leu	Gln	Lys	Asn	Leu	Thr	Cys	Glu	Val	Trp	Gly	Pro	Thr	Ser	Pro
					325					330					335	
55	Lys	Leu	Met	Leu	Ser	Leu	Lys	Leu	Glu	Asn	Lys	Glu	Ala	Lys	Val	Ser
				340					345					350		
	Lys	Arg	Glu	Lys	Ala	Val	Trp	Val	Leu	Asn	Pro	Glu	Ala	Gly	Met	Trp
			355					360					365			
60	Gln	Cys	Leu	Leu	Ser	Asp	Ser	Gly	Gln	Val	Leu	Leu	Glu	Ser	Asn	Ile
		370					375					380				
	Lys	Val	Leu	Pro	Thr	Trp	Ser	Thr	Pro	Val	Gln	Pro	Met	Ala	Leu	Ile
65	385					390					395					400
	Val	Leu	Gly	Gly	Val	Ala	Gly	Leu	Leu	Leu	Phe	Ile	Gly	Leu	Gly	Ile

```

                405                410                415
Phe Phe Cys Val Arg Cys Arg His Arg Arg Arg Gln Ala Glu Arg Met
      420                425                430
5
Ser Gln Ile Lys Arg Leu Leu Ser Glu Lys Lys Glu Cys Gln Cys Pro
      435                440                445
10
His Arg Phe Gln Lys Thr Cys Ser Pro Ile
      450                455

```

## (2) SEQ ID NO: 26的信息:

```

15 (i) 序列特征:
      (A) 长度: 828 氨基酸
      (B) 类型: 氨基酸
      (C) 链型: 单链
      (D) 拓扑结构: 线性
20
(ii) 分子类型: 肽
(xi) 序列描述: SEQ ID NO: 26:
25 Met Asn Ser Gly Val Ala Met Lys Tyr Gly Asn Asp Ser Ser Ala Glu
      1          5          10          15
Leu Ser Glu Leu His Ser Ala Ala Leu Ala Ser Leu Lys Gly Asp Ile
      20          25          30
30 Val Glu Leu Asn Lys Arg Leu Gln Gln Thr Glu Arg Glu Asp Leu Leu
      35          40          45
Glu Lys Lys Leu Ala Lys Ala Gln Cys Glu Gln Ser His Leu Met Arg
      50          55          60
35 Glu His Glu Asp Val Gln Glu Arg Thr Thr Leu Arg Tyr Glu Glu Arg
      65          70          75          80
40 Ile Thr Glu Leu His Ser Val Ile Ala Glu Leu Asn Lys Lys Ile Asp
      85          90          95
Arg Leu Gln Gly Thr Thr Ile Arg Glu Glu Asp Glu Tyr Ser Glu Leu
      100         105         110
45 Arg Ser Glu Leu Ser Gln Ser Gln His Glu Val Asn Glu Asp Ser Arg
      115         120         125
50 Ser Met Asp Gln Asp Gln Thr Ser Val Ser Ile Pro Glu Asn Gln Ser
      130         135         140
Thr Met Val Thr Ala Asp Met Asp Asn Cys Ser Asp Ile Asn Ser Glu
      145         150         155         160
55 Leu Gln Arg Val Leu Thr Gly Leu Glu Asn Val Val Cys Gly Arg Lys
      165         170         175
Lys Ser Ser Cys Ser Leu Ser Val Ala Glu Val Asp Arg His Ile Glu
      180         185         190
60 Gln Leu Thr Thr Ala Ser Glu His Cys Asp Leu Ala Ile Lys Thr Val
      195         200         205
65 Glu Glu Ile Glu Gly Val Leu Gly Arg Asp Leu Tyr Pro Asn Leu Ala
      210         215         220
Glu Glu Arg Ser Arg Trp Glu Lys Glu Leu Ala Gly Leu Arg Glu Glu

```

	225				230					235				240		
	Asn	Glu	Ser	Leu	Thr 245	Ala	Met	Leu	Cys	Ser 250	Lys	Glu	Glu	Glu	Leu	Asn 255
5	Arg	Thr	Lys	Ala 260	Thr	Met	Asn	Ala	Ile 265	Arg	Glu	Glu	Arg	Asp 270	Arg	Leu
10	Arg	Arg	Arg 275	Val	Arg	Glu	Leu	Gln 280	Thr	Arg	Leu	Gln	Ser 285	Val	Gln	Ala
	Thr	Gly 290	Pro	Ser	Ser	Pro	Gly 295	Arg	Leu	Thr	Ser	Thr 300	Asn	Arg	Pro	Ile
15	Asn 305	Pro	Ser	Thr	Gly	Glu 310	Leu	Ser	Thr	Ser	Ser 315	Ser	Ser	Asn	Asp	Ile 320
20	Pro	Ile	Ala	Lys	Ile 325	Ala	Glu	Arg	Val	Lys 330	Leu	Ser	Lys	Thr	Arg	Ser 335
	Glu	Ser	Ser	Ser 340	Ser	Asp	Arg	Pro	Val 345	Leu	Gly	Ser	Glu	Ile 350	Ser	Ser
25	Ile	Gly	Val 355	Ser	Ser	Ser	Val	Ala 360	Glu	His	Leu	Ala	His 365	Ser	Leu	Gln
	Asp	Cys 370	Ser	Asn	Ile	Gln	Glu 375	Ile	Phe	Gln	Thr	Leu 380	Tyr	Ser	His	Gly
30	Ser 385	Ala	Ile	Ser	Glu	Ser	Lys 390	Ile	Arg	Glu	Phe 395	Glu	Val	Glu	Thr	Glu 400
35	Arg	Leu	Asn	Ser 405	Arg	Ile	Glu	His	Leu	Lys 410	Ser	Gln	Asn	Asp	Leu	Leu 415
	Thr	Ile	Thr	Leu 420	Glu	Glu	Cys	Lys	Ser	Asn 425	Ala	Glu	Arg	Met 430	Ser	Met
40	Leu	Val	Gly 435	Lys	Tyr	Glu	Ser	Asn 440	Ala	Thr	Ala	Leu	Arg 445	Leu	Ala	Leu
	Gln	Tyr 450	Ser	Glu	Gln	Cys	Ile 455	Glu	Ala	Tyr	Glu	Leu 460	Leu	Leu	Ala	Leu
45	Ala 465	Glu	Ser	Glu	Gln	Ser	Leu 470	Ile	Leu	Gly	Gln 475	Phe	Arg	Ala	Ala	Gly 480
50	Val	Gly	Ser	Ser 485	Pro	Gly	Asp	Gln	Ser	Gly 490	Asp	Glu	Asn	Ile	Thr 495	Gln
	Met	Leu	Lys	Arg 500	Ala	His	Asp	Cys	Arg 505	Lys	Thr	Ala	Glu	Asn 510	Ala	Ala
55	Lys	Ala	Leu 515	Leu	Met	Lys	Leu	Asp 520	Gly	Ser	Cys	Gly	Gly 525	Ala	Phe	Ala
	Val	Ala	Gly 530	Cys	Ser	Val	Gln 535	Pro	Trp	Glu	Ser	Leu 540	Ser	Ser	Asn	Ser
60	His 545	Thr	Ser	Thr	Thr	Ser	Ser 550	Thr	Ala	Ser	Ser 555	Cys	Asp	Thr	Glu	Phe 560
65	Thr	Lys	Glu	Asp 565	Glu	Gln	Arg	Leu	Lys	Asp 570	Tyr	Ile	Gln	Gln	Leu	Lys 575
	Asn	Asp	Arg	Ala 580	Ala	Val	Lys	Leu	Thr 585	Met	Leu	Glu	Leu	Glu 590	Ser	Ile



```

His Ile Asp Pro Leu Ser Tyr Asp Val Lys Pro Arg Gly Asp Ser Gln
595 600 605
5 Arg Leu Asp Leu Glu Asn Ala Val Leu Met Gln Glu Leu Met Ala Met
610 615 620
Lys Glu Glu Met Ala Glu Leu Lys Ala Gln Leu Tyr Leu Leu Glu Lys
625 630 635 640
10 Glu Lys Lys Ala Leu Glu Leu Lys Leu Ser Thr Arg Glu Ala Gln Glu
645 650 655
Gln Ala Tyr Leu Val His Ile Glu His Leu Lys Ser Glu Val Glu Glu
660 665 670
15 Gln Lys Glu Gln Arg Met Arg Ser Leu Ser Ser Thr Ser Ser Gly Ser
675 680 685
Lys Asp Lys Pro Gly Lys Glu Cys Ala Asp Ala Ala Ser Pro Ala Leu
690 695 700
20 Ser Leu Ala Glu Leu Arg Thr Thr Cys Ser Glu Asn Glu Leu Ala Ala
705 710 715 720
25 Glu Phe Thr Asn Ala Ile Arg Arg Glu Lys Lys Leu Lys Ala Arg Val
725 730 735
Gln Glu Leu Val Ser Ala Leu Glu Arg Leu Thr Lys Ser Ser Glu Ile
740 745 750
30 Arg His Gln Gln Ser Ala Glu Phe Val Asn Asp Leu Lys Arg Ala Asn
755 760 765
35 Ser Asn Leu Val Ala Ala Tyr Glu Lys Ala Lys Lys Lys His Gln Asn
770 775 780
Lys Leu Lys Lys Leu Glu Ser Gln Met Met Ala Met Val Glu Arg His
785 790 795 800
40 Glu Thr Gln Val Arg Met Leu Lys Gln Arg Ile Ala Leu Leu Glu Glu
805 810 815
Glu Asn Ser Arg Pro His Thr Asn Glu Thr Ser Leu
820 825
45

```

## (2) .SEQ ID NO: 27的信息:

```

50 (i) 序列特征:
(A) 长度: 672 氨基酸
(B) 类型: 氨基酸
(C) 链型: 单链
(D) 拓扑结构: 线性
55 (ii) 分子类型: 肽
(xi) 序列描述: SEQ ID NO: 27:
60 Met Ala Asp Val Phe Pro Gly Asn Asp Ser Thr Ala Ser Gln Asp Val
1 5 10 15
Ala Asn Arg Phe Ala Arg Lys Gly Ala Leu Arg Gln Lys Asn Val His
20 25 30
65 Glu Val Lys Asp His Lys Phe Ile Ala Arg Phe Phe Lys Gln Pro Thr
35 40 45

```

	Phe	Cys	Ser	His	Cys	Thr	Asp	Phe	Ile	Trp	Gly	Phe	Gly	Lys	Gly	Gly
		50					55					60				
5	Phe	Gln	Cys	Gln	Val	Cys	Cys	Phe	Val	Val	His	Lys	Arg	Cys	His	Glu
	65					70					75					80
	Phe	Val	Thr	Phe	Ser	Cys	Pro	Gly	Ala	Asp	Lys	Gly	Pro	Asp	Thr	Asp
					85					90					95	
10	Asp	Pro	Arg	Ser	Lys	His	Lys	Phe	Lys	Ile	His	Thr	Tyr	Gly	Ser	Pro
				100				105						110		
	Thr	Phe	Cys	Asp	His	Cys	Gly	Ser	Leu	Leu	Tyr	Gly	Leu	Ile	His	Gln
			115					120					125			
15	Gly	Met	Lys	Cys	Asp	Thr	Cys	Asp	Met	Asn	Val	His	Lys	Gln	Cys	Val
		130					135					140				
	Ile	Asn	Val	Pro	Ser	Leu	Cys	Gly	Met	Asp	His	Thr	Glu	Lys	Arg	Gly
20	145					150					155					160
	Arg	Ile	Tyr	Leu	Lys	Ala	Glu	Val	Ala	Asp	Glu	Lys	Leu	His	Val	Thr
					165					170						175
25	Val	Arg	Asp	Ala	Lys	Asn	Leu	Ile	Pro	Met	Asp	Pro	Asn	Gly	Leu	Ser
				180					185					190		
	Asp	Pro	Tyr	Val	Lys	Leu	Lys	Leu	Ile	Pro	Asp	Pro	Lys	Asn	Glu	Ser
			195					200					205			
30	Lys	Gln	Lys	Thr	Lys	Thr	Ile	Arg	Ser	Thr	Leu	Asn	Pro	Gln	Trp	Asn
		210					215					220				
	Glu	Ser	Phe	Thr	Phe	Lys	Leu	Lys	Pro	Ser	Asp	Lys	Asp	Arg	Arg	Leu
35	225					230					235					240
	Ser	Val	Glu	Ile	Trp	Asp	Trp	Asp	Arg	Thr	Thr	Arg	Asn	Asp	Phe	Met
					245					250					255	
40	Gly	Ser	Leu	Ser	Phe	Gly	Val	Ser	Glu	Leu	Met	Lys	Met	Pro	Ala	Ser
				260					265					270		
	Gly	Trp	Tyr	Lys	Leu	Leu	Asn	Gln	Glu	Glu	Gly	Glu	Tyr	Tyr	Asn	Val
			275					280					285			
45	Pro	Ile	Pro	Glu	Gly	Asp	Glu	Glu	Gly	Asn	Met	Glu	Leu	Arg	Gln	Lys
		290					295					300				
	Phe	Glu	Lys	Ala	Lys	Leu	Gly	Pro	Ala	Gly	Asn	Lys	Val	Ile	Ser	Pro
50	305					310					315					320
	Ser	Glu	Asp	Arg	Lys	Gln	Pro	Ser	Asn	Asn	Leu	Asp	Arg	Val	Lys	Leu
					325					330					335	
55	Thr	Asp	Phe	Asn	Phe	Leu	Met	Val	Leu	Gly	Lys	Gly	Ser	Phe	Gly	Lys
				340					345					350		
	Val	Met	Leu	Ala	Asp	Arg	Lys	Gly	Thr	Glu	Glu	Leu	Tyr	Ala	Ile	Lys
			355					360					365			
60	Ile	Leu	Lys	Lys	Asp	Val	Val	Ile	Gln	Asp	Asp	Asp	Val	Glu	Cys	Thr
		370					375					380				
	Met	Val	Glu	Lys	Arg	Val	Leu	Ala	Leu	Leu	Asp	Lys	Pro	Pro	Phe	Leu
65	385					390					395					400
	Thr	Gln	Leu	His	Ser	Cys	Phe	Gln	Thr	Val	Asp	Arg	Leu	Tyr	Phe	Val



	Ser	Leu	Met	Gln	Leu	Asn	Asp	Asp	Thr	Arg	Leu	Tyr	Ser	Asn	Asp	Phe
				20					25					30		
5	Asn	Ser	Gly	Glu	Ala	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Trp	Thr	Val	Asp
			35					40					45			
	Ser	Glu	Asn	Arg	Thr	Asn	Leu	Ser	Cys	Glu	Gly	Cys	Leu	Ser	Pro	Ser
		50					55					60				
10	Cys	Leu	Ser	Leu	Leu	His	Leu	Gln	Glu	Lys	Asn	Trp	Ser	Ala	Leu	Leu
	65					70					75					80
	Thr	Ala	Val	Val	Ile	Ile	Leu	Thr	Ile	Ala	Gly	Asn	Ile	Leu	Val	Ile
					85					90					95	
15	Met	Ala	Val	Ser	Leu	Glu	Lys	Lys	Leu	Gln	Asn	Ala	Thr	Asn	Tyr	Phe
				100					105					110		
20	Leu	Met	Ser	Leu	Ala	Ile	Ala	Asp	Met	Leu	Leu	Gly	Phe	Leu	Val	Met
			115					120					125			
	Pro	Val	Ser	Met	Leu	Thr	Ile	Leu	Tyr	Gly	Tyr	Arg	Trp	Pro	Leu	Pro
		130					135					140				
25	Ser	Lys	Leu	Cys	Ala	Val	Trp	Ile	Tyr	Leu	Asp	Val	Leu	Phe	Ser	Thr
	145					150					155					160
	Ala	Ser	Ile	Met	His	Leu	Cys	Ala	Ile	Ser	Leu	Asp	Arg	Tyr	Val	Ala
					165					170					175	
30	Ile	Gln	Asn	Pro	Ile	His	His	Ser	Arg	Phe	Asn	Ser	Arg	Thr	Lys	Ala
				180					185					190		
	Phe	Leu	Lys	Ile	Ile	Ala	Val	Trp	Thr	Ile	Ser	Val	Gly	Ile	Ser	Met
			195					200					205			
35	Pro	Ile	Pro	Val	Phe	Gly	Leu	Gln	Asp	Asp	Ser	Lys	Val	Phe	Lys	Glu
		210					215					220				
40	Gly	Ser	Cys	Leu	Leu	Ala	Asp	Asp	Asn	Phe	Val	Leu	Ile	Gly	Ser	Phe
	225					230					235					240
	Val	Ser	Phe	Phe	Ile	Pro	Leu	Thr	Ile	Met	Val	Ile	Thr	Tyr	Phe	Leu
					245					250					255	
45	Thr	Ile	Lys	Ser	Leu	Gln	Lys	Glu	Ala	Thr	Leu	Cys	Val	Ser	Asp	Leu
				260					265					270		
	Gly	Thr	Arg	Ala	Lys	Leu	Ala	Ser	Phe	Ser	Phe	Leu	Pro	Gln	Ser	Ser
			275					280					285			
50	Leu	Ser	Ser	Glu	Lys	Leu	Phe	Gln	Arg	Ser	Ile	His	Arg	Glu	Pro	Gly
		290					295					300				
55	Ser	Tyr	Thr	Gly	Arg	Arg	Thr	Met	Gln	Ser	Ile	Ser	Asn	Glu	Gln	Lys
	305					310					315					320
	Ala	Cys	Lys	Val	Leu	Gly	Ile	Val	Phe	Phe	Leu	Phe	Val	Val	Met	Trp
					325					330					335	
60	Cys	Pro	Phe	Phe	Ile	Thr	Asn	Ile	Met	Ala	Val	Ile	Cys	Lys	Glu	Ser
				340					345					350		
	Cys	Asn	Glu	Asp	Val	Ile	Gly	Ala	Leu	Leu	Asn	Val	Phe	Val	Trp	Ile
			355					360					365			
65	Gly	Tyr	Leu	Ser	Ser	Ala	Val	Asn	Pro	Leu	Val	Tyr	Thr	Leu	Phe	Asn

	370		375		380														
	Lys	Thr	Tyr	Arg	Ser	Ala	Phe	Ser	Arg	Tyr	Ile	Gln	Cys	Gln	Tyr	Lys			
	385					390					395					400			
5	Glu	Asn	Lys	Lys	Pro	Leu	Gln	Leu	Ile	Leu	Val	Asn	Thr	Ile	Pro	Ala			
					405					410					415				
10	Leu	Ala	Tyr	Lys	Ser	Ser	Gln	Leu	Gln	Met	Gly	Gln	Lys	Lys	Asn	Ser			
				420					425					430					
	Lys	Gln	Asp	Ala	Lys	Thr	Thr	Asp	Asn	Asp	Cys	Ser	Met	Val	Ala	Leu			
			435					440					445						
15	Gly	Lys	Gln	His	Ser	Glu	Glu	Ala	Ser	Lys	Asp	Asn	Ser	Asp	Gly	Val			
		450					455					460							
20	Asn	Glu	Lys	Val	Ser	Cys	Val												
	465					470													

## (2) SEQ ID NO: 29的信息:

(i) 序列特征:

(A) 长度: 481 氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 29:

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

	180					185					190					
	Lys	Gly	Ile	Glu	Thr	Asp	Val	Asp	Asn	Pro	Asn	Asn	Ile	Thr	Cys	Val
			195					200					205			
5	Leu	Thr	Lys	Glu	Arg	Phe	Gly	Asp	Phe	Met	Leu	Phe	Gly	Ser	Leu	Ala
		210					215					220				
10	Ala	Phe	Phe	Thr	Pro	Leu	Ala	Ile	Met	Ile	Val	Thr	Tyr	Phe	Leu	Thr
	225					230					235					240
	Ile	His	Ala	Leu	Gln	Lys	Lys	Ala	Tyr	Leu	Val	Lys	Asn	Lys	Pro	Pro
					245					250					255	
15	Gln	Arg	Leu	Thr	Trp	Leu	Thr	Val	Ser	Thr	Val	Phe	Gln	Arg	Asp	Glu
				260					265						270	
	Thr	Pro	Cys	Ser	Ser	Pro	Glu	Lys	Val	Ala	Met	Leu	Asp	Gly	Ser	Arg
20			275					280					285			
	Lys	Asp	Lys	Ala	Leu	Pro	Asn	Ser	Gly	Asp	Glu	Thr	Leu	Met	Arg	Arg
		290					295					300				
25	Thr	Ser	Thr	Ile	Gly	Lys	Lys	Ser	Val	Gln	Thr	Ile	Ser	Asn	Glu	Gln
	305				310						315					320
	Arg	Ala	Ser	Lys	Val	Leu	Gly	Ile	Val	Phe	Phe	Leu	Phe	Leu	Leu	Met
					325					330						335
30	Trp	Cys	Pro	Phe	Phe	Ile	Thr	Asn	Ile	Thr	Leu	Val	Leu	Cys	Asp	Ser
				340					345					350		
	Cys	Asn	Gln	Thr	Thr	Leu	Gln	Met	Leu	Leu	Glu	Ile	Phe	Val	Trp	Ile
35			355					360					365			
	Gly	Tyr	Val	Ser	Ser	Gly	Val	Asn	Pro	Leu	Val	Tyr	Thr	Leu	Phe	Asn
		370					375					380				
40	Lys	Thr	Phe	Arg	Asp	Ala	Phe	Gly	Arg	Tyr	Ile	Thr	Cys	Asn	Tyr	Arg
	385					390					395					400
	Ala	Thr	Lys	Ser	Val	Lys	Thr	Leu	Arg	Lys	Arg	Ser	Ser	Lys	Ile	Tyr
					405					410					415	
45	Phe	Arg	Asn	Pro	Met	Ala	Glu	Asn	Ser	Lys	Phe	Phe	Lys	Lys	His	Gly
				420					425					430		
	Ile	Arg	Asn	Gly	Ile	Asn	Pro	Ala	Met	Tyr	Gln	Ser	Pro	Met	Arg	Leu
50			435					440					445			
	Arg	Ser	Ser	Thr	Ile	Gln	Ser	Ser	Ser	Ile	Ile	Leu	Leu	Asp	Thr	Leu
		450					455					460				
55	Leu	Leu	Thr	Glu	Asn	Glu	Gly	Asp	Lys	Thr	Glu	Glu	Gln	Val	Ser	Val
	465					470					475					480
	Val															

60 (2) SEQ ID NO: 30的信息:

- (i) 序列特征:
- (A) 长度: 2843 氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线性

65

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 30:

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

	Met	Ser	Ser	Ser	Gln	Asp	Ser	Cys	Ile	Ser	Met	Arg	Gln	Ser	Gly	Cys
				340					345					350		
5	Leu	Pro	Leu	Leu	Ile	Gln	Leu	Leu	His	Gly	Asn	Asp	Lys	Asp	Ser	Val
			355					360					365			
	Leu	Leu	Gly	Asn	Ser	Arg	Gly	Ser	Lys	Glu	Ala	Arg	Ala	Arg	Ala	Ser
		370					375					380				
10	Ala	Ala	Leu	His	Asn	Ile	Ile	His	Ser	Gln	Pro	Asp	Asp	Lys	Arg	Gly
	385					390					395					400
	Arg	Arg	Glu	Ile	Arg	Val	Leu	His	Leu	Leu	Glu	Gln	Ile	Arg	Ala	Tyr
					405					410					415	
15	Cys	Ser	Thr	Cys	Trp	Glu	Trp	Gln	Glu	Ala	His	Glu	Pro	Gly	Met	Asp
				420					425					430		
	Gln	Asp	Lys	Asn	Pro	Met	Pro	Ala	Pro	Val	Glu	His	Gln	Ile	Cys	Pro
20			435					440					445			
	Ala	Val	Cys	Val	Leu	Met	Lys	Leu	Ser	Phe	Asp	Glu	Glu	His	Arg	His
		450					455					460				
25	Ala	Met	Asn	Glu	Leu	Gly	Gly	Leu	Gln	Ala	Ile	Ala	Glu	Leu	Leu	Gln
	465					470					475					480
	Val	Asp	Cys	Glu	Met	Tyr	Gly	Leu	Thr	Asn	Asp	His	Tyr	Ser	Ile	Thr
					485					490					495	
30	Leu	Arg	Arg	Tyr	Ala	Gly	Met	Ala	Leu	Thr	Asn	Leu	Thr	Phe	Gly	Asp
				500					505					510		
	Val	Ala	Asn	Lys	Ala	Thr	Leu	Cys	Ser	Met	Lys	Gly	Cys	Met	Arg	Ala
35			515					520					525			
	Leu	Val	Ala	Gln	Leu	Lys	Ser	Glu	Ser	Glu	Asp	Leu	Gln	Gln	Val	Ile
		530					535					540				
40	Ala	Ser	Val	Leu	Arg	Asn	Leu	Ser	Trp	Arg	Ala	Asp	Val	Asn	Ser	Lys
	545					550					555					560
	Lys	Thr	Leu	Arg	Glu	Val	Gly	Ser	Val	Lys	Ala	Leu	Met	Glu	Cys	Ala
					565					570					575	
45	Leu	Glu	Val	Lys	Lys	Glu	Ser	Thr	Leu	Lys	Ser	Val	Leu	Ser	Ala	Leu
				580					585					590		
	Trp	Asn	Leu	Ser	Ala	His	Cys	Thr	Glu	Asn	Lys	Ala	Asp	Ile	Cys	Ala
50			595					600					605			
	Val	Asp	Gly	Ala	Leu	Ala	Phe	Leu	Val	Gly	Thr	Leu	Thr	Tyr	Arg	Ser
		610					615					620				
55	Gln	Thr	Asn	Thr	Leu	Ala	Ile	Ile	Glu	Ser	Gly	Gly	Gly	Ile	Leu	Arg
	625					630					635					640
	Asn	Val	Ser	Ser	Leu	Ile	Ala	Thr	Asn	Glu	Asp	His	Arg	Gln	Ile	Leu
					645					650					655	
60	Arg	Glu	Asn	Asn	Cys	Leu	Gln	Thr	Leu	Leu	Gln	His	Leu	Lys	Ser	His
				660					665					670		
	Ser	Leu	Thr	Ile	Val	Ser	Asn	Ala	Cys	Gly	Thr	Leu	Trp	Asn	Leu	Ser
65			675					680					685			
	Ala	Arg	Asn	Pro	Lys	Asp	Gln	Glu	Ala	Leu	Trp	Asp	Met	Gly	Ala	Val





	Glu Asp Glu Ile Lys Gln Ser Glu Gln Arg Gln Ser Arg Asn Gln Ser	1060	1065	1070
5	Thr Thr Tyr Pro Val Tyr Thr Glu Ser Thr Asp Asp Lys His Leu Lys	1075	1080	1085
	Phe Gln Pro His Phe Gly Gln Gln Glu Cys Val Ser Pro Tyr Arg Ser	1090	1095	1100
10	Arg Gly Ala Asn Gly Ser Glu Thr Asn Arg Val Gly Ser Asn His Gly	1105	1110	1115
	Ile Asn Gln Asn Val Ser Gln Ser Leu Cys Gln Glu Asp Asp Tyr Glu	1125	1130	1135
15	Asp Asp Lys Pro Thr Asn Tyr Ser Glu Arg Tyr Ser Glu Glu Glu Gln	1140	1145	1150
	His Glu Glu Glu Glu Arg Pro Thr Asn Tyr Ser Ile Lys Tyr Asn Glu	1155	1160	1165
20	Glu Lys Arg His Val Asp Gln Pro Ile Asp Tyr Ser Ile Leu Lys Ala	1170	1175	1180
	Thr Asp Ile Pro Ser Ser Gln Lys Gln Ser Phe Ser Phe Ser Lys Ser	1185	1190	1195
25	Ser Ser Gly Gln Ser Ser Lys Thr Glu His Met Ser Ser Ser Ser Glu	1205	1210	1215
	Asn Thr Ser Thr Pro Ser Ser Asn Ala Lys Arg Gln Asn Gln Leu His	1220	1225	1230
30	Pro Ser Ser Ala Gln Ser Arg Ser Gly Gln Pro Gln Lys Ala Ala Thr	1235	1240	1245
	Cys Lys Val Ser Ser Ile Asn Gln Glu Thr Ile Gln Thr Tyr Cys Val	1250	1255	1260
35	Glu Asp Thr Pro Ile Cys Phe Ser Arg Cys Ser Ser Leu Ser Ser Leu	1265	1270	1275
	Ser Ser Ala Glu Asp Glu Ile Gly Cys Asn Gln Thr Thr Gln Glu Ala	1285	1290	1295
40	Asp Ser Ala Asn Thr Leu Gln Ile Ala Glu Ile Lys Glu Lys Ile Gly	1300	1305	1310
	Thr Arg Ser Ala Glu Asp Pro Val Ser Glu Val Pro Ala Val Ser Gln	1315	1320	1325
45	His Pro Arg Thr Lys Ser Ser Arg Leu Gln Gly Ser Ser Leu Ser Ser	1330	1335	1340
	Glu Ser Ala Arg His Lys Ala Val Glu Phe Ser Ser Gly Ala Lys Ser	1345	1350	1355
50	Pro Ser Lys Ser Gly Ala Gln Thr Pro Lys Ser Pro Pro Glu His Tyr	1365	1370	1375
	Val Gln Glu Thr Pro Leu Met Phe Ser Arg Cys Thr Ser Val Ser Ser	1380	1385	1390
55	Leu Asp Ser Phe Glu Ser Arg Ser Ile Ala Ser Ser Val Gln Ser Glu	1395	1400	1405
	Pro Cys Ser Gly Met Val Ser Gly Ile Ile Ser Pro Ser Asp Leu Pro			

		1410				1415					1420									
		Asp	Ser	Pro	Gly	Gln	Thr	Met	Pro	Pro	Ser	Arg	Ser	Lys	Thr	Pro	Pro			
		1425						1430				1435							1440	
5		Pro	Pro	Pro	Gln	Thr	Ala	Gln	Thr	Lys	Arg	Glu	Val	Pro	Lys	Asn	Lys			
						1445					1450					1455				
		Ala	Pro	Thr	Ala	Glu	Lys	Arg	Glu	Ser	Gly	Pro	Lys	Gln	Ala	Ala	Val			
10					1460					1465					1470					
		Asn	Ala	Ala	Val	Gln	Arg	Val	Gln	Val	Leu	Pro	Asp	Ala	Asp	Thr	Leu			
				1475					1480					1485						
15		Leu	His	Phe	Ala	Thr	Glu	Ser	Thr	Pro	Asp	Gly	Phe	Ser	Cys	Ser	Ser			
			1490					1495					1500							
		Ser	Leu	Ser	Ala	Leu	Ser	Leu	Asp	Glu	Pro	Phe	Ile	Gln	Lys	Asp	Val			
20						1510						1515				1520				
		Glu	Leu	Arg	Ile	Met	Pro	Pro	Val	Gln	Glu	Asn	Asp	Asn	Gly	Asn	Glu			
					1525						1530				1535					
		Thr	Glu	Ser	Glu	Gln	Pro	Lys	Glu	Ser	Asn	Glu	Asn	Gln	Glu	Lys	Glu			
25					1540					1545					1550					
		Ala	Glu	Lys	Thr	Ile	Asp	Ser	Glu	Lys	Asp	Leu	Leu	Asp	Asp	Ser	Asp			
				1555					1560					1565						
30		Asp	Asp	Asp	Ile	Glu	Ile	Leu	Glu	Glu	Cys	Ile	Ile	Ser	Ala	Met	Pro			
			1570					1575					1580							
		Thr	Lys	Ser	Ser	Arg	Lys	Ala	Lys	Lys	Pro	Ala	Gln	Thr	Ala	Ser	Lys			
35						1590						1595					1600			
		Leu	Pro	Pro	Pro	Val	Ala	Arg	Lys	Pro	Ser	Gln	Leu	Pro	Val	Tyr	Lys			
					1605						1610					1615				
		Leu	Leu	Pro	Ser	Gln	Asn	Arg	Leu	Gln	Pro	Gln	Lys	His	Val	Ser	Phe			
40					1620					1625					1630					
		Thr	Pro	Gly	Asp	Asp	Met	Pro	Arg	Val	Tyr	Cys	Val	Glu	Gly	Thr	Pro			
				1635					1640					1645						
45		Ile	Asn	Phe	Ser	Thr	Ala	Thr	Ser	Leu	Ser	Asp	Leu	Thr	Ile	Glu	Ser			
			1650					1655					1660							
		Pro	Pro	Asn	Glu	Leu	Ala	Ala	Gly	Glu	Gly	Val	Arg	Gly	Gly	Ala	Gln			
50						1670						1675				1680				
		Ser	Gly	Glu	Phe	Glu	Lys	Arg	Asp	Thr	Ile	Pro	Thr	Glu	Gly	Arg	Ser			
					1685						1690					1695				
		Thr	Asp	Glu	Ala	Gln	Gly	Gly	Lys	Thr	Ser	Ser	Val	Thr	Ile	Pro	Glu			
55					1700					1705					1710					
		Leu	Asp	Asp	Asn	Lys	Ala	Glu	Glu	Gly	Asp	Ile	Leu	Ala	Glu	Cys	Ile			
				1715					1720					1725						
60		Asn	Ser	Ala	Met	Pro	Lys	Gly	Lys	Ser	His	Lys	Pro	Phe	Arg	Val	Lys			
			1730					1735					1740							
		Lys	Ile	Met	Asp	Gln	Val	Gln	Gln	Ala	Ser	Ala	Ser	Ser	Ser	Ala	Pro			
65						1750						1755				1760				
		Asn	Lys	Asn	Gln	Leu	Asp	Gly	Lys	Lys	Lys	Lys	Lys	Pro	Thr	Ser	Pro	Val		
					1765							1770				1775				

	Lys Pro Ile Pro Gln Asn Thr Glu Tyr Arg Thr Arg Val Arg Lys Asn	1780	1785	1790
5	Ala Asp Ser Lys Asn Asn Leu Asn Ala Glu Arg Val Phe Ser Asp Asn	1795	1800	1805
	Lys Asp Ser Lys Lys Gln Asn Leu Lys Asn Asn Ser Lys Asp Phe Asn	1810	1815	1820
10	Asp Lys Leu Pro Asn Asn Glu Asp Arg Val Arg Gly Ser Phe Ala Phe	1825	1830	1835
	Asp Ser Pro His His Tyr Thr Pro Ile Glu Gly Thr Pro Tyr Cys Phe	1845	1850	1855
15	Ser Arg Asn Asp Ser Leu Ser Ser Leu Asp Phe Asp Asp Asp Val	1860	1865	1870
	Asp Leu Ser Arg Glu Lys Ala Glu Leu Arg Lys Ala Lys Glu Asn Lys	1875	1880	1885
20	Glu Ser Glu Ala Lys Val Thr Ser His Thr Glu Leu Thr Ser Asn Gln	1890	1895	1900
	Gln Ser Ala Asn Lys Thr Gln Ala Ile Ala Lys Gln Pro Ile Asn Arg	1905	1910	1915
25	Gly Gln Pro Lys Pro Ile Leu Gln Lys Gln Ser Thr Phe Pro Gln Ser	1925	1930	1935
30	Ser Lys Asp Ile Pro Asp Arg Gly Ala Ala Thr Asp Glu Lys Leu Gln	1940	1945	1950
	Asn Phe Ala Ile Glu Asn Thr Pro Val Cys Phe Ser His Asn Ser Ser	1955	1960	1965
35	Leu Ser Ser Leu Ser Asp Ile Asp Gln Glu Asn Asn Asn Lys Glu Asn	1970	1975	1980
40	Glu Pro Ile Lys Glu Thr Glu Pro Pro Asp Ser Gln Gly Glu Pro Ser	1985	1990	1995
	Lys Pro Gln Ala Ser Gly Tyr Ala Pro Lys Ser Phe His Val Glu Asp	2005	2010	2015
45	Thr Pro Val Cys Phe Ser Arg Asn Ser Ser Leu Ser Ser Leu Ser Ile	2020	2025	2030
	Asp Ser Glu Asp Asp Leu Leu Gln Glu Cys Ile Ser Ser Ala Met Pro	2035	2040	2045
50	Lys Lys Lys Lys Pro Ser Arg Leu Lys Gly Asp Asn Glu Lys His Ser	2050	2055	2060
	Pro Arg Asn Met Gly Gly Ile Leu Gly Glu Asp Leu Thr Leu Asp Leu	2065	2070	2075
55	Lys Asp Ile Gln Arg Pro Asp Ser Glu His Gly Leu Ser Pro Asp Ser	2085	2090	2095
60	Glu Asn Phe Asp Trp Lys Ala Ile Gln Glu Gly Ala Asn Ser Ile Val	2100	2105	2110
	Ser Ser Leu His Gln Ala Ala Ala Ala Cys Leu Ser Arg Gln Ala	2115	2120	2125
65	Ser Ser Asp Ser Asp Ser Ile Leu Ser Leu Lys Ser Gly Ile Ser Leu			



Ser Ser Val Gln Ala Gly Gly Trp Arg Lys Leu Pro Pro Asn Leu Ser  
 2500 2505 2510  
 5 Pro Thr Ile Glu Tyr Asn Asp Gly Arg Pro Ala Lys Arg His Asp Ile  
 2515 2520 2525  
 Ala Arg Ser His Ser Glu Ser Pro Ser Arg Leu Pro Ile Asn Arg Ser  
 2530 2535 2540  
 10 Gly Thr Trp Lys Arg Glu His Ser Lys His Ser Ser Ser Leu Pro Arg  
 2545 2550 2555 2560  
 Val Ser Thr Trp Arg Arg Thr Gly Ser Ser Ser Ser Ile Leu Ser Ala  
 2565 2570 2575  
 15 Ser Ser Glu Ser Ser Glu Lys Ala Lys Ser Glu Asp Glu Lys His Val  
 2580 2585 2590  
 Asn Ser Ile Ser Gly Thr Lys Gln Ser Lys Glu Asn Gln Val Ser Ala  
 2595 2600 2605  
 Lys Gly Thr Trp Arg Lys Ile Lys Glu Asn Glu Phe Ser Pro Thr Asn  
 2610 2615 2620  
 25 Ser Thr Ser Gln Thr Val Ser Ser Gly Ala Thr Asn Gly Ala Glu Ser  
 2625 2630 2635 2640  
 Lys Thr Leu Ile Tyr Gln Met Ala Pro Ala Val Ser Lys Thr Glu Asp  
 2645 2650 2655  
 30 Val Trp Val Arg Ile Glu Asp Cys Pro Ile Asn Asn Pro Arg Ser Gly  
 2660 2665 2670  
 Arg Ser Pro Thr Gly Asn Thr Pro Pro Val Ile Asp Ser Val Ser Glu  
 2675 2680 2685  
 35 Lys Ala Asn Pro Asn Ile Lys Asp Ser Lys Asp Asn Gln Ala Lys Gln  
 2690 2695 2700  
 40 Asn Val Gly Asn Gly Ser Val Pro Met Arg Thr Val Gly Leu Glu Asn  
 2705 2710 2715 2720  
 Arg Leu Asn Ser Phe Ile Gln Val Asp Ala Pro Asp Gln Lys Gly Thr  
 2725 2730 2735  
 45 Glu Ile Lys Pro Gly Gln Asn Asn Pro Val Pro Val Ser Glu Thr Asn  
 2740 2745 2750  
 Glu Ser Ser Ile Val Glu Arg Thr Pro Phe Ser Ser Ser Ser Ser Ser  
 2755 2760 2765  
 50 Lys His Ser Ser Pro Ser Gly Thr Val Ala Ala Arg Val Thr Pro Phe  
 2770 2775 2780  
 Asn Tyr Asn Pro Ser Pro Arg Lys Ser Ser Ala Asp Ser Thr Ser Ala  
 2785 2790 2795 2800  
 Arg Pro Ser Gln Ile Pro Thr Pro Val Asn Asn Asn Thr Lys Lys Arg  
 2805 2810 2815  
 60 Asp Ser Lys Thr Asp Ser Thr Glu Ser Ser Gly Thr Gln Ser Pro Lys  
 2820 2825 2830  
 Arg His Ser Gly Ser Tyr Leu Val Thr Ser Val  
 2835 2840  
 65

## (2) SEQ ID NO: 31的信息:

## (i) 序列特征:

- (A) 长度: 65 碱基对  
 (B) 类型: 核酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

5

## (ii) 分子类型: 其它核酸

10

## (iii) 假设: 否

## (iv) 反义: 否

15

## (xi) 序列描述: SEQ ID NO: 31:

CGGAATTCNN NNNNNNNAAC AGCNNNNNNN NNAATGAANN NCAAAGTCTG NNNTGAGGAT 60  
 CCTCA 65

20

## (2) SEQ ID NO: 32的信息:

## (i) 序列特征:

- (A) 长度: 65 碱基对  
 (B) 类型: 核酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

25

## (ii) 分子类型: 其它核酸

30

## (iv) 反义: 否

## (xi) 序列描述: SEQ ID NO: 32:

CGGAATTCGA CTCAGAANN NNNACTTCA GANNNNNNAT CNNNNNNNN GTCTGAGGAT 60  
 CCTCA 65

35

40

## (2) SEQ ID NO: 33的信息:

## (i) 序列特征:

- (A) 长度: 65 碱基对  
 (B) 类型: 核酸  
 (C) 链型: 单链  
 (D) 拓扑结构: 线性

45

## (ii) 分子类型: 其它核酸

50

## (iv) 反义: 否

## (xi) 序列描述: SEQ ID NO: 33:

CGGAATTCNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNTGAGGAT 60  
 CCTCA 65

55

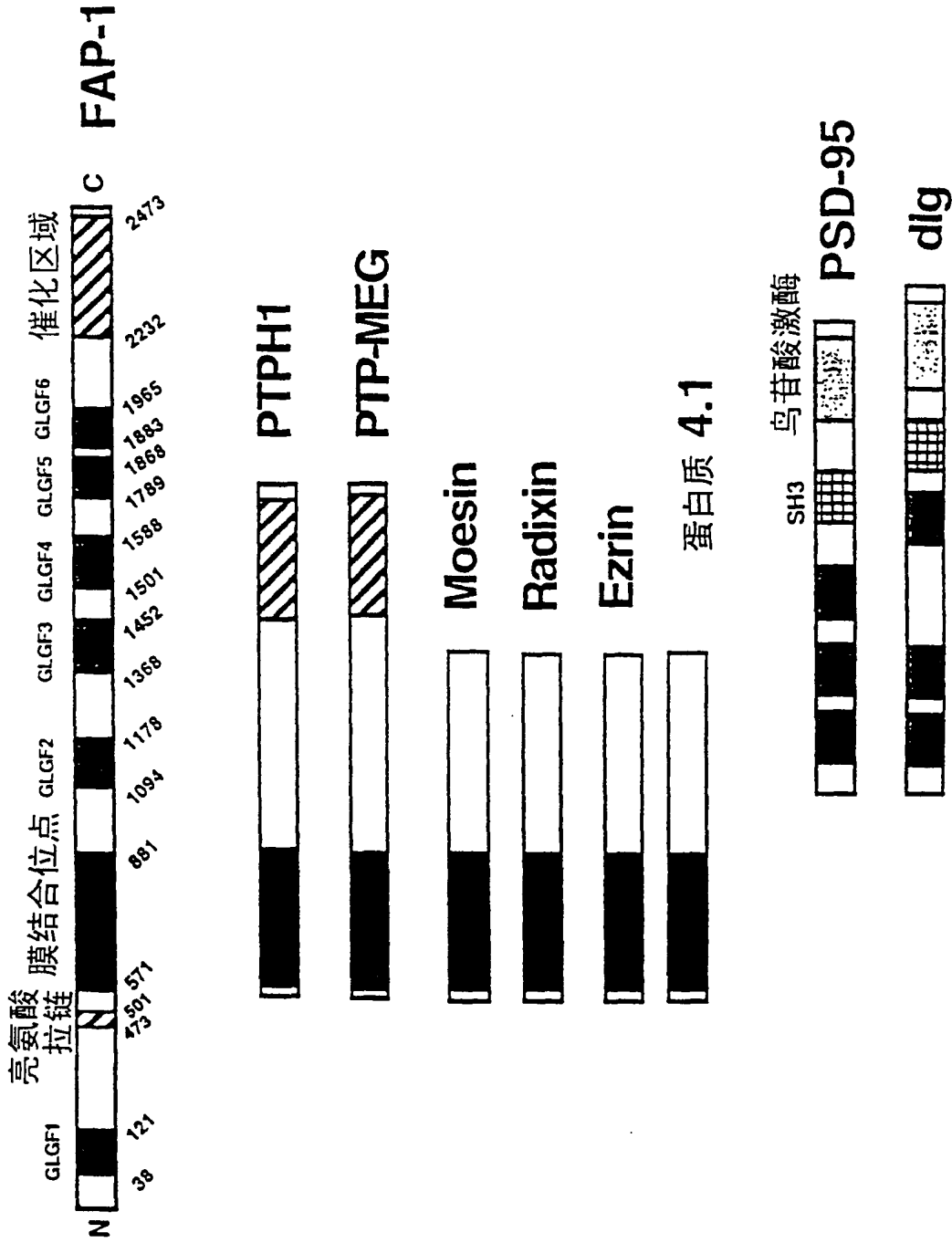


图 1



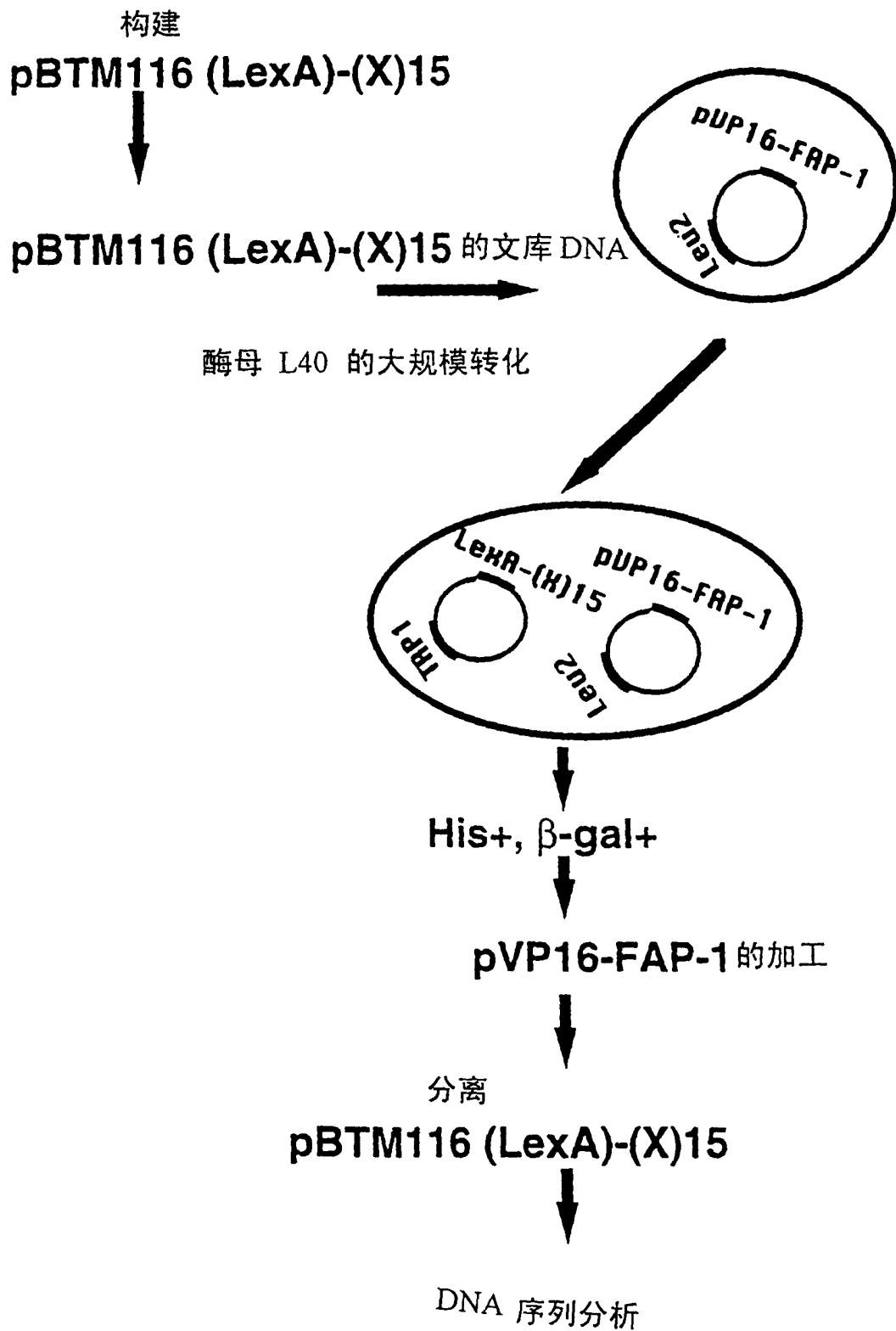


图 2A

人 D S E N S N F R N E I Q S L V  
 大鼠 S I S N S R N E N E G Q S L E  
 小鼠 S T P D T G N E N E G Q C L E

图 2B

- - - N S - - - N E - Q S L -

C	Y	A		A	I	G		L			V	12-0
E	N	A		G	V	S		E			V	5-0
W	W	G		A	T	Q		P			V	13-0
E	H	A		Q	Q	Q		Q			V	20-0
N	S	S		F	H	S		L			V	6-2
G	L	R		L	P	P		D			V	9-5
G	S	D		S	G	V		N			V	18-1
D	K	K		R	P	V		N			V	22-1
T	G	K		D	V	W		A			V	71-1
A	S	R		N	E	E		L			I	14-5

图 2C

I	P	D	S	E	D	G	N	E	E	Q	S	L	V	8-1
D	S	E	M	Y	N	F	R	S	Q	L	A	S	V	9-3
I	D	L	A	S	E	F	L	F	L	S	N	S	F	14-1
P	P	T	C	S	Q	A	N	S	G	R	I	S	T	0-2
S	D	S	N	M	N	M	N	E	L	S	E	V	57-5	
Q	N	F	R	T	Y	I	V	S	F	V	72-1			
R	E	T	I	E	S	T	V	25-9						
R	I	G	F	I	S	S	L	V	16-13					
T	I	Q	S	V	I	6-3								
E	S	L	V	18-1										

图 2D 共有序列: t S-X-V/L/I

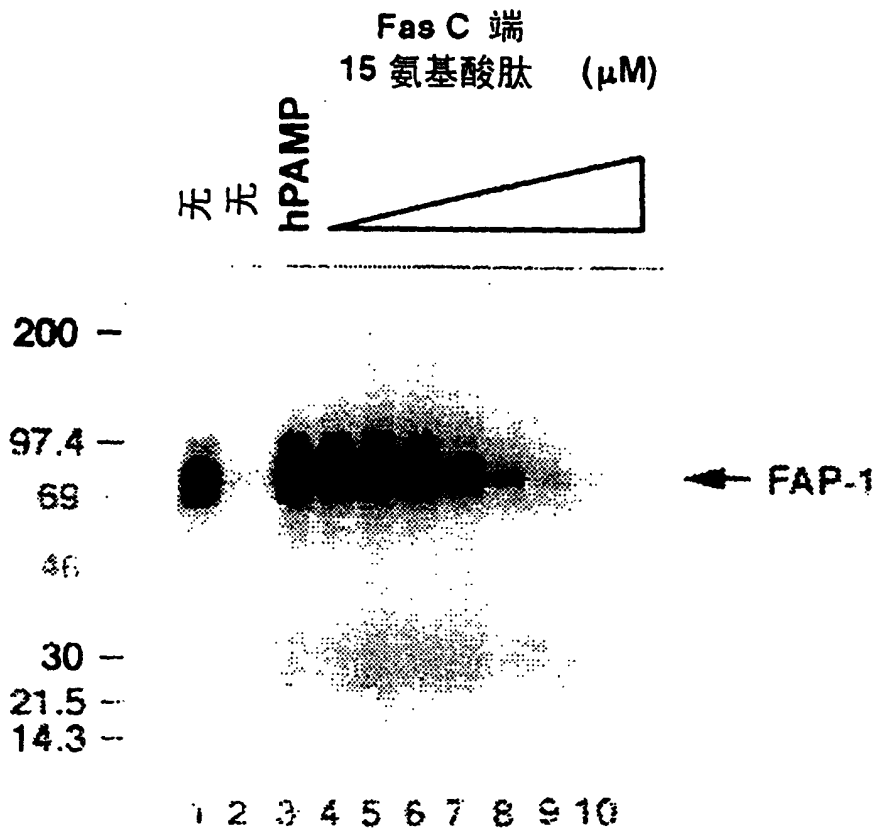


图 3A

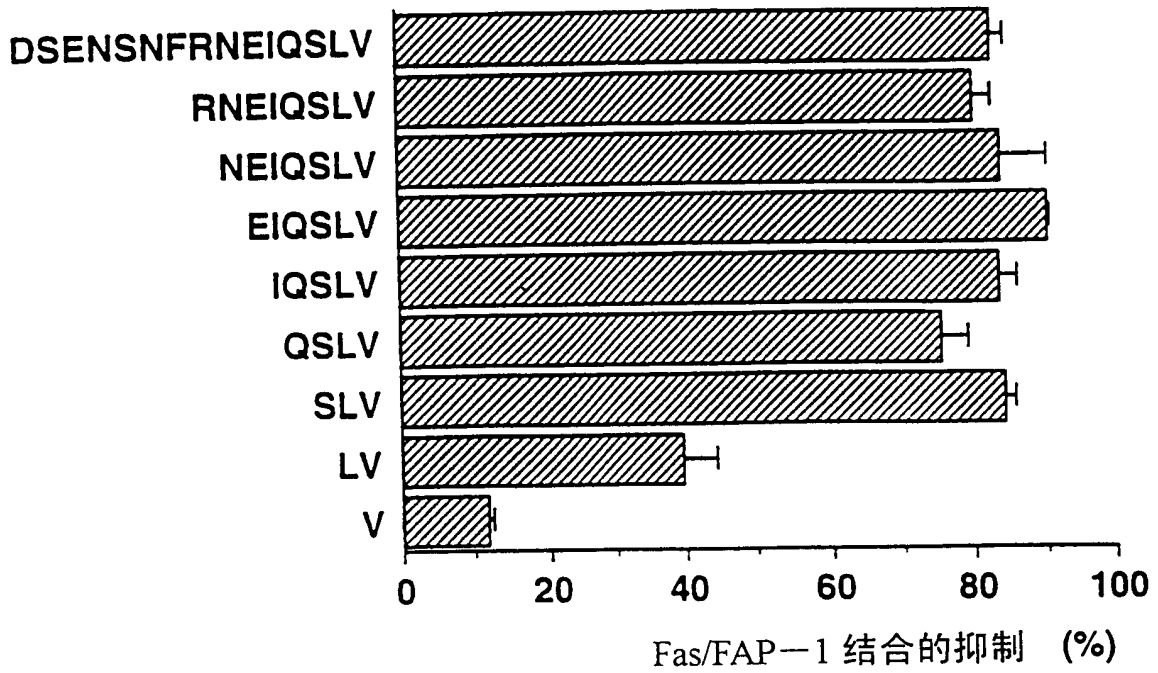


图 3B

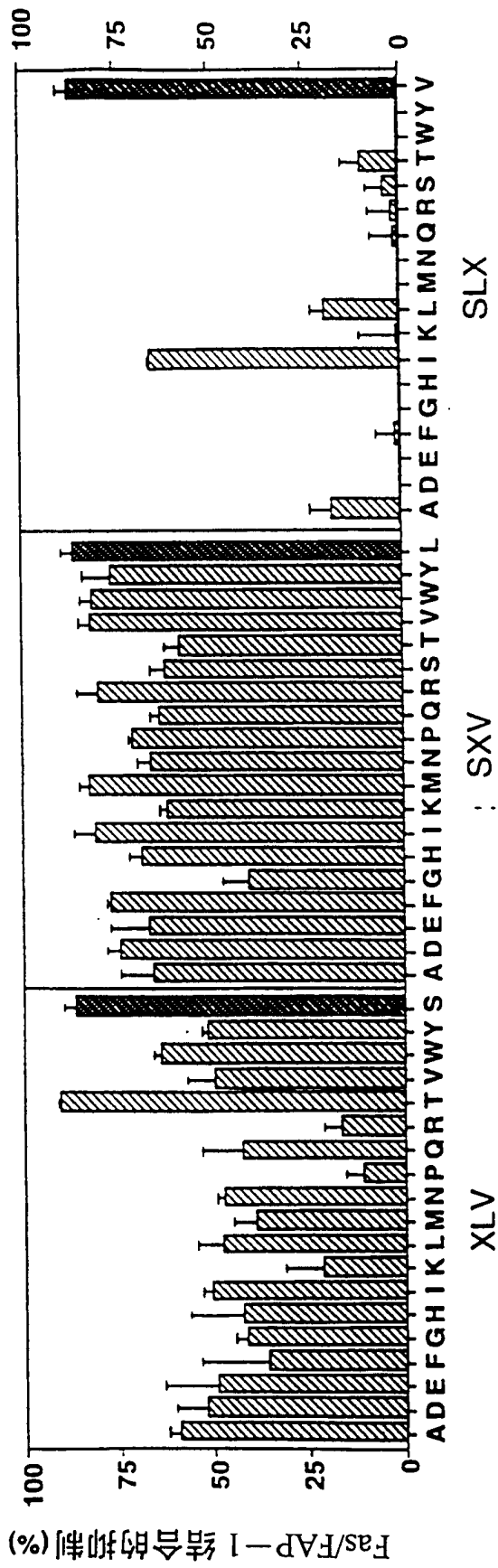


图 3C

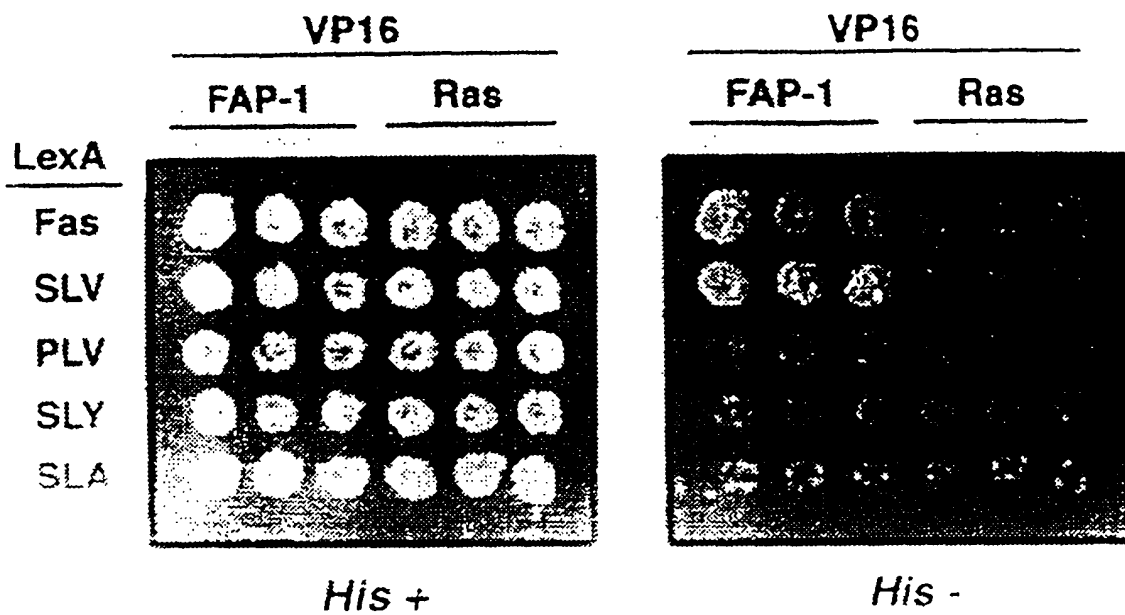


图 4A



图 4B

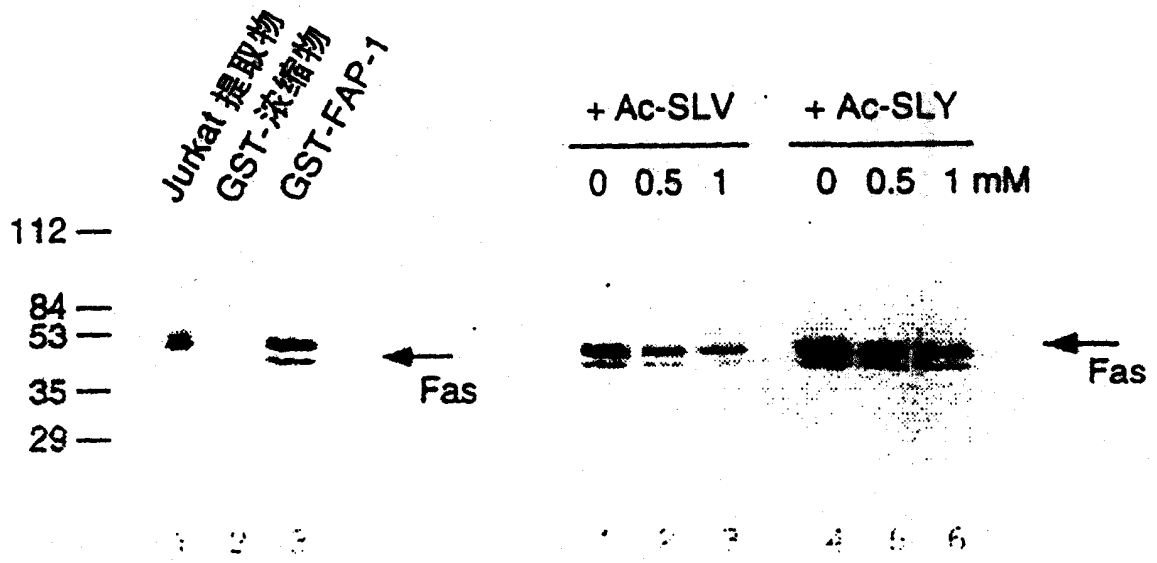
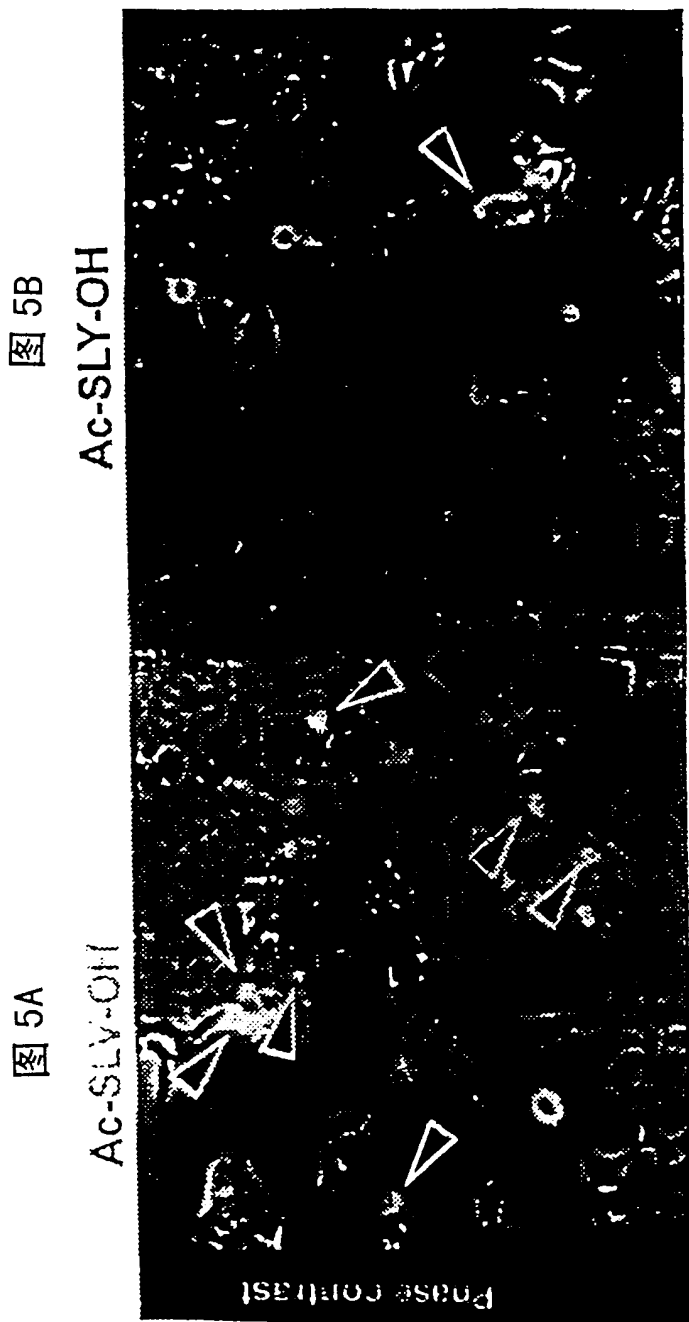
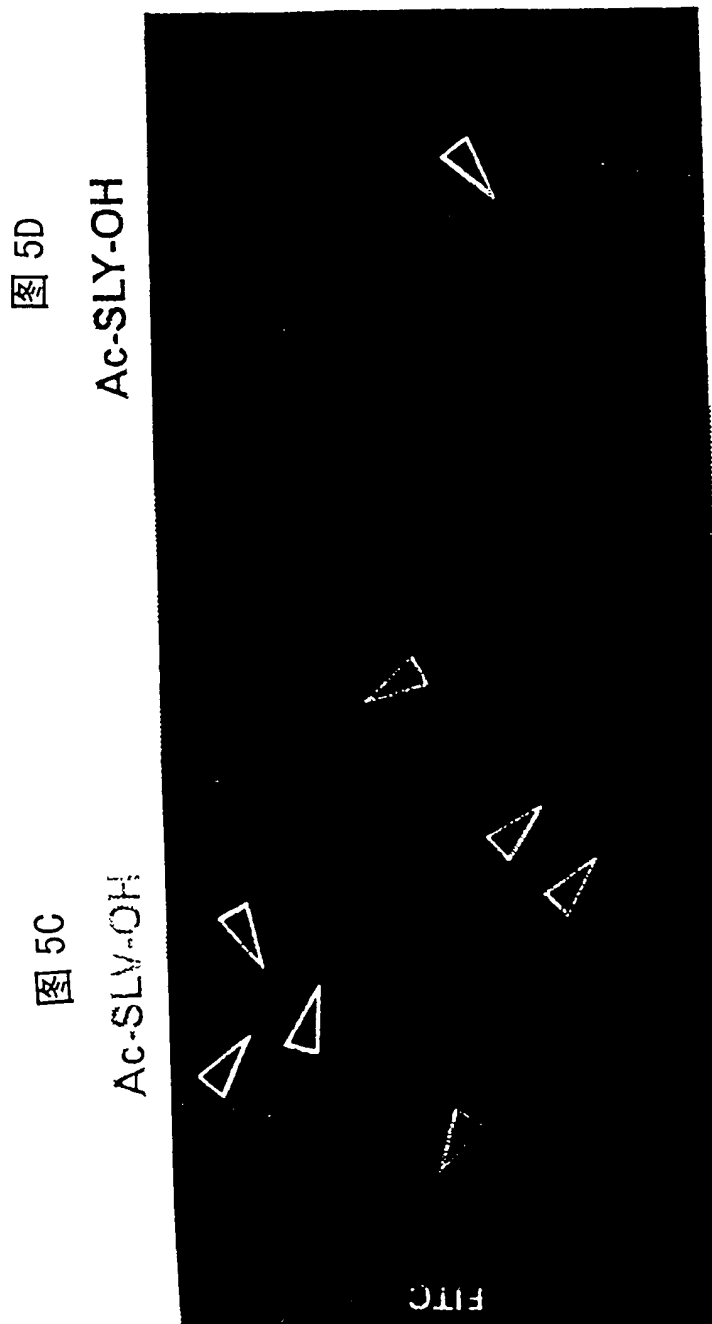


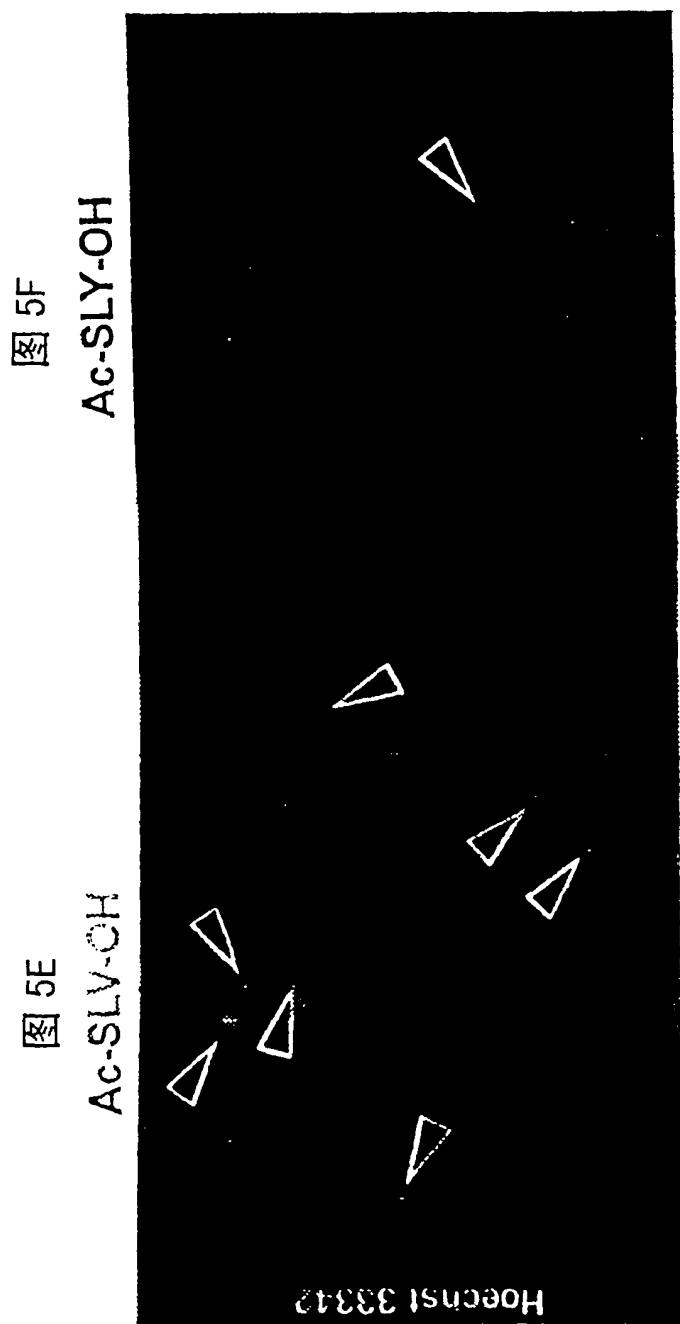
图 4C

图 4D









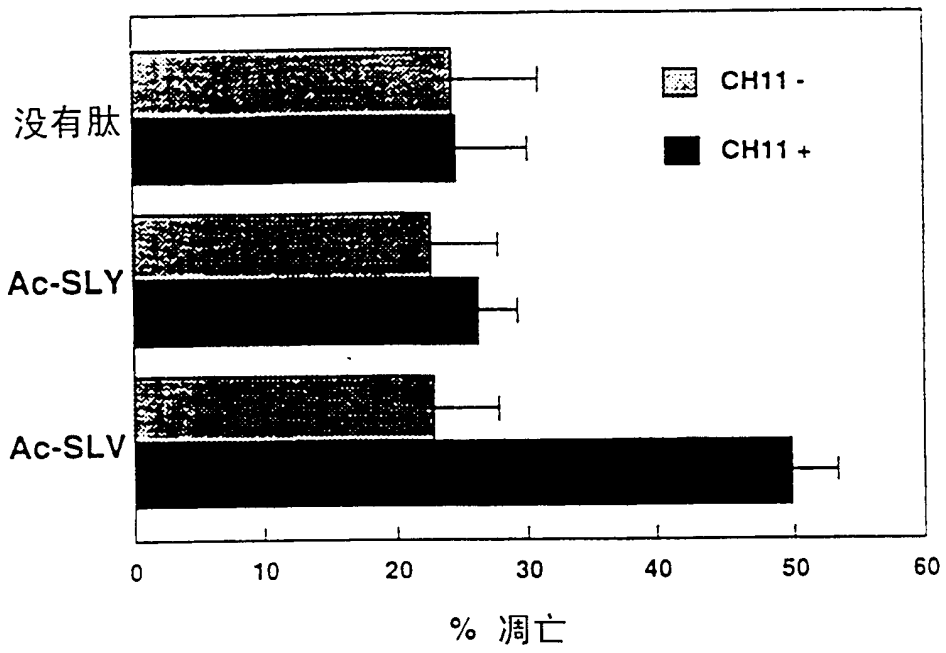


图 6

## NGF 受体

1 mgagatgram dgprlllllll lgvslggake acptglyths gecckacnlg egvaqpcgan  
 61 qtvcepclds vtfsdvvsat epckpctecv glqmsapcv eaddavcrca ygyyqdettg  
 121 rceacrveea gsglvfscqd kqntvceecp dgtysdeanh vdpclpctvc edterqlrec  
 181 trwadaecee ipgrwitrst ppegdsstap stqepeappe qdliastvag vvtvmgssq  
 241 pvvtrgttdn lipvycsila avvvlvayi afkrwnsckq nkqgansrpv nqtpppegek  
 301 lhsdsgisvd sqslhdqqph tqtasgqalk gdgglysslp pakreevekl lngsagdtwr  
 361 hlagelgyqp ehidsfthea cpvrallasw atqdsatlda llaalrriqr adlveslcse  
 421 statspv

图 7A

## CD4 受体

1 mnrgvpfrhl llvlqlallp aatqgkkvvl gkkgdtvelt ctasqkksiq fhwknsnqik  
 61 ilgnqgsflt kgpsklndra dsrrslwdqg nfpliiknlk iedsdtyice vedqkeevql  
 121 lvfgltansd thllqgqslt ltlesppgss psvqcrsprg kniqggkttls vsqlelqdsq  
 181 twtctvlqng kkvefkidiv vlafqkassi vykkegeqve fsfplafte kltgsgelww  
 241 qaerasssks witfdlknke vskrrvtqdp klqmgkklpl htlpqaalp yagsgnltla  
 301 leaktgklhq evnlvumrat qlqknltecew wgptspklml sklkenkeak vskrekavwv  
 361 lnpeagmwqc llsdsgqvll esnikvlptw stpvqpmali vlggvaglll figlgiffcv  
 421 rcrhrrrqaе rmsqikrlls ekkctcqcphr fqktcspi

图 7B

---

种类	C-末端序列 NGFR (p75)	FAP-1 的结合活性
人	tSESTATSPV-COOH	+
大鼠	tSESTATSPV-COOH	+
鸡	tSESTATSPV-COOH	+

图 7C

1 mnsqvamkyg ndssaelsel hsaalaslkg divelnkrlq qtererdlle kklakaqceq  
 61 shlmrehedv qerttlryee ritelhsvia elnkkidrlq gttireedey selrselsqs  
 121 qhevnedrsr mdqdtsvsi penqstmvta dmdncsdlns elqrvltgle nvcgrkks  
 181 cslsvaevdr hieqlttase hcdlaiktve eiegvlgrdl ypnlaeersr wekelaglr  
 241 enesltamc skeelnrtk atmnaireer dlrllrvrel qtrlqsvqat gpsspgrlts  
 301 tnrlpinpstg elstssssnd ipiakiaerv klktrsess ssdrpvlgse issigvsssv  
 361 aehlahslqd csniqEIFQT lyshgsaise skirefevet erlnsriehl ksandlltit  
 421 leecksnaer msmlvgkyes natalrlalq yseqcieaye lllalaeseq slilgqfraa  
 481 gvgsspgdqs gdenitqmlk rahdcrktae naakallmkl dgscggafav agcsvgpwes  
 541 lssnshtstt sstasscdte ftkedeqrk dyiqqlkndr aavkltmlle esihidplys  
 601 dvkprgdsqr ldlenavlmq elmamkeema elkaqlylle kekalekl streaqeqay  
 661 lvhiehlkse veeqeqrnr slsstssgsk dkpgkecada aspalslael rttsenela  
 721 aeftnairre kklkarvqel vsalerltnks seirhqqsae fvndlkrans nlvaayekak  
 781 kkhqnlkkl esqmmamver hetqvrmlkq rialleeens rphtnetsl

图 7D

1 madvfpngds tasqdvannr arkgalrqkn vhevkdhkfi arffkqptfc shctdfiwgfg  
 61 gkqgfcqvc cfvvhkrche fvtfscpgad kqcvinpsl kqcvinpsl kqcvinpsl kqcvinpsl  
 121 llyglihqgm kcdtcdmnhv ipdpkneskq kktirstln eegeyyynvpi pegdeegnme  
 181 knlipmdpng lsdpyvklkl fgvselmkmv asgwykllnq flmvlgkgsf gkvmladrkg  
 241 sveiwdwdrt trndfmgsls gpagnkvisp vect:mvekrv lalldkppfl drlyfvmeyv  
 301 lrqkfeakl kkdvviqddd vfyaaeisig lfflhkrzii yrdlklndvm ldseghikia  
 361 teelyaikil qvqkfkepqa tpdyaiepii ayqpygksvd wwaygvllye mlagqppfdg  
 421 nggdlmyhiq dgttrtfcg lskavsvick glmtkhpakr lgcgpegerd vrehaffrri  
 481 dfgmckehmm mehnvsypks kgaenfdkff trgqpvltpp dqlvianidq sdfegfsyvn  
 541 ededelqasi appfkpkvcg  
 601 dweklenrei pppfkpkvcg  
 661 pqfvhpilqs\_av

图 7E



1 mdilceents lssttnslmq lnddtrlysn dfnsgeants dafnwtvdse nrtnlscegc  
61 lspscslsh lqeknwsall tavviiltia gnilvimavs lekklqnatn yflmslaiad  
121 mllgflvmpv smltilygyr wplpsklcav wiyldvlfst asimhlcais ldryvaiqnp  
181 ihhsrfrnsrt kafliiavw tisvgismpi pvfqlqddsk vfkegsclla ddnfvligsf  
241 vsffipltim vityfltikslqeatlcvs dlgttraklas fsflpqssls seklfqrsh  
301 repgsytgrr tmqsisneqk ackvlgivff lfvvmwcpff itnimavick escnedviga  
361 llnvfwwigy lssavnplvy tlfnktyrsa fsryiqcayk enkkplqlil vntipalayk  
421 ssqlqmgqkk nskqadakttd ndcsmvalgk qhseeaskdn sdgvnekvsc\_v

图 7F

1 malsyrvsel qstipehilq stfvhvisn wsglqtesip eemkqiveeq gnklhwaall  
 61 ilmviptig gntlvilavs lekklqyatn yflmslavad llvglfvmpi alltimfeam  
 121 wplplvlcpa wlfldvlfst asimhlcais vdryiaikkp iqanqynsra tafikityvw  
 181 lisigiaipv pikgietdvd npnnitcvlt kerfgdfmlf gslaafftpl aimivtyflt  
 241 ihalqkayl vknkppqrlt wltvstvfqr detpcsspek vamlmgrkd kalpnsqdet  
 301 lmrrtstigm ksvqtisneq raskvlgiv flflmwcpf fitnitvlc dscnqttlqm  
 361 lleifwigy vssgvnplvy tlfnktrfrda fgryitcnyr atksvktlrk rsskiyfrnp  
 421 maenskkfk hgirnginpa myqspmlrs stiqssiil ldtllltene gdkteeqvsv  
 481 v

图 7G

1 maaasydall kavealkmen snlraeledn snhltklete asnmkevlkq lagsiedeam  
 61 assgaidlle rlkelndss nfpvgklrsk mslrsygsre gsvssrsgec spvpmsgfpr  
 121 rgfvngsres tgyleeleke rsllladldk eekekdwyya qlqnltkrid slpltenfsl  
 181 atdmtrrale yearairvam eeqlgtcqd m ekraqrriar iqqiekdilr irqllasqat  
 241 eaerssankh etgshdaera negqgvgein matsgngqgs ttrmdhetas vlssssthsa  
 301 prrltshlgt kvemvyslls mlgthdkddm srllamsss qdscismraq gclplliqll  
 361 hgndkdsvll gnsrgskear arasaalhui ihsqddkrg rreirvlhll eqiraycetc  
 421 wewqeahepg mdadknpmpa pvehqicpav cvlmklsfde ehrhamnelg glqaiarella  
 481 vacemygltn dhysitlrry agmaltnltf gdvankatlc smkgcmralv aqlksesedl  
 541 aqviasvlrn lswradvnsk ktlrevgsvk almecalevk kestlksvls alwnlsahct  
 601 enkadicava galafvlgtl tyrsqntla iiesgggilr nvssliatne dhrqilrenn  
 661 clatllahlk shsltivnsa cgtlwnlsar npkdqealwd mgavsmknl ihshkhmiam  
 721 gsaaalrnlm anrpakykda nimspgsslp slhvrkqkal eaeldaqls etfdnidnl  
 781 pkashrskqr hkaslygdyv fdtnrhddnr sdnfntgnmt vlspylnttv lpsssssrgrs  
 841 ldsrsekdr slerergigl gnyhpatenp gtsskrglai sttaaqiakv meevsaihts  
 901 aedrsvgstt elhcvtdern alrrssaht hsntynftks ensnrtsmp yakleykrss  
 961 ndslngsvss dgygkrqamk psiesysedd eskfcsygay padlahkihs anhmddndge  
 1021 ldtpinyslk ysdealnsgr aspsanerwa rpkhiiede kaseqrqsrn qsttypvyte  
 1081 stddkhlkfq phfgaaecvs pyrsrgangs etnrvgsnhg inqnvqsalc qeddyeddkp  
 1141 tnyseryseee eqheeeeerpt nysikyneek rhvdqpidys lkyatdipss qkqsfssks  
 1201 ssgqsskteh mssssentst pssnakraq lhpssaqsrs gqpkaatck vssinaeti  
 1261 tyvedtpic fsrccslssl ssaedeign qttqeadsan tlqiaeiekek igtrsaedpv  
 1321 sevpavsqhp rtkssrlqgs slssesarhk avefssgaks psksqaatpk sppehyvqet  
 1381 plmfsrctsv ssldsfers iassvqsepc sgmvsigiisp sdldpspgat mpsrsktpp  
 1441 pppqtaatkr evpknkapt ekresgpkqa avnaavarva vlpdadtlh fatestpdgf  
 1501 scssslsals ldepfiaqdv elrimppvqe nangnetese qpkesenqe keaektidse  
 1561 kdllddsddd dieileeci samptkssrk akkpaatask lpppvarkps qlpvyklps  
 1621 anrlaqkhv sftpgadmp vycvegtpin fstatslsl tiesppnela agegvrsgaa  
 1681 sgefekrdti ptegrstdea aggktsviti pelddnkaee gdilaecins ampkgkshk  
 1741 frvkkimaqv qaasassap nknaldgkk ktpspvkip anteyrtrvr knaasknln  
 1801 aervfsdndk skkanlkns kdfndklpn edrvrgsfaf dspghytpie gtpycfsrnd  
 1861 slssldfddd dvalrekae lrkakenkes eakvtshel tsnqaqsant qaiakapin  
 1921 gqpkpilakq stfpqsskdi pargaatdek lanfaietp vcfshnssl slsdiqenn  
 1981 nkenepiket eppdsqgeps kpqasgyapk sfhvedtpvc fsrnsslssl sidsedllq  
 2041 ecissampkk kkpsrlkgdn ekhsprnmgg ilgedltldl kdiqrpseh glspasenfd  
 2101 wkaiqegans ivsslhaaa aalsrqass dsdsilsks gislgspfh tpdqeeqpf  
 2161 snkgprilpk gekstletk ieseskgik gkkvykslit gkvrnses gamkapian  
 2221 mpsisrgrtm ihipgvrnss sstspvskg pplktpask psegatatts prgakpsvks  
 2281 elspvarqts qiggsskaps rgsrdstps rpaqqlsrp iaspgrnsis pgrngisppn  
 2341 klsalprtss pstastkssg sgkmsytspg ramsqaalk atglsknass iprsesaskg  
 2401 lnqmnngnga nkkvelsrms stkssgsed rserpvlvrq stfikeapsp tlrrkleesa  
 2461 sfeslpsrr pasptrsqaq tpvlspslpd mslsthssvq aggrwklppn lsptieyndg  
 2521 rpakrhdiar shsespsrlp inrsgtwkre hskhssslpr vstwrwtgss ssilsasses  
 2581 sekaksedek hvnsisgtkq skenqvsakg twrkikenef sptnstsqtv ssgatngaes  
 2641 ktliyqmapa vsktedvwr iedcpinnpr sgrsptgntp pvidsvseka npnikdskdn  
 2701 aakqnvngs vpmrtvglen rlnsfiqva pdqkgteikp gqnpvpvse tnessivert  
 2761 pfsssssskh sspsgtvaar vtpfnynp rkssadstsa rpsqiptpvn nntkkrdskt  
 2821 dstessgtas pkrhsgsylv tsv

图 7H

# p75<sup>NGFR</sup>

(低亲和性神经生长因子受体)

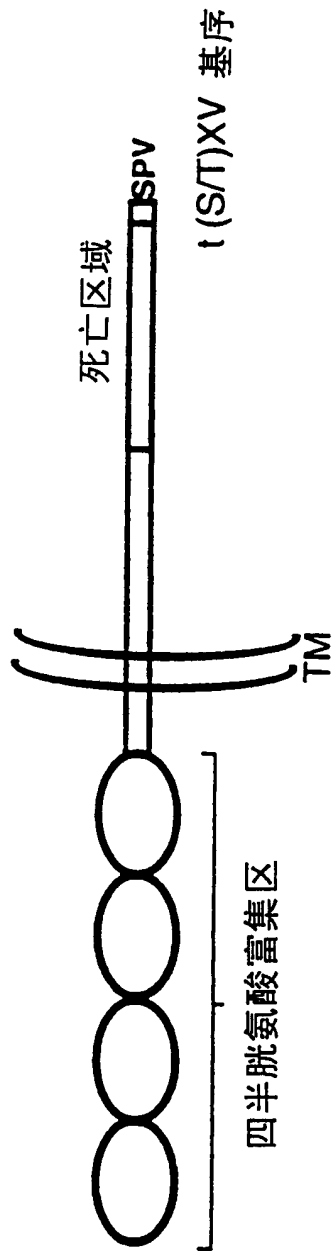
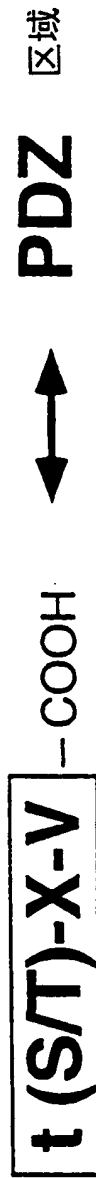


图 8

	C-末端氨基酸序列
Fas	NEIQSLV
p75NGFR	STATSPV



相互作用

图 9

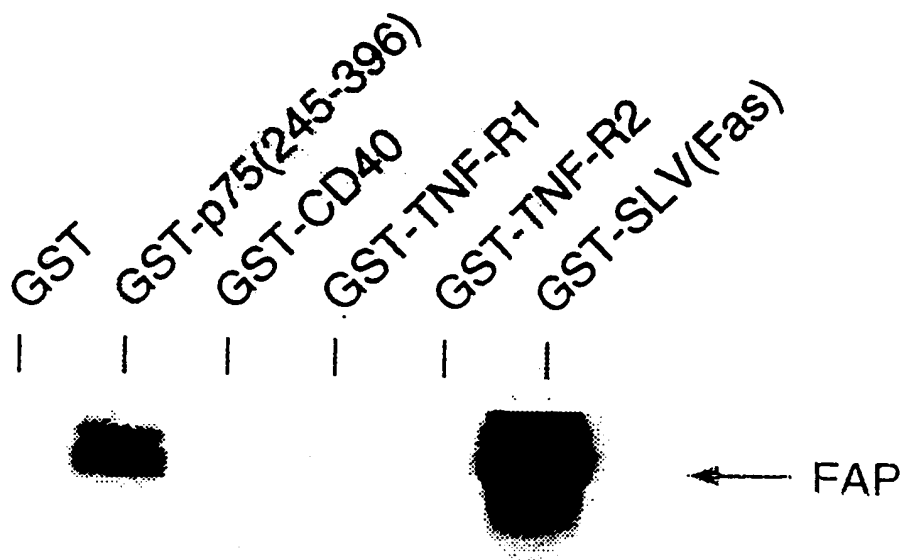


图 10

FAP-1 结合至 p75NGFR 的 C-末端 3 个氨基酸 SPV

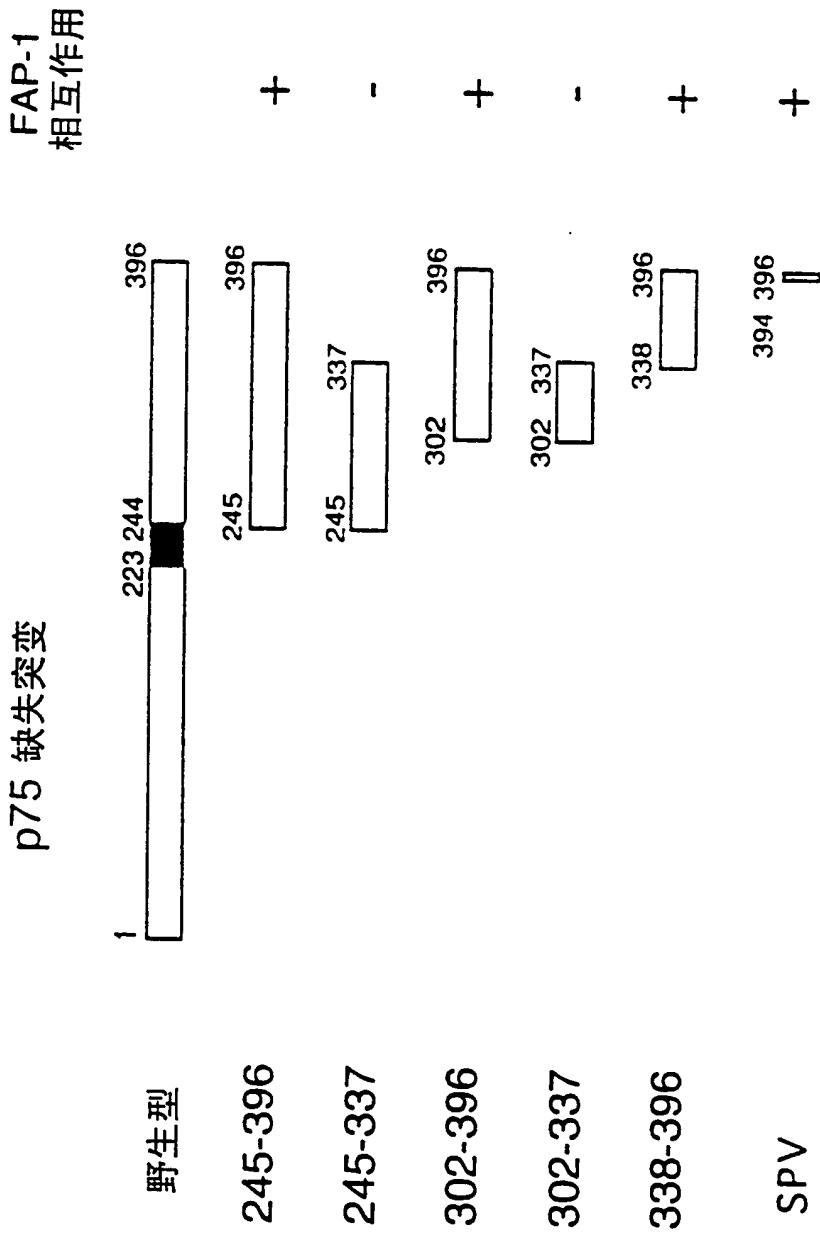


图 11A

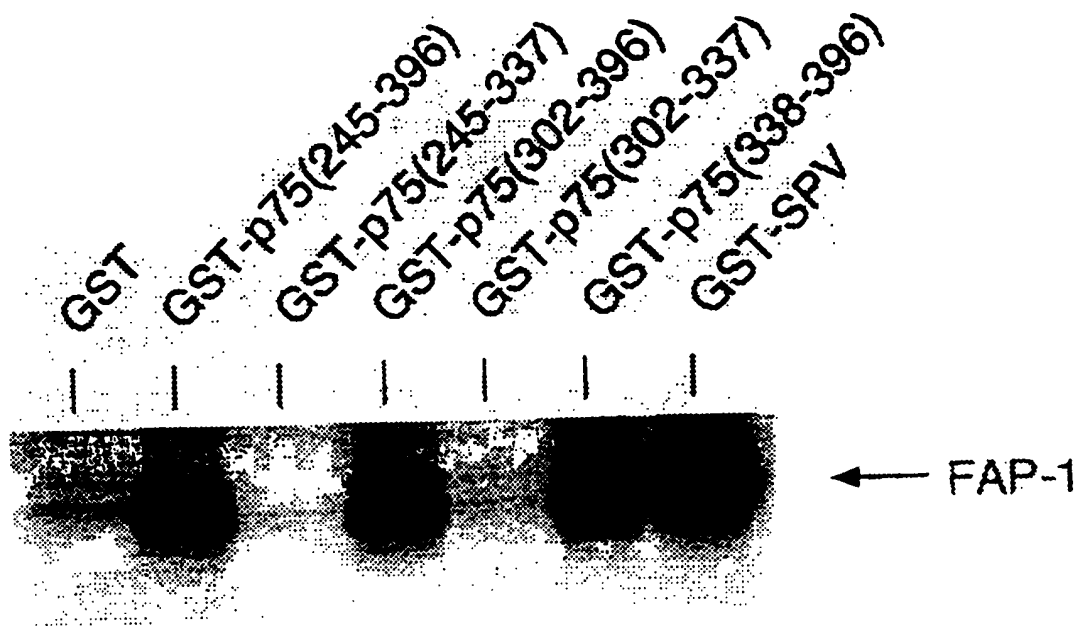


图 11B



在酵母中 FAP-1 结合至 p75NGFR C-末端胞浆区域

	VP16-FAP-1	VP16-cRaf
LexA-p75NGFR(338-396)	+	-
LexA-p75NGFR(365-396)	+	-
LexA-Fas	++	-
LexA-Ras <sup>V12</sup>	-	+
LexA-Lamin	-	-

图 12